Politecnico di Torino

Politecnico di Torino

Dipartimento di Automatica e Informatica
Corso di Studi in Ingegneria del Cinema e dei Mezzi di Comunicazione
Sessione di Laurea luglio 2025

IA e Arte: narrare al pubblico l'impatto emotivo delle performance musicali attraverso video generati con strumenti di intelligenza artificiale generativa.

Tesi di Laurea Magistrale

Relatori:
Prof. Andrea Bottino
Prof.ssa Tatiana Mazali

Candidato:
Mattia Morgut

Anno Accademico 2024/2025

Abstract

Nel panorama contemporaneo, le tecnologie di intelligenza artificiale stanno progressivamente ridefinendo i processi creativi. Il lavoro di tesi esplora il potenziale dell'intelligenza artificiale come agente co-creativo nell'ambito audiovisivo, con l'obiettivo primario di generare un cortometraggio che traduca l'impatto emotivo del pubblico, dopo l'ascolto di performance musicali, in una narrazione visiva. Il progetto si inserisce all'interno di *Narrazioni Parallele*, iniziativa culturale dedicata all'indagine sul rapporto tra intelligenza naturale e artificiale, a partire da contesti musicali.

In particolare, il lavoro svolto propone un flusso produttivo completamente basato su IA, utilizzando i dati provenienti dal pubblico (emozioni, ricordi, immagini mentali) in ogni fase della pre-produzione, dall'ideazione alla formulazione di logline, sinossi, sceneggiatura e storyboard. Nella fase di produzione, un'attenta attività di studio ha permesso la scrittura e l'ottimizzazione dei prompt per ogni singola clip. Sono poi stati generati i video, revisionati manualmente per qualità grafica e fedeltà narrativa. Infine, nella fase di post-produzione, si è proceduto con il montaggio e la sincronizzazione del voiceover.

Attraverso lo sviluppo e l'implementazione di questo flusso, l'elaborato valuta la capacità degli strumenti odierni di mantenere coerenza narrativa e stilistica, affrontare le attuali limitazioni tecnologiche e supportare la creatività umana. La validazione è stata effettuata tramite una proiezione pubblica del prodotto audiovisivo e la successiva analisi dei riscontri ricevuti, che hanno permesso di valutare l'efficacia del video prodotto, attraverso differenti parametri. Il progetto ha affrontato alcune delle attuali principali criticità legate all'uso dell'intelligenza artificiale nella produzione audiovisiva, ad esempio la discontinuità narrativa, dovuta alla frammentazione delle clip, l'incoerenza visiva nella gestione delle immagini della protagonista e i limiti strutturali dei modelli video, spesso incapaci di gestire sequenze professionalmente articolate. Attraverso un processo iterativo di supervisione umana e raffinamento dei prompt, questi aspetti negativi sono stati progressivamente mitigati, ottenendo un risultato soddisfacente.

Il lavoro si configura come un esempio concreto di sperimentazione narrativa con IA in un contesto performativo reale, evidenziando sia le potenzialità di questa tecnologia che i necessari margini di intervento umano per garantire qualità espressiva e una maggiore aderenza agli intenti autoriali. Un importante aspetto esaminato è il rapporto tra autore umano e intelligenza artificiale, fulcro metodologico del progetto: un'interazione continua che ha potenziato l'efficienza e l'accesso alla complessità visiva grazie all'IA, mentre coerenza narrativa, sensibilità estetica e profondità espressiva sono state garantite dalla supervisione umana. Il risultato finale è un flusso di lavoro ibrido, scalabile e adattabile, capace di generare contenuti originali che mantengono integrità emotiva e l'intenzionalità autoriale.

Indice

1 Introduzione **1**
- 1.1 Contesto e motivazione del progetto . . . 1
  - 1.1.1 L'evoluzione dell'IA in campo creativo . . . 1
  - 1.1.2 Obiettivi del progetto . . . 2
  - 1.1.3 Associazione LUCAS e progetto Narrazioni Parallele . . . 2
  - 1.1.4 Tipologia di eventi per la raccolta dei dati e presentazione del prodotto audiovisivo . . . 3
- 1.2 Definizione del problema . . . 4
  - 1.2.1 Difficoltà nella generazione di un video con IA . . . 5
  - 1.2.2 Esigenza di una pipeline automatizzata efficace e riproducibile 6
  - 1.2.3 Domande di ricerca . . . 6
- 1.3 Approccio e metodologia adottata . . . 7
  - 1.3.1 Sviluppo di una pipeline interamente basata su IA . . . 7
  - 1.3.2 Bilanciamento tra automazione e intervento umano . . . 8

2 Stato dell'Arte **10**
- 2.1 L'intelligenza artificiale nella produzione audiovisiva . . . 10
  - 2.1.1 Settori applicativi di maggior impiego . . . 10
  - 2.1.2 Esempi di cortometraggi e film generati con IA . . . 11
  - 2.1.3 IA e stile visivo . . . 12
  - 2.1.4 Impatto della tecnologia nelle arti visive . . . 13
- 2.2 Analisi dei festival e delle iniziative legate all'IA . . . 13
  - 2.2.1 Festival dedicati alla creatività con IA . . . 13
  - 2.2.2 Progetti video rilevanti generati con IA . . . 15
- 2.3 Benchmark degli strumenti IA . . . 16
  - 2.3.1 Strumenti IA per sceneggiatura, regia e produzione . . . 16
  - 2.3.2 Benchmark . . . 17
  - 2.3.3 Modelli text-to-video e image-to-video . . . 19
  - 2.3.4 IA per la generazione di testo . . . 21
  - 2.3.5 IA per la generazione di immagini . . . 23
  - 2.3.6 IA per la generazione di video . . . 26
  - 2.3.7 IA per la generazione di audio . . . 29
  - 2.3.8 IA per la post-produzione . . . 31
- 2.4 Analisi SWOT delle soluzioni IA disponibili . . . 31
  - 2.4.1 Punti di forza . . . 31
  - 2.4.2 Punti di debolezza . . . 32
  - 2.4.3 Opportunità . . . 33
  - 2.4.4 Minacce . . . 35

3 Pipeline di produzione del progetto **37**
3.1 Fase di Pre-produzione . . . 37
3.1.1 Raccolta dati . . . 37
3.1.2 Preparazione dei dati raccolti . . . 38
3.1.3 Analisi dei dati raccolti . . . 39
3.1.4 Elaborazione dei dati . . . 40
3.1.5 Definizione di logline, sinossi e soggetto . . . 42
3.1.6 Tecniche utilizzate nella fase di ideazione . . . 42
3.1.7 Modalità operative di generazione della logline . . . 42
3.1.8 Modalità operative di generazione della sinossi . . . 44
3.1.9 Modalità operative di generazione del soggetto . . . 48
3.1.10 Integrazione dei dati nei concept narrativi . . . 50
3.1.11 Conclusioni sulla fase di ideazione . . . 51
3.1.12 Modalità operative di generazione della sceneggiatura . . . 52
3.1.13 Creazione dello storyboard con IA . . . 53
3.1.14 Creazione delle immagini per lo storyboard con IA . . . 59
3.1.15 Creazione del modello visivo della protagonista . . . 66
3.1.16 Definizione dello stile visivo e sonoro . . . 69
3.1.17 Conclusione della fase di Pre-produzione . . . 70
3.2 Fase di Produzione . . . 71
3.2.1 Studio preliminare per la creazione dei prompt . . . 71
3.2.2 Verifica riferimenti dati narrativi . . . 75
3.2.3 Generazione dei prompt per la produzione video . . . 77
3.2.4 Modifica dei prompt generati . . . 81
3.2.5 Generazione delle clip video . . . 82
3.2.6 Modifiche in seguito alla generazione . . . 83
3.2.7 Creazione delle transizioni . . . 85
3.2.8 Generazione della colonna sonora . . . 87
3.2.9 Conclusione della fase di Produzione . . . 89
3.3 Fase di Post-produzione . . . 90
3.3.1 Montaggio e rifinitura delle sequenze . . . 90
3.3.2 Ottimizzazione del colore e della luce . . . 91
3.3.3 Generazione e sincronizzazione del voiceover . . . 92
3.3.4 Generazione degli effetti sonori . . . 94
3.3.5 Scelta del titolo . . . 95
3.3.6 Revisione del video finale . . . 96
3.3.7 Conclusione della fase di Post-produzione . . . 97
3.4 Sviluppo di un modello di IA per la generazione di prompt per video 97
3.4.1 Problema e motivazioni . . . 97
3.4.2 Creazione dei dataset . . . 98
3.4.3 Fine-tuning del modello di IA . . . 99
3.4.4 Utilizzo all'interno della pipeline . . . 100
3.4.5 Risultati del modello addestrato . . . 102

4 Risultati e conclusioni **106**
4.1 Valutazione dell'efficacia della pipeline con IA . . . 106
4.1.1 Confronto con i metodi di produzione tradizionali . . . 106
4.1.2 Punti di forza e limiti del processo . . . 107

4.1.3 Problemi di coerenza visiva e narrativa . . . 109
4.1.4 Affrontare i problemi di coerenza . . . 113
4.1.5 Problemi di generazione con IA . . . 116
4.1.6 Relazione tra prompt e output . . . 118
4.2 Valutazione dell'estetica generativa . . . 119
4.2.1 Estetica e forma visuale . . . 119
4.2.2 Estetica dell'uncanny valley . . . 120
4.2.3 Lo stile estetico dell'IA . . . 122
4.2.4 Estetica del cortometraggio generato . . . 123
4.2.5 Il carattere didascalico del cortometraggio generato . . . 134
4.2.6 L'emotività nel cortometraggio generato . . . 135
4.2.7 Linguaggio e intenzionalità dell'IA come autore . . . 136
4.2.8 Conclusione . . . 137
4.3 Impatto del video . . . 138
4.3.1 Feedback e reazioni del pubblico . . . 138
4.3.2 Analisi dell'efficacia comunicativa del cortometraggio . . . 141
4.3.3 Corrispondenza tra intenti progettuali e risultati . . . 142
4.3.4 Riflessione critica sull'impatto del cortometraggio . . . 144
4.3.5 Differenze tra aspettativa e ricezione reale . . . 147
4.4 Considerazioni e possibili sviluppi . . . 148
4.4.1 Miglioramenti possibili della pipeline con IA . . . 148
4.4.2 Potenzialità dell'IA per la produzione audiovisiva . . . 150
4.4.3 Implicazioni etiche e artistiche dell'uso di IA nel cinema . . . 152

Appendice **154**
A Domande poste al pubblico . . . 154
B Grafici dei dati raccolti . . . 155
C Dati elaborati . . . 158
D Logline e sinossi . . . 171
E Soggetto cinematografico . . . 171
F Sceneggiatura . . . 173
G Inquadrature in inglese . . . 176
H Storyboard . . . 183
I Tabelle con associazioni dei dati . . . 217
J Prompt generati . . . 220
K Cortometraggio . . . 231
L Feedback del pubblico . . . 232
M Dataset per fine-tuning . . . 236
N Codice per fine-tuning . . . 237

Bibliografia **246**

Capitolo 1

Introduzione

1.1 Contesto e motivazione del progetto

1.1.1 L'evoluzione dell'IA in campo creativo

Negli ultimi anni, l'intelligenza artificiale ha compiuto enormi progressi, passando dall'essere uno strumento puramente analitico e di nicchia fino a diventare un mezzo utilizzato quotidianamente.
In particolare, nel settore audiovisivo, l'IA sta iniziando a essere impiegata non solo per migliorare e snellire i processi già esistenti ma anche per generare nuovi contenuti, sollevando domande fondamentali sul ruolo dell'autore umano e sulle possibilità di automazione della narrazione. Questa rivoluzione apre scenari entusiasmanti ma anche sfide significative, sia dal punto di vista tecnologico che etico, rendendo essenziale un'analisi approfondita sulle potenzialità e sui limiti dell'intelligenza artificiale nella produzione di un contenuto audiovisivo.
L'IA generativa è passata dai primi modelli probabilistici degli anni '50 fino alle moderne architetture neurali capaci di creare contenuti sia testuali che visivi. Il punto di svolta fondamentale è arrivato nel 2014 grazie a Goodfellow et al. [1], con l'introduzione delle *Generative Adversarial Networks*, che hanno permesso per prime di generare immagini realistiche. Nel 2016, il primo esperimento significativo di sceneggiatura interamente generata da IA è stato il cortometraggio *Sunspring* [2], mentre i primi tentativi di creazione video appartengono a Vondrick et al. [3].
Nonostante l'IA abbia dimostrato nel tempo di poter produrre immagini, testi e video in maniera sempre più realistica, l'integrazione di questi strumenti in una pipeline di produzione completa richiede un certo equilibrio tra l'automazione e il controllo creativo umano. Una delle principali difficoltà è garantire la coerenza narrativa e stilistica, in quanto il processo di generazione può portare a contenuti disomogenei da un punto di vista grafico e narrativo. Allo stesso tempo, però, un approccio automatizzato offre la possibilità di accelerare a dismisura il processo creativo e di produzione: i modelli sono oramai in grado di generare molto rapidamente quanto viene richiesto e il livello di precisione raggiunto migliora ad ogni nuova versione. L'evoluzione degli strumenti di IA porterà certamente a una rivoluzione in diversi ambiti, interessante da capire se verso una democratizzazione della produzione video, intesa come prodotto accessibile a tutti, oppure verso la creazione di nuove barriere tecnologiche.

1.1.2 Obiettivi del progetto

A differenza di altre realtà audiovisive, che utilizzano l'intelligenza artificiale solamente in alcune fasi specifiche, questo progetto si propone di elaborare l'intero processo, dalla raccolta dei dati iniziali, passando per la pre-produzione (ideazione e pianificazione), produzione (ripresa ed esecuzione) e post-produzione (montaggio e finalizzazione).
L'importanza di questo studio si inserisce anche in un contesto più ampio. L'utilizzo dell'intelligenza artificiale infatti, non è solo una questione tecnica, legata alla scelta e utilizzo degli strumenti più adatti, ma solleva molti interrogativi sull'originalità, sulla creatività e sul ruolo dell'autore umano.
Il progetto nasce dall'esigenza di esplorare il ruolo dell'intelligenza artificiale nella creazione di contenuti visivi, con l'obiettivo primario di sviluppare una pipeline di produzione interamente basata su IA. Uno degli obiettivi principali del progetto è definire strategie avanzate per l'acquisizione e l'elaborazione dei dati provenienti dall'interazione del pubblico durante eventi musicali.
Un altro scopo è l'ottimizzazione dei processi produttivi di una pipeline di lavoro classica mediante l'utilizzo dell'IA. Dal punto di vista della produzione creativa, il progetto mira a sviluppare una struttura bilanciata, capace di garantire un output audiovisivo che rispetti assolutamente dei vincoli estetici e narrativi definiti, pur sempre mantenendo un equilibrio tra generazione autonoma e controllo artistico umano. L'obiettivo è quindi quello di ottenere un'interazione efficace tra intelligenza artificiale e creatività umana, esplorando nuove modalità di espressione visiva basate sull'IA. Infine, il progetto punta a stabilire delle metriche di valutazione oggettive per misurare l'efficacia comunicativa dei contenuti generati. Attraverso l'analisi del feedback del pubblico, vengono testati parametri come la capacità del video di suscitare emozioni e l'attitudine a mantenere una coerenza narrativa, adattandosi al pubblico e ai vari contesti di fruizione.

1.1.3 Associazione LUCAS e progetto Narrazioni Parallele

Il lavoro è stato sviluppato all'interno dell'iniziativa *Narrazioni Parallele*, promossa dall'associazione *LUCAS*.
LUCAS (Libera Unione di Cultura Arte e Spettacolo) [4] è un'associazione culturale fondata nel 2002 da Roberto Beccaria, con l'obiettivo di sviluppare strategie culturali per la valorizzazione e lo sviluppo sostenibile del territorio, attraverso l'innovazione culturale e la creazione di reti ed sistemi culturali. Tra i progetti promossi da *LUCAS* vi è *Narrazioni Parallele*, un'iniziativa triennale finanziata dal bando *Torino, che cultura!* della Città di Torino.
Il progetto nasce in collaborazione con l'*Accademia Corale Stefano Tempia* e l'associazione *Cultura e Salute* e si sviluppa in partnership con la *Filarmonica TRT*, del Teatro Regio di Torino, le Biblioteche Civiche Torinesi e il Politecnico di Torino.
L'obiettivo dell'iniziativa è coinvolgere un pubblico intergenerazionale attraverso due direttrici, idealmente parallele: la prima è quella della Narrazione-Partecipazione, che si realizza all'interno delle biblioteche civiche tramite concerti e laboratori capaci di attivare un'interazione diretta tra artisti e pubblico; la seconda è quella della Narrazione-Contemporaneità, volta a connettere la musica più tradizionale con quella elettronica, attraverso linguaggi innovativi e tecnologie emergenti. In questo

secondo ambito trovano spazio giovani compositori e compositrici accuratamente selezionati, i cui lavori vengono presentati in concerti speciali dedicati.
Elemento di congiunzione tra queste due dimensioni parallele è l'uso dell'intelligenza artificiale, che diventa un ponte tra partecipazione collettiva e sperimentazione artistica. In particolare, attraverso l'integrazione dell'IA nei processi di raccolta, analisi e trasformazione dei dati provenienti dagli eventi dal vivo, il progetto esplora nuove forme di creazione culturale partecipata, in cui il pubblico contribuisce direttamente all'elaborazione del contenuto creativo finale.
L'associazione *LUCAS* ha ricoperto un ruolo centrale nel progetto di tesi, offrendo supporto concreto in ogni fase. In primo luogo, ha curato l'organizzazione degli eventi pubblici, fornendo così il materiale per l'intera lavorazione attraverso la raccolta dei dati iniziali. Inoltre, l'associazione ha messo a disposizione gli strumenti di intelligenza artificiale, permettendo quindi di elaborare e trasformare i dati in input creativi. *LUCAS* ha promosso e gestito l'evento conclusivo di presentazione del prodotto audiovisivo realizzato, permettendo quindi la raccolta dei feedback, fase fondamentale per la valutazione dell'efficacia del lavoro svolto.

1.1.4 Tipologia di eventi per la raccolta dei dati e presentazione del prodotto audiovisivo

I dati di partenza utilizzati nel progetto sono stati raccolti durante due eventi musicali organizzati da *LUCAS* nell'ambito del progetto *Narrazioni Parallele*:

- **Una passeggiata creativa**: svoltosi il 30 ottobre 2024 presso la Biblioteca Civica Italo Calvino, ha visto la partecipazione del *Quintetto Prestige* della *Filarmonica TRT*.
- **Song Writing: intelligenza naturale generativa**: tenutosi il 22 novembre 2024, sempre presso la Biblioteca Civica Italo Calvino, con la partecipazione dell'artista Arianna Gallo.

Il primo evento è stato condotto dal *Quintetto Prestige* della *Filarmonica TRT*, composto da Federico Giarbella (flauto), Alessandro Cammilli (oboe), Luigi Picatto (clarinetto), Pierluigi Filagna (corno) e Orazio Lodin (fagotto). Il programma ha previsto l'esecuzione della trascrizione per quintetto di fiati, curata da Joachim Linckelmann, di *Quadri di un'esposizione*, composta da Modest Mussorgsky.
Durante il concerto, ma in particolare dopo l'ascolto di tutta l'esibizione, il pubblico è stato invitato a condividere le proprie emozioni, immagini mentali, parole e colori evocati dalla musica. Sono state proposte delle domande, alcune a risposta chiusa, altre a risposta aperta, attraverso la piattaforma *Mentimeter*, strumento utile per la raccolta di dati e la loro visualizzazione in tempo reale.
Questa attività di *passeggiata creativa*, metafora di un percorso immaginario tra quadri sonori, è stata facilitata, per alcuni utenti con difficoltà uditive, dai team delle associazioni *LUCAS* e *Cultura e Salute*, configurandosi come un esercizio di ascolto attivo e restituzione sensoriale collettiva.
Il secondo evento chiave per la raccolta dei dati ha proposto una modalità originale di co-creazione tra artista e pubblico, invertendo le dinamiche tradizionali del concerto. In questo caso, non è stata l'artista a eseguire un brano preesistente, ma il pubblico

stesso ha partecipato attivamente alla scrittura collettiva del testo di una canzone, in un'esperienza ibrida tra performance musicale e laboratorio partecipativo.
Il laboratorio, condotto dalla cantante Arianna Gallo, con l'accompagnamento musicale di Luca Ammazzini (chitarra e loop station) e in collaborazione con l'*Accademia Corale Stefano Tempia*, ha coinvolto i partecipanti in un processo creativo condiviso, facilitato da sessioni di brainstorming e letture collettive. I contributi del pubblico sono stati raccolti in tempo reale e hanno costituito la base testuale di un brano originale, la cui ispirazione tematica è conseguenza delle suggestioni emerse durante la prima parte dell'evento. Inoltre, questi suggerimenti hanno contribuito alla produzione dei dati utilizzati in seguito per la creazione del prodotto audiovisivo con intelligenza artificiale.
Una particolare attenzione è stata dedicata all'accessibilità di entrambe le esperienze, grazie al supporto delle associazioni *APIC* e *MED-EL*. È stato infatti predisposto un sistema di ascolto assistito attraverso cuffie FM e connettività wireless *LOLA*, a beneficio dei partecipanti con disabilità uditive, garantendo così inclusività dell'evento. Come per gli altri appuntamenti di *Narrazioni Parallele*, la facilitazione della partecipazione è stata curata dai team delle associazioni *LUCAS* e *Cultura e Salute*.
I materiali raccolti durante gli eventi sono stati successivamente utilizzati come base di partenza per questo progetto, dando quindi origine a un'opera audiovisiva originale; questi eventi hanno rappresentato, dunque, non solo un momento di fruizione musicale, ma anche un primo atto di co-creazione tra artisti, pubblico e IA.
La presentazione del prodotto audiovisivo finale è avvenuta in un evento organizzato da *LUCAS* a teatro, sempre nell'ambito del progetto *Narrazioni Parallele*. L'evento, dal titolo *Concerto convergente*, si è tenuto il 21 giugno 2025 al Teatro Juvarra e ha rappresentato il momento chiave di un percorso sperimentale in cui musica elettronica e classica, intelligenza artificiale, suono, video e accessibilità si sono fusi in un'unica esperienza performativa. I tre giovani compositori selezionati hanno realizzato opere originali che riflettono la visione ibrida e contemporanea del progetto, combinando linguaggi musicali tradizionali e sperimentali, dando vita a nuove forme di narrazione sonora. Accanto alla musica, anche il lato visivo è stato trattato, rendendo l'esperienza ancora più completa. Il video risultato da questo lavoro è stato infatti proiettato in questa occasione, consentendo la raccolta di feedback del pubblico attraverso strumenti specifici, al fine di valutare l'efficacia e l'impatto del prodotto realizzato.

1.2 Definizione del problema

Creare un video dal nulla, utilizzando l'intelligenza artificiale, è una grande sfida, che richiede studio e sperimentazione. Al giorno d'oggi esistono diversi strumenti che possono essere utilizzati nel processo creativo, ciascuno con le proprie peculiarità e la scelta del mezzo migliore risulta quindi essere fondamentale per la buona riuscita del lavoro. A questo si aggiunge la volontà di utilizzare come partenza, i dati raccolti dal pubblico durante alcuni eventi musicali. Serve dunque trovare una metodologia per permettere all'intelligenza artificiale di comprendere al meglio i dati forniti, utilizzandoli in seguito per realizzare il progetto.

1.2.1 Difficoltà nella generazione di un video con IA

La generazione completa di un video esclusivamente tramite l'utilizzo dell'intelligenza artificiale rappresenta una sfida complessa da diversi punti di vista. Mentre gli strumenti per la generazione di immagini statiche e testi hanno raggiunto dei notevoli livelli di qualità, la generazione di un contenuto video coerente, articolato e narrativamente strutturato richiede una concatenazione di competenze che l'intelligenza artificiale fatica, per il momento, a gestire in autonomia.
Sebbene con gli strumenti attualmente disponibili sia fattibile svolgere le singole fasi classiche della produzione di un contenuto audiovisivo, non si è ancora in grado di automatizzare completamente l'intero processo. Dalla sceneggiatura alla regia, dal montaggio all'integrazione sonora, ogni fase richiede decisioni contestuali e strettamente legate alle attività svolte in precedenza e questa capacità è difficilmente automatizzabile senza una significativa perdita di coerenza e originalità.
Inoltre, la gestione del ritmo narrativo e della progressione emotiva pone ulteriori sfide ai modelli generativi che non riescono ad avere una visione corretta di questi elementi in quanto tendono a lavorare per singole unità, piuttosto che su sequenze più complesse.
Uno dei principali ostacoli nell'utilizzo dell'intelligenza artificiale per la generazione di contenuti audiovisivi risiede nella difficoltà di mantenere una coerenza narrativa e visiva su scala temporale estesa. I modelli generativi attuali riescono a creare, con ottimi risultati, immagini o clip isolate, ma faticano a garantire una continuità semantica e narrativa tra scene successive. Pertanto, le sequenze prodotte possono risultare frammentate, con bruschi salti logici o repentini cambiamenti di inquadratura, che compromettono l'efficacia comunicativa del prodotto.
A livello visivo, invece, persistono problematiche legate alla stabilità degli elementi rappresentati, come volti o movimenti, che spesso variano in modo incoerente all'interno della stessa generazione o ancor di più da un'inquadratura all'altra. Inoltre, l'assenza di una reale comprensione della struttura narrativa elementare da parte dell'IA limita la sua capacità di costruire archi narrativi coerenti, rendendo quindi necessario l'intervento umano, l'unico che può garantire un risultato globale efficace.
Integrare dati reali all'interno di un processo gestito dall'intelligenza artificiale rappresenta una sfida aggiuntiva. I dati raccolti in contesti reali, in questo caso performance musicali dal vivo, sono spesso eterogenei e fortemente contestuali. Tradurre queste informazioni in input significativi per un sistema generativo richiede un processo di verifica, normalizzazione, astrazione e codifica che mantenga però l'essenza del contenuto originario, senza compromettere la leggibilità da parte del modello.
È importante sottolineare che la qualità e la coerenza dei risultati finali dipendono altamente dalla capacità del sistema di IA di interpretare correttamente questi dati, evitando generalizzazioni o letture distorte. Per ovviare a questo problema è necessario sviluppare una metodologia di mediazione tra dato reale e IA, assicurandosi che il contenuto generato mantenga un legame significativo con la realtà da cui trae origine.

1.2.2 Esigenza di una pipeline automatizzata efficace e riproducibile

Nel contesto della produzione audiovisiva basata su intelligenza artificiale, emerge con forza la necessità di sviluppare una sequenza operativa che sia al tempo stesso efficace, modulare e replicabile. Un flusso di lavoro automatizzato consente di sistematizzare le diverse fasi della produzione (in questo caso dalla raccolta dei dati alla sintesi visiva), riducendo l'intervento manuale e aumentando la scalabilità del processo. Tuttavia, bisogna tenere anche conto dell'efficacia della pipeline, misurata non solo in termini di produttività e tempistiche, ma anche nella capacità di mantenere alta la coerenza narrativa e la qualità estetica rispetto agli input iniziali forniti.
La riproducibilità è un altro requisito fondamentale per garantire la validità del processo: una sequenza operativa ben progettata deve essere in grado di restituire risultati consistenti a parità di input. Questo obiettivo è difficile da raggiungere visto il funzionamento interno dei modelli di intelligenza artificiale generativa, incapaci di visione temporale ampia.
Infine, la pipeline proposta deve dare la possibilità di combinare il rigore tecnico alla flessibilità artistica, permettendo di adattarla a diversi contesti di utilizzo, senza compromettere la sua struttura portante.
Questo progetto si inserisce esattamente in questa ricerca, cercando di definire una pipeline di lavoro basata interamente sull'intelligenza artificiale ma che sia allo stesso tempo adattabile, efficiente e capace di produrre risultati qualitativamente elevati in maniera sistematica.

1.2.3 Domande di ricerca

La rapida evoluzione degli strumenti di intelligenza artificiale e il loro sempre maggiore impiego nei processi creativi e produttivi pongono una serie di domande rilevanti, sia a livello tecnico che teorico. Questo progetto si propone di affrontarle attraverso un approccio sperimentale. Si delineano di seguito le principali domande di ricerca che guideranno l'intero percorso progettuale e alle quali si cercherà di fornire una risposta attraverso l'analisi, l'implementazione e la valutazione di una pipeline automatizzata per la produzione audiovisiva basata su IA.
La principale questione da affrontare riguarda l'identificazione e, successivamente, la scelta del migliore strumento di IA attualmente disponibile per la produzione di video. Non bisogna però considerare solamente la fase di generazione delle singole clip, ma anche tutto quello che riguarda la fase di pre-produzione, come ad esempio la scrittura della sceneggiatura, la generazione delle immagini per lo storyboard, fino ad arrivare alla fase di post-produzione, che comprende fasi come il montaggio e l'elaborazione sonora.
Esistono al momento numerose soluzioni specializzate in segmenti specifici della filiera produttiva. Tuttavia, uno degli obiettivi del progetto è valutare la possibilità di integrare questi strumenti all'interno di un flusso operativo il più possibile omogeneo, valutandone compatibilità, performance, qualità dell'output e grado di automazione.
Definire un flusso di lavoro efficace e riproducibile è essenziale per strutturare una produzione audiovisiva completamente basata sull'IA. Si vuole quindi individuare

con precisione una sequenza ottimale di operazioni, strumenti e passaggi che consentano di passare in modo fluido dai dati presi in input (testi, descrizioni e dati emotivi raccolti dal pubblico) alla generazione delle singole clip, fino al montaggio del video finale. Il progetto si propone quindi di testare diversi flussi di lavoro, valutando la loro efficienza, scalabilità e capacità di garantire coerenza tra le fasi produttive. Il confronto tra le diverse soluzioni alternative permetterà di trovare il miglior flusso di lavoro, analizzando per ogni fase i punti critici e definendo in maniera specifica tutti gli accorgimenti necessari a replicarlo in ambiti diversi.
Infine, il progetto esplorerà il bilanciamento tra generazione automatica e intervento umano in ognuna delle fasi di produzione di un contenuto audiovisivo.
Con gli strumenti attualmente disponibili, la difficoltà nel mantenere un'alta coerenza narrativa e stilistica lungo l'intero sviluppo del video è elevata. Questo è dovuto al fatto che i modelli non sono ancora in grado di valutare il progetto nel suo insieme, ma sono invece estremamente abili nelle singole attività.
Una parte fondamentale del progetto analizzerà quindi le possibili strategie per assicurare una maggiore continuità tra le varie inquadrature e sequenze, garantendo uniformità estetica nelle immagini e un arco narrativo coerente. Verranno valutati sia approcci basati sulla scrittura, sia tecniche di condizionamento dei modelli generativi con immagini che possono essere usate come riferimento per la produzione delle singole clip.
La generazione audiovisiva con IA è un campo in rapida crescita ma è ancora fortemente limitata sia dal punto di vista tecnico che concettuale. I modelli attuali mostrano vincoli nella gestione di contenuti temporali lunghi e nella capacità di interpretare contesti complessi. Il progetto identificherà tali limiti in relazione alla pipeline proposta, con il principale obiettivo di delineare strategie correttive e integrative. Verranno esplorati approcci differenti, dall'integrazione di riscontri e correzioni umane nel ciclo generativo, all'impiego di input curati per migliorare la pertinenza e la qualità dei risultati.

1.3 Approccio e metodologia adottata

In questa sezione vengono delineati i principali approcci metodologici adottati nel progetto, con l'obiettivo di offrire una visione d'insieme sul processo che ha portato allo sviluppo del prodotto audiovisivo finale. Ciascuno degli aspetti descritti sarà poi approfondito in modo più sistematico e dettagliato nei capitoli successivi, dove verranno analizzate le singole fasi operative, gli strumenti utilizzati e le scelte progettuali che hanno guidato l'intero flusso di lavoro.

1.3.1 Sviluppo di una pipeline interamente basata su IA

Alla base del progetto si colloca la costruzione di una pipeline produttiva interamente basata su strumenti di intelligenza artificiale. L'obiettivo è quello di testare ed esplorare le reali potenzialità di un sistema automatizzato capace di generare un prodotto audiovisivo completo, partendo da input testuali ed emotivi, passando alla ideazione della sceneggiatura e delle immagini per lo storyboard, fino ad arrivare alla creazione delle singole clip e al montaggio.

Il flusso di lavoro si compone di più fasi interconnesse. In primo luogo abbiamo la pre-produzione dove, attraverso l'intelligenza artificiale generativa, si ottengono principalmente documenti testuali, fondamentali per lo sviluppo del progetto. Si passa poi alla generazione delle singole inquadrature, utilizzando diversi approcci per ottenere una maggiore coerenza. In questa fase è inoltre importante validare il prodotto sia da un punto di vista estetico che tecnico. Infine, si conclude con la fase di post-produzione, dove l'IA non gioca più un ruolo fondamentale ma rimane cruciale per la creazione della colonna sonora e di tutte le voci dei personaggi.
La sperimentazione su questa struttura permette di valutarne la flessibilità, la coerenza narrativa e l'adattabilità in un contesto creativo non standardizzato. In letteratura, sono già presenti pipeline di produzione simili che mirano a sfruttare l'IA per creare contenuti a partire da dati esterni. Ad esempio, si propone di seguito un'immagine tratta dallo studio di Wei et al. [5], che illustra in linea di massima le fasi da seguire per ottenere un prodotto audiovisivo finale generato con l'intelligenza artificiale.

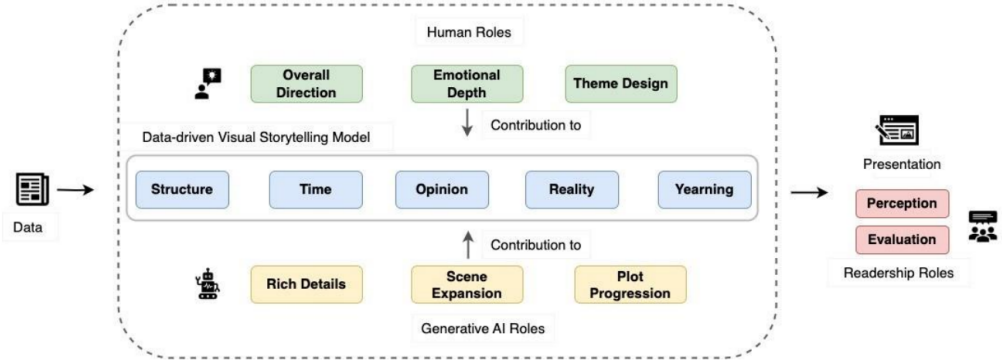


Figura 1.1: Immagine rappresentante il flusso utilizzato durante una ricerca condotta da Wei et al. [5]. Sono evidenziati i ruoli distinti tra umano e IA, suddividendo il processo in diverse fasi intermedie.

La prima fase del progetto ha previsto la raccolta e l'elaborazione di dati provenienti dagli eventi promossi dall'associazione *LUCAS*. Questi, principalmente di natura descrittiva, emozionale e partecipativa, sono stati trasformati in input testuali da utilizzare per la generazione della sceneggiatura del prodotto audiovisivo. Lo scopo del progetto, infatti, è proprio quello di far emergere, attraverso l'IA, narrazioni ispirate alle suggestioni, emozioni e idee provenienti direttamente dal pubblico.
La raccolta dei dati è avvenuta in modo strutturato attraverso l'utilizzo della piattaforma dedicata *Mentimeter* per facilitare la loro integrazione nella pipeline di lavoro, garantendo una base solida su cui costruire il prodotto creativo.

1.3.2 Bilanciamento tra automazione e intervento umano

Un aspetto centrale del progetto ha riguardato il bilanciamento tra l'automazione offerta dall'intelligenza artificiale e il controllo umano all'interno del processo creativo. Il lavoro è stato impostato più come una forma di collaborazione con l'IA, dove la generazione ha rappresentato uno stimolo per l'esplorazione, ma non un vincolo: l'IA non è mai stata lasciata completamente libera in quanto la supervisione umana era sempre presente nel processo creativo.

Nella fase di pre-produzione, la prima attività è stata quella di effettuare un brainstorming guidato che ha permesso di generare differenti opzioni di sviluppo narrativo, dalle quali sono state poi selezionate le idee più interessanti e coerenti con la visione progettuale. Ad una prima idea, particolarmente adatta, è stata poi aggiunta una seconda parte, derivata da un suggerimento completamente differente. Questo approccio ha portato alla definizione di una versione iniziale della sceneggiatura, seguita dallo sviluppo dello storyboard. In questa fase, ogni singola scena veniva dettagliata dall'IA e, successivamente, modificata e perfezionata dall'intervento umano. Per quanto riguarda lo storyboard, ogni inquadratura è stata concepita lasciando che l'IA generasse le immagini sulla base delle indicazioni fornite, intervenendo manualmente dove necessario con modifiche, sostituzioni o correzioni visive, per garantire uniformità stilistica e pertinenza rispetto alla narrazione complessiva. In questo stadio preliminare del progetto non si è data estrema importanza a tale aspetto, considerando come accettabili anche immagini non pienamente coerenti tra loro.
Si è poi passati alla fase di generazione delle singole clip che compongono il prodotto audiovisivo dove l'intervento umano di valutazione e la conseguente richiesta di apportare alcune modifiche sono stati fondamentali per permettere di ottenere esattamente quanto richiesto dal progetto. L'intelligenza artificiale è stata spesso poco precisa, soprattutto nei movimenti e nella generazione di ambienti coerenti e per questo motivo spesso è stato necessario intervenire per modificare e migliorare quanto inizialmente proposto.
Infine, per quanto riguarda la post-produzione, l'IA non è stata molto utile, in quanto non sono ancora disponibili strumenti efficaci per il montaggio. In questa fase l'intervento umano ha permesso di finalizzare il prodotto audiovisivo.
Il risultato finale del lavoro è un cortometraggio generato con il supporto dell'intelligenza artificiale, proiettato durante un evento organizzato dall'associazione *LUCAS*.
Il momento della visione su schermo è servito non solo per restituire al pubblico il lavoro svolto, ma anche per valutare l'efficacia della pipeline proposta. In questa occasione, infatti, è stato possibile raccogliere le opinioni e i feedback qualitativi degli spettatori che, attraverso osservazioni e valutazioni relative alla coerenza narrativa, al valore estetico e alla capacità del prodotto di comunicare emozioni, hanno fornito i dati per trarre le conclusioni finali. In questo modo è stato possibile valutare l'impatto del progetto e le sue potenzialità di evoluzione in futuri contesti simili.

Capitolo 2

Stato dell'Arte

2.1 L'intelligenza artificiale nella produzione audiovisiva

2.1.1 Settori applicativi di maggior impiego

L'intelligenza artificiale è oggi utilizzata in maniera trasversale, ma si osservano con chiarezza almeno quattro ambiti d'impiego maggiormente sfruttati. Il primo è certamente la generazione visiva di clip video attraverso modelli come *Runway*, *Sora* o *Veo*, che traducono prompt testuali in sequenze video. Il secondo è la scrittura narrativa, in cui modelli linguistici come *ChatGPT* o *Claude* sono impiegati per sviluppare soggetti e sceneggiature complete. Il terzo ambito, in forte crescita, riguarda il doppiaggio e la clonazione di voci, che permette di creare dialoghi sintetici ma espressivi. Il quarto, infine, è il montaggio, ancora ampiamente in fase di sviluppo, in cui l'IA assiste o automatizza la selezione e composizione delle inquadrature.
Accanto a questi settori più tecnici si sviluppano approcci concettuali in cui l'IA viene utilizzata anche come elemento narrativo o performativo, introducendo nuove forme di interattività, anche collettiva. Si delinea quindi un panorama in cui il linguaggio audiovisivo è in piena evoluzione, dove l'IA non rappresenta un semplice strumento, ma un nuovo paradigma operativo culturale e diversi prodotti audiovisivi ne sono un chiaro esempio. Questa trasformazione si riflette concretamente nei processi produttivi di numerosi studi e artisti indipendenti, che adottano l'IA non solo per ottimizzare tempi e costi, ma anche per esplorare possibilità creative altrimenti raggiungibili solamente da esperti del settore. La capacità di generare scenari, volti, movimenti e voci a partire da semplici descrizioni testuali sta ridisegnando i confini di ideazione e realizzazione, abbattendo le barriere tradizionali presenti tra pre-produzione e post-produzione. Alcuni cortometraggi mostrano come sia già possibile produrre un intero contenuto audiovisivo senza l'intervento di troupe fisiche o attrezzature convenzionali, affidandosi a una catena creativa interamente basata su IA.

2.1.2 Esempi di cortometraggi e film generati con IA

L'analisi di una selezione di opere contemporanee, studiate in questo progetto di tesi, ha permesso di delineare un quadro critico e aggiornato sullo stato dell'arte nella generazione di video e cortometraggi con l'IA. È stata posta una particolare attenzione ai nuovi linguaggi visivi introdotti, ai processi creativi automatizzati e a tutte le sperimentazioni fatte sia da case di produzione rinomate che da artisti indipendenti.

Tra i primi esempi di produzione generativa completamente italiana si trova *Miss Polly Had a Dolly*, realizzato da Pietro Lafiandra, Andrea Rossini e Flavio Pizzorno [6], in cui sono stati utilizzati *Runway Gen-2*, e *MusicGen* per creare un cortometraggio senza riprese reali. A livello internazionale, invece, *DreadClub: Vampire's Verdict* di Hooroo Jackson [7], rappresenta un esempio importante, essendo il primo lungometraggio interamente animato e prodotto con IA, premiato a diversi festival ma fortemente criticato, come riportato dall'autore stesso. L'utilizzo dell'IA ha coinvolto però anche le maggiori istituzioni cinematografiche: già nel 2016, infatti, la *20th Century Studios*, ha sperimentato un montaggio promozionale interamente automatizzato per il film *Morgan* [8], affidandosi a *IBM Watson* per la selezione delle scene da utilizzare.

In ambito maggiormente sperimentale spicca *Love*, di Francesca Fini [9], performance audiovisiva interamente generata da IA, mentre *What's Next?*, di Cao Yiwen [10], presentato alla *Berlinale*, riflette sulle forme narrative frammentate e allucinate, a partire da immagini generate con IA. *The First Humans* [11] rappresenta invece il primo cortometraggio prodotto interamente utilizzando la piattaforma *Sora* di *OpenAI* e mette in luce alcune significative carenze dello strumento, risultando comunque un buon prodotto audiovisivo finale. Similmente, *The Last Screenwriter* di Peter Luisi [12], interamente scritto da *ChatGPT*, solleva interrogativi cruciali sul ruolo della scrittura automatizzata nella narrazione filmica. Il contributo di figure come Paul Trillo, regista e sceneggiatore americano, è stato fondamentale per la diffusione dell'estetica generata da IA nei videoclip musicali e spot pubblicitari a dimostrazione di come l'IA possa essere impiegata con coerenza formale anche in ambito commerciale. *Promise Studios*, casa di produzione californiana, si impone come una realtà indipendente che lavora esclusivamente con intelligenza artificiale, offrendo non solo opere audiovisive sperimentali ma anche una costante attività divulgativa. L'esperimento collettivo di *Salt*, di Fabian Stelzer [13], è un altro esempio notevole. Si tratta di un film interattivo realizzato con *Midjourney*, *Synthesia* e *DALL·E*, dove la trama si sviluppa attraverso le decisioni del pubblico sui social media. *NeoCinemaAi*, fondato da Jeff Synthesized [14], ha ottenuto un grande riconoscimento per il corto *THE PAPERCLIP MAX - REIMAGINED*, premiato al *Czech AI Film Festival*. Infine, il duo creativo formato da Aaron McGurk e Chris Lapham ha realizzato *Ai P's & Q's* [15], opera più volte premiata, che riflette con intelligenza ed ironia sul futuro della regia.

La visione e l'analisi di queste opere hanno rappresentato una fase essenziale del percorso di ricerca, permettendo non solo di mappare i principali settori in cui si concentra l'utilizzo dell'IA ma anche di comprendere la qualità e le possibilità offerte da questa tecnologia. Particolare attenzione è stata posta nel capire le limitazioni dei modelli e in generale i campi dove i difetti emergono maggiormente. Ciò che spicca dall'osservazione di queste opere è un panorama in rapida e costante evoluzione, in

cui l'intelligenza artificiale non si limita più a supportare i processi produttivi, ma diventa co-autrice dell'intero ciclo creativo. Come spiegato da diversi autori, rimane fondamentale la collaborazione tra uomo e macchina che apre spazi di sperimentazione ancora più ampi e nuovi, sfidando i concetti di autenticità, controllo creativo e partecipazione collettiva.

2.1.3 IA e stile visivo

La recente letteratura accademica ha analizzato in profondità l'estetica visiva dei video generati con intelligenza artificiale in ambito cinematografico, evidenziando vari parametri chiave tra cui il realismo, la consistenza temporale, il fenomeno dell'*uncanny valley*, l'espressività e la fluidità narrativa. Diversi studi, tra cui quello di Zhang et al. [16], segnalano importanti progressi nella generazione video: i sistemi generativi odierni possono raggiungere un notevole realismo in molte scene, ad esempio nel riprodurre fenomeni naturali o complesse riprese ambientali, producendo video che nulla invidiano a quelli girati in modo tradizionale. Inoltre, l'IA offre nuove possibilità stilistiche prima impraticabili, con aspetti visivi originali, in particolare quelli di impronta onirica o astratta che ampliano il panorama cinematografico. Allo stesso tempo, però, molti creatori scelgono volutamente estetiche non completamente realistiche, come ad esempio uno stile cartoon o surreale, proprio per mascherare le imperfezioni che risultano più evidenti in approcci foto realistici. Nonostante questo, rimangono criticità tecniche ricorrenti; ottenere infatti un'elevata consistenza temporale è difficile e i video generati manifestano spesso incoerenze tra fotogrammi consecutivi. Analogamente, la composizione delle inquadrature e il controllo del movimento di camera non risultano ancora pienamente affidabili, soprattutto in sequenze più lunghe e complesse. Uno dei punti più critici è sicuramente la sintesi di volti e figure umane credibili. Se nelle immagini statiche l'ostacolo è stato quasi del tutto superato, nei video i personaggi spesso difettano di quelle sfumature espressive e movimenti naturali indispensabili per conferire realismo e autenticità. Di conseguenza tali figure tendono a suscitare nello spettatore un senso di estraneità riconducibile al noto effetto dell'*uncanny valley*, un'inquietudine percepita di fronte a simulazioni umane non perfette. Anche la fluidità narrativa risulta una sfida aperta: sempre dallo studio di Zhang et al. [16] risulta che diversi autori riscontrano il grande problema di mantenere un racconto cinematografico pienamente coeso e comprensibile sul lungo periodo, pur riuscendo a generare singole scene visivamente valide. Queste ricerche delineano uno scenario estetico in rapida evoluzione, in cui le straordinarie potenzialità visive offerte dall'IA si accompagnano a limiti tecnici ed espressivi ancora irrisolti. Alla luce di tali riferimenti teorici, il presente progetto cerca dunque di valorizzare gli elementi estetici più efficaci tratti dalla letteratura, come l'uso di stili visuali onirici e di inquadrature creative rese possibili dall'IA, evitando le aree maggiormente problematiche. Si è scelto di realizzare un prodotto in linea di massima foto realistico che metta seriamente alla prova gli strumenti generativi e che permetta di sperimentare e allo stesso tempo ricercare il modo migliore per generare video coerenti, naturali e in linea con i dati raccolti dal pubblico.

2.1.4 Impatto della tecnologia nelle arti visive

Un esempio approfondito sull'impatto dell'IA nelle arti visive emerge da recenti studi multidisciplinari che vanno oltre al semplice entusiasmo portato necessariamente da una nuova tecnologia, esplorando le vere contraddizioni tra innovazione e crisi dei processi. Ad esempio, Wadinambiarachchi et al. [17], in un esperimento con 60 partecipanti, hanno dimostrato che l'uso dell'IA durante la fase di ideazione favorisce la determinazione concettuale sui modelli iniziali, riducendo la varietà e l'originalità delle proposte creative. Doshi e Hauser [18] confermano che sebbene l'IA stimoli una produttività percettivamente più alta, ne snatura l'autenticità, generando output simili e meno innovativi. Bisogna quindi porre particolare attenzione all'uso esclusivo di questi modelli, che portano con sé, tra le altre cose, tanti bias dovuti ai dati di partenza su cui sono stati allenati, come riportato da Zhou et al. [19]. L'adozione dell'IA nelle arti visive ha scatenato la definizione di nuovi processi creativi, come evidenziato in numerose ricerche. Un'indagine sistematica condotta da Ibrahim [20] ha documentato come strumenti basati su GAN e visione artificiale abbiano trasformato sia la produzione che la fruizione artistica, introducendo forme performative immersive, interattive e collettive ma evidenziando al contempo i rischi legati a bias, copyright e trasparenza nel processo progettuale.
A seguito di questa analisi emerge che l'IA da una parte potenzia l'immaginazione estetica e le possibilità creative, stimolando l'efficienza ideativa e democratizzando l'accesso a forme artistiche altrimenti non alla portata di tutti; dall'altra, però, mette a rischio l'originalità, accentua processi cognitivi fermi sulle medesime idee, scardina meccanismi di valutazione estetica e mette a dura prova i principi etici e legali della professione artistica.

2.2 Analisi dei festival e delle iniziative legate all'IA

2.2.1 Festival dedicati alla creatività con IA

Negli ultimi anni si è assistito a un crescente interesse del panorama culturale e cinematografico nei confronti dell'intelligenza artificiale, non solo come oggetto di rappresentazione narrativa, ma come vero e proprio strumento creativo. Questo cambiamento ha portato all'organizzazione di numerosi festival e iniziative che riconoscono, promuovono e mettono in mostra opere audiovisive in cui l'IA riveste un ruolo centrale nella produzione, nella regia o nella scrittura. L'attenzione non è più riservata solamente a circuiti sperimentali o accademici, ma si sta progressivamente estendendo a manifestazioni di rilievo e a spazi di riflessione di livello professionale, con un'ampia varietà di approcci e differenti sensibilità culturali, legati sia alla geografia che ai linguaggi visivi adottati.
Negli Stati Uniti si colloca l'*AI International Film Festival*, che si svolge a Hollywood, in California e rappresenta uno dei principali appuntamenti mondiali dedicati all'intreccio tra cinema e intelligenza artificiale. Questo evento propone una selezione di cortometraggi, documentari e lungometraggi provenienti da tutto il mondo. Nel 2024, sempre negli Stati Uniti, il *Tribeca Festival* ha segnato un momento di svolta, introducendo per la prima volta una sezione dedicata ai cortometraggi generati interamente con IA, grazie a una collaborazione con *OpenAI*. Degno di nota è sicuramente il *New Media Film Festival* di Los Angeles che include una sezione de-

dicata all'IA riconoscendo il valore di formati innovativi come cortometraggi, realtà virtuale e NFT. Questa apertura da parte di un festival storicamente legato all'industria tradizionale ha segnato un'importante svolta nel panorama cinematografico professionale internazionale, favorendo anche altre manifestazioni di questo tipo. In Europa, ad esempio, il panorama si è arricchito di eventi mirati come il *+RAIN Film Festival*, organizzato dalla *Universitat Pompeu Fabra* di Barcellona, che è stato il primo festival europeo interamente dedicato a video generati con intelligenza artificiale. L'iniziativa ha unito ricerca tecnologica, sperimentazione artistica e riflessione critica, con un'attenzione specifica alle pratiche emergenti nel settore audiovisivo. A Praga, il *Czech International AI Film Festival* ha riunito nella sua edizione del 2024 esperti, registi, ricercatori e artisti per discutere delle implicazioni teoriche e pratiche dell'intelligenza artificiale nel linguaggio filmico, configurandosi come un importante polo di scambio e approfondimento. Per quanto riguarda l'Italia, invece, l'*AI.motion Festival*, ospitato dall'Università *IULM* di Milano durante la *Milano Digital Week*, ha introdotto una piattaforma dedicata a cortometraggi, videoclip e spot pubblicitari generati con l'ausilio dell'IA, riservando anche una sezione tematica ai personaggi digitali e agli *avatar* realistici. Sempre in ambito italiano, il *Burano Artificial Intelligence Film Festival* si concentra specificamente sui cortometraggi e promuove la sperimentazione narrativa e visiva, con una particolare attenzione alle nuove forme di espressione rese possibili dai sistemi generativi.
Oltre ai classici eventi legati al mondo cinematografico, anche le aziende stesse sono particolarmente interessate a promuovere la creatività legata a questa nuova tecnologia: *Runway*, una delle compagnie leader nello sviluppo di tecnologie video, ha fondato l'*AI Film Festival*, che dal 2022 premia ogni anno le migliori produzioni realizzate con strumenti generativi. I lavori selezionati vengono poi presentati anche in altri festival internazionali, contribuendo a dare visibilità ai nuovi linguaggi. Accanto a questi appuntamenti, emergono format più orientati alla competizione creativa, come il *Reply AI Film Festival*, organizzato da *Reply*, azienda italiana di consulenza informatica. Questa opportunità è in particolare una sfida internazionale che invita videomaker e designer a produrre opere utilizzando le più recenti tecnologie di intelligenza artificiale, stimolando così la produzione sperimentale e il confronto tra giovani creativi. Altre manifestazioni come il *Festival of Creative AI* promosso dalla *University for the Creative Arts* in Regno Unito favoriscono il dialogo tra ricerca e produzione attraverso demo e workshop dedicati ai video generati con IA.
La varietà geografica e tematica di questi festival dimostra come l'intelligenza artificiale sia ormai percepita come una componente legittima, seppur ancora in via di definizione, del discorso artistico e produttivo contemporaneo. Le opere selezionate e premiate in questi contesti non solo mettono in mostra le capacità tecniche dell'IA, ma anche la volontà di ridefinire i confini tra autore, macchina e pubblico, aprendo nuove prospettive sulla creatività audiovisiva. Questa rete di festival, osservata nella sua evoluzione, rappresenta per il presente progetto di tesi un quadro di riferimento fondamentale, per comprendere le tendenze estetiche produttive attualmente riconosciute e per orientare il lavoro in modo consapevole e critico, valorizzando le direzioni più innovative e coerenti con il nuovo linguaggio cinematografico emergente.

2.2.2 Progetti video rilevanti generati con IA

Negli ultimi anni la produzione di video generati principalmente o interamente con l'ausilio dell'IA ha dato vita a diversi progetti di rilievo, sia per innovazione estetica sia per l'impatto critico e culturale suscitato. Ad esempio, l'azienda *Campari* ha realizzato *Fellini Forward*, il primo cortometraggio creato con strumenti di intelligenza artificiale ispirati all'estetica di Federico Fellini. Un team di esperti ha utilizzato algoritmi di *machine learning* per analizzare le sceneggiature, le inquadrature e i movimenti di macchina tipici del regista italiano, generando così una nuova opera dal marcato stile felliniano. Il cortometraggio, presentato con successo alla *Mostra del Cinema di Venezia* e accompagnato da un documentario sul processo creativo, ha riscosso un notevole interesse per la sua esplorazione del rapporto tra uomo e macchina. Sempre sul versante dell'arte generata attraverso IA si colloca *The Crow* di Glenn Marshall [21], un cortometraggio che racconta la metamorfosi di una danzatrice in corvo, all'interno di un paesaggio desolato. Il risultato, una visione onirica e pittorica in movimento, ha vinto il *Jury Award* al *Cannes Short Film Festival*, rendendolo uno dei primi film generati da intelligenza artificiale premiati in un contesto cinematografico internazionale.
Non mancano però i casi in cui l'impiego dell'IA ha sollevato critiche. Il collettivo *Corridor Digital* ha suscitato un dibattito acceso con *Anime Rock, Paper, Scissors* [22], un cortometraggio girato in live action e successivamente trasformato tramite reti generative per replicare l'estetica dei film d'animazione giapponesi anni '90. I creatori hanno presentato questa tecnica innovativa, basata sui filmati girati mediante l'utilizzo del green screen e successivamente ripassati dall'IA fotogramma per fotogramma. Numerosi animatori e critici contemporanei hanno riscontrato una bassa qualità e diversi problemi etici, accusando il modello generativo utilizzato di aver plagiato lo stile di disegni preesistenti senza alcun apporto creativo. Anche il recente caso di *Star Wars: Field Guide* [23], ha suscitato non poche critiche. Si tratta di un cortometraggio dimostrativo in cui *Lucasfilm* ha adottato l'IA per immaginare un documentario naturalistico ambientato in un pianeta della galassia *Star Wars*. La fauna aliena creata interamente dall'IA, presenta caratteristiche banali e ripetitive, oltre ad essere palesemente frutto dell'unione di più specie di animali terrestri, gli unici conosciuti dai modelli di intelligenza artificiale. Presentato come un esperimento tecnico dal direttore degli effetti speciali Rob Bredow, il video è stato realizzato in due settimane da un singolo artista con l'ausilio di strumenti generativi. La sua pubblicazione ha scatenato reazioni negative da parte dei fan, che lo hanno trovato esteticamente poco originale e hanno valutato i contenuti generati da IA di qualità inferiore rispetto a quelli tradizionali.
Questi esempi recenti, tra sperimentazioni premiate, passi falsi e collaborazioni tra uomo e macchina, testimoniano le nuove possibilità creative offerte dall'IA nella produzione video ma anche le resistenze e i dibattiti critici che tale innovazione porta con sé.

2.3 Benchmark degli strumenti IA

2.3.1 Strumenti IA per sceneggiatura, regia e produzione

Nel processo di creazione, una fase preliminare ma fondamentale è l'analisi comparativa, critica e sperimentale dei molteplici strumenti disponibili per la scrittura, la regia e la produzione. Il sistema degli strumenti di IA oggi disponibili, infatti, si presenta estremamente frammentato e in continua espansione: si passa da modelli di tipo generativo a modelli discriminativi, da software multimodali a sistemi altamente specializzati.
Negli ultimi anni si è assistito a una crescita senza precedenti del numero di strumenti basati su IA, in particolare nel campo delle tecnologie generative, con coinvolgimento diretto e tangibile anche del settore audiovisivo. Secondo l'*AI Index 2025 Annual Report* pubblicato dallo *Stanford Institute for Human-Centered Artificial Intelligence* [24], il numero complessivo di strumenti generativi disponibili pubblicamente cresce a dismisura ogni anno, raddoppiando rispetto all'anno precedente. All'interno della classificazione riportata dal report, si evidenzia infatti che buona parte degli strumenti sono oggi attivi in ambiti direttamente collegati alla produzione audiovisiva, tra cui si trovano applicazioni per la generazione video da testo e l'animazione di immagini statiche. Tale proliferazione non è soltanto quantitativa ma segna un'evoluzione nella natura stessa degli strumenti: si è passati da soluzioni generiche e poco ottimizzate a piattaforme verticali, pensate per rispondere a compiti specifici che riguardano tutta la pipeline creativa, come ad esempio la scrittura di sceneggiature tramite linguaggio naturale, il montaggio, la creazione di effetti speciali o la generazione di intere sequenze visuali coerenti. Viene inoltre osservato come questa espansione sia accompagnata da un processo di convergenza multimodale, in cui strumenti come *Pika Labs*, *Sora* e *Veo* puntano a unificare, all'interno di un'unica piattaforma, la scrittura narrativa, la direzione registica e la realizzazione visiva. Anche il valore economico complessivo del comparto legato all'IA per i media audiovisivi è stimato essere in forte crescita; questo fatto riflette una trasformazione profonda e accelerata del sistema produttivo, in quanto l'intelligenza artificiale non costituisce più una tecnologia accessoria o sperimentale, ma si configura sempre più come uno standard operativo da integrare all'interno dei processi produttivi. In questa eterogeneità di piattaforme risiede sicuramente la possibilità di un'alta sperimentazione ma allo stesso tempo una complessità metodologica. È dunque emersa fin da subito la necessità di procedere a una mappatura funzionale dei vari strumenti, individuando non solo le loro caratteristiche ma anche le modalità con cui i modelli analizzati affrontano compiti simili utilizzando talvolta approcci differenti. In una prima fase, si è fatto affidamento a benchmark resi pubblici, che hanno permesso di effettuare una selezione importante, prendendo in analisi solamente i modelli con i punteggi più alti. Successivamente, si è adottata una logica di analisi trasversale: a ogni modello rimasto è stato assegnato lo stesso compito creativo (come ad esempio la scrittura di una scena, la generazione di un'immagine dello storyboard, la produzione di una sequenza visiva) e le risposte ottenute sono state valutate secondo criteri di coerenza, precisione, capacità di tradurre l'intento autoriale e dare continuità visuale. Questo processo di comparazione, condotto in una fase preliminare del progetto, ha permesso di cogliere differenze sostanziali non solo a livello di output, ma soprattutto in termini di logiche interne, modalità

di gestione del contesto e grado di controllo offerto all'utente. L'obiettivo non era semplicemente identificare lo strumento migliore in senso assoluto, quanto piuttosto comprendere, per ogni fase della pipeline produttiva, quale fosse il modello o la piattaforma più adatta a svolgere una funzione specifica in modo efficace e coerente con la visione complessiva, ponendo particolare attenzione anche alla semplicità di utilizzo dello strumento e alle possibilità offerte per trasferire i risultati ottenuti alla piattaforma utilizzata successivamente.

2.3.2 Benchmark

Durante il processo di valutazione dei vari modelli di intelligenza artificiale, ricoprono un ruolo chiave i benchmark, standard di riferimento solitamente costituiti da un insieme di test o compiti, utilizzati per valutare in modo sistematico e comparabile le prestazioni di un sistema. Molti dei benchmark più noti stanno però raggiungendo la saturazione: il rapido miglioramento dei modelli rende spesso obsoleti anche i test più recenti nel giro di pochi anni. Inoltre, quando vengono utilizzate tecniche non standard di prompting, come nel caso della *chain-of-thought*, il processo di valutazione si complica ulteriormente, vanificando di fatto il ruolo dei benchmark. Esistono poi ambiti dell'intelligenza artificiale che, per natura, sfuggono a una facile misurazione, come i sistemi multi agente e quelli che utilizzano principalmente l'interazione tra uomo e macchina. Nella moderna ricerca sul *machine learning*, molti benchmark vengono gestiti da terze parti senza che i dati di test vengano resi pubblici, aumentando così l'attendibilità dei risultati. Infine, un ulteriore problema, emerso con modelli di linguaggio di grandi dimensioni, è il fenomeno della contaminazione dei benchmark, ovvero la presenza di dati dei test già inclusi nei dataset di addestramento. Questo insieme di osservazioni evidenzia quanto la costruzione, l'utilizzo e l'interpretazione dei benchmark siano oggi elementi importanti per un'analisi affidabile dello stato dell'arte dell'intelligenza artificiale e per evitare conclusioni affrettate o fuorvianti rispetto alle reali capacità dei modelli.

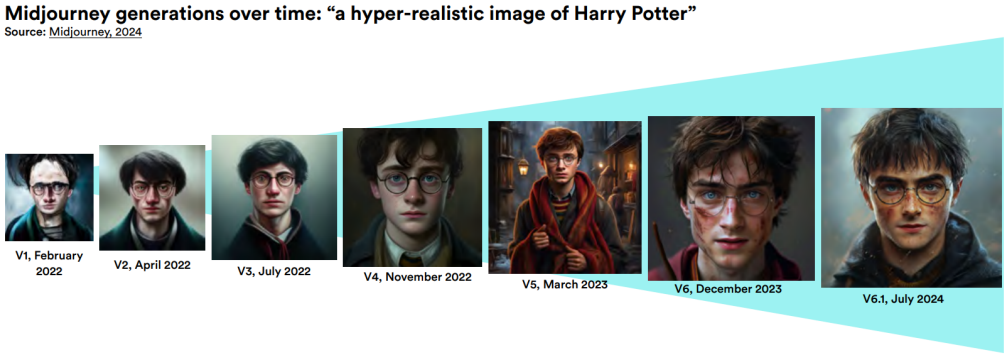


Figura 2.1: Evoluzione delle capacità generative del modello *Midjourney*. Negli anni il risultato è notevolmente migliorato.

Tra i benchmark più importanti, emblematici ed efficaci, troviamo il celebre *Will Smith eating spaghetti*. Il prompt, semplice ma allo stesso tempo estremamente funzionale da un punto di vista scientifico, nato a metà 2023 all'interno degli ambienti di ricerca e sviluppo legati ai modelli di generazione video tramite IA, è diventato nel giro di poche settimane un vero e proprio caso di studio, sia in ambienti accademici

sia nella divulgazione tecnica. L'origine va ricondotta a un esperimento eseguito con un prototipo iniziale di modello *text-to-video* che, sollecitato con la richiesta di generare la clip specifica, produceva una sequenza visivamente inquietante e deformata, in cui i movimenti, le espressioni facciali e la relazione tra il soggetto e il cibo risultavano fortemente disturbanti e surreali. Proprio questa imperfezione, unita alla scelta di un soggetto universalmente riconoscibile e di un'azione ordinaria ma ricca di implicazioni visive, ha reso il prompt particolarmente indicativo per mettere alla prova le capacità dei modelli generativi emergenti. A partire da quel momento, *Will Smith eating spaghetti* è stato adottato come benchmark per testare qualità visiva, sincronizzazione temporale e gestione dell'unione tra riconoscibilità facciale e azione dinamica. L'immagine dell'attore, già codificata in innumerevoli dataset disponibili pubblicamente, ha rappresentato un punto di riferimento visivo forte, ma allo stesso tempo ha rivelato le fragilità intrinseche dei modelli nella resa realistica di interazioni corporee complesse e di oggetti debolmente strutturati come la pasta. In particolare, il prompt ha permesso di sintetizzare in un solo test molte delle sfide tecniche che la generazione video automatica ancora affronta. Nel tempo, il prompt ha assunto anche una funzione metanarrativa e ironica; il suo utilizzo ricorrente è servito a sottolineare non solo i progressi, ma anche gli errori, le allucinazioni visive e le problematiche legate all'*uncanny valley*. In molti casi, i risultati sono stati ampiamente condivisi come elementi virali, trasformando un semplice test in un esempio simbolo della difficoltà di creare un prodotto il più possibile realistico.

Figura 2.2: Confronto tra gli output generati da differenti modelli dello strumento *Pika* nel corso degli anni. È notevole l'avanzamento tecnologico mostrato nel susseguirsi dei vari frame.

Dal punto di vista metodologico, questo prompt si è imposto come uno stress test informale, ma estremamente utile, in grado di comparare rapidamente i risultati tra diverse architetture e versioni successive dello stesso sistema. La sua longevità come riferimento è anche dovuta al fatto che riguarda un compito apparentemente semplice ma con implicazione di dettagli importanti come il coordinamento tra mani e bocca, la consistenza e il moto del cibo, le variazioni espressive, la parziale visibilità del volto, il contesto ambientale coerente. Per queste ragioni, questo benchmark si è progressivamente trasformato in un punto di passaggio quasi obbligato per la dimostrazione delle capacità visive dei nuovi modelli *text-to-video*, fornendo un equilibrio tra familiarità del contenuto e complessità tecnica. In un panorama in rapida evoluzione, in cui i benchmark formali sono spesso in ritardo rispetto ai

rilasci tecnologici, questo prompt rimane un punto di riferimento per la valutazione qualitativa immediata della generazione video basata su IA.
La cosa più sorprendente è l'incredibile miglioramento avvenuto in meno di due anni. Se inizialmente i risultati prodotti da questo test nel 2023 erano praticamente inguardabili e inutilizzabili, già nel 2024 si è visto un netto miglioramento da parte di vari modelli nella realizzazione della clip video, come mostrato in Figura 2.3. Nello specifico, si è notato un miglioramento notevole ogni circa sei mesi. Ad oggi, diverse testate giornalistiche si trovano d'accordo nell'affermare che il test è stato pienamente superato dai modelli di IA più recenti, come ad esempio *Google Veo 3*, rappresentato in Figura 2.4.

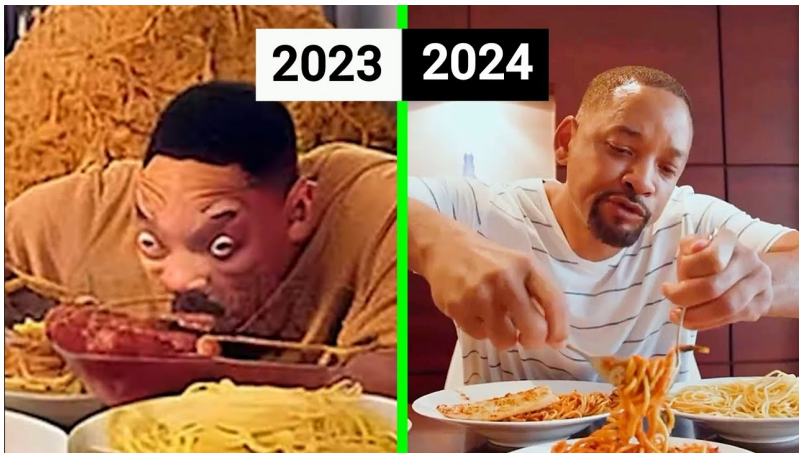


Figura 2.3: Confronto tra due frame generati rispettivamente nel 2023 e nel 2024. Si nota chiaramente un notevole miglioramento nella generazione della clip video.

Figura 2.4: Frame relativo al test fatto con la piattaforma *Google Veo 3*.

2.3.3 Modelli text-to-video e image-to-video

Una parte centrale dell'analisi preliminare ha riguardato il confronto tra modelli di generazione *text-to-video* e *image-to-video*, poiché la scelta di uno o dell'altro approccio comporta la definizione di due pipeline produttive profondamente diverse, ciascuna delle quali caratterizzata da specifici punti di forza e vincoli metodologici. I

modelli *text-to-video* (T2V), come suggerisce il nome, si basano sulla capacità di generare un output a partire da un input esclusivamente testuale. Questo implica una forma di traduzione dal linguaggio testuale a quello visivo in cui la descrizione, strutturata attraverso prompt dettagliati, viene interpretata e trasformata in sequenze animate. L'impiego di questi strumenti richiede un elevato grado di competenza nel prompt engineering, ovvero nella costruzione precisa, modulata e strategica del testo che guiderà la generazione. Qualsiasi omissione, ambiguità o genericità viene interpretata liberamente dal modello che, in assenza di vincoli forti, può generare sequenze incoerenti, visivamente scollegate o lontane dall'intento autoriale. È quindi fondamentale generare dei prompt altamente dettagliati e precisi, in modo tale da garantire l'effettiva buona riuscita della generazione dei video. Un punto a favore di questa modalità è certamente la rapidità con cui è possibile esplorare alternative stilistiche o narrative, senza dover aspettare a lungo per la generazione di nuove immagini, caratteristica invece dei modelli *image-to-video*. Si possono inoltre realizzare scene complesse in maniera immediata, andando quindi a creare in fase di pre-produzione degli esempi di visualizzazione del contenuto finale. Quanto sperimentato è supportato anche dalla ricerca scientifica: Singh [25] descrive come i modelli T2V impieghino reti neurali complesse addestrate su coppie testo–video, valutando prestazioni e metriche di qualità. Questi modelli mostrano flessibilità narrativa ma soffrono spesso di incoerenze temporali e di linguaggio, oltre ad avere un'alta richiesta computazionale.
L'approccio *image-to-video* (I2V) si fonda sulla possibilità di fornire al modello uno o più frame che vengono interpretati come stato iniziale e finale della sequenza, costituendo una base visuale da cui il modello genera un'animazione coerente. A differenza del T2V il controllo autoriale si sposta interamente sulla componente visiva: le immagini di partenza devono essere realizzate con estrema precisione, poiché costituiscono il riferimento formale per la generazione dell'intera clip. La loro accuratezza in termini di composizione, rapporto tra soggetti e sfondo, prospettiva, azione implicita e coerenza narrativa risulta fondamentale per il successo del processo generativo. La creazione dei frame iniziali e finali di ogni sequenza rappresenta una delle operazioni più delicate in quanto una buona definizione di queste due componenti riduce notevolmente la possibilità che il modello generi delle allucinazioni con elementi indesiderati o si discosti tanto dal significato progettuale. L'I2V consente quindi un controllo visivo molto più diretto ma risulta più oneroso in termini di produzione: la generazione delle immagini richiede tempo e competenze per la scrittura dei prompt e serve un'attenta pianificazione dei movimenti all'interno della clip.
Complessivamente, la ricerca evidenzia che i modelli T2V offrono una maggiore libertà generativa ma richiedono prompt molto precisi e risorse computazionali elevate, mentre gli I2V richiedono accuratezza nella fase visiva iniziale e conferiscono un controllo superiore sui contenuti, risultando quindi particolarmente adatti quando si vuole preservare una forte identità estetica. L'intera fase di analisi, ricerca e sperimentazione preliminare, condotta attraverso lo studio di benchmark pubblici e mediante test pratici, ha rappresentato un momento metodologicamente fondamentale per lo sviluppo del progetto. La possibilità di confrontare in modo strutturato le potenzialità e i limiti di ciascuna tecnologia ha permesso poi di costruire una pipeline produttiva coerente e consapevole dei limiti tecnici.

2.3.4 IA per la generazione di testo

Durante la fase di pre-produzione del progetto audiovisivo, una delle scelte metodologiche più rilevanti è stata l'individuazione del modello di intelligenza artificiale generativa più adatto alla gestione di tutte le attività legate alla scrittura, alla progettazione narrativa e alla strutturazione dei documenti testuali preliminari. Questa fase ha riguardato nello specifico la redazione e la rifinitura di materiali fondamentali quali logline, sinossi e sceneggiatura. Si è trattato di un processo esteso, iterativo e creativo, in cui l'IA è stata impiegata come un interlocutore nello sviluppo dei contenuti. Per garantire un risultato soddisfacente dal punto di vista sia espressivo che funzionale, è stato necessario analizzare inizialmente i principali benchmark di riferimento nell'ambito della generazione testuale e di linguaggio: in particolare, si è fatto riferimento a indicatori come l'*MMLU* (Massive Multitask Language Understanding) [26], il *WildBench* [27] e i più recenti standard di valutazione come *HELMET* [28]. A partire da questi risultati, visibili chiaramente in Figura 2.5, 2.6, 2.7, la selezione è stata progressivamente ristretta a due modelli: *Claude* di *Anthropic* e *ChatGPT* di *OpenAI*.

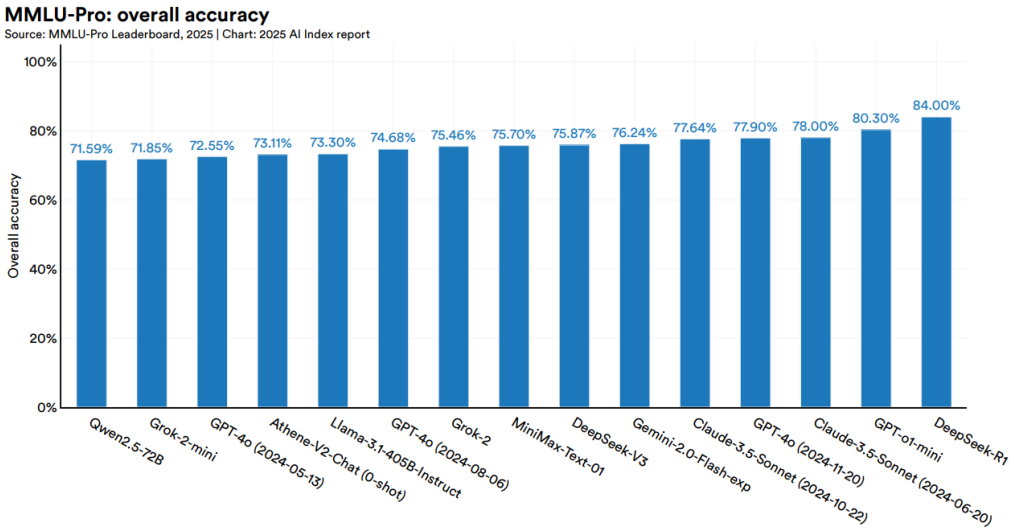


Figura 2.5: Risultati del benchmark *MMLU*. *Claude* si posiziona tra due modelli differenti di *ChatGPT*, tutti comunque nelle quattro posizioni migliori.

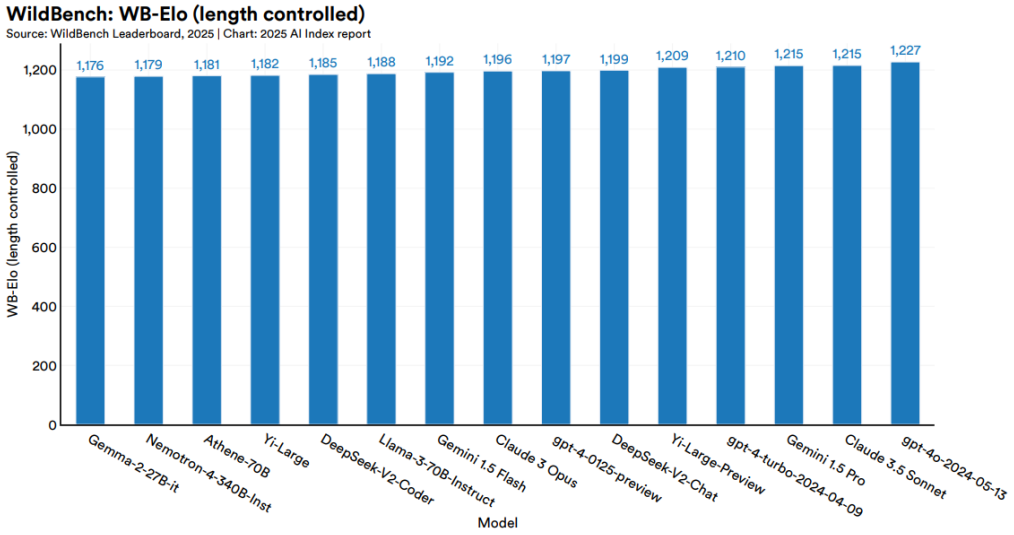


Figura 2.6: Risultati del benchmark *WildBench*. *ChatGPT* si posiziona al primo posto, seguito immediatamente da *Claude*.

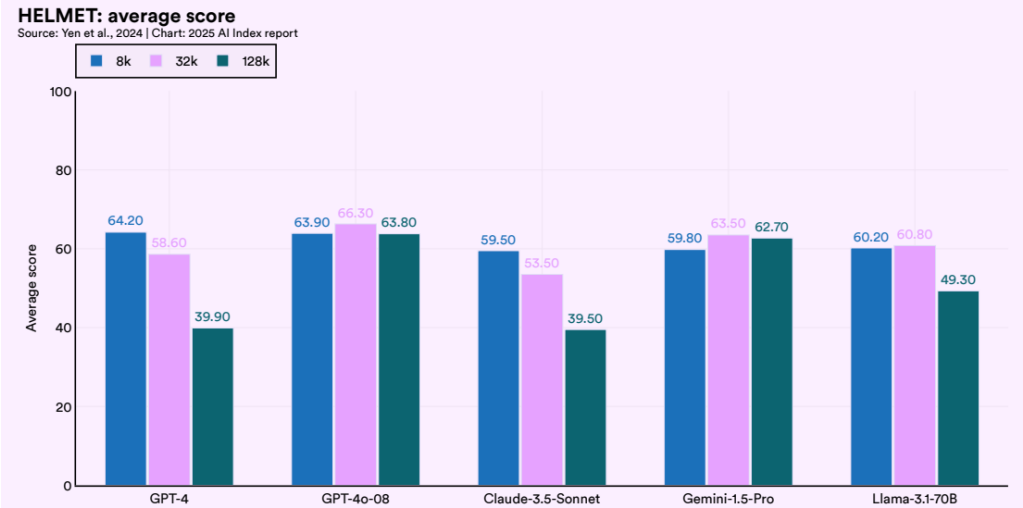


Figura 2.7: Risultati del benchmark *HELMET*. *ChatGPT* ottiene nuovamente il primato, staccando gli altri modelli.

Entrambi i modelli si sono dimostrati altamente competitivi, con prestazioni notevoli nei task di generazione creativa. Tuttavia, durante la fase di test e sperimentazione in situazioni operative reali, sono emerse differenze significative. *Claude* si è rivelato particolarmente efficiente nella gestione di documenti lunghi e strutturati, grazie a una finestra di contesto ampia e a una capacità di ragionamento logico più esplicita. In molti casi, ha mostrato un approccio riflessivo, orientato alla cautela e all'analisi delle richieste, risultando utile per operazioni di revisione e sintesi. Questo stesso atteggiamento conservativo si è rivelato meno funzionale nei momenti in cui era necessario generare materiale creativo, scenico o narrativamente efficace, dove l'esigenza non era solo quella di capire profondamente il testo ma di amplificarne l'immaginario, proponendo soluzioni alternative. *ChatGPT*, al contrario, si è dimostrato complessivamente più reattivo, adattabile e, soprattutto, in grado di

assecondare la direzione autoriale del progetto, facilitando la co-creazione con l'utente. L'interazione, da questo punto di vista, è stata spesso più semplice e naturale, portando a dei risultati più efficaci; il modello si è dimostrato capace di intervenire in diverse fasi con efficacia costante. Gli elementi appena descritti sono ovviamente puramente soggettivi e l'autore umano ha preferito proseguire in questa direzione visto l'approccio positivo. Inoltre, un elemento che ha pesato in maniera decisiva nella scelta finale è stata la possibilità, offerta dal piano *Plus* di *ChatGPT*, di sfruttare la funzionalità *Progetto*, che consente di organizzare, memorizzare e richiamare file, chat e istruzioni complesse attraverso sessioni persistenti. Questa funzione si è rivelata essenziale nella gestione della scrittura e di tutte le fasi successive del progetto, non rendendo necessaria la spiegazione continua.
È necessario infine spiegare che, durante la fase di valutazione degli strumenti da impiegare nello stadio di pre-produzione, sono stati presi in esame diversi modelli e piattaforme specificamente progettate per la generazione automatica di sceneggiature, soggetti cinematografici o trattazioni narrative. Questi strumenti, spesso pubblicizzati come soluzioni verticali pensate per il settore audiovisivo, offrivano interfacce semplificate, strutture predefinite e suggerimenti fortemente ispirati a modelli narrativi consolidati (come il *viaggio dell'eroe* o i *tre atti*). Sottoponendoli però ad alcuni brevi test operativi, tali modelli si sono rivelati significativamente limitanti sotto diversi aspetti: i risultati ottenuti sono apparsi estremamente altalenanti dal punto di vista qualitativo, alternando parti ben strutturate ad altre incoerenti o ripetitive, senza la possibilità di intervenire manualmente per personalizzare in profondità la struttura del testo generato, rendendo così difficile effettuare in maniera coerente le successive revisioni. Per questi motivi si è scelto di non adottare tali soluzioni, preferendo strumenti più generali ma altamente configurabili, che garantiscono maggiore coerenza, adattabilità e controllo effettivo sul processo di creazione.

2.3.5 IA per la generazione di immagini

Nel corso della fase di pre-produzione del progetto audiovisivo, l'utilizzo dell'intelligenza artificiale generativa per la produzione di immagini ha rappresentato un punto importante, soprattutto per due ambiti di applicazione in particolare: la realizzazione dello storyboard visivo e la costruzione di un modello della protagonista del cortometraggio, attraverso la generazione di un *character sheet*. Inizialmente, è stata condotta un'analisi delle principali piattaforme di generazione di immagini disponibili. Partendo da un esame dei risultati riportati dal progetto *LMSYS Chatbot Arena for LLMs*, di Chiang et al. [29], che permette agli utenti di votare le migliori generazioni di immagini, si è avuto modo di avere un elenco delle migliori piattaforme dedicate alla creazione di immagini tramite IA.
A partire da questi risultati, che hanno valutato *Gemini* come modello migliore, seguito a poca distanza da *ChatGPT* (Figura 2.8), la fase operativa è stata strutturata attraverso una serie di test comparativi condotti su diverse piattaforme di IA generativa, ciascuna delle quali è stata interrogata con lo stesso prompt, al fine di valutarne le capacità stilistiche e la precisione semantica. Le piattaforme sperimentate includono: *Freepik*, *Hailuo AI*, *Ideogram*, *DreamStudio*, *Leonardo AI*, *Pixverse*, *DeepAI* e *TensorArt*, oltre al modello integrato in *ChatGPT*, che all'epoca utilizzava ancora *DALL·E*.

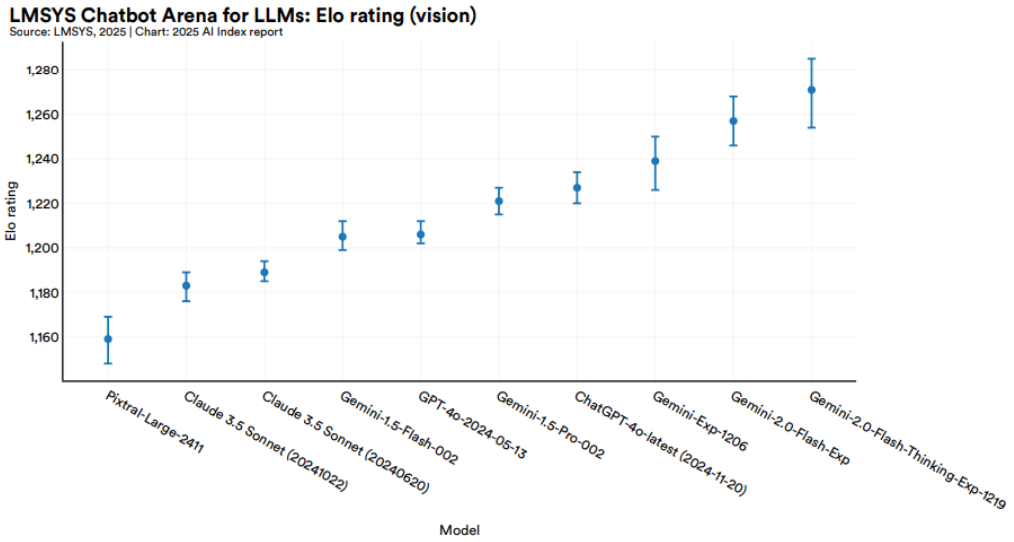


Figura 2.8: Risultati dalla piattaforma *Chatbot Arena for LLMs*.

In riferimento specifico alla generazione delle immagini necessarie alla costruzione del modello del personaggio (*character sheet*), sono stati condotti numerosi test tra i modelli citati, al fine di ottenere la migliore immagine possibile, mantenendo un completo controllo sui tratti somatici, abbigliamento e contesto stilistico forniti in input attraverso un prompt specifico, visionabile al Prompt 3.8. Si è poi scelto di proseguire con l'immagine generata da *Freepik*, che si è dimostrata essere più vicina alle indicazioni fornite.

Figura 2.9: Alcune immagini della protagonista generate su diverse piattaforme, utilizzando lo stesso prompt. In particolare, partendo dalla prima in alto a sinistra e proseguendo in senso orario: *Freepik*, *Hailuo AI*, *Leonardo AI* e *Ideogram*.

Per quanto riguarda invece la generazione dello storyboard, è stato deciso di concentrare la produzione su un unico modello, nello specifico *DALL·E*, integrato all'interno dell'ambiente di lavoro di *ChatGPT*. Tale scelta, pur con limiti qualitativi del modello in termini di dettaglio visivo, è risultata motivata da una serie di fattori operativi. In primo luogo, la piena integrazione con il sistema di gestione del progetto già attivo sulla piattaforma *OpenAI*, che permette il mantenimento del contesto narrativo precedentemente generato e la gestione del linguaggio. Questa caratteristica ha semplificato notevolmente il flusso di lavoro, permettendo di generare rapidamente le immagini visive dello storyboard, a partire direttamente dalla parte testuale già sviluppata in precedenza, permettendo così di rafforzare la coerenza interna tra narrazione e rappresentazione.

Va inoltre considerato che, successivamente alla fase di generazione dello storyboard, alcuni dei modelli più recenti hanno effettuato importanti aggiornamenti, migliorando sensibilmente la qualità e la stabilità degli output. Di conseguenza, si è riflettuto a lungo, valutando l'ipotesi di rigenerare interamente lo storyboard con i nuovi strumenti disponibili, optando alla fine per mantenere il materiale già creato, in quanto la fase di produzione era già stata avviata, coerentemente con le scelte progettuali. Riprogettare nuovamente questa parte avrebbe sicuramente migliorato la pre visualizzazione del contenuto audiovisivo finale, senza però andare a migliorare direttamente le clip prodotte.

2.3.6 IA per la generazione di video

Nel processo di individuazione del modello di intelligenza artificiale generativa più adatto alla creazione delle sequenze visive del cortometraggio è stata condotta inizialmente un'attenta analisi dei benchmark accademici più autorevoli, al fine di comprendere le prestazioni dei modelli disponibili. Tra i più rinominati troviamo *VBench*, proposto da Huang et al. [30], che si distingue per l'utilizzo di criteri valutativi umani e la capacità di fornire indicazioni approfondite e versatili sul comportamento dei sistemi in analisi. La versione estesa, denominata *VBench++*, amplia ulteriormente il campo di indagine, includendo anche la valutazione dei modelli I2V. I modelli di nuova generazione vengono costantemente integrati nella classifica ufficiale, permettendo così un confronto aggiornato e continuo tra soluzioni differenti. Le valutazioni condotte attraverso *VBench++* evidenziano numerosi aspetti critici e dinamiche complesse nella generazione di video con IA. Innanzitutto si osserva una contrapposizione tra la coerenza temporale e il grado di dinamicità delle clip: modelli come *LaVie* risultano eccellenti nella stabilità visiva ma generano contenuti tendenzialmente statici, mentre altri come *VideoCrafter* mostrano animazioni più vivaci ma con maggiori incoerenze temporali. L'analisi rivela anche che alcuni modelli hanno un alto potenziale all'interno di specifiche categorie. Ad esempio, *CogVideo* eccelle nella resa visiva degli alimenti ma non in quella di animali e veicoli. Questo porta a pensare che ciascun modello potrebbe specializzarsi in un particolare tipo di contenuto per migliorare le prestazioni globali unendo i singoli risultati ottenuti. In generale, categorie complesse dal punto di vista temporale e spaziale (come *esseri umani* o *veicoli*), rappresentano ancora una sfida, in particolare per quanto riguarda la qualità e la gestione dei movimenti. Sempre lo studio di Huang et al. [30], suggerisce che un aumento della quantità di dati nei parametri di addestramento di un modello, non garantisce per forza un miglioramento qualitativo. Dai risultati si evince poi che i modelli I2V non garantiscono una completa aderenza tra l'immagine di input e il video, dove spesso una corrispondenza eccessiva al frame iniziale tende a compromettere la fluidità temporale tra le immagini. Infine, la valutazione portata avanti dal gruppo di ricerca, evidenzia che i modelli industriali performano meglio in termini di affidabilità e consistenza rispetto a quelli sviluppati in ambito accademico. Il principale motivo potrebbe risiedere nell'utilizzo di dataset interni più curati e rigorosi.
Si analizzano ora nel dettaglio alcune delle piattaforme più importanti per quanto riguarda la generazione di video, considerando anche quanto riportato nella valutazione su *VBench-2.0*, benchmark ideato e portato avanti da Zheng et al. [31]. Grazie a un set strutturato di prompt tematici e ad un sistema di valutazione ibri-

da (automatica e umana), i ricercatori sono riusciti a valutare attentamente i vari modelli di IA generativa.
In questo progetto, sono state scartate da subito le piattaforme che richiedevano installazioni locali ad alta intensità computazionale, come *ComfyUI* o le implementazioni avanzate di *Stable Diffusion* per video, a causa dell'incompatibilità con le risorse hardware a disposizione.

Models	Human Anatomy	Human Clothes	Human Identity	Composition	Diversity	Mechanics	Material	Thermotics	Multi-view Consistency
HunyuanVideo [72]	**88.58%**	82.97%	75.67%	43.96%	39.73%	76.09%	64.37%	56.52%	43.80%
CogVideoX-1.5 [86]	59.72%	87.18%	69.51%	44.70%	42.61%	**80.80%**	**83.19%**	**67.13%**	21.79%
Sora [53]	86.45%	**98.15%**	**78.57%**	**53.65%**	**67.48%**	62.22%	64.94%	43.36%	58.22%
Kling 1.6 [69]	86.99%	91.75%	71.95%	43.89%	53.26%	65.55%	68.00%	59.46%	**64.38%**
Models	**Dynamic Spatial Relationship**	**Dynamic Attribute**	**Motion Order Understanding**	**Human Interaction**	**Complex Landscape**	**Complex Plot**	**Camera Motion**	**Motion Rationality**	**Instance Preservation**
HunyuanVideo [72]	**21.26%**	22.71%	26.60%	67.67%	19.56%	10.11%	33.95%	34.48%	73.79%
CogVideoX-1.5 [86]	19.32%	**24.18%**	26.94%	**73.00%**	**23.11%**	**12.42%**	33.33%	33.91%	71.03%
Sora [53]	19.81%	8.06%	14.81%	59.00%	14.67%	11.67%	27.16%	34.48%	74.60%
Kling 1.6 [69]	20.77%	19.41%	**29.29%**	72.67%	18.44%	11.83%	**61.73%**	**38.51%**	**76.10%**

Figura 2.10: Risultati della ricerca condotta da Zheng et al. [31]. In grassetto è evidenziato il risultato migliore ottenuto in ciascuna categoria analizzata.

Sora, di *OpenAI*, emerge come uno dei modelli più performanti per la creatività, dimostrando una forte capacità immaginativa che lo rende particolarmente adatto a progetti di storytelling artistico. Tuttavia il modello risulta carente in ambiti fondamentali come la controllabilità e la coerenza fisica, generando talvolta scene che non rispettano le istruzioni testuali.
Kling 1.6 si posiziona tra i modelli più solidi per quanto riguarda la consistenza tra video, la gestione del movimento di camera e la controllabilità narrativa, mostrando una buona aderenza ai prompt e una costruzione visiva razionale. Le sue prestazioni equilibrate in tutte le categorie, suggeriscono un addestramento su dataset ampi e vari, rendendolo un ottimo riferimento per sviluppi futuri, sebbene il tempo di generazione resti spesso più elevato rispetto a quello di altri modelli.
HunyuanVideo è significativamente debole in diverse dimensioni valutative ma mostra eccellenti risultati nella gestione dell'umano, in particolare per quanto riguarda la fedeltà anatomica e del movimento.
LaVie mostra punteggi alti nella coerenza del soggetto e nella consistenza dello sfondo ma si conferma poco dinamico, portando quindi ad un'elevata stabilità poco adatta a un'animazione complessa.
VideoCrafter-2.0 si distingue per una buona performance nel movimento e nella dinamicità, sacrificando in parte la coerenza tra i frame.
Nel campo I2V, *I2VGen-XL* e *DynamicCrafter* ottengono punteggi molto alti mostrando però una forte influenza del frame iniziale fornito, che può compromettere a volte la fluidità temporale successiva.
A livello globale, modelli come *Google Veo* e *Runway* (non inclusi direttamente nei dati *VBench* ma ampiamente documentati) si posizionano come tra i più completi in termini di risoluzione, realismo fisico e audio coerente, affermandosi come benchmark tecnici di riferimento, mentre *Sora* resta preferito in contesti narrativi grazie alla sua accessibilità, alla profonda integrazione con il proprio modello e all'efficienza della sua pipeline testuale.
Il confronto tra questi sistemi da un punto di vista accademico non permette di individuare il migliore, ma consente invece di orientare la scelta progettuale verso

il modello più adatto in funzione delle necessità specifiche di produzione. L'integrazione di benchmark come *VBench* consente oggi una valutazione strutturata, permettendo di aggiornarla con le ultime novità di mercato ed evidenziando anche le criticità ancora aperte nella generazione video con intelligenza artificiale.

Models	Subject Consistency	Background Consistency	Temporal Flickering	Motion Smoothness	Dynamic Degree	Aesthetic Quality	Imaging Quality	Object Class
LaVie [26]	91.41%	97.47%	98.30%	96.38%	49.72%	54.94%	61.90%	91.82%
ModelScope [20], [27]	89.87%	95.29%	98.28%	95.79%	66.39%	52.06%	58.57%	82.25%
VideoCrafter-0.9 [24]	86.24%	92.88%	97.60%	91.79%	89.72%	44.41%	57.22%	87.34%
CogVideo [19]	92.19%	96.20%	97.64%	96.47%	42.22%	38.18%	41.03%	73.40%
VideoCrafter-1.0 [62]	95.10%	98.04%	98.93%	95.67%	55.00%	62.67%	65.46%	78.18%
Show-1 [25]	95.53%	98.02%	99.12%	98.24%	44.44%	57.35%	58.66%	93.07%
VideoCrafter-2.0 [36]	96.85%	98.22%	98.41%	97.73%	42.50%	63.13%	67.22%	92.55%
Gen-2 [166]	97.61%	97.61%	99.56%	99.58%	18.89%	66.96%	67.42%	90.92%
AnimateDiff-v2 [86]	95.30%	97.68%	98.75%	97.76%	40.83%	67.16%	70.10%	90.90%
Latte-1 [26]	88.88%	95.40%	98.89%	94.63%	68.89%	61.59%	61.92%	86.53%
Pika-1.0 [167]	96.94%	97.36%	99.74%	99.50%	47.50%	62.04%	61.87%	88.72%
Kling [168]	98.33%	97.60%	99.30%	99.40%	46.94%	61.21%	65.62%	87.24%
Gen-3 [169]	97.10%	96.62%	98.61%	99.23%	60.14%	63.34%	66.82%	87.81%
CogVideoX-2B [170]	96.78%	96.63%	98.89%	97.73%	59.86%	60.82%	61.68%	83.37%
CogVideoX-5B [170]	96.23%	96.52%	98.66%	96.92%	70.97%	61.98%	62.90%	85.23%
Empirical Min	14.62%	26.15%	62.93%	70.60%	0.00%	0.00%	0.00%	0.00%
Empirical Max	100.00%	100.00%	100.00%	99.75%	100.00%	100.00%	100.00%	100.00%

Models	Multiple Objects	Human Action	Color	Spatial Relationship	Scene	Appearance Style	Temporal Style	Overall Consistency
LaVie [26]	33.32%	96.80%	86.39%	34.09%	52.69%	23.56%	25.93%	26.41%
ModelScope [20], [27]	38.98%	92.40%	81.72%	33.68%	39.26%	23.39%	25.37%	25.67%
VideoCrafter-0.9 [24]	25.93%	93.00%	78.84%	36.74%	43.36%	21.57%	25.42%	25.21%
CogVideo [19]	18.11%	78.20%	79.57%	18.24%	28.24%	22.01%	7.80%	7.70%
VideoCrafter-1.0 [62]	45.66%	91.60%	93.32%	58.86%	43.75%	24.41%	25.54%	26.76%
Show-1 [25]	45.47%	95.60%	86.35%	53.50%	47.03%	23.06%	25.28%	27.46%
VideoCrafter-2.0 [36]	40.66%	95.00%	92.92%	35.86%	55.29%	25.13%	25.84%	28.23%
Gen-2 [166]	55.47%	89.20%	89.49%	66.91%	48.91%	19.34%	24.12%	26.17%
AnimateDiff-v2 [86]	36.88%	92.60%	87.47%	34.60%	50.19%	22.42%	26.03%	27.04%
Latte-1 [26]	34.53%	90.00%	85.31%	41.53%	36.26%	23.74%	24.76%	27.33%
Pika-1.0 [167]	43.08%	86.20%	90.57%	61.03%	49.83%	22.26%	24.22%	25.94%
Kling [168]	68.05%	93.40%	89.90%	73.03%	50.86%	19.62%	24.17%	26.42%
Gen-3 [169]	53.64%	96.40%	80.90%	65.09%	54.57%	24.31%	24.71%	26.69%
CogVideoX-2B [170]	62.63%	98.00%	79.41%	69.90%	51.14%	24.80%	24.36%	26.66%
CogVideoX-5B [170]	62.11%	99.40%	82.81%	66.35%	53.20%	24.91%	25.38%	27.59%
Empirical Min	0.00%	0.00%	0.00%	0.00%	0.00%	0.09%	0.00%	0.00%
Empirical Max	100.00%	100.00%	100.00%	100.00%	82.22%	28.55%	36.40%	36.40%

Models	I2V Subject	I2V Background	Camera Motion	Subject Consistency	Background Consistency	Temporal Flickering	Motion Smoothness	Dynamic Degree	Aesthetic Quality	Imaging Quality
DynamiCrafter-1024 [56]	96.71%	96.05%	35.44%	95.69%	97.38%	97.63%	97.38%	47.40%	66.46%	69.34%
SEINE-512x320 [57]	94.85%	94.02%	23.36%	94.20%	97.26%	96.72%	96.68%	34.31%	58.42%	70.97%
I2VGen-XL [64]	96.74%	95.44%	13.32%	96.36%	97.93%	98.48%	98.31%	24.96%	65.33%	69.85%
Animate-Anything [63]	98.54%	96.88%	12.56%	98.90%	98.19%	98.14%	98.61%	2.68%	67.12%	72.09%
ConsistI2V [60]	94.69%	94.57%	33.60%	95.27%	98.28%	97.56%	97.38%	18.62%	59.00%	66.92%
VideoCrafter-I2V [62]	90.97%	90.51%	33.58%	97.86%	98.79%	98.19%	98.00%	22.60%	60.78%	71.68%
SVD-XT-1.1 [55]	97.51%	97.62%	-	95.42%	96.77%	99.17%	98.12%	43.17%	60.23%	70.23%

Figura 2.11: Risultati della ricerca svolta da Huang et al. [30]. Rispettivamente sopra sono presentati i risultati dei confronti tra modelli T2V mentre sotto si trovano quelli relativi ai modelli I2V.

Una volta individuati, i modelli più adatti sono stati sistematicamente valutati inserendo per ognuno lo stesso prompt, col fine di confrontarne la resa visiva, la coerenza narrativa, la fluidità del movimento e la qualità dei dettagli. Di seguito vengono riportati i risultati.
Pictory AI ha mostrato un'impostazione più adatta a contenuti informativi piuttosto che a video narrativi. *Descript*, centrato su immagini statiche e montaggio, non è risultato utile per scene dinamiche. *InVideo*, sebbene completo, è risultato molto più commerciale e meno adatto a gestire sequenze cinematografiche. *Pika Labs* ha mostrato ottime potenzialità ma anche grande instabilità nella coerenza tra frame, mentre *Pika* ha presentato costi molto elevati per i piani avanzati, rendendola di fatto inutilizzabile. *Sea Art* e *Hailuo AI*, operanti a crediti, sono state escluse per la difficoltà di pianificazione dei consumi. Altre piattaforme, come *Kling* e *Luma*, sono state valutate positivamente ma scartate per l'estrema lentezza nella generazione dei

contenuti. I tempi di attesa rendevano impossibile una gestione fluida del processo produttivo. Questo fattore, combinato con l'imprevedibilità dei risultati, ha influito negativamente sul giudizio.
La scelta finale è ricaduta su *Sora*, non solo per l'affidabilità associata all'azienda *OpenAI*, ma anche per l'efficienza tecnica e la qualità visiva costante. A differenza di altri strumenti che limitavano la durata dei video, imponevano *watermark* o funzionavano a consumo, *Sora* offre una generazione fluida, senza vincoli tecnici particolari e con tempi di risposta validi per una pratica iterativa. La precisione nella comprensione dei prompt e l'accurata generazione dei video hanno reso questa piattaforma la più adatta alle finalità progettuali.

2.3.7 IA per la generazione di audio

Nel processo di realizzazione di un cortometraggio mediante l'utilizzo dell'IA, la gestione della componente audio ha assunto un ruolo centrale per garantire un'esperienza audiovisiva completa e immersiva. Tale processo ha previsto tre ambiti principali di intervento: la generazione della colonna sonora, la creazione del voiceover dell'artista protagonista e l'inserimento di effetti sonori. Ciascuna di queste parti è stata affrontata attraverso l'impiego di strumenti di IA generativa specifici, selezionati dopo un'analisi approfondita di funzionalità, efficacia operativa e aderenza estetica. L'aspetto sonoro non viene considerato solo come un completamento visivo, ma come una componente strutturale della narrazione, essenziale per definire atmosfera, ritmo e tensione emotiva.
Per quanto riguarda la colonna sonora, la scelta dello strumento di IA da utilizzare è ricaduta sulla piattaforma *Suno*, estremamente efficace nella generazione automatizzata di tracce musicali personalizzabili. La piattaforma consente di inserire parole chiave tematiche, definire se la traccia debba essere vocale o solamente strumentale e, successivamente, selezionare alcuni parametri tra cui il genere musicale, il tempo, la dinamica e l'intensità desiderata. L'interfaccia semplice e intuitiva rende la piattaforma adatta sia a utenti esperti sia a chi si approccia per la prima volta alla creazione musicale tramite IA. Un aspetto particolarmente vantaggioso è rappresentato dal fatto che, per ogni richiesta, la piattaforma genera due versioni differenti dello stesso brano, permettendo all'utente di effettuare una scelta tra le due varianti.
Dal lato accademico, sono presenti sicuramente progetti degni di nota, come lo studio condotto da Sulun et al. [32]. Si introduce *EMSYNC*, un modello di generazione musicale basato su video che allinea la musica alle emozioni e ai confini temporali delle scene. Il sistema opera in due fasi: un classificatore del video estrae caratteristiche emotive e un modello di generazione musicale condizionato produce una colonna sonora in formato *MIDI*, guidata sia da indici emotivi sia dai cambi di scena del video. Una novità rispetto i precedenti studi è l'uso di *boundary offsets* per anticipare i cambi di scena inserendo accordi musicali in corrispondenza dei tagli video, migliorando notevolmente la sincronizzazione tra musica e immagine. In test di ascolto soggettivi, *EMSYNC* ha superato gli studi precedenti su tutti i criteri valutati, rendendolo attualmente il migliore nel settore.
Non potendo utilizzare strumenti simili a causa dell'alta richiesta computazionale necessaria, si è optato per comparare altre piattaforme presenti sul mercato, come *Suno*, *Soundraw* e *Aiva*. *Suno* ha mostrato maggiore precisione nell'aderenza al prompt e una resa qualitativa maggiore. *Aiva* tende a privilegiare uno stile orche-

strale classico poco flessibile, difficilmente adattabile a contesti narrativi non epici, mentre *Soundraw* spesso genera loop musicali ripetitivi, meno adatti a strutture narrative complesse.
Per quanto riguarda il voiceover dell'artista, l'obiettivo era ottenere una voce femminile espressiva e coerente con il personaggio, mantenendo al contempo la lingua italiana per favorire l'immediata comprensione da parte del pubblico. L'italiano, tuttavia, è ancora supportato parzialmente nei principali modelli *text-to-speech* (TTS), i quali risultano generalmente ottimizzati per l'inglese e solo parzialmente adattabili a fonetiche differenti. Di conseguenza, si è optato per una soluzione ibrida: la voce è stata inizialmente registrata dall'autore umano in lingua italiana, per poi essere trasformata da maschile a femminile tramite la clonazione della voce.
Anche in questo campo la ricerca ha prodotto benchmark dedicati. Christop et al. [33] propongono *ClonEval*, che definisce un protocollo standard per valutare quantitativamente vari modelli di clonazione della voce. I risultati mostrano dati interessanti. Su set di test in inglese, ad esempio, diverse piattaforme ottengono punteggi di similarità molto alti, a prova del fatto che riescono effettivamente a trasferire, con buona fedeltà, il timbro del parlato di riferimento nella voce sintetizzata. In particolare, il modello *XTTS-v2* ha ottenuto il punteggio medio complessivo più alto, mentre su alcuni set specifici il modello *SpeechT5* di *Microsoft* era leggermente superiore. Si è osservato però che clonare voci emotive è più difficile. Sui dataset di parlato emozionale i punteggi sono leggermente più bassi rispetto al parlato neutro, suggerimento del fatto che i modelli attuali funzionano meglio nel riprodurre timbri in condizioni standard, mentre certe inflessioni emotive o espressive restano una sfida.
Tra le varie piattaforme testate, *ElevenLabs* si è distinta per la qualità elevata dei suoi modelli di clonazione vocale e per la capacità di mantenere l'espressività originaria della voce sorgente, oltre ad offrire un ampio supporto per lingue differenti. Il sistema consente inoltre una personalizzazione accurata delle caratteristiche timbriche e delle emozioni veicolate, con opzioni avanzate. Altri strumenti simili hanno mostrato limitazioni rilevanti per quel che riguarda l'accuratezza nel cambio di genere o la resa sintetica, compromettendo l'effetto di naturalezza della voce.
Per la generazione degli effetti sonori si è nuovamente fatto ricorso a *ElevenLabs*, che tra le sue funzionalità include anche la possibilità di generare brevi segmenti audio, perfettamente sincronizzabili con la componente visiva.
Gli effetti sonori sono essenziali per sottolineare passaggi narrativi chiave, enfatizzare determinati movimenti o evocare ambienti sonori specifici, andando a rafforzare stati emotivi o transizioni visive, guidando l'attenzione dello spettatore.
L'integrazione dell'intelligenza artificiale nei processi audio del cortometraggio ha comportato una fase iniziale di sperimentazione estensiva, seguita da una selezione ponderata degli strumenti più adatti. L'approccio combinato tra input umano e generazione automatica ha garantito sia il controllo artistico che la flessibilità produttiva, confermando l'efficacia dell'IA nelle fasi di post-produzione sonora. L'intervento umano ha continuato a rivestire un ruolo importante nella supervisione, nel raffinamento e nell'adattamento del materiale prodotto, confermando la natura collaborativa e non sostitutiva dell'intelligenza artificiale nel contesto creativo.

2.3.8 IA per la post-produzione

Fino a questo punto l'IA ha avuto un ruolo centrale nella produzione del cortometraggio. Per quanto riguarda la post-produzione, nonostante siano stati presi in esame diversi strumenti, si è scelto di non utilizzarli per la realizzazione del prodotto finale. In particolare, il processo di selezione, taglio e montaggio delle clip, potrebbe essere supportato da soluzioni come *Adobe Premiere Pro AI* e *Descript* che offrono algoritmi avanzati, capaci di individuare i momenti salienti, suggerire sequenze coerenti e perfino generare versioni multiple dello stesso contenuto a seconda dell'output desiderato. Si è osservato che tali strumenti, sebbene utili per contenuti verticali o a scopo commerciale, faticano a sostenere una logica narrativa più complessa come quella di un cortometraggio strutturato. In molti casi, la selezione automatica delle clip si è rivelata priva di coerenza e incapace di restituire le giuste pause, i ritmi e le alternanze visive pensate dall'autore umano. Per questo motivo, l'IA è stata usata principalmente in fase esplorativa e nella creazione della prima bozza, ma non per il montaggio definitivo, preferendolo manuale.
Un ulteriore ambito applicativo è stato quello della *color correction e grading*. L'utilizzo di strumenti dotati di algoritmi predittivi per la correzione automatica del colore, ha permesso di normalizzare le clip, correggere gli eventuali sbilanciamenti cromatici e uniformare l'estetica complessiva. Sebbene la resa finale sia stata in parte modificata manualmente per ottenere un risultato più personale e omogeneo, l'uso di queste tecnologie ha fornito certamente una base solida per la continuazione del processo.
L'impiego dell'IA in post-produzione non ha avuto lo scopo di sostituire l'operatore umano, ma piuttosto di affiancarlo e supportarlo in compiti ripetitivi o che richiedono tempo. La possibilità di testare diverse soluzioni in tempi brevi, combinata con il crescente livello qualitativo degli strumenti disponibili, ha reso la post-produzione più efficiente e flessibile, senza compromettere la visione autoriale umana del progetto.

2.4 Analisi SWOT delle soluzioni IA disponibili

L'impiego dell'intelligenza artificiale sta rivoluzionando la filiera produttiva audiovisiva, dalla pre-produzione fino alla post-produzione. Questa evoluzione offre numerosi vantaggi ma anche limiti intrinseci, oltre a nuove opportunità creative e problemi di carattere tecnico, sociale ed etico. Di seguito si presenta un'analisi *SWOT* completa dell'IA nella produzione audiovisiva, evidenziando punti di forza, debolezze, opportunità e minacce di questa tecnologia.

2.4.1 Punti di forza

I principali punti di forza sono sicuramente la possibilità di automatizzare e rendere efficienti i processi e ridurre i costi della produzione di un prodotto audiovisivo. L'IA eccelle in generale nell'automatizzare i compiti maggiormente ripetitivi, accelerando i flussi di lavoro. Ad esempio, algoritmi avanzati possono svolgere in pochi minuti operazioni di editing video che richiederebbero ore di lavoro manuale, come tagliare scene o applicare correzioni colore dettagliate. Questo comporta tempistiche di produzione più rapide e costi operativi inferiori, permettendo agli autori umani di

concentrarsi maggiormente sugli aspetti creativi. La possibilità di sfruttare strumenti basati sul *machine learning* permette di recuperare filmati a bassa risoluzione e generare fotogrammi mancanti, attività chiave nel campo del restauro di vecchi video, spesso parte integrante di produzioni odierne. Il potenziamento della creatività è un altro aspetto importante: l'IA non si limita ad accelerare i processi, ma aiuta sempre più gli autori e artisti nella parte di ideazione.
Da un punto di vista visivo invece, generatori di immagini e video basati su IA consentono di realizzare storyboard, concept art, effetti o animazioni complesse a partire da semplici descrizioni testuali, fornendo spunti visivi immediati ai registi. L'IA amplia quindi le capacità creative umane, offrendo nuovi strumenti per sperimentare con rapidità idee narrative e stilistiche. Un altro effetto positivo dell'IA è la riduzione delle barriere d'accesso alla produzione di prodotti audiovisivi. Pratiche un tempo riservate a grandi studi o specialisti altamente qualificati del settore sono oggi alla portata di piccoli gruppi o singoli autori grazie proprio a strumenti di intelligenza artificiale. Questo significa che effetti speciali avanzati, animazioni e altre tecniche normalmente molto costose possono essere realizzate a costi inferiori e con meno competenze specialistiche rispetto al passato. Una diretta conseguenza è la maggiore varietà di contenuti, poiché più artisti provenienti da contesti differenti, possono realizzare la propria visione creativa grazie all'ausilio dell'IA.
Questa tecnologia può essere utilizzata anche per prevedere trend e preferenze, supportando i produttori nelle scelte artistiche e di marketing. Ad esempio, alcuni studi cinematografici impiegano sistemi di IA per valutare la probabilità di successo di una sceneggiatura o di un progetto. Ad esempio, *Warner Bros* ha adottato un modello di IA per stimare l'incasso potenziale di un film analizzandone lo script, così da facilitare le decisioni sui nuovi progetti. Queste applicazioni guidate dai dati rappresentano un punto di forza altamente strategico: l'IA fornisce informazioni predittive e automatizza la gestione dei contenuti, contribuendo a rendere l'industria audiovisiva più informata, mirata ed efficiente nelle scelte.

2.4.2 Punti di debolezza

Il principale punto di debolezza riscontrato è la scarsa originalità creativa e l'alto tasso di contenuti derivativi. Questo è dovuto al fatto che i modelli utilizzati non possiedono una reale creatività o intuito, ma generano output basandosi su schemi appresi dai dati di addestramento. Questa mancanza di originalità può costituire un grosso svantaggio soprattutto in ambiti pubblicitari dove l'unicità e l'inventiva sono fondamentali per catturare l'attenzione del consumatore.
Un altro aspetto critico rilevato è la totale mancanza di comprensione contestuale ed emozionale: i modelli operano infatti elaborando pattern statistici nei dati, ma senza comprendere il contesto culturale e narrativo. Questa carenza si traduce in output che talvolta mancano di profondità e sfumature. Studi del settore evidenziano come le componenti generate dall'IA spesso faticano a trasmettere le sfumature emotive, la profondità e il contesto culturale, tipici invece dei prodotti interamente creati da autori.
Risulta inoltre difficile controllare e prevedere i risultati in quanto questi sistemi operano spesso come scatole nere di cui è difficile conoscere esattamente il funzionamento. Questo comporta una significativa mancanza di strumenti di controllo sul risultato generato e chi utilizza l'IA per creare dei contenuti può solo influenzare

il modello attraverso prompt testuali o parametri generali, ricevendo però spesso esiti discutibili. Questa imprevedibilità rende l'uso dell'IA parecchio inefficiente se utilizzata per effettuare piccole correzioni. È complesso infatti applicare modifiche puntuali visto che il controllo dell'utente sull'output è indiretto e limitato. Inoltre, i modelli possono avere comportamenti non intenzionali come glitch visivi o incoerenze nel video, difficili da eliminare a priori.
Bisogna inoltre essere consapevoli del fatto che l'IA è affidabile tanto quanto lo sono i dati su cui è stata allenata. Se i dataset contengono pregiudizi, stereotipi o semplicemente non sono rappresentativi di certe realtà, l'IA tenderà a riprodurre quei bias nei contenuti che genera. C'è inoltre un tema di qualità generale dei dati utilizzati: per ottenere risultati validi servono grandi quantità di informazioni audiovisive di qualità. Non tutte le aziende che si occupano di IA hanno accesso a ricchi dataset per addestrare modelli personalizzati, limitando di conseguenza l'efficacia dell'IA in contesti specifici.
Nonostante i rapidi progressi degli ultimi anni, molte applicazioni di IA per la generazione di video sono ancora in una fase embrionale. Gli strumenti disponibili spesso non hanno funzionalità avanzate e presentano limitazioni significative. Clip brevi, di qualità visiva modesta, lontane dalla complessità di un prodotto cinematografico sono quanto riescono a produrre alcuni modelli. Anche le piattaforme più avanzate non consentono ancora di generare automaticamente una scena coerente, composta da più inquadrature, lasciando al massimo la possibilità di produrre brevi sequenze sperimentali. Chi adotta oggi l'IA durante qualsiasi fase della produzione deve accettare di usarla principalmente come tecnologia di collaborazione, mettendo in conto che dovrà sicuramente fare del lavoro aggiuntivo per sopperire a ciò che gli strumenti non sono ancora in grado di fare. Questa fase, che verrà sicuramente superata col tempo, oggi costituisce certamente un punto debole di questa tecnologia.

2.4.3 Opportunità

L'IA sta aprendo nuove opportunità per rendere i contenuti audiovisivi fruibili con facilità su scala internazionale, permettendo di doppiare o sottotitolare interi prodotti audiovisivi in lingue diverse, a costi contenuti. Anche le produzioni minori possono quindi raggiungere Paesi e utenti differenti con estrema facilità. Un altro effetto positivo è anche l'inclusione di lingue meno diffuse e di culture poco conosciute. Il progetto europeo *MOSAIC* [34] mira proprio a sfruttare l'IA per promuovere la diversità linguistica nei media, automatizzando doppiaggi e sottotitoli di alta qualità per lingue poco rappresentate. In questo modo è chiara l'opportunità di ampliare il mercato, arricchendo l'offerta culturale globale. Non da ultimo, le stesse tecnologie possono migliorare l'accessibilità a persone con disabilità: generando sottotitoli per non udenti o descrizioni audio per non vedenti in maniera automatica e scalabile è possibile rendere fruibili sempre più contenuti.
L'adozione dell'IA apre le porte a modalità inedite di creazione di contenuti audiovisivi, favorendo sperimentazioni prima impraticabili. I creativi possono usare modelli generativi per prototipare rapidamente idee di sceneggiatura, visualizzare concept di scene o provare diversi stili visivi senza dover per forza girare realmente parte del filmato. Ad esempio, un regista può descrivere a un generatore di immagini l'ambientazione di una scena e ottenere in pochi secondi uno storyboard disegnato automaticamente, utile per valutare l'idea e comunicarla alla produzio-

ne. Questa rapidità di iterazione creativa consente di esplorare molte più opzioni in fase di sviluppo, portando potenzialmente a storie più originali o soluzioni visive innovative.
Un altro aspetto altamente innovativo risiede nella possibilità di rendere effettivi contenuti narrativi emergenti come quelli interattivi o personalizzati. Basti pensare alla possibilità di film o episodi di serie TV il cui montaggio si adatta in tempo reale ai gusti dello spettatore, rilevati in tempo reale attraverso algoritmi di raccomandazione, creando così esperienze narrative su misura. Alcune sperimentazioni hanno già avuto luogo in ambito pubblicitario, dove video generati dall'IA vengono mostrati con elementi diversi a seconda del pubblico presente.
Dal punto di vista industriale, un'enorme opportunità offerta dall'IA è la scalabilità della produzione di contenuti. Le piattaforme di streaming, in particolare, hanno un bisogno crescente di nuovi prodotti per soddisfare un pubblico sempre più esigente e vario e l'IA può aiutare a generare i video, permettendo di completare i cataloghi di offerte, ottimizzando tempi e costi. Questo potrebbe tradursi in una maggiore varietà di contenuti disponibili sul mercato, con benefici sia per gli autori sia per il pubblico, che andrebbe a beneficiare di un'offerta più ampia. Ovviamente, l'aumento di produttività deve essere gestito senza compromettere la qualità e la creatività e quindi anche per questo motivo un sistema collaborativo uomo-macchina è sicuramente il più adatto. Questa sinergia, se ben sfruttata, costituirà un vantaggio competitivo e un'opportunità di crescita significativa per le aziende del settore audiovisivo.
Questa forte integrazione nei processi produttivi sta richiedendo collaborazioni tra professionisti di settori molto diversi, come ad esempio sviluppatori software, esperti di dati, creativi e artisti del video, andando a formare gruppi di lavoro eterogenei. Questo aspetto può portare ad una contaminazione reciproca, con la conseguente possibilità di trovare soluzioni innovative e far emergere nuovi ruoli professionali. Investire quindi in queste competenze rappresenta un'opportunità chiave per far evolvere l'industria nel suo complesso. Ad esempio, festival dedicati all'IA mettono in contatto creativi tradizionali con esperti di *machine learning*, permettendo la proliferazione di progetti audiovisivi complessi e maggiormente strutturati da un punto di vista di modelli utilizzati. Questa tendenza suggerisce che l'IA può fungere da catalizzatore di innovazione tecnologica anziché sostituire l'uomo e può stimolare gli autori ad apprendere nuovi strumenti, mescolare linguaggi e trovare modi originali di raccontare storie attraverso strumenti informatici mai usati prima.
Infine, un altro campo di enorme applicazione è quello del restauro e della valorizzazione del patrimonio audiovisivo. Algoritmi basati su reti neurali sono in grado di eliminare graffi e rumore da vecchie pellicole, stabilizzare immagini che tremolano, ricostruire fotogrammi danneggiati e persino rendere a colori film originariamente in bianco e nero. Queste sono operazioni che normalmente richiederebbero un lungo lavoro manuale di esperti restauratori ma che possono essere svolte dall'IA semi automaticamente, ridando nuova vita a contenuti altrimenti deteriorati o difficilmente fruibili. È possibile così valorizzare cataloghi già esistenti, raggiungendo sia nostalgici che le nuove generazioni curiose di scoprire capolavori del passato. L'IA offre l'opportunità non solo di guardare al futuro con nuovi contenuti, ma anche di riscoprire il passato, arricchendo l'offerta culturale e proteggendo la memoria audiovisiva.

2.4.4 Minacce

Una delle minacce più discusse legate all'IA nel settore audiovisivo riguarda la possibile perdita di posti di lavoro o la trasformazione radicale di ruoli professionali tradizionali. L'automazione di diverse mansioni solleva il timore che molti tecnici e creativi possano essere sostituiti dalle macchine. Già oggi alcune produzioni impiegano meno montatori o grafici perché determinati compiti vengono svolti dall'IA. Figure come assistenti, montatori, tecnici del suono e tante altre potrebbero vedere una significativa riduzione della domanda man mano che gli strumenti automatici diventano più sofisticati. Questo scenario genera comprensibili resistenze da parte degli addetti del settore. Negli ultimi anni associazioni, come ad esempio la *Writers Guild of America*, hanno messo in atto proteste e negoziati per regolamentare l'uso dell'IA, nel tentativo di tutelare il lavoro umano. Se non gestita con attenzione, l'adozione su larga scala dell'IA potrebbe dunque portare a conflitti e instabilità nell'industria audiovisiva, con professionisti che si sentono minacciati e un possibile impoverimento del mercato del lavoro. Sarà necessario monitorare come evolverà il rapporto tra umano e IA, predisponendo misure per mitigare l'impatto negativo, favorendo ad esempio la riqualificazione professionale, attraverso corsi specializzati.
L'omologazione dei contenuti e la conseguente riduzione dell'originalità artistica che potrebbe ridurre significativamente la diversità stilistica e creativa, è una preoccupazione largamente diffusa. Se produttori e piattaforme si affideranno sempre più ai modelli predittivi per decidere quali storie raccontare o come strutturarle, c'è il rischio di ottenere sempre gli stessi schemi narrativi già collaudati, penalizzando sperimentazione e innovazione. Per quanto riguarda l'impatto sul pubblico i contenuti generati da IA rischiano di avere tutti la stessa estetica, dettata dai modelli. La minaccia di standardizzazione riguarda infatti anche l'estetica visiva: se tutti utilizzano gli stessi modelli pre-addestrati sugli stessi dati di partenza, è possibile che molti prodotti finiscano per condividere uno stile visivo comune, riducendo la varietà dei linguaggi cinematografici.
Un altro grande problema riguarda l'etica; deepfake e disinformazione saranno sempre più al centro del dibattito pubblico in quanto l'IA ha ormai dimostrato di poter riprodurre la realtà con grande efficacia. Questa capacità è ovviamente un'arma a doppio taglio; se da un lato offre opportunità creative maggiori, dall'altro costituisce una minaccia seria in termini di possibili abusi. Video falsi potrebbero essere usati per diffondere disinformazione oppure per diffamare o violare la privacy. Tutto questo rappresenta una minaccia anche per l'intero sistema mediatico. In un mondo saturo di contenuti artefatti, il pubblico potrebbe perdere fiducia in ciò che vede, mettendo a rischio l'affidabilità del video.
Ci si domanda poi se sia giusto utilizzare queste tecnologie per far rivivere, in nuovi prodotti audiovisivi, attori del passato oppure modificarne l'età di quelli in vita. Questi e altri interrogativi stanno emergendo prepotentemente man mano che la tecnologia progredisce.
L'uso dell'IA nella creazione audiovisiva pone problemi senza precedenti in ambito legale, in particolar modo nella determinazione di chi possieda effettivamente i diritti delle opere generate e nella gestione di eventuali violazioni di copyright per i dati utilizzati nell'addestramento. Attualmente la legge sul diritto d'autore spesso non riconosce un autore legale, lasciando un vuoto normativo. Questa situazione crea insicurezza sia per gli artisti che utilizzano l'IA sia per i produttori, che rischiano

controversie legali sulla distribuzione di contenuti generati da modelli di IA. La mancanza di regole chiare su quanto si possa considerare originale e su come evitare violazioni di copyright rappresenta dunque una minaccia legale significativa e le case di produzione che utilizzano intensivamente l'IA potrebbero trovarsi coinvolte in lunghe battaglie giudiziarie.
Un aspetto da non sottovalutare è la possibile reazione negativa degli spettatori a contenuti percepiti troppo artificiali. Quando in un film compare un personaggio digitale molto realistico, ma non perfettamente identico a un attore tradizionale, parte del pubblico avverte un senso di straniamento noto come *uncanny valley*. Se l'uso dell'IA porterà a un utilizzo maggiore di volti sintetici o ricostruzioni di attori, c'è il rischio che il pubblico non li accolga positivamente, soprattutto qualora la tecnologia non riesca a superare completamente la barriera tra realistico e artificiale. Di fronte a questa problematica, l'industria dovrà prestare attenzione a come presentare l'utilizzo dell'IA, accettando che un film interamente girato con le nuove tecnologie non venga accolto allo stesso modo di un prodotto tradizionale. Risulta quindi importante educare il pubblico al valore aggiunto che l'IA può portare, permettendo di realizzare cose altrimenti difficili o impossibili.
Infine, una particolare attenzione deve essere posta alla questione della dipendenza tecnologica. Le produzioni audiovisive potrebbero diventare vulnerabili ai cambiamenti dei costi di questi servizi di IA, oppure a problematiche tecniche come le interruzioni nei servizi online. Per provare a mitigare queste minacce, sarà meglio investire in soluzioni di IA proprie, adeguandosi da un punto di vista della sicurezza informatica.

Capitolo 3

Pipeline di produzione del progetto

3.1 Fase di Pre-produzione

3.1.1 Raccolta dati

La fase di pre-produzione del progetto è stata avviata da un'attività di analisi e confronto preliminare con gli enti promotori del progetto, in particolare l'associazione *LUCAS*. L'obiettivo era quello di definire un quadro metodologico valido per la raccolta dei dati durante gli eventi pubblici, assicurandosi che le informazioni raccolte fossero utili sia a *Narrazioni Parallele* che alla specifica finalità della tesi.
Sono stati inizialmente presentati gli eventi che avrebbero permesso la raccolta dei dati e successivamente è stata spiegata la scelta dell'utilizzo della piattaforma interattiva *Mentimeter* per la raccolta delle risposte del pubblico in tempo reale. Si è reso necessario quindi esaminare attentamente le domande che sarebbero state poste ai partecipanti, con la possibilità di suggerirne di nuove o modificarne alcune già presenti, in modo tale da orientarle maggiormente alle necessità narrative e progettuali della tesi. Dopo un'attenta valutazione, è stato aggiunto solamente un nuovo quesito ritenuto particolarmente significativo per la costruzione della sceneggiatura. La domanda inserita è la seguente:

Quale esperienza di vita personale passata ti ha ricordato l'ascolto?

Questa scelta è stata motivata dalla volontà di esplorare la dimensione autobiografica e mnemonica dell'ascolto musicale, con l'intento di raccogliere input che potessero stimolare una narrazione più profonda ed emotivamente coerente con la struttura di un cortometraggio. La raccolta dati serve a fornire una serie di contenuti e spunti significativi che vanno ad alimentare la fase di scrittura e sviluppo della sceneggiatura.
Durante il primo dei due eventi in cui sono stati raccolti i dati, dal titolo *Una passeggiata creativa*, le domande sono state poste al termine dell'esibizione musicale, senza alcuna interruzione, attraverso la piattaforma *Mentimeter*. L'esibizione è stata concepita non solo come performance, ma come esperienza immersiva per il pubblico, al termine della quale i partecipanti sono stati invitati a condividere le proprie emozioni, immagini mentali e ricordi evocati.
Vengono riportate di seguito alcune domande rivolte al pubblico, mentre si rimanda alla Sezione A dell'appendice per l'elenco completo.

Quali pensieri sono emersi durante l'ascolto. Quali immagini hai visualizzato?
Hai notato, prima e dopo l'ascolto, delle differenze nel tuo umore e stato d'animo?
Scegli un colore che rappresenta le tue emozioni in questo momento.

Passando invece al secondo evento, dal titolo *Song Writing: intelligenza naturale generativa*, la gestione delle domande è stata differente, in quanto alcune sono state poste all'inizio dell'incontro, diverse durante l'esperienza e altre ancora alla fine. L'evento ha trasformato il pubblico da semplice ascoltatore a creatore. Attraverso sessioni di brainstorming e letture condivise, i partecipanti hanno collaborato alla scrittura del testo di una canzone, esplorando temi personali e collettivi. In questa occasione è stato nuovamente richiesto di integrare, a partire da una base già ben strutturata, alcune domande da porre al pubblico. Ci si è concentrati sull'ottenere maggiori informazioni che potessero risultare poi utili per ambientare il video finale. Di seguito vengono riportati i quesiti proposti.

Se dovessi rappresentare la canzone con una stagione, quale sarebbe?
Quale colore assoceresti al tema trattato oggi?
Se la canzone fosse un viaggio, verso quale destinazione ti porterebbe?

Porre al pubblico domande orientate a evocare ambientazioni, sensazioni cromatiche o spaziali è stato essenziale per fornire all'intelligenza artificiale dati di input su cui basare la costruzione ambientale del video. Attraverso le risposte fornite è stato possibile estrarre indicazioni fondamentali per determinare l'atmosfera visiva del prodotto finale, partendo dall'ambientazione delle scene, passando per la determinazione della palette cromatica da utilizzare, fino ad arrivare al tono emotivo da trasmettere. In mancanza di questi elementi, il processo generativo avrebbe rischiato di produrre risultati generici o incoerenti rispetto a quanto è stato percepito durante l'evento. I dati raccolti hanno invece fornito elementi chiari che hanno orientato le scelte creative automatiche dell'IA, permettendo di ottenere una traduzione visiva più significativa e legata profondamente all'esperienza vissuta dal pubblico.

3.1.2 Preparazione dei dati raccolti

Una volta completata la fase di raccolta delle risposte del pubblico tramite la piattaforma *Mentimeter*, è stato necessario preparare i dati, per renderli idonei alla successiva fase. Le informazioni raccolte, seppur estremamente ricche dal punto di vista qualitativo, presentavano una varietà significativa in termini di forma, lunghezza, coerenza linguistica e pertinenza. Per questa ragione, una volta ricevuti i dati grezzi dalla piattaforma, si è proceduto a leggere quanto prodotto, proseguendo con un breve intervento di pulizia, correzione e standardizzazione degli elementi. La prima operazione ha riguardato l'eliminazione di tutte quelle risposte che risultavano fuori contesto, probabilmente a causa di un errore di inserimento da parte del pubblico. Analogamente, riscontri privi di senso compiuto se relazionati alla domanda in esame, sono stati scartati. Successivamente, sono state identificate e rimosse eventuali duplicazioni. In alcuni casi infatti, lo stesso partecipante ha inserito risposte identiche o molto simili, probabilmente nel timore che la prima risposta non fosse stata salvata correttamente dalla piattaforma e in questi casi si è quindi optato per un'aggregazione ragionata, evitando ripetizioni inutili ma mantenendo la rappresentatività del dato. Inoltre, un'attenzione particolare è stata dedicata in

questa fase alla revisione ortografica e grammaticale. Essendo input scritti spesso in tempo reale e in forma spontanea, alcuni testi contenevano refusi, poca o mancanza di punteggiatura, uso incoerente delle maiuscole e piccole sviste linguistiche. Tutti questi elementi sono stati corretti manualmente, con un intervento puntuale ma che ha ovviamente rispettato il tono e l'intenzione espressiva dell'autore originale, senza alterare in alcun modo il contenuto concettuale.

3.1.3 Analisi dei dati raccolti

Dopo la fase di preparazione e normalizzazione dei dati, questi sono stati analizzati qualitativamente con l'obiettivo di avere una visione nell'insieme di quanto ottenuto. I dati, raccolti in forma anonima tramite la piattaforma *Mentimeter*, includevano risposte aperte e chiuse a diverse domande poste al pubblico durante gli eventi del progetto *Narrazioni Parallele*. Le informazioni più rilevanti per il progetto possono essere racchiuse in tre categorie principali: preferenze stilistiche per il video, associazioni emotive e sensoriali (colori o parole chiave) e immagini evocate dall'ascolto musicale.
Le diverse tipologie di grafici prodotti dalla piattaforma a partire dai dati forniti dagli utenti sono di grande aiuto per visualizzare, a primo impatto, le risposte maggiormente significative; d'altra parte però, non rappresentano un buon punto di partenza per l'intelligenza artificiale in quanto, al momento, ancora fatica a comprendere pienamente quanto riportato nei grafici.
La tipologia di grafico più funzionale è certamente quella a nuvola, che permette di visualizzare le risposte più frequenti con un font più grande e una posizione maggiormente centrale rispetto alle altre. Sono poi ovviamente presenti grafici a barre e a torta, più classici, che permettono di capire rapidamente la distribuzione dei dati. Vediamo di seguito alcuni esempi interessanti, riportando il loro relativo grafico.
Per quanto riguarda le preferenze stilistiche del video, le risposte mostrano una sorprendente varietà di indicazioni: da *documentario musicale o astratto*, a *video con immagini reali non astratte*, passando per *esperienze 3D*, *fantascienza* o *narrazione emotiva*. Questo ampio spettro di risposte ha creato una grande difficoltà nel definire il tono, il ritmo visivo e il linguaggio visivo del cortometraggio, evidenziando un'alta aspettativa verso l'interazione tra realtà, simbolismo e immaginazione. Nei capitoli successivi vedremo come l'intelligenza artificiale ha saputo rispondere a questa richiesta.
Un secondo set di dati presi in esempio riguarda l'associazione tra emozioni e colori. I risultati aggregati hanno mostrato una netta prevalenza del blu e dell'azzurro come colori più rappresentativi delle emozioni provate, seguiti da turchese, giallo, arancio e verde. Tali risposte si sono rivelate fondamentali nella definizione della palette cromatica del video, suggerendo una dominante visiva legata anche a particolari stati d'animo.

Figura 3.1: Grafico a nuvola riportante le risposte alla domanda *Quale COLORE associeresti al tema trattato oggi?*

Infine, domande aperte come *Scrivi una, due o tre parole per descrivere le emozioni e le sensazioni che l'ascolto ha stimolato* hanno fornito una serie di parole chiave, proposte poi sotto forma di nuvola, tra cui spiccano *gioia*, *nostalgia*, *malinconia*, *pace*, *relax* e *serenità*, tutte ripetute più volte, seguite da termini più unici ma altrettanto evocativi come *bosco*, *danza* e *montagna*. Questi dati sono stati profondamente utili per identificare concetti e simboli ricorrenti da trasformare in ambientazioni, scene visive o transizioni all'interno del video.

Figura 3.2: Grafico a nuvola riportante le risposte alla domanda: *Scrivi una, due o tre parole per descrivere le emozioni e le sensazioni che l'ascolto ha stimolato.*

È possibile visualizzare tutti i grafici completi nella Sezione B dell'appendice.

3.1.4 Elaborazione dei dati

Questa fase è stata necessaria per rendere le informazioni ottenute maggiormente fruibili all'intelligenza artificiale. Nonostante la tecnologia sia infatti ormai in grado di leggere i grafici e interpretarli, si è preferito formulare i dati in maniera alternativa e sicuramente maggiormente comprensibili dall'IA. La piattaforma *Mentimeter*

produce infatti output molto efficaci per la comprensione umana, come grafici a barre o nuvole di parole; questi formati risultano però poco funzionali se utilizzati direttamente come input per sistemi generativi di intelligenza artificiale.
Bisogna pertanto riportare le informazioni in una forma più strutturata e schematica. Si è scelto così di riformulare i dati in formato tabellare. Tutti i grafici generati automaticamente dalla piattaforma sono stati rimossi, in quanto sono utili solamente all'utente umano ma potrebbero risultare fuorvianti per un sistema di IA. Sono stati invece mantenuti tutti i dati testuali, disposti ordinatamente in tabelle, una per ciascuna domanda proposta al pubblico. Le risposte uguali sono state raggruppate aggiungendo accanto a ogni voce il conteggio della frequenza con cui quella risposta si presentava. Questo ha permesso di trasformare dati sparsi e ridondanti in una tabella ordinata, leggibile e quantificabile.

Parola	**Conteggio**
Blu	10
Azzurro	9
Turchese	4
Bianco	4
Giallo	4
Arancio	3
Celeste	3
Verde	3
Rosa	3
Verde corallino	2
Arancione	2
Ghiaccio	1

Tabella 3.1: Esempio di tabella per riportare le risposte alla domanda: *Quale colore associeresti al tema trattato oggi?*

Per facilitare ulteriormente l'elaborazione delle informazioni nelle successive fasi, si è deciso di ordinarli secondo logiche di frequenza: più un tema o un concetto erano ricorrenti, più venivano posti in evidenza, costituendo delle vere e proprie priorità narrative da proporre all'intelligenza artificiale. In questo modo si è cercato di influenzare l'IA generativa, in modo tale da poter restituire un prodotto maggiormente coerente con le risposte raccolte durante gli eventi.
L'intera operazione di rielaborazione è stata condotta inizialmente tramite un file *Excel*, che ha consentito di aggregare, correggere e ordinare facilmente i dati. Già in questa fase, però, è stata d'aiuto l'IA, in particolare, è stato utilizzato *ChatGPT* per contare in maniera automatica tutte le ripetizioni e ordinarle per frequenza, cercando di inserire direttamente in un'unica corrispondenza tutti i valori simili tra loro. Una volta completata l'organizzazione in *Excel*, si è scelto di esportare il file finale in formato *PDF*. È stata presa questa decisione per garantire una compatibilità più immediata con i sistemi di intelligenza artificiale che gestiscono meglio la lettura di questi documenti rispetto ai formati nativi dei fogli di calcolo. Il file finale, visionabile interamente nell'apposita Sezione C dell'appendice, rappresenta quindi una versione leggibile sia da operatori umani che da sistemi automatizzati, pronta per essere utilizzata nella fase successiva di generazione creativa delle idee.

3.1.5 Definizione di logline, sinossi e soggetto

Conclusa l'elaborazione e la preparazione dei dati raccolti durante gli eventi, si è avviato uno stadio centrale della pre-produzione: l'ideazione narrativa e visiva del prodotto finale. Questa fase rappresenta il naturale collegamento tra i dati forniti dal pubblico e la loro traduzione creativa in un'opera audiovisiva coerente e originale, generata tramite strumenti di intelligenza artificiale. L'obiettivo principale è quello di trasformare gli input raccolti in un'idea strutturata, capace di sostenere una narrazione, mantenendo al tempo stesso la connessione con le suggestioni offerte dai partecipanti. Lo strumento di generazione testuale utilizzato in questo fase è *ChatGPT*, piattaforma di IA fornita da *OpenAI*.

3.1.6 Tecniche utilizzate nella fase di ideazione

La metodologia seguita per l'ideazione ha combinato approcci di brainstorming creativo con tecniche di storytelling avanzato, integrando tutto nei modelli di intelligenza artificiale generativa testuale.
La prima attività si è sviluppata attraverso una sessione intensiva di brainstorming basato sull'analisi dei dati raccolti: colori prevalenti, emozioni e immagini mentali più evocate sono state trasformate in possibili linee narrative. Ad esempio, temi come il *viaggio*, la *natura*, l'*infanzia* e la *trasformazione* sono emersi come i principali cardini attorno ai quali costruire l'universo visivo del cortometraggio.
In particolare, fin da subito si è dimostrata fondamentale una funzionalità messa a disposizione dallo strumento *ChatGPT*: la possibilità di creare una cartella, dedicata interamente al progetto di tesi. All'interno di questo spazio è stato possibile caricare documenti di vario formato, impartire istruzioni precise e salvare delle indicazioni che sarebbero sempre state disponibili all'intelligenza artificiale. I documenti, in particolare, hanno permesso di rendere disponibile una grande quantità di dati in ogni momento, senza dover spiegare continuamente concetti ed evitando di riportare ogni volta le stesse informazioni.
Successivamente, sono stati definiti i primi prompt da fornire all'intelligenza artificiale per generare delle idee che potessero comprendere in parte i temi maggiormente ricorrenti. L'uso di prompt accuratamente descrittivi ha permesso di orientare la generazione delle prime idee in linea con l'immaginario emerso dal pubblico, riducendo il rischio di deviazioni stilistiche o semantiche. In una seconda fase, invece, è stato scelto un approccio maggiormente improntato sul dialogo, per definire meglio l'idea che più si avvicinava a quella collettiva del pubblico. Questa tecnica ha permesso all'intelligenza artificiale di comprendere meglio le parti della storia che dovevano essere modificate e produrre analisi approfondite sul lavoro fin a quel momento svolto.

3.1.7 Modalità operative di generazione della logline

Nel corso della fase di ideazione del progetto, uno degli aspetti fondamentali è stato verificare la reale capacità dell'intelligenza artificiale di leggere, comprendere e analizzare i materiali allegati. Questa verifica è avvenuta inizialmente attraverso la formulazione di domande specifiche, pensate per testare l'effettiva comprensione delle risposte provenienti dal pubblico e contenute all'interno del file preparato. Questo è stato fatto, ad esempio, chiedendo quale fosse il colore maggiormente votato tra

quelli proposti nel sondaggio. L'intelligenza artificiale ha risposto correttamente individuando il colore più scelto e dimostrando così di saper accedere correttamente ai dati e restituire informazioni pertinenti. Questa verifica ha rappresentato un passaggio preliminare necessario per poi passare alla fase creativa vera e propria.
È stato definito infatti un primo prompt orientato alla generazione di logline e idee per un video da realizzare, riportato di seguito.

> Bene, ora devi iniziare a ragionare e agire come se fossi uno sceneggiatore di alto livello, per scrivere assieme una sceneggiatura per un cortometraggio. Dovrai inizialmente dare tante idee, che andremo poi a sviluppare assieme. Iniziamo sicuramente scrivendo 10 log line differenti con 10 diverse idee da sviluppare in un secondo momento.

Prompt 3.1: Prompt utilizzato per la fase di brainstorming iniziale.

A partire da questo prompt, si è chiesto all'intelligenza artificiale di sviluppare diversi cicli creativi distinti. Ogni ciclo consisteva nella generazione autonoma di nuove logline e suggestioni narrative, ciascuna potenzialmente ispirata ai dati contenuti nel file con tutte le risposte date durante i concerti partecipativi. In questo modo è stato possibile ottenere una gamma diversificata di idee, esplorando più strade narrative e tematiche ma, allo stesso tempo, aumentando le probabilità di individuare, in modo coerente, le suggestioni emotive e visive espresse dal pubblico.
In parallelo, per ogni idea generata, l'IA è stata incaricata di indicare esplicitamente quale riferimento all'interno del file di risposte del pubblico avesse ispirato la proposta: se una parola chiave, immagini ricorrenti, preferenze estetiche o stati emotivi. Questo lavoro di ricondurre ogni output creativo a un input partecipativo ha consentito di mantenere una coerenza profonda tra i contenuti generati e il materiale di partenza, rafforzando il legame tra dimensione collettiva e produzione narrativa.
Tuttavia, la generazione di idee non si è fermata a questo. In uno stadio successivo, infatti, si è deciso di invertire il flusso creativo, procedendo con un'altra fase di sperimentazione: è stato proposto all'intelligenza artificiale un nuovo approccio in cui era l'umano a fornire direttamente una logline o un'idea narrativa, mentre l'IA doveva cercare nel materiale partecipativo i riferimenti a supporto delle proposte pervenute. In questo scenario, l'intelligenza artificiale ha assunto il ruolo di controllore, verificando in che modo ciascuna idea umana potesse ritrovare corrispondenze autentiche nei dati raccolti. Questo cambio di paradigma ha permesso di sperimentare una modalità collaborativa più simmetrica, in cui la creatività umana e la capacità analitica dell'intelligenza artificiale operavano in modo complementare.
A completamento di questo processo, è stata attivata una fase di selezione manuale che ha avuto il compito di scegliere tra le idee emerse quelle più interessanti e promettenti da sviluppare ulteriormente. Al termine di alcune iterazioni di questo doppio processo, è stata individuata tra le altre una logline considerata la più efficace, nata dall'unione e dalla sintesi delle due modalità descritte. In Output 3.1 si riporta il filone narrativo scelto per sviluppare il cortometraggio.

Una giovane artista, ispirata dalla vitalità della città, si perde tra murales pulsanti di vita e opere che sembrano respirare, per poi lasciarsi chiamare dalla vastità incontaminata della natura, intraprendendo un viaggio che fonde l'arte urbana con il richiamo selvaggio dell'ambiente.

Output 3.1: Logline scelta per proseguire con la fase di pre-produzione.

3.1.8 Modalità operative di generazione della sinossi

Tale logline è stata quindi inserita all'interno di un nuovo prompt, Prompt 3.2, più articolato e preciso, che è servito a definire la sinossi del cortometraggio:

Sviluppo del concept: Definisci chiaramente la trama centrale, il conflitto e i personaggi principali. Concentrati su un'idea semplice, chiara e potente. Ricorda di utilizzare tutte le tecniche di storytelling che conosci per coinvolgere lo spettatore. Mi piacerebbe che siano in qualche modo sfruttate a pieno le opere con cui l'artista interagisce. Agisci come un importante regista cinematografico che applica tutta la propria conoscenza.

Prompt 3.2: Prompt preciso per la creazione del concept a partire dalla logline che si è scelto di utilizzare.

Questo prompt ha segnato l'avvio di una fase di sviluppo narrativo approfondito, nella quale l'intelligenza artificiale è stata incaricata di espandere l'idea inizialmente scelta, in tutte le sue componenti strutturali. Sono stati così elaborati i temi centrali della storia, i conflitti narrativi, la caratterizzazione dei personaggi principali, nonché le scelte stilistiche e tecniche utili alla resa audiovisiva del racconto. Questo processo ha trasformato una suggestione narrativa in una vera e propria architettura audiovisiva, coerente e contestualizzata rispetto al patrimonio immaginativo del pubblico. Da notare come sia stato chiesto all'IA, nel prompt, di comportarsi come un regista cinematografico in quanto questa richiesta aumenta la consapevolezza dell'ambito in cui si sta muovendo e permette di reperire in maniera più precisa le informazioni di cui ha bisogno. Un altro aspetto importante è sicuramente la scelta di inserire fin da subito l'elemento chiave all'interno della storia, ovvero il rapporto tra l'artista protagonista e le sue opere d'arte. Il concept prodotto ha costituito una solida base su cui si è avviato un nuovo ciclo di dialogo e revisione. In questa fase, si è lavorato per perfezionare lo sviluppo della storia, intervenendo sulle parti che risultavano meno efficaci o meno coerenti con l'impostazione desiderata. Uno degli interventi più significativi ha riguardato la scelta di rendere anonimo il personaggio protagonista, evitando l'attribuzione di elementi troppo specifici. Tale scelta ha risposto a una precisa esigenza di universalità e accessibilità emotiva, permettendo in questo modo allo spettatore di identificarsi più facilmente nella figura protagonista del racconto, con le proprie esperienze e vissuti.
Proseguendo con lo sviluppo narrativo, c'è stato un grande lavoro di esplorazione e perfezionamento per ridefinire la sinossi precedentemente generata. Per facilitare

questo passaggio e consentire un'interazione più diretta con l'intelligenza artificiale, è stata creata una nuova chat operativa sempre all'interno dello stesso progetto, utilizzando *ChatGPT*. A partire da questa, è stato sfruttato lo strumento di riconoscimento vocale integrato per effettuare qualche modifica alla storia già esistente: questa operazione ha permesso di importare molto rapidamente il testo del prompt direttamente in formato testuale. L'IA ha provveduto a una prima correzione della narrazione, organizzandone la struttura e rendendo la storia raccontata più coerente, scorrevole ed equilibrata. Il continuo dialogo con l'umano è stato fondamentale per arrivare al risultato finale: attraverso una serie di suggerimenti, idee e proposte di miglioramento, sono state effettuate varie modifiche al concept inizialmente proposto, controllando ogni volta che la coerenza, il tono e l'evoluzione venissero rispettati con quanto già impostato.
Questa interazione è proseguita in modo iterativo, portando alla proposta e rielaborazione di un finale concepito in origine esclusivamente dalla mente umana. Anche in questo caso, lo sviluppo è stato portato avanti non solo da considerazioni personali ma anche dai vari spunti emersi durante la fase di brainstorming.
In questo momento si è posta una particolare attenzione all'analisi del *viaggio dell'eroe*, uno degli schemi archetipici più utilizzati nella costruzione di storie dal forte impianto narrativo. La richiesta posta all'intelligenza artificiale di esaminare con attenzione la corrispondenza della trama con le tappe canoniche del *viaggio dell'eroe* è nata dal desiderio di garantire solidità alla struttura narrativa ma anche profondità al percorso della protagonista. Dall'analisi effettuata è emerso che la prima parte della storia risultava ben definita, con una *chiamata all'avventura* riconoscibile ma, allo stesso tempo, è stato rilevato che il *rifiuto della chiamata* era presente ma implicito. Sono state quindi suggerite in tal senso alcune modifiche per rinforzare la parte di racconto più debole. Le altre fasi dello schema narrativo, invece, sono apparse meglio marcate e stilisticamente coerenti, suggerendo che la struttura complessiva fosse già sufficientemente efficace.
Successivamente sono stati presi in esame il ritmo interno della storia, l'equilibrio tra momenti di tensione e distensione e l'originalità delle scelte narrative adottate. Durante questo stadio, è stato anche richiesto di analizzare i punti di forza e di debolezza della trama fino a quel momento sviluppata, con il principale obiettivo di preservare la coerenza stilistica all'interno del video.
Si lasciano di seguito due concept, rispettivamente Output 3.2, ovvero quello inizialmente generato a partire dalla logline e Output 3.3, l'ultimo prodotto in seguito alle varie richieste. La versione finale della sinossi viene invece riportata nella Sezione D dell'appendice.

Titolo provvisorio: L'Arte che Respira
Trama Centrale:
La storia segue Elena, una giovane e brillante artista urbana, che si trova a un bivio esistenziale: la frenesia e l'energia vibrante della città, con le sue strade animate e le opere d'arte che pulsano di vita, contrastano con il richiamo irresistibile della natura incontaminata. Mentre attraversa quartieri caratteristici, Elena interagisce con murales e installazioni che, in modo quasi magico, sembrano respirare e raccontare storie. Queste opere, animate dalla sua immaginazione e da una misteriosa sinergia, divengono veri e propri personaggi che riflettono i suoi stati d'animo, i suoi sogni e le sue paure.

Output 3.2: Primo concept generato dall'IA. In seguito è stato sviluppato grazie alla continua interazione con l'autore umano.

Titolo provvisorio: L'Artista e la Città
Nell'atelier di un quartiere artistico della città, un'artista sui trentacinque anni si muove tra le sue tele, sculture e schizzi sparsi ovunque. Il disordine del suo studio riflette il suo stato d'animo: è irrequieta, combattuta. Da tempo avverte una crescente insoddisfazione, una sensazione di distacco dalla vita che ha costruito. Lavora, espone, riceve apprezzamenti, ma dentro di sé qualcosa le manca.
Davanti a lei, una tela incompiuta cattura la sua attenzione. È un paesaggio, ma ancora indefinito, come se mancasse un pezzo essenziale. Lo osserva a lungo, quasi ipnotizzata, come se fosse lì che risiedesse la risposta ai suoi dubbi. Sa che quel quadro è importante, ma ancora non capisce perché.
Sopraffatta dai pensieri, decide di uscire e camminare per la città, ripercorrendo le tracce delle sue opere. I suoi murales colorano i muri grigi, le sue sculture moderne popolano le piazze, le sue installazioni interagiscono con il paesaggio urbano. È la sua firma, il suo segno nella città. Ma mentre cammina, qualcosa di straordinario accade: le sue opere iniziano a prendere vita.
Un murale di una foresta si anima, i rami si muovono come scossi dal vento. Una scultura di un fiume comincia a scorrere, trasformandosi in acqua vera. Un dipinto che raffigura una montagna si espande, avvolgendola completamente. Il confine tra arte e realtà si dissolve. Lei si ritrova immersa nelle sue creazioni, trasportata da un'opera all'altra in un viaggio che è al tempo stesso fisico e interiore.
Naviga tra i suoi stessi ricordi, tra immagini della sua infanzia in montagna e la vita urbana che ha scelto. Ogni opera sembra parlare con lei, mostrandole frammenti di sé stessa. La città svanisce del tutto, lasciando il posto a un mondo fluido, un universo fatto di colori, forme e memorie. È un percorso di esplorazione e confronto, un dialogo silenzioso tra la sua arte e il suo spirito.
Dopo questo viaggio onirico e trasformativo, improvvisamente si ritrova nel suo atelier, come se nulla fosse accaduto. Ma dentro di lei tutto è cambiato. Si volta e osserva di nuovo la tela incompiuta. Ora capisce. Quel paesaggio sfocato, quelle linee ancora incerte... è la sua casa d'infanzia, il paesino di montagna dove è cresciuta.
Senza esitazione, inizia a raccogliere le sue cose. Scatoloni, tele, strumenti, tutto viene imballato.
La sua decisione è presa: lascerà la città. Non perché rinnega la sua arte, ma perché ha capito che per continuare a creare deve ritrovare le sue radici.
L'ultima scena la mostra in un paesaggio completamente diverso: le montagne, il cielo limpido, la natura che la circonda. È seduta all'aperto, la tela davanti a sé. Osserva per un attimo il paesaggio reale, poi abbassa lo sguardo sul quadro e ricomincia a dipingere.
Con un ultimo colpo di pennello, l'opera è finalmente completa.

Output 3.3: Concept finale generato in seguito a diverse richieste di approfondimento fatte all'intelligenza artificiale.

3.1.9 Modalità operative di generazione del soggetto

A questo punto si è ripetuta una fase di brainstorming per individuare con più precisione quali potessero essere i quadri esplorati dall'artista all'interno del suo viaggio onirico tra le opere. È stato particolarmente significativo per sviluppare al meglio la trama, in quanto il viaggio attraverso l'arte è centrale all'interno del cortometraggio. Si è inoltre chiesto all'intelligenza artificiale di associare ogni idea di quadro con i precisi termini o immagini visualizzate dal pubblico, in modo tale da assicurarsi di centrare a pieno l'obiettivo di utilizzare quanto raccolto nella prima fase del progetto. Di seguito alcuni dei risultati prodotti dopo alcune interazioni.

- Un paesaggio montano con ruscelli e neve: Realismo Romantico, con dettagli nitidi e giochi di luce naturali, legato a "una passeggiata in montagna, tra ruscelli e prati" e "l'infanzia e le nevicate in montagna".
- Una città notturna con strade lucide e riflessi colorati: Cyberpunk, con contrasti netti e luci al neon, basato su "città di notte piove strade lucide guizzare di topi vita notturna dei vicoli".
- Un cielo pieno di mongolfiere: Illustrazione Onirica, con tinte pastello e linee morbide, tratto da "un cielo pieno di mongolfiere".
- Un mare in tempesta che si trasforma in un campo di grano: Espressionismo Astratto, con linee fluide e pennellate movimentate, ispirato a "cielo e mare" e "pensieri di serenità, immagini di fiabe".
- Un castello immerso in un paesaggio fiabesco: Pittura Fiamminga, con dettagli minuziosi e luce dorata, tratto da "un castello immerso in un paesaggio fiabesco" e "c'era una volta in un castello...".
- Un sogno astratto con forme geometriche: Astrattismo Geometrico, con forme dinamiche e colori accesi, ispirato a "un sogno astratto con forme geometriche che si muovono".
- Un torrente che nasce da una scultura vivente: Scultura Organica, con rocce scolpite in forme fluide, basato su "il fiume che si trova da qualche parte in montagna".

Output 3.4: Raccolta di alcune idee generate per ampliare il soggetto. Sono riportati virgolettati gli esatti termini utilizzati dal pubblico durante gli eventi musicali.

Successivamente si è reso necessario rendere più strutturata la sessione di brainstorming effettuata per la parte relativa al viaggio. Sempre utilizzando la funzione di riconoscimento vocale integrato, è stato chiesto all'intelligenza artificiale di analizzare meglio le varie possibilità, scegliendo con cura quali opere includere nel racconto e quali tralasciare. Comunicando vocalmente è stato più semplice indicare con precisione le tipologie di scene più adatte e le transizioni da utilizzare. L'IA ha quindi prodotto la parte di soggetto specifico per la sezione della narrazione in cui l'artista viaggia all'interno delle sue opere, a cui è seguita una richiesta di ridurre leggermente quanto prodotto. Si riporta di seguito il soggetto cinematografico, reso testuale per essere maggiormente fruibile, frutto di quest'ultima parte di lavoro.

L'artista cammina lungo le strade della città. L'aria è ferma, il rumore del traffico lontano. Le sue opere sono sparse ovunque, segni visibili della sua presenza nel tessuto urbano. Arriva davanti a un murale, dipinto su una parete di mattoni. L'opera raffigura "una strada acciottolata di paese". L'artista si sofferma, lo osserva attentamente, poi riprende a camminare.
Poco più avanti, nella piazza centrale, si erge una scultura moderna ispirata a "il fiume che si trova da qualche parte in montagna". L'acqua scolpita sembra quasi reale. Improvvisamente, la scultura prende vita: i solchi rigidi si sciolgono in un flusso d'acqua vero, che scorre via. L'artista si blocca, sorpresa. Per un istante resta in silenzio, poi si allontana, meravigliata.
In fondo a un vicolo, su un muro largo, c'è un'altra sua opera. È un murale che rappresenta "un bosco fantastico con creature misteriose". Ma mentre lo guarda, qualcosa cambia. Dai contorni del disegno iniziano a emergere radici reali, poi alberi veri che si innalzano fino a coprire parte della strada. Il cemento si crepa, la città svanisce. L'artista si ritrova immersa nella sua creazione. Attorno a lei, dai rami iniziano a sbocciare "fiori grandi felci sottobosco". Il suolo sotto i suoi piedi si scioglie in "rivoli d'acqua", e l'artista viene trasportata dalla corrente. Ora è in un paesaggio montano, circondata da ruscelli e neve. Raccoglie un po' di neve tra le mani, la osserva scivolare tra le dita, ma mentre cade, cambia forma: la neve si trasforma in "cielo e mare". L'acqua si espande intorno a lei, il paesaggio muta ancora. Cammina lungo la riva del mare, ma la sabbia cambia consistenza sotto i suoi passi. Le sue impronte affondano e la sabbia si trasforma in "spighe di grano", che ondeggiano leggere. L'artista allunga la mano e raccoglie una spiga. Tra le sue dita, si irrigidisce e si trasforma in una scultura, una struttura che rappresenta "una danza tra foglie che volteggiano nel vento". Le foglie scolpite iniziano a distaccarsi e a fluttuare nell'aria. L'artista le segue con lo sguardo: si librano sempre più in alto fino a fondersi nel cielo, diventando parte di "un cielo pieno di mongolfiere". Le mongolfiere diventano sempre più grandi. Una di esse inizia a dissolversi, rivelando al suo interno un nuovo scenario: "un castello immerso in un paesaggio fiabesco". L'artista avanza verso l'ingresso, ma non appena supera il portone, il castello si deforma, le pareti mutano colore, le finestre si fondono. L'architettura si trasforma in "un sogno astratto con forme geometriche che si muovono". Le forme si sovrappongono, poi si ricompongono fino a delineare i profili di una città. L'artista si ritrova ora in una "città di notte piove strade lucide", i neon pulsano, i riflessi dell'acqua brillano sui marciapiedi. Si guarda intorno, ma tutto inizia a dissolversi. Le immagini si sovrappongono, come se il tempo scorresse a una velocità irreale. Il paesaggio cambia ancora e l'artista si accorge di camminare attraverso "un quadro in cui si alternano le stagioni". Ogni passo la porta avanti in un ciclo continuo: inverno, primavera, estate, autunno. Poi, all'improvviso, tutto svanisce. L'artista si ritrova nel suo atelier, lo stesso da cui era partita. Davanti a lei, ancora sulla tela, c'è il quadro incompiuto. Ora sa cosa fare. Prende il pennello, si avvicina alla tela e aggiunge l'ultimo dettaglio: il paesaggio della sua infanzia.

Output 3.5: Soggetto cinematografico comprendente la parte di viaggio attraverso le opere personali dell'artista. Sempre virgolettate le esatte parole utilizzate dalle persone durante la raccolta dati in occasione degli eventi.

Il tutto è stato poi unito anche con la restante parte del racconto per ottenere il soggetto finale, riportato nella Sezione E, in appendice, dove sono sempre virgolettate le esatte parole utilizzate dal pubblico durante la raccolta dati. È interessante notare come in questa seconda parte della definizione del soggetto ci sia stato un ulteriore cambiamento, soprattutto per quanto riguarda il finale del cortometraggio, che è stato integrato e diversificato, anche grazie all'intervento umano, guidando l'intelligenza artificiale e suggerendo un finale alternativo, più strutturato e d'impatto.

3.1.10 Integrazione dei dati nei concept narrativi

Dopo aver ottenuto una versione del concept che riusciva a rappresentare in modo efficace e coerente ciò che si desiderava esprimere nel cortometraggio, si è ritenuto opportuno procedere con un'ulteriore fase di verifica incrociata, volta a consolidare ulteriormente la coerenza tra la narrazione costruita e le informazioni provenienti dai materiali raccolti durante gli eventi musicali. È stato pertanto richiesto all'intelligenza artificiale di individuare e rendere espliciti tutti i collegamenti tra la versione finale della sinossi e le risposte fornite dal pubblico durante le fasi partecipative, confermando che quanto espresso nel cortometraggio fosse realmente radicato nelle emozioni, nei ricordi e nelle immagini condivise durante gli eventi musicali.
In particolare, l'analisi restituita ha evidenziato diverse corrispondenze tematiche e simboliche. Parole chiave e argomenti ricorrenti come *natura*, *montagna*, *infanzia*, *passeggiata*, *rivoli d'acqua*, *viaggio* sono tutti elementi presenti sia nella struttura narrativa che nei dati raccolti dal pubblico. Il testo e l'immaginario audiovisivo hanno così trovato una forte corrispondenza nella memoria, nelle atmosfere evocate e nelle immagini interiori che avevano preso forma durante gli eventi.
Parallelamente, è stato anche chiesto di verificare se le scelte stilistiche e narrative operate nello script rientrassero nei desideri espressi dal pubblico in merito alla tipologia di video preferita. L'IA ha confermato che l'estetica e la struttura del cortometraggio si collocavano perfettamente tra le categorie maggiormente richieste dal pubblico, tra cui troviamo: *animazione*, *paesaggi naturali*, *stile narrativo evocativo e simbolico*. Anche da questo punto di vista, dunque, si è potuto constatare un notevole allineamento tra il progetto audiovisivo e l'evento partecipativo.
Si lascia di seguito in Output 3.6 parte dell'analisi fatta dall'IA per quanto riguarda la tipologia di video.

La sinossi è ben allineata alle preferenze generali del pubblico, ma si posiziona più verso il lato artistico, narrativo ed emozionale, piuttosto che verso il realismo puro o l'azione. Se il target è un pubblico che apprezza storie simboliche e visivamente immersive, allora il progetto è perfettamente in linea con le aspettative.
Se invece si vuole allargare l'audience, si potrebbero considerare:
- Un mix di sequenze realistiche e oniriche, per accontentare chi preferisce video più tangibili.
- Una maggiore variazione di ritmo, per rendere il viaggio visivamente più dinamico.
Nel complesso, però, direi che il cortometraggio risponde bene ai desideri del pubblico!

Output 3.6: Conclusioni tratte dall'IA per quanto riguarda l'allineamento della sinossi con quanto espresso dal pubblico.

3.1.11 Conclusioni sulla fase di ideazione

Al termine della fase di ideazione, il lavoro creativo ha portato alla formulazione di un soggetto cinematografico dettagliato, coerente con i dati raccolti e profondamente ispirato dalle suggestioni del pubblico. Questo soggetto, frutto di un processo di analisi, interpretazione e sintesi, definisce in modo chiaro la trama, l'arco narrativo e la progressione tematica dell'opera audiovisiva. La protagonista, un'artista in crisi creativa, intraprende un viaggio onirico e di trasformazione attraverso paesaggi che rievocano i suoi ricordi, le sue emozioni e le immagini simboliche suggerite dal pubblico durante gli eventi.
Il soggetto ha raggiunto un buon livello di dettaglio, delineando in linea generale già le scene principali e anticipando alcune delle trasformazioni visive più rilevanti durante il viaggio onirico della protagonista. Questa strutturazione si è rivelata fondamentale nella successiva stesura della sceneggiatura, che verrà affrontata nel prossimo capitolo. In questa fase, l'approfondimento della sequenza narrativa, il ritmo e le transizioni visive sono già state pensate e realizzate, sempre in continuità con le indicazioni raccolte.
È fondamentale sottolineare che tutto il processo creativo fino a questo punto è stato condotto in stretta collaborazione con uno strumento di intelligenza artificiale generativa, *ChatGPT*. L'accesso al file contenente le risposte e i dati raccolti durante gli eventi (comprese le tabelle riassuntive con i risultati delle elaborazioni) si è rivelato fondamentale per procedere in maniera più spedita con la creazione dei prompt adatti a generare quanto voluto.
L'IA è stata determinante nel trovare e suggerire collegamenti concettuali, ipotesi narrative e suggestioni coerenti con i temi emersi dai dati raccolti. L'interazione tra umano e IA ha seguito un flusso creativo iterativo: l'intelligenza artificiale proponeva spunti sulla base dei dati ricevuti, l'autore selezionava e raffinava, chiedendo di riadattare continuamente i risultati in funzione del progetto. Questo ha permesso di mantenere la coerenza con le fonti originarie e di esplorare soluzioni creative altrimenti difficilmente accessibili in tempi contenuti.

Con la generazione del soggetto definitivo, corredato da logline e sinossi, si conclude la prima fase progettuale. Il passo successivo sarà la stesura della sceneggiatura completa, in cui le scene verranno articolate nel dettaglio, assieme alle azioni intraprese e al ritmo narrativo, per definire completamente il corpo del cortometraggio, tenendo sempre al centro la simbiosi tra creatività umana e intelligenza artificiale. Tutti i documenti finali prodotti (logline, sinossi e soggetto) possono essere consultati in appendice, rispettivamente in Sezione D e Sezione E.

3.1.12 Modalità operative di generazione della sceneggiatura

Un documento fondamentale per lo sviluppo di un progetto audiovisivo è la sceneggiatura; nel contesto della progettazione narrativa per un cortometraggio artistico, le indicazioni lì contenute sono fondamentali per la buona riuscita del progetto. Questo documento è il punto di riferimento in ogni situazione. In questo progetto, l'interazione con il sistema di intelligenza artificiale generativa è stata sfruttata con il fine di trasformare un soggetto cinematografico precedentemente elaborato in una sceneggiatura formalmente strutturata secondo gli standard professionali dell'industria audiovisiva. In un primo prompt, è stato condiviso il soggetto in questione, articolato come un racconto lineare ma ricco di elementi visivi e simbolici, utili per mantenere un'alta coerenza tra le diverse fasi della produzione. In particolare, è stato chiesto all'IA di analizzare il testo ricevuto, chiedendo una valutazione del suo impianto narrativo e delle sue potenzialità filmiche. In risposta, l'IA ha confermato la solidità del soggetto dal punto di vista tematico e visivo, suggerendo alcune linee di potenziale approfondimento, come l'arricchimento della caratterizzazione psicologica del personaggio e l'integrazione di un legame più evidente tra le opere artistiche descritte e la vita della protagonista. In seguito, è stata richiesta la generazione della sceneggiatura in maniera professionale, specificando l'esigenza di evitare l'introduzione di dialoghi non presenti nel testo originale e l'aggiunta di transizioni visive non previste.
In questa fase, l'IA ha svolto un ottimo lavoro, procedendo in primo luogo alla segmentazione del testo originale in scene, utilizzando la sintassi convenzionale della sceneggiatura: intestazioni di scena con indicazione dell'ambiente (interno o esterno), del luogo e del momento della giornata, descrizioni brevi e dirette dell'azione, focalizzazione sulle componenti visive e sulla sequenzialità degli eventi. Il testo generato è stato fornito in un formato continuo, con scene chiaramente identificabili, mantenendo uno stile descrittivo adatto alla produzione filmica.
Il lavoro si è poi articolato in ulteriori momenti di rifinitura, attraverso richieste mirate da parte dell'autore umano. Una delle modifiche significative ha riguardato la revisione di alcune intestazioni di scena, inizialmente descritte in maniera errata, per allinearsi meglio con l'atmosfera narrativa immaginata.
L'interazione, anche in questo caso, ha seguito una logica iterativa: da una prima generazione testuale a una serie di verifiche e adattamenti semantici e formali, fino alla definizione di una versione finale della sceneggiatura. La funzione dell'intelligenza artificiale è stata quella di eseguire operazioni di formattazione, trasposizione e adattamento tecnico del linguaggio, partendo da istruzioni discorsive fornite manualmente.
Al termine del processo, si è giunti alla redazione definitiva di una sceneggiatura completa, conforme ai requisiti tecnici e linguistici del settore cinematografico. Il

risultato è un documento che rappresenta una trasposizione fedele e formalizzata dell'idea originaria, consultabile nella Sezione F dell'appendice.

3.1.13 Creazione dello storyboard con IA

Questa fase del lavoro ha avuto inizio con una richiesta mirata, ovvero la generazione di un prompt efficace e dettagliato da fornire all'intelligenza artificiale per creare uno storyboard completo. Questo approccio, estremamente particolare, si distingue non solo per la sua efficacia, ma anche per la sua versatilità, permette infatti di ottenere in output prompt molto specifici e accurati in pochissimo tempo. Creare gli stessi prompt manualmente richiederebbe un grosso sforzo dell'autore umano; al contrario, chiedendo di farlo all'IA si può risparmiare tempo che può essere investito nel perfezionare maggiormente il prompt generato. In questo modo si ottengono direttive ancora più precise, senza però dover riscrivere ogni indicazione tecnica. È una strategia che, come verrà trattato nei capitoli successivi, può essere applicata anche nella produzione di immagini, concept art e video. Di seguito, nel Prompt 3.3 si riporta la richiesta di generazione iniziale e la conseguente risposta fornita in Output 3.7.

> Okay, ora passiamo a fare lo storyboard, da dove si parte? Vorrei avere un prompt da dare all'intelligenza artificiale per aiutarmi a farlo.

Prompt 3.3: Prompt iniziale per generare un testo adatto da usare a sua volta come prompt per generare uno storyboard.

Basandoti sulla sceneggiatura fornita, crea uno storyboard dettagliato per un cortometraggio visivamente evocativo. Lo stile dovrà rispecchiare il tono narrativo del film, integrando elementi naturalistici e simbolici con una fotografia curata e un montaggio fluido. Il video dovrà alternare momenti di introspezione e dinamismo, evocando emozioni come nostalgia, meraviglia, rinascita e viaggio interiore.
Ogni scena dovrà essere costruita con un forte impatto visivo:
Ambientazioni suggestive, come paesaggi naturali, città illuminate dalla pioggia, stanze avvolte dalla luce calda del tramonto.
Personaggi espressivi, catturati in primi piani intensi e movimenti simbolici.
Elementi chiave della narrazione messi in evidenza attraverso l'uso del colore e della composizione dell'inquadratura.
Lo storyboard dovrà seguire il ritmo del racconto, con transizioni cinematiche fluide che enfatizzano il passaggio del tempo e l'evoluzione emotiva dei protagonisti. Le scene più intense potranno essere rappresentate con tagli dinamici, mentre i momenti contemplativi useranno inquadrature più ampie e statiche.
L'atmosfera cromatica sarà dominata da una palette di colori coerente con il messaggio del corto, con una particolare attenzione ai dettagli e all'uso della luce per creare profondità e coinvolgimento emotivo.
Il risultato dovrà essere uno storyboard che serva come guida visiva chiara e ispiratrice per la produzione del video, assicurando che ogni inquadratura trasmetta la potenza narrativa della sceneggiatura.

Output 3.7: Risposta contenente un prompt specifico da usare per generare uno storyboard dettagliato e professionale.

In seguito a qualche richiesta di modifica e inserendo manualmente qualche informazione aggiuntiva, il prompt finale per generare lo storyboard è stato strutturato come si può vedere nel Prompt 3.4.

Basandoti sulla sceneggiatura fornita, genera uno storyboard dettagliato per il cortometraggio, creando un'immagine per ogni inquadratura (shot). Ogni immagine dovrà rappresentare visivamente l'inquadratura, includendo dettagli tecnici e creativi per guidare la produzione del film. Lo stile dell'immagine deve essere uno schizzo, senza troppo realismo.

Per ogni shot, fornire le seguenti informazioni, tenendo ordinata la numerazione delle scene e delle inquadrature (1.1, 1.2; 2.1, 2.2, 2.3 e così via):
Azione: Descrizione chiara di ciò che accade nella scena (movimenti dei personaggi, oggetti in movimento, azioni principali).
Focale: Specifica la lunghezza focale della camera (es. Grandangolo, Medio, Teleobiettivo) e il tipo di inquadratura (Close-up, Medium Shot, Wide Shot, Over-the-shoulder, POV, ecc.).
Movimento Macchina: Indica se la camera è statica o in movimento (Pan, Tilt, Dolly In/Out, Tracking, Handheld, Drone, ecc.).
Personaggi: Chi è presente nella scena? Qual è la loro posizione, espressione, postura? Se necessario, specifica il blocking (posizionamento nello spazio).
VFX (Effetti Visivi): Specifica se sono presenti effetti visivi (es. nebbia, pioggia, esplosioni, transizioni digitali, effetti onirici, distorsioni visive, ecc.).
Indicazioni Aggiuntive: Qualsiasi altro dettaglio utile alla composizione della scena, come illuminazione (soft, dura, controluce), colori dominanti, atmosfera emotiva, riferimenti stilistici.

Lo stile visivo dello storyboard deve essere coerente con il tono della sceneggiatura, rappresentando ogni scena con chiarezza cinematografica e dettagli sufficienti per la pre-visualizzazione del film. Le immagini devono essere ben composte, con attenzione all'illuminazione e alla profondità di campo per trasmettere al meglio l'atmosfera di ogni sequenza.
Ora sto per inviarti la sceneggiatura, sei pronto?

Prompt 3.4: Prompt finale utilizzato per generare lo storyboard del progetto audiovisivo.

Il prompt così elaborato è uno strumento altamente specifico, pensato per rispondere alle particolari esigenze di un prodotto audiovisivo di questo tipo. L'obiettivo era costruire un documento in grado di istruire l'IA non solo sulla natura del compito, ma anche sulle modalità operative da utilizzare e sulle regole del linguaggio da rispettare. La costruzione segue un'organizzazione logica: si apre con un'introduzione generale che spiega chiaramente cosa dovrà essere eseguito, specificando che ogni clip dovrà anche essere visualizzata e sottolineando l'importanza della composizione, della profondità di campo e della coerenza stilistica rispetto al tono del cortometraggio. Viene fornito poi un esempio concreto di numerazione ordinata delle inquadrature da seguire, elemento fondamentale che permette all'IA di comprendere il formato richiesto e di organizzare logicamente i contenuti generati.
L'uso di esempi espliciti rappresenta uno degli strumenti più efficaci nell'interazio-

ne con i sistemi di intelligenza artificiale, in quanto riduce l'ambiguità e migliora l'accuratezza della risposta. La parte centrale del prompt è strutturata come una tabella che specifica, per ogni clip, sei voci che devono essere obbligatoriamente presenti nello storyboard: *Azione*, *Focale*, *Movimento Macchina*, *Personaggi*, *VFX* e *Indicazioni Aggiuntive*. Questa struttura permette di ottenere uno storyboard tecnicamente completo, vicino agli standard usati nella produzione cinematografica. Per ciascuna voce, viene indicato chiaramente cosa deve essere descritto, contribuendo a un output uniforme, coerente e preciso. Anche in questo caso, il fatto di fornire esempi concreti è di fondamentale importanza per ottenere esattamente quanto richiesto. L'inclusione della voce *Indicazioni Aggiuntive* amplia ulteriormente le possibilità espressive, permettendo di aggiungere elementi come l'illuminazione, i colori dominanti, l'atmosfera emotiva, o i riferimenti stilistici più adatti. Questa sezione viene strategicamente posta alla fine, garantendo così che l'IA attribuisca a questa un peso superiore nel momento della generazione. Il prompt si conclude con una domanda rivolta direttamente all'IA, utile non solo a segnalare il passaggio alla fase operativa, ma anche a comprendere se lo strumento ha ben capito il contesto attuale in cui sta lavorando.

Sebbene il prompt sia stato originariamente pensato anche per la generazione visuale delle immagini inerenti lo storyboard, in questa fase è stato impiegato esclusivamente per la definizione della sola parte testuale, rivelandosi in ogni caso estremamente utile. Il livello di dettaglio, l'organizzazione logica e la ricchezza tecnica del risultato ottenuto hanno costituito un punto di riferimento fondamentale per le fasi successive del progetto. La creazione delle immagini, invece, è stata rimandata a un momento successivo, poiché si è notato essere piuttosto complicato generare contemporaneamente le due cose.

Solo dopo aver ultimato e quindi approvato il prompt, è stata introdotta la sceneggiatura, fornita in forma completa. Da quel momento, l'intero lavoro si è concentrato sulla traduzione cinematografica, scena per scena, inquadratura per inquadratura, della narrazione contenuta nella sceneggiatura. La collaborazione tra uomo e macchina ha incluso l'inserimento sistematico e progressivo di inquadrature, la scelta accurata della lunghezza focale, dei movimenti di macchina, delle luci e così via. Ogni scena generata è stata poi analizzata e rivisitata in modo ciclico: partendo dalla proposta iniziale, l'autore umano ha suggerito modifiche creative e registiche, che hanno portato a uno sviluppo sempre più articolato dello storyboard. Anche in questo caso, la modalità di riconoscimento vocale è stata particolarmente utile, spiegando a voce cosa fosse necessario modificare, aggiungere o cambiare in alcune inquadrature, è stato possibile effettuare le modifiche in tempi molto rapidi.

La fase di ridefinizione si è rivelata fondamentale per ottenere un risultato finale ottimale: sono state fornite nuove indicazioni, richieste modifiche, specificate meglio le intenzioni narrative e suggeriti alcuni riferimenti visivi. Il lavoro si è concentrato sull'individuazione delle giuste angolazioni, dei movimenti macchina più adatti, dei livelli di profondità, nonché sull'introduzione di elementi simbolici e visivi che rafforzassero il senso delle scene.

Questo approccio è stato utilizzato in maniera iterativa per ognuna delle tredici scene che compongono il cortometraggio. Ogni passaggio è stato descritto con estrema cura, corredato dalla propria scheda tecnica e indicazioni stilistiche dettagliate, arrivando alla generazione di uno storyboard testuale completo.

Creare prompt per l'IA in funzione del lavoro creativo e utilizzarli come strumenti a

monte per dirigere processi più complessi e strutturati è uno dei modelli più avanzati di collaborazione uomo-macchina. Il fatto che questo storyboard sia stato costruito in modo interattivo, con costante supervisione di regia e attenzione autoriale umana, lo rende un prodotto altamente qualificato.
È stata poi avviata una prima fase esplorativa per testare la generazione di immagini dello storyboard. Si è tentato di visualizzare alcune inquadrature attraverso il modello di intelligenza artificiale generativa di *OpenAI*. Tuttavia, i risultati ottenuti non si sono rivelati sufficientemente soddisfacenti. In particolare, non vi era ancora corrispondenza tra lo stile richiesto e le immagini prodotte, spesso troppo incentrate sull'ambito cinematografico quando in realtà dovevano essere parte dello storyboard. Questa notevole discrepanza ha suggerito di sospendere momentaneamente la componente visiva per continuare, almeno momentaneamente, a rafforzare l'aspetto testuale. Di seguito un'immagine generata in questa fase, completamente incoerente con quanto richiesto.

Figura 3.3: Immagine generata che dovrebbe rappresentare un'inquadratura ma, come è evidente, rappresenta in parte uno storyboard.

È importante precisare che in questo momento particolare si è consolidato un metodo di lavoro più fluido e naturale; le indicazioni infatti venivano fornite in forma discorsiva, come se si stesse raccontando a parole un'inquadratura a un collaboratore. Questo approccio, basato sul linguaggio naturale, ha richiesto l'utilizzo costante del riconoscimento vocale integrato all'interno di *ChatGPT* ma ha permesso di effettuare modifiche precise e puntuali. Inquadratura dopo inquadratura, lo storyboard si è ampliato e consolidato, mantenendo sempre una struttura tecnica precisa ma arricchita dalle sfumature emotive e stilistiche imposte dall'autore umano, che ha sempre avuto il controllo del flusso di lavoro. Un esempio di prompt utilizzato per descrivere queste modifiche, nato dall'inserimento vocale, è quello riportato in Prompt 3.5.

Bene, allora per il 4.1 farei un carrello laterale che segue la protagonista che cammina nei vicoli, finché appunto arriva davanti a quest'altro murale, e a quel punto anche il tracking shot laterale si ferma. Poi, la seconda inquadratura, lei che osserva il murale, farei un medium shot, magari prendendolo un po' dal basso verso l'alto, poi va bene fare un detail shot come 4.3, lascialo così, poi farei un medium shot su di lei che si guarda i piedi come 4.4, per invogliare lo spettatore a capire cosa sta succedendo, col 4.5 farei uno shot di dettaglio del pavimento che si crepa con le radici che iniziano a salire, dopodiché farei quello che hai scritto tu, un wide shot tilt up, dopodiché tornerei su un full body shot, con il cemento sotto ai piedi dell'artista che si crepa, poi farei un close up su di lei, mentre tutto continua a trasformarsi, per poi finire la scena che è tutta trasformata accanto a lei, e quindi facciamo prima un extreme close up per mostrare la sua espressione estremamente sorpresa, e poi torniamo su un medium wide shot comunque per mostrare un po' l'ambiente di dove si trova con lei che si gira attorno, magari separalo in due inquadrature diverse, prima un medium full shot e poi un campo largo proprio.

Prompt 3.5: Trascrizione dell'interazione vocale dell'autore umano, usata come prompt per ridefinire alcune parti dello storyboard.

Questo tipo di interazione ha dimostrato l'efficacia dell'intelligenza artificiale nel seguire istruzioni complesse, anche quando espresse in modo colloquiale e riportate in maniera automatica dal riconoscimento vocale. In questo modo ogni scena dello storyboard è stata progressivamente costruita e ottimizzata, attraverso un linguaggio tecnico ma allo stesso tempo creativo e umano.
Il processo di costruzione dello storyboard ha dunque seguito, con grande coerenza metodologica, un iter preciso e strutturato, ripetuto per ogni singola scena del cortometraggio. All'inizio, si richiedeva all'intelligenza artificiale di mostrare nuovamente la parte di sceneggiatura corrispondente alla scena in questione. Questo passaggio era fondamentale per assicurarsi che il sistema operasse con piena consapevolezza del contesto narrativo specifico prima di procedere alla generazione dello storyboard, reso possibile dal fatto che l'IA avesse accesso ai file di progetto, tra i quali si trovava ovviamente anche la sceneggiatura. Una volta esaminato il testo della scena, si avviava la richiesta di generazione dello storyboard, sfruttando il prompt principale, creato e poi modificato dall'autore umano. L'intelligenza artificiale forniva quindi una prima versione della sequenza, composta da una serie di inquadrature, ciascuna corredata da parametri tecnici e descrittivi. Su questa base, venivano poi fornite nuove indicazioni tramite input vocale, con una lunga e dettagliata descrizione delle modifiche desiderate. L'intelligenza artificiale rispondeva quindi alle indicazioni riformulando lo storyboard della scena, aggiornando le inquadrature secondo i nuovi parametri e implementando ogni modifica. In alcuni casi, era necessario effettuare ulteriori richieste di aggiustamenti, in seguito ai quali si raggiungeva la versione definitiva della scena. A quel punto, veniva data approvazione e si procedeva con la richiesta di sceneggiatura successiva, avviando un nuovo ciclo.
Il metodo si è rivelato estremamente efficace per una serie di motivi. In primo

luogo, ha consentito di mantenere un altissimo livello di dettaglio nella costruzione visiva della narrazione. Ogni singola inquadratura è stata descritta con parametri tecnici completi, creando così un documento coerente con gli standard professionali di uno storyboard cinematografico. In secondo luogo, questa modalità di lavoro ha permesso di costruire un impianto visivo perfettamente allineato all'identità del film, rispettandone il tono e il ritmo interno. L'uso dell'intelligenza artificiale non ha sostituito l'intenzione creativa, ma si è rivelato uno strumento potente per tradurla con precisione e velocità in una struttura operativa. Infine, si è potuto constatare, soprattutto nel bilancio finale del progetto, quanto questa fase sia stata cruciale per lo sviluppo complessivo. Pur essendo, al momento, uno storyboard unicamente testuale, il suo livello di completezza e ricchezza di informazioni ha posto basi solide per i successivi momenti di produzione.

3.1.14 Creazione delle immagini per lo storyboard con IA

Nel contesto della pre-produzione del cortometraggio uno degli aspetti più complessi del lavoro è stato quello legato alla costruzione visiva dello storyboard attraverso l'uso dell'intelligenza artificiale. Questo percorso ha rappresentato una vera e propria esplorazione sperimentale, che ha messo alla prova non solo le capacità tecniche degli strumenti disponibili, ma anche la possibilità concreta di automatizzare un processo creativo tradizionalmente manuale, codificato e fortemente legato all'esperienza e abilità umane.

La prima è stata una fase di impostazione concettuale e di prova, in cui si è iniziato a generare immagini utilizzando prompt costruiti manualmente con grande precisione. L'obiettivo era quello di ottenere singole immagini che rappresentassero una specifica inquadratura, coerente con la sceneggiatura e con lo sviluppo visivo del cortometraggio. I prompt iniziali sono stati pensati per evocare lo stile tipico degli storyboard professionali usati nel cinema: immagini in bianco e nero, realizzate con uno stile a schizzo, linee forti e ben delineate, rapporto d'aspetto 16:9 e, quando possibile, frecce direzionali per indicare i movimenti della macchina da presa o degli attori all'interno della scena. L'intento era quello di avvicinarsi quanto più possibile alla realtà delle produzioni audiovisive, in cui lo storyboard è uno strumento estremamente funzionale e narrativo.

È importante precisare che durante questa fase l'intelligenza artificiale aveva a disposizione tutte le descrizioni testuali generate in precedenza. Era quindi in grado di ottenere in qualsiasi momento la descrizione di ogni inquadratura, sapere con quale focale dovesse essere girata e leggere le eventuali annotazioni aggiuntive.

Il primo prompt utilizzato per testare questa modalità, inserito nel sistema, è riportato nel Prompt 3.6.

Genera un'immagine in bianco e nero, in stile storyboard cinematografico, con forti contrasti e linee definite, simile agli storyboard utilizzati nel mercato del cinema. Ogni scena deve essere rappresentata con un'attenzione particolare alla composizione e alla direzione della luce. L'immagine deve includere:
Personaggi e ambientazione chiaramente distinguibili.
Movimenti della macchina da presa indicati con frecce direzionali (es. carrello avanti, zoom-in, panoramica).
Movimenti degli attori segnalati con frecce di direzione.
Profondità di campo evidenziata tramite sfocature o linee prospettiche.
Dinamismo e inquadrature cinematografiche che enfatizzino l'azione.
Usa uno stile schizzato a mano, con tratti forti e decisi, simile agli storyboard tradizionali del cinema.
Usa per forza un'inquadratura 16:9, mi raccomando, ora iniziamo con la prima scena, prima inquadratura.

Prompt 3.6: Prompt utilizzato per generare le immagini dello storyboard.

Con questa descrizione si è cercato di produrre un'immagine coerente con l'inquadratura di apertura del film, ovvero una vista aerea della città nel pomeriggio, ripresa con un drone. Tuttavia, il primo risultato ha subito evidenziato numerose criticità. In diversi tentativi, il drone compariva visivamente all'interno dell'inquadratura invece di essere utilizzato solamente come punto di vista della clip. Questo errore, ripetuto più volte, ha rivelato una comprensione superficiale da parte dell'IA della distinzione tra soggetto dell'immagine e dispositivo di ripresa. Quando il prompt suggeriva che la scena fosse *ripresa da un drone*, il modello generava spesso un drone visibile, come se fosse un oggetto da rappresentare e non una fonte della prospettiva.

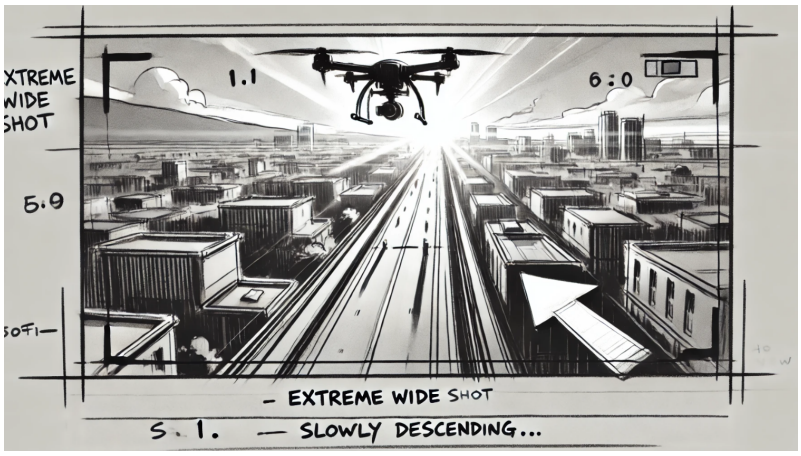


Figura 3.4: Immagine generata rappresentante la prima inquadratura, ma sono visibilmente presenti degli errori, a partire dalla presenza del drone.

Un secondo problema, altrettanto ricorrente, riguarda l'inserimento arbitrario di scritte e diciture testuali direttamente sulle immagini generate. Nonostante nel prompt non fosse mai stato richiesto, molte immagini includevano intestazioni co-

me *Scene 1.1*, *Wide shot*, *Tracking* oppure etichette generiche completamente fuori contesto. A volte queste indicazioni erano persino errate rispetto alla scena descritta, creando confusione e compromettendo l'uso dell'immagine all'interno di un flusso di lavoro professionale. L'IA, presumibilmente tentando di emulare un'estetica da storyboard, aggiungeva elementi grafici che però non solo non erano utili, ma ostacolavano la pulizia e la chiarezza del risultato finale.

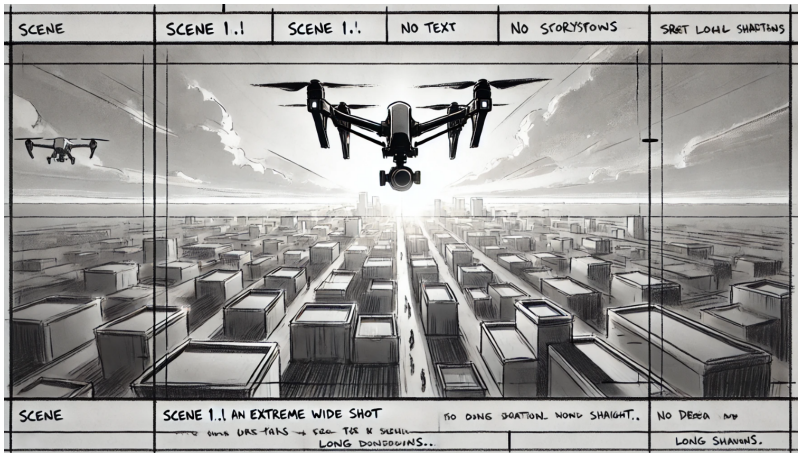


Figura 3.5: Immagine generata rappresentante la prima inquadratura, ma sono presenti ulteriori scritte inutili, anche sbagliate.

Un ulteriore tentativo ha previsto la modifica del prompt per evitare questi difetti, ma spesso i problemi si ripresentavano, anche se in forme diverse. In alcuni casi, invece di produrre una sola immagine, come richiesto, il sistema generava intere tavole composte da molte vignette consecutive, come se stesse cercando di sintetizzare una sequenza narrativa in più fotogrammi. Questo comportamento, seppure interessante da un punto di vista visivo, non era compatibile con la struttura impostata precedentemente nello storyboard che prevedeva invece una singola immagine per inquadratura, da inserire parallelamente alla descrizione tecnica precedente.

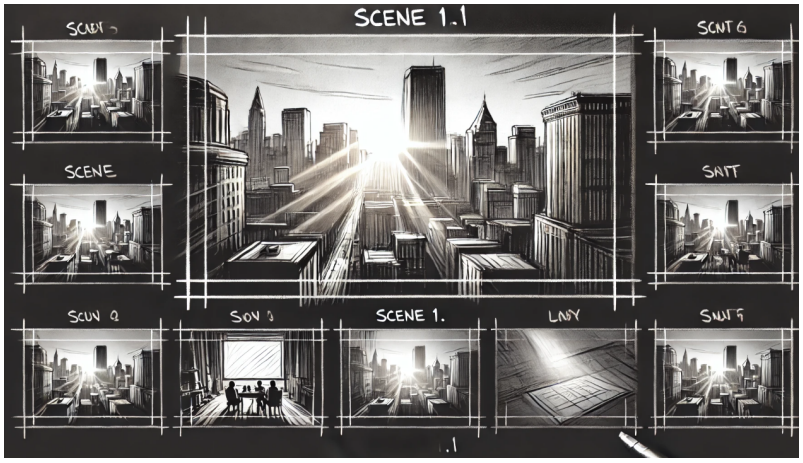


Figura 3.6: Immagine generata rappresentante la prima inquadratura, ma sono presenti ulteriori vignette non richieste relative ad altre inquadrature.

Ciononostante, alcune immagini prodotte sono risultate convincenti e coerenti con

quanto richiesto. In quelle rare occasioni l'output era ben composto, l'inquadratura rispettava l'indicazione fornita nel prompt, la rappresentazione della scena risultava fedele e la sensazione generale richiamava davvero l'estetica degli storyboard tradizionali. Nonostante questi successi parziali, il tempo richiesto per raggiungere risultati validi non era compatibile con le esigenze di produzione di un progetto come questo. Ogni immagine richiedeva vari tentativi, aggiustamenti nei prompt, valutazioni visive e operazioni di post-produzione per rimuovere elementi superflui o correggere in parte la composizione. Questo comportava, paradossalmente, un rallentamento nella produzione generale del progetto, rendendolo meno efficiente rispetto a una soluzione manuale o mista, a causa dell'uso dell'intelligenza artificiale, che non era in grado di soddisfare le richieste. Inoltre, le immagini, anche quando corrette, spesso necessitavano di ulteriori ritocchi con programmi di grafica come *Photoshop*, per essere adattate in modo preciso all'impaginazione e allo stile desiderato. L'uso di questo strumento esterno implicava però un ulteriore investimento di tempo, che andava contro l'idea iniziale di semplificare il flusso produttivo.

Figura 3.7: Immagine generata pienamente soddisfacente, rappresentante un'inquadratura della prima scena. Nonostante l'ottimo risultato, il tempo richiesto per arrivare a questo risultato è troppo elevato.

A un certo punto, è stato quindi necessario riconsiderare l'approccio per la generazione delle immagini per lo storyboard. Il bilancio finale di questa prima fase sperimentale infatti, non è stato positivo. L'intelligenza artificiale si è dimostrata uno strumento interessante, in grado di generare soluzioni visive sorprendenti, estremamente inerenti e, in alcuni casi, efficaci. D'altra parte però, le sue limitazioni rendevano necessario un continuo intervento umano per guidare, correggere e completare il lavoro, annullando parte del risparmio di tempo e risorse.
Nonostante il primo tentativo non avesse portato al risultato sperato, si è scelto di proseguire utilizzando sempre l'interazione testuale, partendo però questa volta da un prompt specifico in lingua inglese, per poi analizzare progressivamente la coerenza e la fedeltà alle indicazioni fornite e valutare infine l'efficacia del metodo di lavoro. Il Prompt 3.7 contiene l'ultima versione utilizzata in questa fase.

> Create a four-panel storyboard based on the descriptions provided in the file 'scriptEng.' For each panel, use the corresponding scene description to generate the composition, lighting, and setting. When depicting the main protagonist, reference the image file 'artist.jpeg' to maintain consistency in appearance. Ensure that the storyboard follows a cinematic visual style, with dynamic angles and expressive framing that convey the emotions and atmosphere of each scene. The final result should be a cohesive sequence of four images that visually represent the first four shots described in the script.

Prompt 3.7: Prompt finale utilizzato per generare le immagini dello storyboard del progetto audiovisivo.

Una svolta metodologica applicata rispetto ai precedenti approcci è l'utilizzo della lingua inglese, maggiormente compatibile con i modelli linguistici dell'IA. In questo modo si è puntato a ottenere risultati più allineati agli standard cinematografici e al linguaggio tecnico reale della sceneggiatura. Per realizzare tutto ciò, però, è stato necessario tradurre quanto fatto riguardo lo storyboard testuale: i prompt precedenti, infatti, avevano generato il testo in italiano, incompatibile con questo nuovo approccio. Fermo restando che l'IA è in grado di tradurre testi e quindi non avrebbe avuto problemi in tal senso, si è preferito evitare di effettuare un passaggio in più con la relativa possibilità di errori.
Il file *scriptEng* citato nel prompt, consultabile nella sua interezza in appendice nella Sezione G, contiene tutte le inquadrature tradotte in inglese. Anche in questo caso l'intelligenza artificiale ha permesso di velocizzare le tempistiche, effettuando tutto il lavoro in poco tempo.
Tuttavia, nella precedente fase, era stato riscontrato un ulteriore ostacolo: la coerenza visiva del personaggio protagonista. Infatti, nonostante fosse stato chiesto in modo esplicito che ogni sua raffigurazione fosse uguale, la generazione automatica non sempre rispettava questa direttiva. L'artista veniva spesso rappresentata con tratti o posture diverse, creando una discontinuità visiva che rischiava di compromettere l'omogeneità del racconto per immagini.
Si è quindi reso necessario creare un modello visivo che potesse essere utilizzato come riferimento per tutte le generazioni che richiedessero la presenza dell'artista, domandando di utilizzarlo sempre come base di partenza per la creazione del personaggio. Grazie a queste indicazioni e sfruttando il file di progetto contenente il modello è iniziata la generazione delle immagini per lo storyboard.
Fin dalle prime immagini, è apparso evidente come l'interpretazione dell'intelligenza artificiale fosse perfettamente in linea con quanto descritto nella sceneggiatura ma soprattutto nello storyboard testuale. Le inquadrature e le atmosfere create mostravano una coerenza molto soddisfacente al documento di partenza, favorendo così un proseguimento nella generazione delle immagini di altre inquadrature.

Figura 3.8: Immagine generata utilizzando il nuovo prompt, rappresentante un'inquadratura della prima scena. Il risultato è molto buono.

Un ulteriore elemento critico ha riguardato la presenza, assolutamente non richiesta, di oggetti e figure appartenenti all'ambito cinematografico: macchine da presa, treppiedi, cineprese e perfino operatori, venivano inspiegabilmente inseriti in molte delle immagini, nonostante fosse stato chiesto espressamente di evitare qualunque riferimento alla messa in scena filmica.

Figura 3.9: Immagine generata utilizzando il nuovo prompt, ma contenente elementi non voluti. Si è reso necessario modificare attentamente le singole immagini.

Per risolvere questo problema, è stato utilizzato lo strumento di modifica visiva integrato, che permette di selezionare con un pennello le aree da correggere e successivamente scrivere un prompt specifico per guidare l'intervento nell'area desiderata. Questo approccio si è rivelato particolarmente efficace, offrendo una soluzione veloce e pratica rispetto all'utilizzo di software esterni come *Photoshop*. Nei casi in cui le modifiche da effettuare erano più complesse, si preferiva invece rigenerare l'intera immagine utilizzando un nuovo prompt migliorato, andando a descrivere, completamente dall'inizio, l'immagine da realizzare.

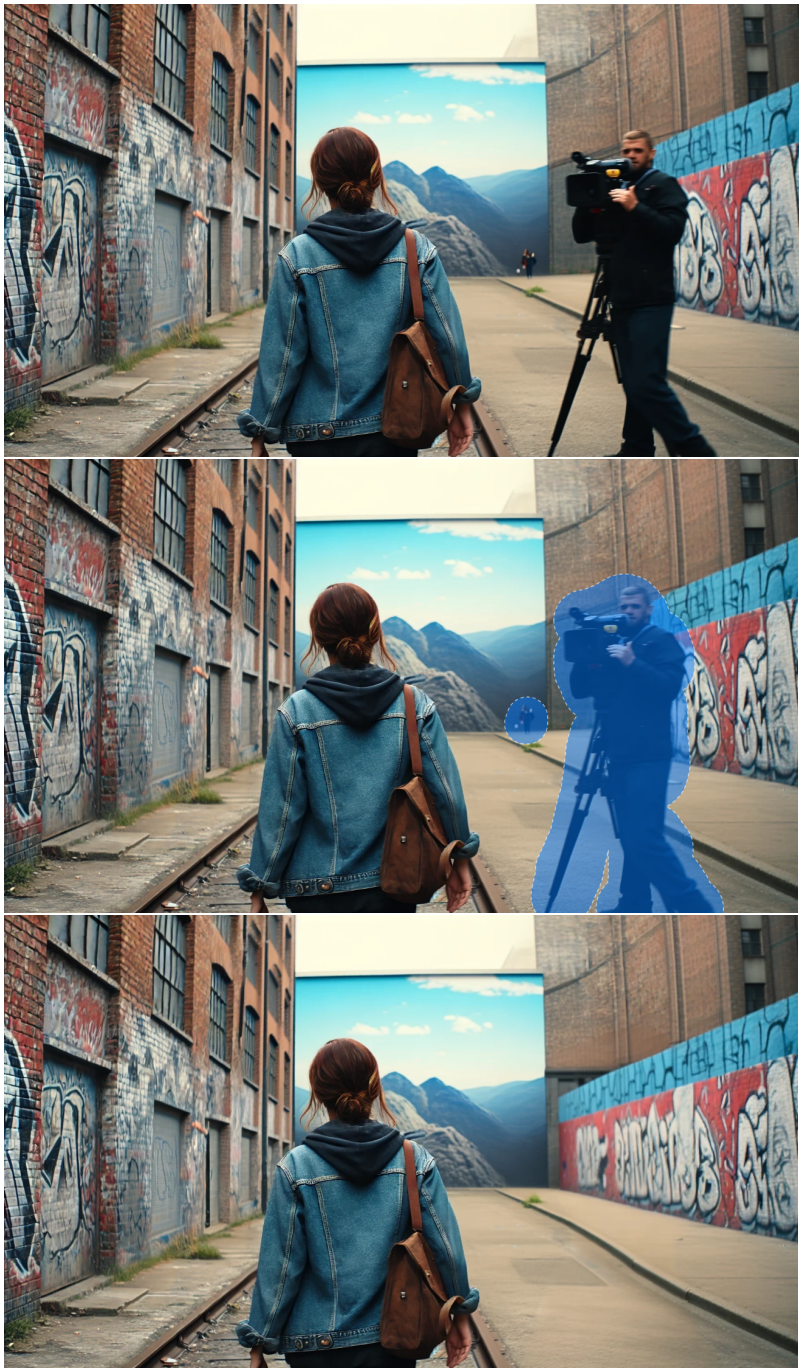


Figura 3.10: In ordine, dall'alto in basso: immagine generata utilizzando il nuovo prompt, contenente elementi estranei. Selezione degli elementi da rimuovere. Risultato finale con gli elementi rimossi dall'intelligenza artificiale.

Durante la creazione delle immagini per lo storyboard, si è posta particolare attenzione al criterio estetico utilizzato per ottenere le varie immagini: gli ambienti naturali o magici, come foreste e paesaggi onirici, venivano prodotti con uno stile prettamente pittorico e artistico, mentre le inquadrature urbane, reali, venivano generate in stile ultra fotorealistico.
Questo criterio ha consentito di rafforzare la distinzione tra realtà e immaginazione, in coerenza con le trasformazioni narrative presenti nel soggetto cinematografico. Di

seguito sono presenti alcune immagini generate utilizzando questo approccio.

Figura 3.11: Esempi di immagini generate utilizzando l'approccio descritto in questo capitolo.

Il lavoro documentato in questa fase ha messo in luce il vero potenziale delle tecnologie generative. In particolare, è stato possibile generare uno storyboard completo, professionale e altamente dettagliato a partire da indicazioni chiare e ben strutturate. Questo risultato non sarebbe stato possibile nello stesso tempo e con la stessa qualità. È importante sottolineare però che questi risultati sono fortemente condizionati dalla chiarezza dei prompt, dal riferimento visivo del personaggio principale e dalla capacità di intervenire con precisione sulle immagini prodotte per correggerne le incongruenze. Lo storyboard completo è possibile consultarlo nell'apposita Sezione H dell'appendice.

3.1.15 Creazione del modello visivo della protagonista

Per la definizione dell'identità visiva della giovane artista protagonista del cortometraggio è stato fondamentale elaborare un modello coerente, completo e tecnicamente funzionale alla generazione dello storyboard illustrato e, in prospettiva, delle sequenze video. Questo modello corrisponde a ciò che in ambito visivo viene comunemente denominato *character sheet*, ovvero una scheda illustrativa che mostra il personaggio in diverse pose, angolazioni e talvolta con diverse espressioni, al fine di cristallizzarne

l'aspetto e assicurare una completa coerenza in tutte le rappresentazioni successive. La protagonista, una giovane artista bionda dai lineamenti dolci ma decisi, è stata definita visivamente attraverso una serie di iterazioni mirate a esplorarne le sembianze e la postura. Dal punto di vista stilistico, è stata concepita come una figura contemporanea, credibile e vicina alla realtà, ma molto semplice: indossa una camicia bianca, ampia ma curata, jeans chiari vissuti e un paio di scarpe da ginnastica chiare. Tutto questo è stato fatto con diversi strumenti di intelligenza artificiale generativa ma alla fine i risultati migliori sono stati quelli prodotti da *Freepik*, modello particolarmente interessante per la produzione di immagini fotorealistiche di persone fisiche. Di seguito i risultati ottenuti.

Figura 3.12: Immagini dell'artista generate utilizzando il modello di *Freepik*.

Il modello finale generato non si limita a fornire una visione frontale del personaggio, ma mostra il corpo in pose multiple: frontale, laterale, di tre quarti e retro. In questo modo, l'IA può comprendere in maniera completa le proporzioni e l'atteggiamento corporeo della protagonista. Per generare questa scheda, è stato utilizzato un prompt dettagliato (Prompt 3.8).
Attraverso diversi tentativi di generazione e variando parametri legati all'illuminazione da utilizzare, sono emerse numerose versioni del modello. Ciascuna è stata valutata con particolare attenzione alla leggibilità dell'abbigliamento, alla riconoscibilità del volto e alla naturalezza delle pose. Il processo di selezione si è quindi concentrato non solo sull'estetica, ma anche sull'usabilità tecnica dell'immagine: un buon *character sheet*, infatti, deve essere un chiaro riferimento per i sistemi generativi utilizzati in seguito.

A highly detailed, ultra-realistic digital character sheet of a contemporary woman in her mid-30s. She has fair skin, a slender build, and naturally soft facial features. Her wavy blonde hair is tied back into a relaxed, modern chignon. She has light blue or green eyes and a few natural freckles on her cheeks.
She's dressed in effortlessly casual, contemporary clothing: a slightly oversized, off-white linen shirt with rolled sleeves, faded high-rise jeans, and neutral-toned canvas sneakers. Minimalist accessories like a thin silver ring or subtle stud earrings may be visible.
The character sheet includes five angles: front view, side profile, back view, 3/4 pose, and a close-up facial portrait. All views are rendered in a clean, soft-lit studio setting with a neutral beige or soft gray background.
The style is hyper-realistic and cinematic — natural skin textures, ambient soft shadows, and lifelike lighting are emphasized. Her expression is reflective and calm, like a modern-day artist, designer, or actress in a contemporary urban setting.
Designed with the visual aesthetics of a modern character study — suitable for casting, film preproduction, or editorial concept art. Natural posture, relaxed hands, realistic eye focus, and modern body proportions.

Prompt 3.8: Prompt utilizzato per la definizione visiva della protagonista.

Alla fine è stata scelta una versione specifica. Il modello è stato successivamente integrato nel flusso di produzione del cortometraggio, diventando lo standard su cui sono state prodotte tutte le immagini dello storyboard. Si è quindi inserito il file all'interno del progetto, rendendolo così accessibile da qualunque interazione. Di seguito si può osservare il modello della protagonista, generato e utilizzato nel progetto.

Figura 3.13: *Character sheet* dell'artista protagonista del cortometraggio.

3.1.16 Definizione dello stile visivo e sonoro

Nel processo di sviluppo del cortometraggio, la definizione dello stile visivo e sonoro è strettamente legata all'intero impianto narrativo ed è stata costruita attraverso un dialogo costante tra l'autore umano e l'intelligenza artificiale generativa, configurandosi come un processo iterativo in cui sensibilità estetica e potenzialità tecnologiche hanno lavorato assieme. In particolare, lo stile visivo è stato elaborato mediante una serie di interazioni mirate: l'autore ha inizialmente fornito all'IA una successione di indicazioni testuali che chiedevano di fare riferimento ai colori e alle descrizioni contenute nel file con i dati raccolti durante gli eventi musicali.
Successivamente l'autore umano ha descritto nel dettaglio l'atmosfera desiderata, le cromie predominanti e le scelte compositive da trasmettere ad ogni scena. Queste richieste non sono state formulate tutte assieme, in un unico prompt, ma si sono affinate nel tempo, in relazione ai risultati generati. Questo approccio ha permesso di ottenere una coerenza visiva crescente, allineata con lo stile narrativo e contemplativo del cortometraggio.

Per quanto riguarda lo stile sonoro, invece, è importante sottolineare il ruolo centrale della collaborazione con i sistemi generativi. Le tracce audio, così come le componenti ambientali del cortometraggio e gli effetti speciali sono stati concepiti attraverso un meccanismo simile a quello visivo, attraverso la scrittura di prompt descrittivi che ha guidato la successiva generazione delle sequenze audio. In sintesi, la costruzione dello stile visivo e sonoro del cortometraggio si configura come un esempio paradigmatico di co-creazione uomo-macchina, in cui la direzione artistica umana si è intrecciata con la capacità generativa dell'intelligenza artificiale in un percorso di progressiva convergenza espressiva.

3.1.17 Conclusione della fase di Pre-produzione

La fase di pre-produzione del cortometraggio ha rappresentato la base strutturale indispensabile per l'intera realizzazione del progetto. Questa parte del lavoro ha previsto una successione ordinata di passaggi, ciascuno finalizzato a definire e consolidare gli elementi portanti del prodotto audiovisivo.
In principio è stato prodotto un documento ordinato e facilmente fruibile contenente tutti i dati raccolti durante gli eventi musicali iniziali. Questo ha rappresentato la base di partenza e la conclusione di tutto il progetto, anche per verificare i risultati raggiunti.
È stata effettuata poi una fase di brainstorming che ha portato all'elaborazione dell'idea generale e la conseguente stesura del soggetto (con le relative sinossi e logline) e della sceneggiatura. In seguito, si è costruita l'identità del personaggio protagonista attraverso la generazione di un modello visivo dedicato, fondamentale per garantire coerenza e riconoscibilità nelle fasi successive. Il passo successivo è stato lo sviluppo dello storyboard, elemento centrale per la pianificazione visiva delle scene, per il quale sono stati prodotti inizialmente numerosi prompt testuali, specifici per ciascuna inquadratura. Questo ha portato alla creazione di uno storyboard prettamente testuale mentre, solo in seguito, sono state generate le immagini per ogni inquadratura.
Accanto a questi materiali sono stati prodotti altri documenti, fondamentali per il supporto e lo sviluppo del progetto. In particolare, il documento contenente tutte le indicazioni testuali dello storyboard in lingua inglese. Una sua versione ordinata, pulita e priva di indicazioni tecniche ha costituito la base per le varie attività di elaborazione delle immagini e pianificazione delle scene. Questo si è reso necessario in quanto i sistemi di IA utilizzati ragionano e funzionano meglio se interpellati in lingua inglese, da loro supportata in origine.
La completezza e la precisione dei materiali preliminari hanno permesso di impostare le successive fasi operative con grande chiarezza e semplicità, soprattutto per tutta la parte di creazione dei prompt di generazione delle singole clip video. La fase di pre-produzione si è dunque rivelata non essere solo una fase esplorativa e organizzativa ma un momento fondamentale per la buona riuscita del progetto, dimostrando il valore strategico della pre-produzione anche in un procedimento in cui creatività umana e intelligenza artificiale si fondono in un sistema unico e integrato.

3.2 Fase di Produzione

3.2.1 Studio preliminare per la creazione dei prompt

Prima di iniziare a creare i prompt da usare per la generazione delle singole clip video, una fase cruciale dello sviluppo progettuale è stata quella relativa all'analisi metodica e approfondita per comprendere al meglio come strutturare un prompt destinato alla generazione di clip mediante intelligenza artificiale. Questa attività si è rivelata centrale per l'intero progetto poiché ogni elemento successivo era legato alla qualità e alla costruzione dei prompt stessi, determinanti per il successo espressivo e narrativo del prodotto audiovisivo finale.
Si è scelto di raccogliere e studiare diverse risorse informative e guide tecniche reperibili online, con lo scopo di ottenere una panoramica esaustiva sui principi teorici e operativi per la scrittura di prompt efficaci. Sono state selezionate fonti diverse, autorevoli e complementari tra loro, in grado di offrire indicazioni dettagliate, esempi concreti e strutture facilmente replicabili. Tra queste si è dato particolare rilievo alle guide utente ufficiali della piattaforma *Hailuo AI*, strumento specializzato nella generazione video tramite input testuale. I documenti analizzati hanno illustrato le differenze tra prompt base e prompt avanzati, evidenziando l'importanza di indicare con precisione elementi come il movimento di camera, l'atmosfera visiva, la posizione del soggetto e la gestione della luce, fornendo al contempo una vasta gamma di esempi positivi e negativi.
Analogamente, data la scelta di utilizzare *Sora* di *OpenAI* come motore principale per la realizzazione del cortometraggio, è stato fondamentale integrare e confrontare le informazioni accademiche, come quelle riportate da Liu et al. [35], con le linee guida fornite dall'azienda produttrice [36]. Si è quindi fatto uso della documentazione ufficiale per quanto riguarda la generazione di testo, immagini e video, cogliendo le raccomandazioni specifiche per l'ottimizzazione dei prompt destinati ai modelli multimodali. In aggiunta, è stato utilizzato inizialmente un *GPT* personalizzato denominato *Prompt Engineering and Elevator* [37], progettato per fornire supporto nella scrittura avanzata di prompt per video attraverso un approccio fondato su ventisei principi operativi. Questo strumento si è rivelato particolarmente utile per identificare le principali logiche di produzione video e per creare i primi prompt coerenti e specifici.
Ulteriori elementi sono stati tratti da una repository *GitHub*, contenente una vasta collezione altamente specializzata di esempi di prompt cinematici per *Sora* [38]. In questa raccolta, oltre a modelli strutturati e suggerimenti operativi, erano presenti anche prompt completi già utilizzati in contesti reali di produzione, corredati da note tecniche e riflessioni sul linguaggio descrittivo da adottare in relazione alla resa visiva desiderata.
Una volta raccolto e selezionato il materiale, si è proceduto con una fase di elaborazione e sintesi svolta in collaborazione con l'intelligenza artificiale. A quest'ultima è stato affidato il compito di analizzare in profondità i documenti forniti, estrapolando in maniera autonoma tutti gli accorgimenti rilevanti. Il lavoro è stato avviato attraverso un prompt dettagliato che chiedeva di isolare le regole fondamentali presenti all'interno dei file, di confrontare esempi buoni con esempi poco efficaci e di restituire un quadro il più possibile esaustivo e operativo.

> Leggi, analizza, dettaglia il più possibile questo documento, come se fossi un esperto di cinematografia. Fai in modo di ottenere e capire tutti i migliori accorgimenti per ottenere il prompt perfetto come viene descritto in questo documento. In output ti chiedo di mandarmi tutti i principali accorgimenti. Tu dovrai ricordare anche le varie differenze che puoi notare tra un prompt fatto bene e uno fatto male.

Prompt 3.9: Prompt utilizzato per dare il via all'analisi dei documenti trovati online. L'IA ha prodotto quindi un report specifico, indicando tutti i punti chiave che ha trovato nei documenti.

Il risultato è una sintesi organizzata, chiara ed esaustiva in cui sono stati classificati gli elementi fondamentali in un prompt efficace:

- Presenza di dettagli visivi
- Strutturazione narrativa
- Impiego di verbi dinamici
- Controllo della coerenza del soggetto
- Equilibrio tra espressività e chiarezza

L'intelligenza artificiale ha di fatto individuato con precisione quali aspetti fossero essenziali all'interno di un prompt e quali invece risultassero fuorvianti, proponendo esempi comparativi tra prompt ben scritti, mediamente efficaci e completamente errati. Questo confronto ha permesso di acquisire un sistema di riferimento interno che ha guidato le successive fasi creative. Questa prima parte si è conclusa con la formulazione di una lista strutturata di dieci regole fondamentali per la scrittura di un prompt cinematografico di qualità, in seguito capaci di adattarsi sia alla produzione sperimentale che a quella narrativa lineare. Queste regole sono state sempre imposte durante le richieste di generazione dei prompt specifici per le varie scene del cortometraggio.
Successivamente alla prima fase di studio, è stato avviato un secondo stadio di approfondimento, ancora più tecnico, mirato ad affinare ogni singolo aspetto della generazione video. L'intelligenza artificiale è stata quindi invitata a dettagliare ulteriormente le tecniche di movimento di camera, precisando per ciascuna la durata ideale e il suo corretto utilizzo narrativo. È stata inoltre analizzata la gestione dell'atmosfera, sono stati cioè definiti i criteri per descrivere correttamente luce, colore, tono emotivo e tipo di ambiente, tenendo conto delle sfumature che questi elementi devono evocare. Ulteriori approfondimenti hanno riguardato la gestione dei prompt in presenza di più soggetti, la costruzione di ambientazioni complesse e l'ottimizzazione dei riferimenti fotografici. Sono stati infine raccolti suggerimenti sulla descrizione dell'aspetto fisico e psicologico dei personaggi.
In seguito a questa solida preparazione e alla successiva definizione di una serie di linee guida, è stata effettuata una prima prova applicativa. L'intelligenza artificiale è stata invitata a utilizzare tutte le conoscenze apprese, per sviluppare un prompt completo per un cortometraggio realistico. L'obiettivo era quello di realizzare un sistema di scrittura automatico in cui l'intelligenza artificiale si occupasse

della costruzione integrale e dettagliata del prompt, a partire da una semplice frase descrittiva fornita dall'autore umano. Questo approccio è stato pensato per velocizzare l'interazione con il sistema di generazione video *Sora*, garantendo al contempo risultati di alto livello qualitativo e coerenza stilistica. Di seguito si lasciano i risultati della prima prova, presentati in Output 3.8, a partire dall'indicazione fornita con il Prompt 3.10.

> Prova a creare un prompt perfetto per un cortometraggio ultra realistico con protagonista un'artista sui 35 anni, con abiti casual.

Prompt 3.10: Prompt utilizzato come prima prova base per valutare se l'IA avesse appreso quanto indicato nei documenti.

> In uno studio d'artista minimalista, illuminato dalla luce naturale che filtra da una grande finestra a sinistra, una donna di circa 35 anni siede davanti a una tela bianca. Indossa una camicia oversize color lino e jeans sbiaditi, i capelli raccolti in uno chignon disordinato. La stanza è silenziosa, con solo il ticchettio lontano di un orologio.
> La camera parte da un'inquadratura over-the-shoulder che mostra la tela, poi scivola lentamente di lato in dolly, rivelando il suo volto assorto. Un raggio di luce si posa sulla guancia.
> Si alza, prende un pennello e cammina lentamente verso la parete dove sono appesi schizzi a carboncino. Il movimento è fluido, quasi danzante. La camera la segue lateralmente, mantenendo un'inquadratura a mezza figura.
> Poi, un dettaglio stretto sulla sua mano mentre tocca la superficie ruvida del disegno, seguita da uno zoom lento sul suo volto, ora assorto, mentre sospira profondamente.
> L'estetica è ultra-realistica: colori neutri, texture materiche, luci morbide e naturali. L'atmosfera è intima, contemplativa, come in un film indipendente europeo.

Output 3.8: Prompt testuale generato per la realizzazione di quanto richiesto nel Prompt 3.10

Componente	**Contenuto**
Soggetto	Donna 35enne, artista, look casual
Scena	Studio d'artista, luce naturale, arredamento minimale
Movimenti camera	Over-the-shoulder → dolly laterale → close-up mano → zoom lento volto
Atmosfera	Realismo emotivo, toni neutri, silenzio, ritmo lento
Estetica	Ultra realistico, texture visibili, luce ambientale, mood contemplativo

Tabella 3.2: Sintesi descrittiva della scena generata sempre con il Prompt 3.10

Successivamente, l'attenzione si è focalizzata sulla creazione di un prompt fisso e replicabile per quanto riguarda la definizione della protagonista del cortometraggio. Si è compreso fin da subito che la coerenza visiva e narrativa del personaggio principale avrebbe costituito un elemento imprescindibile per garantire la continuità stilistica all'interno delle diverse scene. È stato quindi richiesto all'intelligenza artificiale di generare una descrizione completa, accurata e utilizzabile ripetutamente, capace di restituire sempre lo stesso soggetto, con le medesime caratteristiche somatiche, espressive e stilistiche. Sono state effettuate più iterazioni per raffinare il risultato, definendo in maniera stabile tratti come l'età apparente, la carnagione, l'acconciatura, l'espressione, l'abbigliamento e la postura. Di seguito vengono riportati i risultati ottenuti.

> okay definisci ora un prompt in modo tale che una Ai possa generare sempre lo stesso stile di ragazza. Fallo in inglese.

Prompt 3.11: Prompt scritto per generare una descrizione coerente e dettagliata della protagonista.

> Generate a hyper-realistic digital portrait or full-body image of a contemporary woman in her mid-30s with a slender build, fair skin, and soft natural facial features. Her hair is blonde, lightly wavy, and styled in a relaxed, modern low chignon. She has light blue or green eyes, a calm expression, and a few freckles on her cheeks.
> She is always dressed in the same outfit: an oversized off-white linen shirt with rolled sleeves, high-waisted faded blue jeans, and beige canvas sneakers. Subtle silver earrings or a thin ring may be visible.
> Her look should be consistent across different angles and scenes. Keep a cinematic, neutral-toned aesthetic with soft, diffused studio lighting. Backgrounds should be minimal or softly blurred (beige or gray tones preferred) to focus on the subject.
> Ensure consistent facial structure, eye color, hairstyle, clothing, and body proportions across all images. The tone should remain natural, reflective, and grounded in realism — like a character from a European independent film.

Output 3.9: Descrizione dettagliata dell'artista utilizzabile per ottenere video coerenti tra loro in fase di generazione delle clip video.

Il prompt è stato pensato per essere utilizzato come base ricorrente, impiegandolo come standard descrittivo per la generazione e reinserimento del personaggio principale in nuove clip video.
Tutto questo processo ha permesso di strutturare un sistema di produzione assistita dei prompt estremamente dettagliato, in un contesto nel quale la componente descrittiva testuale è fondamentale per la realizzazione del prodotto audiovisivo. Questo approccio ha reso possibile una sintesi originale tra metodo tecnico e slancio

creativo, ponendo le basi per una scrittura filmica ibrida, a metà tra umano e artificiale, in cui la precisione linguistica del prompt diventa il veicolo principale per ottenere un risultato professionale.

3.2.2 Verifica riferimenti dati narrativi

Prima di iniziare a creare tutti i prompt da utilizzare per generare i video, si è reso necessario procedere con un'analisi approfondita finalizzata a verificare e garantire un'alta corrispondenza tra i dati raccolti durante gli eventi e gli elementi narrativi proposti nel cortometraggio.
È stato innanzitutto effettuato un esame del file contenente la riorganizzazione in tabelle ordinate dei dati raccolti durante gli spettacoli, identificando i concetti principali per costituire la base di riferimento della sceneggiatura e dello storyboard. È stato quindi chiesto all'intelligenza artificiale di creare una prima tabella di riferimenti, visualizzabile in Tabella 3.3, identificando in particolare le corrispondenze presenti nei file di sceneggiatura e storyboard.

> Ciao, prendendo in esame il file RaccoltaDatiNarrazioniParalleleEsportato.pdf, Storyboard Tesi Light.pdf e Sceneggiatura.txt, ti chiedo di dirmi e farmi un elenco preciso di tutti i riferimenti che trovi negli ultimi due file che hanno un riscontro con qualche dato presente nel primo file. Fai quindi un match tra i dati presenti nel primo file e come sono stati usati nei secondi due file. Produci come output una tabella molto semplice, con due colonne, che metta in relazione quanto ti ho chiesto.

Prompt 3.12: Prompt utilizzato per richiedere la generazione di una prima tabella contenente le relazioni trovate.

Dopo aver consolidato una tabella esaustiva di citazioni e riscontri visivi, la richiesta si è spostata sull'integrazione delle preferenze del pubblico espresse all'interno del file dei dati raccolti. Sempre attraverso prompt, sono stati quindi individuati i termini più rilevanti, confrontandoli con lo stile complessivo del corto. Anche in questo caso sono state ricercate le esatte stringhe presenti nel file, avendo cura di non includere i riferimenti numerici presenti (come percentuali o punteggi), poco utili al confronto tematico.
Le iterazioni successive hanno visto un ulteriore approfondimento: sono state inglobate frasi più lunghe ed elaborate, evidenziando come concetti così evocativi potessero diventare di ispirazione per scene particolari. Anche questi passaggi sono stati ricercati testualmente nei file di progetto e ogni istanza è stata puntualmente associata a un preciso momento narrativo. I risultati di queste ulteriori analisi sono presentati nella Sezione I, in appendice.

Dato dal documento	Come è stato ritrovato
"Il fiume che si trova da qualche parte in montagna"	Si trasforma in una scultura urbana che diventa un vero fiume nel racconto visivo dell'artista.
"Passeggiata in montagna, tra ruscelli e prati"	Rievocata nella scena finale in cui l'artista torna alle origini e dipinge nel paesaggio montano.
"Una strada acciottolata di paese"	Ripresa come murale che rappresenta il passato dell'artista e si anima nel suo percorso.
"Bosco dai colori autunnali"	Tradotto nella scena in cui il murale boschivo prende vita e trasporta l'artista in un mondo naturale fantastico.
"Cielo e mare"	Comparsa durante la trasformazione della neve che si scioglie e si espande in cielo e acqua attorno all'artista.
"Rivoli d'acqua"	Visualizzati quando il suolo si scioglie e l'artista viene trasportata dalla corrente in uno stato onirico.
"Mongolfiera"	Elemento trasformativo: le foglie scolpite si sollevano diventando mongolfiere che trasportano l'artista.
"Ogni quadro ha fatto emergere una scena diversa... tra castelli medievali..."	Esplicitato nel passaggio in cui l'artista attraversa un castello che si dissolve in architetture geometriche.
"Il viaggio può essere un percorso di formazione, di scoperta di sé..."	Riflesso nell'intero arco narrativo, in cui l'artista attraversa trasformazioni per riscoprire la propria identità.

Tabella 3.3: Prima associazione tra le citazioni del pubblico, i dati presenti nel documento e i corrispondenti utilizzi narrativi nel cortometraggio

L'intero processo è stato caratterizzato da un ciclo di verifica portato avanti dall'IA: ogni citazione o preferenza del pubblico veniva inizialmente individuata nel file dei dati, ricercata nei testi di sceneggiatura e storyboard, poi annotata nella tabella e infine controllata manualmente per confermare la corrispondenza esatta. Questo percorso di controllo incrociato ha rappresentato un punto di monitoraggio indispensabile, volto a garantire che i dati raccolti non fossero utilizzati in modo approssimativo o parziale, ma che fossero integrati in modo rigoroso, coerente e completo nell'ideazione del cortometraggio.
In termini tecnici, si è trattato di un lavoro di estrazione di testo a partire dai documenti del progetto e di abbinamento semantico tra contenuti testuali e scene visive.
Questa parte di verifica e integrazione è risultata essenziale per diverse ragioni. In primo luogo, ha creato un'ulteriore garanzia di coerenza della sceneggiatura e dello storyboard con i dati raccolti nelle fasi preliminari del progetto, evitando discrepanze o derivazioni arbitrarie. In secondo luogo, ha offerto la certezza che ogni preferenza del pubblico fosse considerata e tradotta adeguatamente, creando un legame diretto tra le aspettative dei destinatari e le scelte stilistiche, compositive e di animazione del cortometraggio.

3.2.3 Generazione dei prompt per la produzione video

La generazione dei prompt per l'intelligenza artificiale generativa è stata centrale nella realizzazione del cortometraggio e ha richiesto un approccio meticoloso, iterativo e integrato con le fasi precedenti del progetto. L'obiettivo di questa fase è la traduzione puntuale e controllata dei contenuti finora prodotti in istruzioni operative chiare e ben definite.
Tutto è iniziato dalla definizione di una struttura standardizzata del prompt, seguendo quanto appreso in precedenza. Questa struttura, costruita su più livelli, serviva come modello di riferimento rigido per ogni singola inquadratura, assicurando omogeneità e chiarezza nella generazione dei vari prompt. Le sezioni fondamentali sono:

- *Obiettivo / Messaggio principale*: che cosa comunica l'inquadratura, la funzione narrativa, simbolica o emotiva.
- *Ambientazione / Location*: dove si svolge la scena, riferimenti ambientali, geografici e atmosferici.
- *Personaggi*: chi è presente nell'inquadratura, con descrizione precisa di età, look, atteggiamento, postura e ruolo.
- *Atmosfera e musica*: tono emotivo, presenza o assenza di suono o musica, e sensazione trasmessa.
- *Stile visivo*: definizione dello stile richiesto (es. realistico, pittorico, onirico, cinematografico, ecc.).

Inizialmente si era scelto di includere un'ulteriore sezione denominata *Tipo di video*, che in seguito però è stata rimossa, perché si è compreso che l'informazione risultava molto più efficace se riportata direttamente nella prima riga del prompt, in forma diretta e sintetica. Nei Prompt 3.13 e Prompt 3.14 si può quindi visualizzare quanto richiesto all'IA e in Output 3.11 la risposta che ha portato alla generazione del prompt da usare successivamente per generare l'inquadratura in questione.
Fin dall'inizio è stato richiesto che ogni prompt dovesse essere scritto in inglese. Questa scelta non era solo legata a una questione linguistica o stilistica, ma a una necessità tecnica. L'intelligenza artificiale, infatti, viene allenata e addestrata a comprendere meglio la lingua inglese, perciò è sicuramente preferibile ottenere dei prompt in inglese per evitare eventuali errori di trascrizione o traduzione.

Ciao, iniziamo ora a generare assieme i prompt che darò all'IA per generare i vari shot che andranno a comporre il cortometraggio finale. Ti chiedo di mettere sempre un richiamo al fatto che i video devono essere ultra realistici e che devono sempre fare riferimento al mockup che allego per quanto riguarda l'artista (quindi questa parte mettila solo nei prompt in cui compare effettivamente l'artista). Ti chiedo di essere molto dettagliato e di usare una struttura come la seguente per descrivere il prompt:
Struttura del Prompt Video:
Tipo di video (es: pubblicitario, motivazionale, educativo, ecc.)
Obiettivo / Messaggio principale (cosa deve comunicare)
Ambientazione / Location (dove si svolge la scena)
Personaggi (chi appare, età, look, ruolo)
Atmosfera e musica (tono emotivo, tipo di colonna sonora)
Stile visivo (realistico, animato, cinematografico, ecc.)
Ricordati di rendere sempre coerenti tra loro i personaggi nelle descrizioni che effettuerai, in più l'IA avrà sempre un'immagine di reference alla sua portata.

Prompt 3.13: Primo prompt utilizzato per iniziare la generazione dei testi da utilizzare per fornire le indicazioni all'IA generativa su come creare le clip video del cortometraggio. In seguito la struttura è stata modificata per essere più efficace.

Il processo di generazione è stato attentamente supervisionato: l'intelligenza artificiale produceva uno o due prompt contemporaneamente e, di volta in volta, si procedeva alla verifica della coerenza, alla correzione di eventuali errori o mancanze e all'approvazione finale prima di passare alle inquadrature successive. Una delle richieste chiave, ribadita costantemente, riguardava la descrizione del personaggio dell'artista. Ogni volta che la protagonista era presente nell'inquadratura, doveva essere descritta in maniera identica e dettagliata, per garantire un riconoscimento visivo coerente e costante in tutte le clip. Di seguito un esempio di testo, generato in tutti i prompt in cui l'artista appare.

The artist is a woman in her mid-30s, slender and fair-skinned, with soft natural facial features. Her blonde, lightly wavy hair is tied into a low, relaxed chignon. She wears an oversized off-white linen shirt with rolled-up sleeves, high-waisted faded blue jeans, and beige canvas sneakers.

Output 3.10: Descrizione dell'artista protagonista, presente in questa forma nella sezione del pompt *Personaggi* ogni qualvolta compare in un'inquadratura.

Allora, ti richiederei di ricominciare a fare questo lavoro ma dall'inizio, ovvero shot 1.1 Con delle leggere variazioni: non serve che scrivi nella struttura del prompt la sezione type of video. Però devi sempre scrivere all'inizio che il video deve essere ultra realistico e in stile cinematico. Inoltre ricorda sempre di includere i movimenti di macchina che devono essere fatti all'interno dell'inquadratura. Infine ricordati che se viene nominata l'artista come personaggio presente in scena, questa deve essere sempre uguale come descrizione! Infatti l'attrice sarà sempre la stessa e deve essere coerente tra le varie clip.
Ti lascio di seguito quindi un esempio di prompt che serve per descrivere l'artista ogni volta che viene nominata:
Generate a hyper-realistic digital portrait or full-body image of a contemporary woman in her mid-30s with a slender build, fair skin, and soft natural facial features. Her hair is blonde, lightly wavy, and styled in a relaxed, modern low chignon. She has light blue or green eyes, a calm expression, and a few freckles on her cheeks.
She is always dressed in the same outfit: an oversized off-white linen shirt with rolled sleeves, high-waisted faded blue jeans, and beige canvas sneakers. Subtle silver earrings or a thin ring may be visible.
Her look should be consistent across different angles and scenes. Keep a cinematic, neutral-toned aesthetic with soft, diffused studio lighting. Backgrounds should be minimal or softly blurred (beige or gray tones preferred) to focus on the subject.
Ensure consistent facial structure, eye color, hairstyle, clothing, and body proportions across all images. The tone should remain natural, reflective, and grounded in realism — like a character from a European independent film.
okay bene ora puoi ricominciare dal primo shot della prima scena e fare i prompt per tutte le inquadrature.

Prompt 3.14: Prompt di rifinitura della struttura da utilizzare, inviato in seguito a una serie di tentativi di prime generazioni testuali. Da osservare la richiesta di utilizzare un testo predefinito per descrivere l'artista e la richiesta iniziale di generare sempre quella specifica tipologia di video.

In un secondo momento si è aggiunta la descrizione dettagliata dei movimenti di macchina. Questa informazione ha consentito all'intelligenza artificiale di interpretare l'inquadratura non solo come un'immagine statica ma come un vero e proprio shot cinematografico.

This video must be ultra-realistic and cinematic in tone and execution. The scene should look like it was captured on a high-end digital cinema camera, with soft, natural lighting and carefully composed visuals.
Main Goal / Message: Introduce the city environment from above, creating a warm and organic sense of space. The city is vibrant yet distant, filled with movement and rhythm.
Setting / Location: Aerial view of the city in the afternoon. The skyline is a mix of historical rooftops and modern buildings. Cars and people move like tiny dots. The light bathes the urban textures in gold.
Characters: No visible characters in this shot — the city itself is the subject.
Mood and Music: Contemplative and subtly melancholic. The audio consists of soft urban ambience: distant traffic, faint wind. Optionally accompanied by a minimal ambient score with long piano notes or light strings.
Visual Style: Hyper-realistic and cinematic. The shot is an Extreme Wide Shot captured with a drone, featuring a slow, descending movement from a high vantage point toward the urban grid. Lighting is natural, warm, and directional, casting long shadows and enhancing building volumes. The palette leans into soft oranges, light grays, and sun-washed concrete. The shot should convey scale and emotional distance simultaneously, as if the viewer is watching from a detached, contemplative perspective.

Output 3.11: Output definitivo utilizzato successivamente come prompt per generare la prima inquadratura del cortometraggio.

In parte del cortometraggio è stato rappresentato un viaggio onirico tra le opere dell'artista che riprendono degli stili particolarmente famosi del passato. Per alcune scene si è dovuto effettuare un particolare cambiamento nel prompt utilizzato per generare le clip video in quanto lo stile richiesto non doveva essere più ultra realistico ma diventare affine ad una particolare corrente pittorica.
In particolare, per ogni scena interessata da cambio di stile (scene 6, 7, 8, 9, 10 e 11), è stata assegnata una corrente pittorica differente, variando dal Puntinismo al Fauvismo, dal Cubismo al Futurismo, fino all'Impressionismo. È stato quindi sviluppato un sistema di frasi introduttive specifiche da inserire nella prima riga del prompt, con formulazioni uguali a quelle riportate nel Prompt 3.15.

This video must be created in the style of Vincent van Gogh, with swirling brushwork, emotional intensity, and rich, vibrant textures that give the scene a dreamlike, poetic resonance.

This video must be designed in the visual language of Italian Futurism, using dynamic compositions, angular fragmentation, and rhythmic repetition to convey energy, perception, and transformation.

Prompt 3.15: Prompt utilizzati per definire in maniera differente lo stile della clip da generare nelle scene in cui veniva richiesto.

Queste istruzioni stilistiche, perfettamente integrate all'interno della struttura dei prompt già elaborata, hanno garantito che ogni scena avesse una sua identità visiva, pur rimanendo all'interno del linguaggio complessivo del cortometraggio.
Questa fase del progetto ha portato alla generazione di tutti i prompt da utilizzare nella successiva creazione delle clip video. Nello specifico, sono state generate in totale 27.877 parole, per un totale di 120 pagine di documento. Si riportano in appendice, nella Sezione J, i prompt così generati, uno per ogni scena.

3.2.4 Modifica dei prompt generati

Durante il processo di generazione video tramite intelligenza artificiale, sono emerse diverse criticità nella resa visiva di alcune scene rispetto alle intenzioni narrative e di regia previste. In particolare, alcuni prompt inizialmente utilizzati non producevano esattamente i risultati attesi: i video generati non rispecchiavano pienamente le richieste in termini di contenuto e atmosfera.
Per correggere tali discrepanze, si è proceduto con una revisione mirata. I prompt ritenuti non funzionali sono stati selezionati direttamente dal documento originale e reinseriti in *ChatGPT*. A quel punto, è stato richiesto all'intelligenza artificiale di modificarli affinché risultassero più precisi, espliciti e coerenti con quanto richiesto dalla narrazione. È stato fondamentale, in questo passaggio, mantenere la struttura originale del prompt, assicurando uniformità nello stile e nella continuità tra le clip.
Le scene che hanno richiesto un intervento più sostanziale sono state:

- Scena 4: riguardante la transizione tra il mondo reale e quello immaginario che segna l'inizio del viaggio dell'artista. Il prompt originario non enfatizzava adeguatamente la trasformazione visiva del paesaggio. Per migliorarlo, è stato richiesto di rendere più chiara e sequenziale l'azione, riducendo movimenti generici e rendendo i cambiamenti visivi più evidenti e controllati.
- Scena 7: ambientata in un campo di grano accanto al mare, comprendeva elementi eterogenei che necessitavano di maggiore coerenza visiva. I prompt originali presentavano troppi movimenti casuali o poco giustificati. Si è intervenuti rendendo l'azione più mirata e leggibile, specificando il ruolo degli elementi naturali e richiedendo un effetto di trasformazione chiaro e continuo.
- Scena 11: inizialmente risultava poco identificabile come momento estivo. È stata richiesta una riscrittura del prompt per aumentare la percezione della stagione, introducendo elementi legati all'estate.

- Scena 12: la versione generata non risultava sufficientemente chiara nel rappresentare il concetto di ritorno alla realtà e la chiusura del ciclo narrativo. Sono stati richiesti approfondimenti nella descrizione del gesto, della postura e dello spazio per rendere più chiare le azioni dell'artista.
- Scena 13: volendo inserire non solo clip statiche, questa scena è stata trasformata con l'intenzione di inserire qualche timelapse per rappresentare in modo più efficace il passare del tempo nei momenti dedicati alla pittura e rafforzare l'idea conclusiva del viaggio dell'artista.

Un ulteriore aspetto emerso nel processo è stato il rischio di omissione della figura dell'artista nei prompt. Per questo motivo è stato più volte ribadito che la descrizione del personaggio dovesse venir ripetuta sistematicamente in ogni prompt in cui l'artista compare, al fine di garantire coerenza lungo tutta la sequenza.
Attraverso questo processo di revisione e riscrittura, si è giunti a una serie di prompt più precisi e funzionali alla generazione delle clip.

3.2.5 Generazione delle clip video

L'intero processo di creazione del materiale audiovisivo è stato condotto tramite l'utilizzo della piattaforma *Sora*, sviluppata da *OpenAI*, uno strumento progettato per la produzione di contenuti video a partire sia da prompt testuali che visivi.
Per tutta questa fase si è scelto di adottare una struttura organizzativa rigida e ben ordinata. Sfruttando la funzionalità della piattaforma, per ciascuna scena del progetto, è stata creata una cartella dedicata, all'interno della quale sono state poi salvate tutte le clip generate con esito soddisfacente. Questa modalità di archiviazione ha garantito un tracciamento puntuale della produzione, consentendo una facile consultazione e comparazione tra le diverse versioni generate, permettendo inoltre di scaricare facilmente le clip migliori.
Ogni clip è stata generata a partire dai prompt appositamente creati, tutti raccolti in un unico documento. I prompt sono stati quindi inseriti all'interno di *Sora*, configurata per produrre video con risoluzione 1080p (Full HD), in formato 16:9 e di durata pari a cinque o dieci secondi. La scelta di durate brevi non è stata casuale, ma frutto di una sperimentazione iniziale: le clip superiori ai dieci secondi, infatti, risultavano spesso incoerenti, meno accurate nella fisica, nella narrazione visiva e nella gestione dell'ambiente o dei personaggi; producevano spesso dei cambi di inquadratura improvvisi e non coerenti, andando a modificare pesantemente la struttura narrativa precedentemente sviluppata. Le brevi durate, invece, garantivano maggiore controllo, precisione e correttezza visiva.
Considerando che il piano *PRO* di *Sora* consente cinque generazioni simultanee, si è elaborata una strategia che ha permesso di ottimizzare il tempo di attesa e al contempo aumentare la varietà delle proposte generate per ciascun prompt. Nella maggior parte dei casi si è deciso di generare due clip da cinque e due clip da dieci secondi con lo stesso prompt, tenendo infine una prima clip relativa all'inquadratura successiva come quinta generazione simultanea. In questo modo, si è potuto mantenere un flusso continuo di produzione e parallelizzare le fasi di generazione. Questo approccio si è rivelato particolarmente efficace per le sequenze complesse, in cui erano richiesti elementi dinamici come movimenti di camera articolati, trasformazioni visive o particolari spostamenti dei personaggi.

Per le scene più semplici, invece, la procedura è stata semplificata. In questi casi venivano generate solamente due clip, una da cinque e una da dieci secondi, per poi passare direttamente al prompt successivo. La valutazione del grado di complessità era effettuata manualmente, in base alle caratteristiche descritte nei prompt e alla tipologia di inquadratura prevista.
Una volta completata la generazione, le clip venivano analizzate attentamente per determinare la loro qualità e idoneità alla buona riuscita del progetto audiovisivo. Le valutazioni comprendevano considerazioni sulla coerenza visiva rispetto al prompt, la resa grafica, l'illuminazione, la profondità di campo e l'assenza di difetti tecnici come distorsioni, errori di prospettiva o imperfezioni morfologiche. Le clip che rispondevano positivamente a questi criteri venivano salvate nella cartella della scena corrispondente, quelle che invece risultavano incoerenti o difettose venivano eliminate. In molti casi, nessuna delle clip create nel primo ciclo era soddisfacente, rendendo necessaria la rielaborazione.
Durante questo processo è stato spesso necessario apportare modifiche ai prompt in corso d'opera. Tali variazioni avvenivano in due forme principali: le correzioni minori riguardavano piccoli dettagli, come l'aggiunta o la rimozione di elementi descrittivi e la precisazione di movimenti; le modifiche più rilevanti richiedevano invece la riscrittura totale del prompt, specialmente nei casi in cui la scena prodotta risultava del tutto disallineata rispetto all'idea iniziale.
Il processo si è rivelato lungo e intensivo, ogni clip impiegava un tempo di generazione variabile, tra un minimo di trenta secondi fino ad arrivare a diversi minuti, a seconda della complessità del prompt e della disponibilità delle risorse. Inoltre, l'esperienza ha dimostrato che la maggior parte dei prompt richiede più cicli di generazione per ottenere un risultato soddisfacente, rendendo la reiterazione una componente indispensabile della pipeline creativa.
In totale sono state generate circa 1000 clip. Di queste, 614 sono state valutate soddisfacenti e sono state ammesse ad una successiva selezione che ha portato alla scelta definitiva di 248 clip da scaricare ed utilizzare per la realizzazione del prodotto audiovisivo.
In conclusione, l'utilizzo di *Sora* non ha comportato una completa sostituzione dell'azione autoriale ma ne ha richiesto un'integrazione continua, fatta di controllo, adattamento e selezione. L'approccio adottato ha consentito di ottenere un archivio delle clip ordinato e coerente, estremamente utile nella successiva fase di post-produzione.

3.2.6 Modifiche in seguito alla generazione

Un aspetto interessante emerso durante il processo di generazione dei video ha riguardato la fase di modifica e perfezionamento delle clip già prodotte. In alcune circostanze, infatti, è accaduto che i video generati dai prompt testuali risultassero globalmente validi, ma presentassero elementi specifici indesiderati o comunque migliorabili, come ad esempio difetti nel movimento, distorsioni visive, incoerenze minori nello sfondo, oppure semplicemente scelte estetiche non del tutto coerenti con la direzione artistica desiderata. In questi casi non si procedeva necessariamente con una rielaborazione completa della clip, ma si adottava una strategia di intervento mirato tramite l'utilizzo della funzione *Remix* offerta dalla piattaforma *Sora*. Questo strumento permette di intervenire su un singolo video già esistente, utilizzandolo

come base di partenza e applicando su di esso un nuovo prompt descrittivo, per indicare con precisione i cambiamenti richiesti.

Figura 3.14: Un esempio di utilizzo dello strumento *Remix*. La clip in alto, generata inizialmente, presenta un'incoerenza riguardante la presenza di lentiggini sul volto dell'artista. Utilizzando la funzione di *Remix* è stato possibile chiedere di toglierle, ottenendo la clip rappresentata in basso, perfettamente in linea con il modello della protagonista sempre utilizzato.

I prompt del *Remix*, pur mantenendo la stessa struttura testuale degli input originari, venivano redatti in forma più sintetica e focalizzata, con l'obiettivo di agire su alcuni dettagli specifici. A completamento della procedura, lo strumento *Remix* consentiva anche la selezione del livello di intensità della modifica. Questo parametro, espresso attraverso una scala numerica, influenzava la radicalità dell'intervento: un valore elevato generava una trasformazione più marcata rispetto alla clip originale, mentre valori bassi producevano cambiamenti minimi. Tuttavia, l'esperienza ha mostrato che l'effettivo impatto visivo del *Remix* non era sempre proporzionale al valore scelto. Per questo motivo, nella quasi totalità dei casi, si è optato principalmente per l'utilizzo di valori alti, in grado di garantire un cambiamento maggiormente percettibile e più facilmente controllabile a posteriori.
Il risultato di questa operazione non era sempre prevedibile. Talvolta la nuova clip generata attraverso il *Remix* rappresentava una versione effettivamente migliorata e rispondente alle esigenze creative, risolvendo i problemi presenti nella versione originale. In altri casi, invece, la modifica non correggeva gli errori precedenti o ne

introduceva di nuovi, generando versioni meno coerenti o addirittura visivamente peggiorative. Quando ciò accadeva ripetutamente, si tralasciava la clip originaria e tutte le sue varianti, procedendo con la produzione ex novo di clip alternative per la stessa inquadratura.
Questo approccio ha permesso di ottimizzare tempi e risorse, potendo contare su video già di buona qualità che necessitavano soltanto di una leggera rifinitura per essere utilizzabili. La funzione *Remix* ha quindi costituito uno strumento efficace per raffinare la resa visiva del progetto, contribuendo al controllo del risultato finale, senza però compromettere il flusso produttivo complessivo.

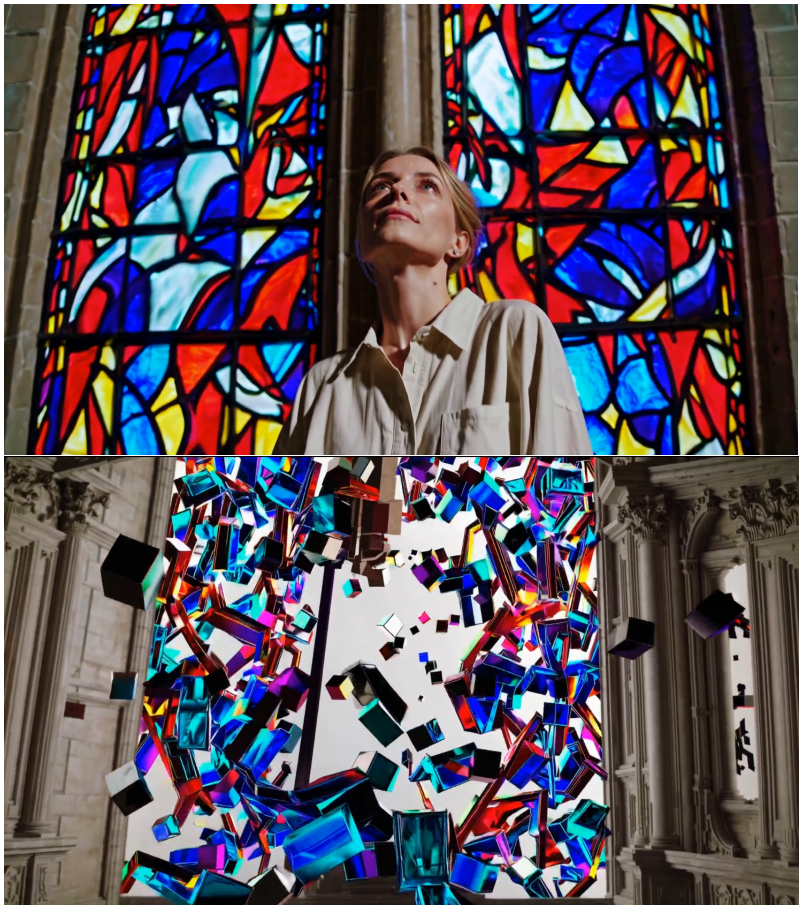


Figura 3.15: Un esempio di utilizzo dello strumento *Remix*. A partire da una clip precedentemente generata (figura in alto) è stato creato un nuovo prompt che richiedeva di cambiare completamente la clip, mantenendo solo l'ambientazione. Il risultato ottenuto (figura in basso) è ottimo, in quanto ha permesso di mantenere un'alta coerenza visiva cambiando però gran parte della clip.

3.2.7 Creazione delle transizioni

Un'ulteriore funzione messa a disposizione dalla piattaforma e utilizzata nella fase finale del processo di produzione dei video è la modalità *Blend*. Questo strumento si è rivelato particolarmente utile nella costruzione di transizioni fluide tra clip differenti, contribuendo in maniera sostanziale alla continuità narrativa dell'intero cortometraggio. A differenza delle generazioni effettuate per creare le varie clip,

basate esclusivamente sull'elaborazione di prompt testuali, la funzione *Blend* non richiede input descrittivi: si basa infatti unicamente su due elementi visivi di partenza, ovvero due clip precedentemente generate oppure, alternativamente, una generata e l'altra caricata dal dispositivo in uso. L'obiettivo dello strumento è quello di generare una clip di transizione che connetta visivamente le due inquadrature selezionate, trasformando progressivamente l'immagine iniziale in quella finale.
La configurazione della transizione avviene attraverso la scelta di una curva di interpolazione che determina la modalità con cui le due clip verranno fuse. Tra le opzioni disponibili figurano: *transition* (transizione progressiva canonica), *sample* (basata su campionamento) e *mix* (interpolazione visiva più marcata). In rari casi è stata presa in considerazione la modalità *custom*, in quanto produceva risultati poco controllabili sul piano visivo.
Dal punto di vista operativo, la strategia adottata ha previsto l'impiego delle cinque generazioni simultanee offerte dal piano *PRO*, ripartite tra le varie curvature disponibili. In genere, per ciascuna coppia di clip, venivano selezionate le tre curve principali, destinando le ultime due generazioni rimanenti a un'ulteriore variazione della curva ritenuta più promettente.
I risultati ottenuti tramite la funzione *Blend* si sono rivelati particolarmente interessanti sotto il profilo estetico. A differenza delle tradizionali transizioni usate nel montaggio classico, come dissolvenze incrociate, tagli a nero o entrate con *fade in*, venivano prodotte trasformazioni visive particolari, in cui gli elementi presenti nella prima clip si deformavano o si riconfiguravano gradualmente fino a fondersi con quelli della seconda. Questo effetto dinamico non solo aumentava il livello di coesione visiva tra le sequenze, ma rafforzava anche la narrazione, suggerendo una continuità tematica ed emotiva tra momenti altrimenti distinti.
L'integrazione di questa funzione ha dunque arricchito il progetto dal punto di vista narrativo, introducendo un nuovo livello di espressività visiva. Grazie a questa modalità, le clip non sono state semplicemente accostate, ma letteralmente rivisitate e completate insieme in modo armonico. Il valore aggiunto di tale approccio è evidente non solo in termini estetici, ma anche sul piano percettivo: la transizione si è infatti trasformata da un semplice momento tecnico di raccordo a un potenziale elemento di maggiore significato.
In Figura 3.16 è possibile vedere un esempio di transizione creata utilizzando la funzione *Blend*.

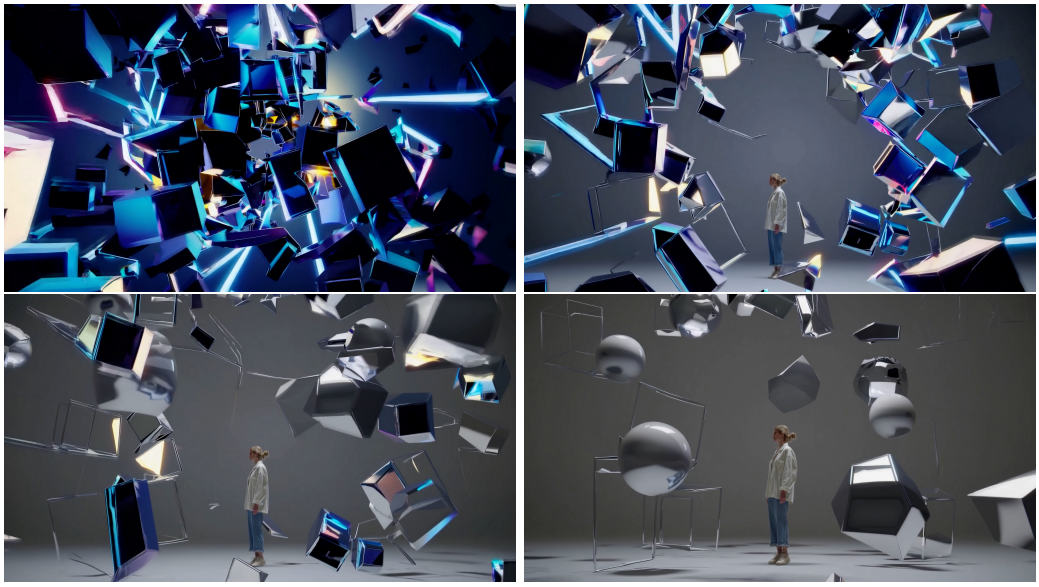


Figura 3.16: Esempio di transizione creata con la funzione *Blend*. In alto a sinistra è presente un frame della prima clip, seguito da due immagini intermedie che ben rappresentano la transizione creata. In basso a destra, invece, è presente un frame della seconda clip, punto di arrivo e conclusione del passaggio.

3.2.8 Generazione della colonna sonora

L'ultima fase nel processo di produzione ha riguardato la creazione della colonna sonora del cortometraggio, interamente supportata anche in questo caso dall'intelligenza artificiale. Il flusso di lavoro ha avuto inizio con la definizione di prompt dettagliati, necessari per guidare in modo accurato la generazione delle musiche. Per questo motivo è stata aperta una nuova conversazione su *ChatGPT*, sempre all'interno dello stesso progetto, dedicata esclusivamente alla definizione dei prompt per la colonna sonora del cortometraggio.
Attraverso un prompt iniziale, visionabile in Prompt 3.16, è stato chiesto al modello di creare, in modo strutturato, dei prompt specifici per ogni singola scena del cortometraggio. L'istruzione chiave era quella di garantire una forte coerenza stilistica e narrativa tra l'audio generato e il contenuto visivo di ciascuna scena. Prima di procedere alla generazione vera e propria tramite *Suno*, si è resa necessaria una fase di raffinamento: alcuni prompt sono stati mantenuti pressoché invariati, mentre altri hanno richiesto modifiche sostanziali per adattarsi meglio al formato di input richiesto dalla piattaforma. Ogni singola traccia è stata progettata e selezionata con l'intento preciso di essere pienamente coerente con la scena associata, rispettandone il tono, dinamiche narrative e ritmo visivo. Si lascia in Output 3.12 un esempio di questo procedimento.

Ciao, dobbiamo passare adesso alla generazione di tutta la colonna sonora del cortometraggio. Quindi quello che ti chiedo di fare è prendere la sceneggiatura e o lo storyboard e o tutti i vari prompt che abbiamo già generato e creare dei prompt precisi e puntuali che possano servire agli strumenti di intelligenza artificiale per generare in maniera coerente e corretta le colonne sonore e tutte le basi musicali. Ovviamente dovranno essere tutte e quante solo strumentali senza voce e dovranno seguire il ritmo della narrazione adattandosi di scena in scena a quello che accade. Ovviamente la musica è fondamentale, dovrà trasportare lo spettatore attraverso tutte le fasi del viaggio dell'artista quindi deve essere più concitata in alcuni momenti, più veloce, più incalzante e più rilassata e morbida in altri. Ricordati di descrivere bene lo stile che deve avere la canzone, le modalità, i tipi di note e quello che deve anche un po' suscitare la canzone stessa. In questo modo riusciremo a generare correttamente tutto quanto. Genera quindi vari prompt ordinati scena per scena e ovviamente puoi anche decidere di cambiare musica più volte nella stessa scena se necessario. L'importante è che spieghi bene come e quando viene utilizzata qualsiasi cosa. Ricorda infine che i prompt dovranno essere scritti in lingua inglese. Grazie!

Prompt 3.16: Prompt utilizzato inizialmente per avviare la generazione di testo dedicato a creare le varie tracce audio del cortometraggio.

Scene 6 – Mountain and Snow Metamorphosis
Cold, crystalline soundscape with glass harmonica, sparse piano notes, and granular textures. Transition into warmth as snow turns to water—use slow legato strings and subtle orchestral swells. The transformation should feel serene, fragile, yet uplifting.

Output 3.12: Prompt generato dall'IA per creare la colonna sonora di una scena del cortometraggio.

Una volta pronti i prompt, si è passati alla fase di generazione su *Suno*, ottenendo per ciascun input due alternative musicali differenti. Questo ha fornito un notevole margine di scelta, consentendo di selezionare la traccia più adatta alla scena. Veniva scaricata solo la versione migliore, tenendo talvolta la seconda opzione come riserva o potenziale adattamento per altre parti del progetto.
Parallelamente alla colonna sonora principale, sono state generate alcune tracce audio extra, sempre strumentali, da utilizzare in caso di necessità. Queste, non collegate a scene specifiche, sono state archiviate come riserva per eventuali criticità in fase di montaggio o in caso di tracce principali non perfettamente efficaci.

3.2.9 Conclusione della fase di Produzione

Al termine della fase di produzione, il progetto ha raggiunto un livello di completezza tale da permettere di predisporre in maniera rapida la parte di post-produzione del cortometraggio. La produzione, durata diverse settimane, ha visto l'alternarsi e l'integrazione di numerose attività, con l'attenta supervisione dell'autore umano e con il supporto continuo dell'intelligenza artificiale. Ogni passaggio è stato affrontato in modo sequenziale con la possibilità di rivedere il lavoro in ogni momento, per migliorare, adattare o perfezionare i contenuti, secondo un approccio dinamico e iterativo.
Il primo passo operativo ha riguardato la stesura dei prompt per la generazione video. L'intero progetto si è fondato su una solida fase preliminare di studio dei prompt, grazie alla quale è stato possibile comprendere le logiche, le tecniche e le strategie migliori per costruire istruzioni testuali efficaci. L'analisi della documentazione tecnica e delle guide ufficiali si è poi concretizzata in una collaborazione diretta con l'intelligenza artificiale, che ha elaborato una sintesi dei principi fondamentali emersi. Successivamente, si è passati alla redazione dei prompt attraverso un modello descrittivo articolato che prevedeva sia sezioni dedicate a indicazioni più tecniche che sezioni maggiormente concentrate sull'estetica del prodotto audiovisivo. La generazione dei prompt ha riguardato ogni singola inquadratura del cortometraggio, con una grande attenzione alla coerenza del personaggio principale, la cui descrizione veniva ripetuta in forma identica ogni volta che appariva in scena, per mantenere costante il suo riferimento estetico.
In seguito è stata avviata la fase di generazione video tramite l'intelligenza artificiale. Questa attività è stata svolta interamente attraverso la piattaforma *Sora* di *OpenAI*, che consente la produzione di video realistici a partire da prompt testuali. La scelta delle impostazioni da utilizzare, relativa alla durata delle clip, non è stata casuale, ma frutto di un'attenta sperimentazione che ha evidenziato come le clip più lunghe tendano a perdere coerenza, qualità e precisione. Il sistema utilizzato metteva a disposizione la generazione simultanea di cinque clip. Normalmente, per ogni prompt venivano prodotte due versioni da cinque secondi, due da dieci e una aggiuntiva riferita già all'inquadratura successiva. Questo sistema ha permesso di mantenere un flusso costante di produzione, limitando i tempi di attesa e incrementando la varietà dei risultati. Ogni video prodotto veniva valutato attentamente sulla base di criteri come correttezza visiva, aderenza al prompt, assenza di errori e rispetto dei movimenti di macchina. Solo le clip che soddisfacevano questi requisiti venivano salvate e archiviate all'interno della cartella dedicata alla scena di riferimento. In molti casi, però, le clip prodotte non erano apprezzabili, rendendo necessario generare nuove versioni a partire dallo stesso prompt o, nei casi più complessi, procedere a una riscrittura completa dello stesso.
Si è sviluppata infine la colonna sonora, realizzata per intero con l'ausilio dell'intelligenza artificiale. È stata creata una conversazione dedicata esclusivamente alla stesura dei prompt musicali, con l'obiettivo di generare una base strumentale per ogni scena del cortometraggio. *ChatGPT* ha prodotto per ogni scena un prompt musicale differente, calibrato sulle esigenze narrative, sul tono emotivo e sulla dinamica visiva della scena. I prompt, dopo una fase iniziale di scrittura, sono stati in diversi casi modificati per adattarsi meglio al formato richiesto da *Suno*. Per ciascun prompt, la piattaforma ha restituito due alternative audio, offrendo una

doppia opzione per ogni sequenza. Nella maggior parte dei casi è stata scaricata la versione ritenuta migliore, anche se in alcuni episodi sono state mantenute entrambe le tracce, con l'intenzione di poter eventualmente utilizzare la seconda in caso di necessità.
Anche durante la fase di produzione si è reso necessario un intervento costante di revisione e modifica. Molti prompt sono stati riadattati sulla base dei risultati ottenuti. In alcuni casi si è trattato di modifiche minime, come la precisazione di un termine o l'aggiunta di un elemento visivo, mentre in altri casi si è dovuto intervenire in modo radicale, riscrivendo completamente il prompt o riformulando l'intera sequenza. In aggiunta, sono stati utilizzati strumenti avanzati, messi a disposizione dalla piattaforma *Sora*, come la funzione *Remix* che ha consentito di modificare in maniera selettiva le clip già esistenti, correggendo solo gli aspetti maggiormente problematici. In fase conclusiva è stata adottata anche la modalità *Blend*, che ha permesso di creare transizioni visive fluide e coerenti tra le clip.
Complessivamente, nella fase di produzione, sono state generate oltre 1000 clip video, di cui 614 ritenute valide e 248 effettivamente selezionate per essere scaricate e impiegate nel montaggio finale. Sono state redatte 120 pagine di prompt testuali, per un totale di 27.877 parole, utili a generare tutte le clip video. A tutto questo si aggiungono decine di tracce audio originali, create per dare vita alla colonna sonora del cortometraggio.

3.3 Fase di Post-produzione

3.3.1 Montaggio e rifinitura delle sequenze

La fase di montaggio del prodotto audiovisivo finale è iniziata con un'accurata organizzazione del girato, già precedentemente suddiviso in cartelle distinte, ognuna corrispondente a una specifica scena del progetto. Questa modalità di archiviazione, impostata sin dalle fasi iniziali della produzione, ha permesso di gestire in modo fluido il montaggio preliminare. Attraverso questa operazione è stato possibile identificare rapidamente quali porzioni delle clip fossero realmente funzionali alla narrazione e quali, invece, si rendessero superflue o non sufficientemente coerenti sul piano narrativo.
Una volta completata questa prima selezione visiva, sono state importate anche tutte le tracce audio. Ogni traccia è stata scelta in relazione alla scena a cui era destinata, sia per coerenza emotiva che per compatibilità strutturale. Il montaggio è quindi proseguito con l'unione progressiva degli elementi visivi e sonori, con specifica attenzione al ritmo musicale. In questo stadio, particolare cura è stata posta alla gestione delle pause, degli stacchi e delle transizioni tra le varie sequenze. La colonna sonora, con le sue dinamiche interne, ha fornito alcune linee guida per definire al meglio i tempi di taglio e i momenti di enfasi, permettendo così una costruzione narrativa efficace.
Tuttavia, il processo ha richiesto una certa flessibilità e capacità di adattamento, in quanto il materiale visivo a disposizione, per quanto numeroso, era comunque limitato e non sempre esaustivo rispetto alle esigenze del montaggio. In alcune circostanze, infatti, è stato necessario rinunciare a soluzioni creative più articolate, possibili solo con una maggiore varietà di clip.

Una scelta importante, perseguita lungo tutta la fase di montaggio, è stata quella di non sovraccaricare le clip con effetti visivi o interventi post-produttivi importanti. L'intento autoriale era quello di preservare l'integrità e l'identità visiva originaria delle clip così come erano state generate dalla piattaforma di intelligenza artificiale. Gli unici interventi sono stati quelli minimi e funzionali alla narrazione o alla transizione tra le sequenze, come leggeri *fade in* o *fade out*, eventuali correzioni di colore o adattamenti di durata per questioni di ritmo. La volontà di evitare manipolazioni eccessive ha contribuito a costruire un prodotto visivamente autentico, in cui il linguaggio dell'intelligenza artificiale emerge con chiarezza, dialogando con le scelte di montaggio dell'autore umano in un equilibrio compositivo calibrato e consapevole.
Nonostante l'intero progetto sia stato profondamente guidato e caratterizzato dall'impiego di strumenti basati su intelligenza artificiale, la fase di montaggio non ha rappresentato una particolare innovazione tecnica o metodologica rispetto agli standard tradizionali. Il lavoro ha seguito infatti le modalità operative consuete tipiche del montaggio audiovisivo.
È importante dunque sottolineare che, in un progetto in cui l'intelligenza artificiale ha avuto un ruolo centrale nella generazione dei contenuti visivi e sonori, la fase di montaggio ha rappresentato un momento in cui l'intervento umano è stato predominante, confermando l'importanza dell'esperienza manuale nel risultato dell'opera finale.

3.3.2 Ottimizzazione del colore e della luce

Nel corso della post-produzione, un'attenzione particolare è stata dedicata alla fase di correzione del colore e all'uniformazione dell'aspetto visivo tra le diverse clip del progetto. Il lavoro principale svolto è stato effettuare un processo di corrispondenza cromatica e luminosa quanto più possibile precisa, finalizzata a garantire una maggiore continuità estetica e coerenza percettiva tra le varie scene. Data la natura del materiale video generato, che non presenta le caratteristiche di profondità colore e flessibilità proprie dei file girati in scala logaritmica o in formati ad alta gamma dinamica, si è scelto consapevolmente di non apportare modifiche troppo decise. Le operazioni si sono limitate a leggere correzioni locali, intervenendo su parametri come la temperatura colore, la tonalità dominante, il contrasto e la luminosità, sempre con l'obiettivo di omogeneizzare il materiale in funzione della sequenza narrativa di riferimento.
Non si è trattato dunque di un'operazione di *color grading* in senso stretto, volta a trasformare radicalmente l'aspetto delle immagini o ad attribuire loro una precisa identità stilistica indipendente. L'intento principale è stato quello di armonizzare le clip tra loro, eliminando eventuali discontinuità evidenti, causate dalla piattaforma di generazione video. In alcuni casi è stato necessario agire per compensare piccole oscillazioni tra clip consecutive, come nelle scene in cui il ritmo visivo è sostenuto da una progressione musicale o da un flusso di immagini particolarmente omogeneo. Questa fase ha contribuito in modo rilevante alla qualità complessiva del progetto, rafforzando l'unitarietà visiva e migliorando l'esperienza di fruizione. Il risultato è una coerenza visiva che non rivoluziona l'origine sintetica delle immagini ma anzi la valorizza, presentando una continuità stilistica discreta ed efficace, ottenuta attraverso un lavoro puntuale e calibrato sulla luce e sul colore, effettuato scena per scena, inquadratura per inquadratura.

3.3.3 Generazione e sincronizzazione del voiceover

Il processo di generazione dei voiceover, destinati a dare voce ai pensieri e alle parole dell'artista all'interno del cortometraggio, ha richiesto una fase preliminare preparatoria.
Inizialmente è stato redatto un messaggio in formato vocale molto lungo e complesso, elaborato direttamente dall'autore umano, fornendo all'intelligenza artificiale tutte le istruzioni necessarie affinché il contenuto prodotto risultasse coerente, efficace e perfettamente integrato con la struttura narrativa già definita (Prompt 3.17). Nel dettaglio, il messaggio includeva una panoramica completa dell'intero progetto, richiamando esplicitamente i documenti sviluppati (sceneggiatura e storyboard). Sulla base di questi materiali, è stato richiesto all'intelligenza artificiale di produrre un voiceover articolato, adatto ad accompagnare le immagini con riflessioni, pensieri e osservazioni espressi direttamente dalla voce narrante dell'artista. L'obiettivo era quello di rendere percepibili le emozioni interne del personaggio, i momenti di dubbio, scoperta o consapevolezza ma allo stesso tempo fornire al pubblico una chiave di lettura utile per comprendere l'evoluzione narrativa del cortometraggio.
Un'attenzione specifica è stata dedicata alla redazione del monologo finale, elemento conclusivo del prodotto audiovisivo, pensato per lasciare allo spettatore un messaggio. Su questo passaggio si è lavorato in maniera più insistente: sono state richieste all'intelligenza artificiale numerose versioni differenti del testo, ognuna con variazioni stilistiche ed enfasi narrative più o meno marcate. L'intelligenza artificiale ha prodotto differenti variazioni, fornendo più bozze. In questo caso, l'autore umano ha selezionato le parti più efficaci emerse nelle varie versioni, per poi procedere a una riscrittura manuale mirata, che ha portato alla stesura definitiva del monologo conclusivo. Questa fase di composizione ha rappresentato un momento di sintesi tra input automatici e intervento autoriale, confermando ancora una volta la necessità di una presenza umana attiva per garantire qualità e intensità comunicativa.
Nel complesso, il voiceover prodotto attraverso questo metodo ha permesso di conferire profondità e interiorità al racconto, trasformando le immagini in sequenze emotivamente partecipate, guidate da una voce narrante che accompagna lo spettatore per l'intero arco della storia. Il lavoro svolto in questa fase ha dimostrato come l'intelligenza artificiale, se opportunamente indirizzata, possa costituire uno strumento potente per l'elaborazione di contenuti verbali complessi, lasciando comunque aperto lo spazio per l'intervento umano decisivo nei momenti più delicati e significativi del progetto.

Ciao, adesso dobbiamo passare alla prossima fase. Dobbiamo andare a costruire tutti i sottotitoli, tutte le parole che l'artista pensa all'interno di questo cortometraggio. Infatti tutte le cose parlate saranno semplicemente dei pensieri dell'artista che fa mano a mano che andiamo avanti con il cortometraggio. E' quindi fondamentale che tu riprenda la sceneggiatura, lo storyboard che trovi tra i file del progetto e agisca come un esperto sceneggiatore per creare al meglio questi dialoghi interiori che poi andranno messi come sottofondo all'interno del cortometraggio. Nella prima scena per esempio, di far ben capire qual è il suo problema principale, il fatto che non sa come continuare questo quadro, che è molto indecisa se rimanere in città oppure tornare verso il suo paese di montagna. Durante la scena 2, devi ovviamente mettere in forte contrasto le cose positive della città, la fama, le prime mostre, contro invece la tranquillità, la pace che aveva a casa. Nella scena 3 il viaggio onirico, diciamo, sta per iniziare, quindi lei deve essere molto confusa, non deve ben capire cosa stia succedendo attorno a lei, cosa che continua nella scena 4, quando l'albero prende vita e deve iniziare questo viaggio in maniera quasi sorpresa. Poi dalla scena 5 in poi, quando il viaggio entra nel suo culmine, dovrai ovviamente far capire che lei all'interno di questo viaggio riconosce le sue varie opere, riconosce questi mondi che lei ha dipinto, ha scolpito in passato. è fondamentale poi fare un po' di alternanza, un po' di contrasto all'interno del viaggio tra lei, che comunque ha piacere di rivedere le sue opere, ma vorrebbe comunque tornare alla normalità, vorrebbe capire, vorrebbe trovare una soluzione al suo problema principale. Man mano che si va avanti nel viaggio, lei può comunque spiegare i vari stili pittorici usati, può prendere ispirazione, per esempio anche poi nella scena del castello potrebbe finalmente essere contenta e sperare che il viaggio stia per finire, ma no, anche qua continua, parte un altro viaggio assurdo tra le forme geometriche, arriva in una città nella scena 10, dove c'è tanta confusione, dove anche la musica sarà molto forte, e invece c'è un forte contrasto con la scena 11 dove viaggia attraverso le stagioni e dove c'è più tranquillità, dove si troverà sicuramente più a suo agio e poi finalmente finisce questo viaggio, lei è contenta nella scena 12, quindi deve far capire che lei finalmente ha trovato una quadra per fare in modo di poter continuare a finire questo quadro, con un po' di malinconia ovviamente quando prepara tutti gli scatoloni per andarsene. Nell'ultima scena deve partire un monologo finale con un forte messaggio di speranza, di spinta nel tornare alle proprie origini senza paura, lei deve essere comunque contenta di aver fatto la propria scelta, quindi il pubblico deve essere sicuramente molto attratto da questo messaggio, deve condividerlo e ovviamente l'artista poi deve dire che è riuscita a completare il suo quadro. Ora agisci quindi come uno sceneggiatore, utilizzando tutte le sue competenze per creare al meglio questi dialoghi, ricordandoti che saranno tutti pensieri dell'artista e che devi dividerli scena per scena in modo tale che io possa poi aggiustarli e tenerli in maniera più semplice possibile. Grazie.

Prompt 3.17: Prompt fornito tramite input vocale per generare nel dettaglio tutti i voiceover del cortometraggio.

Il voiceover creato come descritto precedentemente è stato successivamente doppiato dall'autore umano. Questa scelta è stata fatta per conferire al progetto una maggiore coerenza espressiva e un controllo preciso dell'intonazione, del ritmo e delle sfumature vocali richieste da ciascun segmento narrativo. Tuttavia, poiché il cortometraggio ha come protagonista una figura femminile, si è reso necessario trasformare la voce maschile dell'autore in una più adatta al personaggio rappresentato nel prodotto audiovisivo. A tal fine, è stato scelto di utilizzare la piattaforma *Eleven Labs*, strumento avanzato per l'elaborazione vocale, in grado di trasformare la voce dandole caratteristiche differenti. Volendo mantenere intatta l'intonazione e la carica empatica delle registrazioni originali, si è scelto di non procedere con un approccio più tradizionale come quello del *text-to-speech* che, seppur più rapido e automatizzato, tende a produrre risultati più piatti e meno coinvolgenti, specialmente in lingua italiana.
Sebbene la soluzione adottata su *Eleven Labs* abbia notevolmente complicato la pipeline tecnica del progetto, richiedendo di fatto una doppia fase di lavorazione, si è ritenuto che questa modalità potesse offrire una resa interpretativa più aderente alla sensibilità narrativa richiesta. Tuttavia, i risultati ottenuti non sono sempre stati soddisfacenti. La resa della voce femminile generata, pur risultando tecnicamente corretta, non ha raggiunto i livelli di naturalezza e profondità emotiva auspicabili per un contesto altamente espressivo come quello di un cortometraggio introspettivo. Nonostante ciò, si è comunque scelto di proseguire con questa metodologia, valorizzando, quanto più possibile, la coerenza vocale e la fedeltà interpretativa delle versioni finali. Una volta generato ciascun file audio è stato progressivamente sincronizzato con le rispettive scene del progetto audiovisivo. Il processo di montaggio della componente vocale del cortometraggio è stata completata quindi al termine di questa fase di sincronizzazione.

3.3.4 Generazione degli effetti sonori

La realizzazione degli effetti sonori del cortometraggio è avvenuta attraverso un procedimento semplice, distribuito lungo tutta la fase di montaggio. Mentre il lavoro procedeva, si presentavano progressivamente, di scena in scena, diverse necessità legate alla richiesta di suoni che potessero accompagnare azioni specifiche, enfatizzare passaggi emotivi o contribuire alla costruzione di ambientazioni realistiche. A differenza delle altre fasi progettuali, in cui era stato adottato un approccio ampiamente pianificato, in questo caso la creazione degli effetti sonori è stata gestita in maniera manuale: ogni volta che si presentava un'esigenza concreta, si procedeva subito alla sua produzione.
Per la generazione degli effetti sonori è stata utilizzata la piattaforma *Eleven Labs*, già impiegata per la realizzazione del voiceover. Tale strumento dispone infatti di una sezione dedicata alla creazione di effetti sonori che sfrutta tecnologie di intelligenza artificiale generativa. In questo contesto, l'autore umano ha avuto il compito di scrivere in modo diretto e conciso dei prompt descrittivi, mirati a istruire la generazione dell'effetto desiderato: si trattava di comandi semplici, spesso costituiti da poche parole chiave, come *walking on snow* oppure *wind blowing*.
Ogni richiesta generava automaticamente quattro diverse versioni dello stesso effetto sonoro, tra le quali scegliere e scaricare la migliore per la scena in questione. Questa varietà di alternative ha rappresentato un vantaggio significativo, in quanto ha per-

messo di valutare più opzioni per ciascun effetto, rendendo più snella la procedura di generazione degli effetti sonori.
Il lavoro è stato eseguito con attenzione per tutte le scene del cortometraggio e ha richiesto una supervisione costante, pur mantenendo una gestione facile e intuitiva grazie alla semplicità della piattaforma. La possibilità di generare rapidamente effetti sonori personalizzati ha evitato il ricorso a librerie sonore preesistenti e ha permesso di costruire un ambiente acustico completamente creato dall'IA.

3.3.5 Scelta del titolo

Durante la prima fase di sviluppo, il progetto è stato provvisoriamente intitolato *L'artista e la città*. Tale scelta, dettata da esigenze pratiche e organizzative, si è rivelata fin da subito temporanea. Prevedendo di affrontare in una fase finale la questione relativa al titolo, con un approccio più approfondito e consapevole e attraverso una riflessione che tenesse conto dell'evoluzione narrativa e simbolica del cortometraggio, la revisione del titolo è stata inserita intenzionalmente all'interno della fase conclusiva della lavorazione, proprio per valutare tutti gli elementi del progetto. A tal fine è stato attivato un percorso co-creativo articolato, basato sull'analisi del voiceover, della sceneggiatura, dello storyboard e dei materiali visuali, sfociato in una sessione di brainstorming partecipata tra l'autore e l'intelligenza artificiale. Sono state sperimentate diverse formule linguistiche, da titoli più descrittivi a strutture più simboliche, da costruzioni narrative a titoli metaforici. Alcune opzioni riprendevano elementi naturali o visivi, altre evocavano stati d'animo o condizioni percettive. Alcune varianti ricorrenti includevano costruzioni come *Dove X respira* oppure *La tela X* (dove X rappresenta diverse opzioni), mentre altre si avvicinavano a strutture binarie più compatte. Tutte le soluzioni sono state valutate in relazione all'identità visiva e narrativa del film.
Tra le proposte, ha progressivamente preso forma il titolo *La Tela Oscura*, selezionato poi come titolo definitivo. La scelta è avvenuta per la sua efficacia simbolica, per la coerenza con il percorso della protagonista e per il richiamo al mistero. La *tela* rappresenta sia l'opera artistica ma anche la superficie su cui si specchia il viaggio interiore della protagonista; l'aggettivo *oscura* ne sottolinea la dimensione ambigua, nascosta, ancora in attesa di essere compresa.
A seguito della definizione del titolo, si è aperto uno stadio di progettazione grafica volta alla sua visualizzazione. Questa fase ha incluso la disposizione del testo su due righe, l'uso di caratteri in maiuscolo, la selezione di un font coerente con l'atmosfera del cortometraggio e la scelta di una colorazione chiara, tendente al bianco, con uno sfondo trasparente. Dopo alcune iterazioni, è stata individuata la combinazione tipografica definitiva, equilibrata e visivamente integrabile con le immagini del progetto.
Il titolo *La Tela Oscura* rappresenta quindi l'esito di un processo intenzionale e collaborativo. Non è stato semplicemente un compito assegnato, ma costruito, discusso e formalizzato a partire dall'interazione tra la struttura narrativa e l'identità estetica del cortometraggio.
Di seguito il risultato finale.

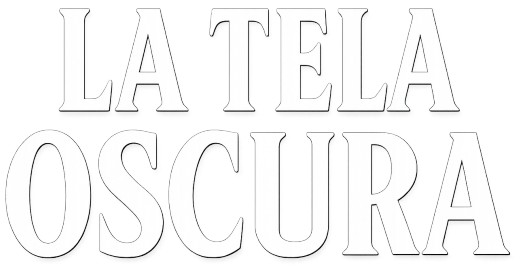


Figura 3.17: Rappresentazione grafica del titolo, scelta in seguito a diverse interazioni e prove effettuate con l'intelligenza artificiale generativa.

3.3.6 Revisione del video finale

La fase finale del lavoro di post-produzione ha previsto un'attenta revisione complessiva del video, condotta attraverso numerose visualizzazioni ripetute su dispositivi differenti, al fine di verificare la qualità del prodotto audiovisivo in ogni suo aspetto tecnico ed estetico. Questo passaggio è stato fondamentale per individuare eventuali discrepanze nei livelli audio, regolati poi con precisione per garantire un ascolto uniforme e bilanciato, sia nei passaggi più intensi che nei momenti più quieti. Sono stati inoltre inseriti i titoli di coda e il titolo iniziale, definendo così in modo completo la struttura formale del cortometraggio. Questa fase conclusiva ha previsto un ulteriore controllo della coerenza visiva tra le varie sequenze: si è verificato che la resa dei colori, dei toni e delle luci fosse omogenea e che non vi fossero variazioni improvvise o incoerenze percettive che potessero compromettere l'unitarietà stilistica del progetto.
Tutti questi interventi hanno portato all'elaborazione del prodotto audiovisivo nella sua versione definitiva, risultato di una costruzione attentamente guidata, in cui ogni fase è stata pensata per contribuire alla migliore realizzazione del cortometraggio finale.

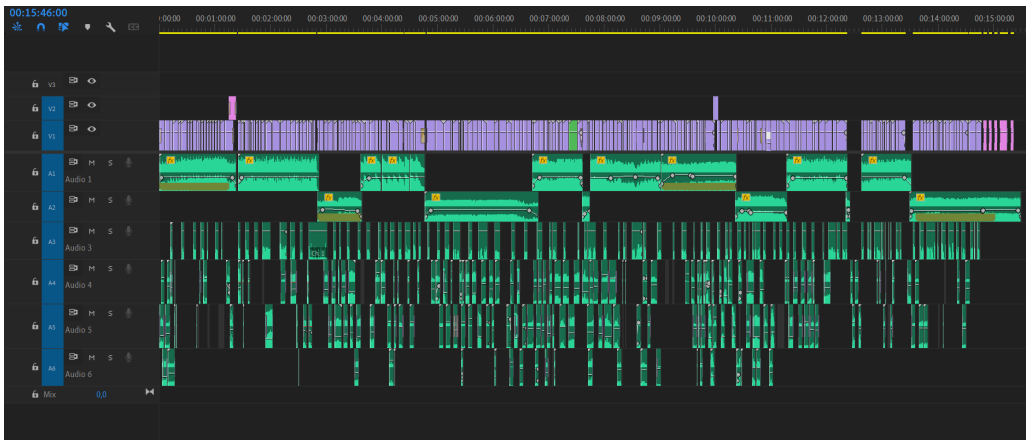


Figura 3.18: Timeline finale del progetto audiovisivo.

3.3.7 Conclusione della fase di Post-produzione

L'intera fase di post-produzione ha consentito di trasformare una raccolta di materiali generati in un prodotto audiovisivo compiuto, coeso e narrativamente strutturato. Il primo passaggio è stato rappresentato dal montaggio preliminare delle sequenze, possibile grazie a una strutturazione ordinata del materiale visivo già suddiviso per scena. Questo ha permesso di avviare una fase di selezione e taglio, durante la quale si è identificato cosa mantenere e cosa eliminare. In seguito, sono state progressivamente aggiunte le tracce audio, generando un primo livello di sincronia narrativa tra suono e immagine.
Si è poi svolto un lavoro di rifinitura, legato al colore e alla luce. Pur non trattandosi di una *color grading* approfondita, sono stati effettuati interventi mirati per garantire omogeneità cromatica tra le diverse clip, soprattutto in termini di tonalità, contrasto e bilanciamento colore. La stessa attenzione è stata posta nella stesura del voiceover, elaborato a partire da un messaggio fornito all'intelligenza artificiale, contenente tutte le indicazioni narrative necessarie. Il testo è stato generato scena per scena e successivamente doppiato dall'autore umano, prima di essere trasformato in voce femminile attraverso *Eleven Labs* e infine sincronizzato con precisione sulla timeline del progetto.
A queste fasi si è aggiunta la generazione degli effetti sonori, gestita in tempo reale durante il montaggio: ogni effetto necessario veniva identificato, descritto con un breve prompt e generato tramite intelligenza artificiale in quattro alternative tra cui scegliere. La fase finale è stata infine dedicata alla revisione completa dell'opera: verifica dei livelli audio su dispositivi diversi, inserimento dei titoli, controllo dell'omogeneità visiva e messa a punto degli ultimi dettagli grafici e ritmici.
Il risultato di tutte queste attività ha portato alla realizzazione del video conclusivo, costruito in modo progressivo ma con una visione unitaria e coerente. Il cortometraggio è frutto di un processo ibrido, in cui l'apporto umano e quello dell'intelligenza artificiale hanno dialogato in maniera complementare. In appendice, nella Sezione K, sono disponibili tutte le indicazioni necessarie per accedere alla visione del cortometraggio nella sua forma finale.

3.4 Sviluppo di un modello di IA per la generazione di prompt per video

3.4.1 Problema e motivazioni

Una parte molto impegnativa del progetto è stata la scrittura dei prompt testuali necessari per la generazione dei video con l'intelligenza artificiale. Ogni descrizione deve infatti essere precisa, lunga e dettagliata per ottenere risultati visivi coerenti e di qualità. Questo processo richiede all'autore umano molto tempo e diventa un'operazione sempre più complessa man mano che il numero delle scene aumenta. Durante la fase di produzione è stato possibile rendere parzialmente automatica anche questa fase, mantenendo però manuale l'inserimento dei vari prompt che trasformavano le indicazioni contenute nello storyboard in comandi specifici e altamente specializzati per la generazione di video sulle piattaforme di riferimento. Successivamente, ogni

risposta contenente i prompt generati, doveva essere copiata e inserita in un file testuale a parte, rendendo questo procedimento ripetitivo e altamente macchinoso. Per questo motivo, si è pensato a una diversa modalità di creazione, maggiormente automatizzata e più adatta allo scopo di generare un intero prodotto audiovisivo tramite intelligenza artificiale. È stato deciso quindi di allenare un modello di IA già esistente e pre-addestrato attraverso una fase di *fine-tuning*, ovvero una personalizzazione basata sull'impartizione di contenuti specifici.
Questo sistema è stato poi utilizzato in locale, grazie a uno script in *Python* che ha permesso di leggere automaticamente da un file le descrizioni delle singole inquadrature di ogni scena e, interrogando il suddetto modello, ha consentito di generare i prompt in modo continuo e veloce, oltre che in maniera dettagliata. Così facendo, non si è reso più necessario inserire manualmente i comandi nella piattaforma ma è stato il sistema stesso a leggerli e trasformarli nei prompt complessi indispensabili per la creazione delle clip video, salvandoli automaticamente e in maniera ordinata nel file dedicato.
È stato interessante analizzare e mettere in pratica questa modifica non prevista fin da subito, andando a confrontare poi i risultati ottenuti, anche in termini di tempistiche.

3.4.2 Creazione dei dataset

Un passaggio essenziale per poter procedere con la fase di *fine-tuning* del modello di intelligenza artificiale è stata la preparazione del dataset, ovvero il materiale testuale su cui il sistema sarebbe poi stato addestrato. Se non adeguatamente progettato, l'intero processo di allenamento risulterebbe inefficace o addirittura fuorviante rispetto agli obiettivi prefissati. Per questa ragione, la costruzione del dataset è stata affrontata in due modalità distinte. La prima ha previsto l'utilizzo di una raccolta di prompt già esistente, disponibile pubblicamente tramite una repository su *GitHub* [38]. Tale dataset contiene una serie di esempi che, a fronte di input testuali generici o concettuali, fornisce una versione più articolata e sviluppata del prompt, concepita per guidare i modelli di generazione video. Questo approccio ha consentito di dare al sistema una base iniziale ampia e varia, utile per familiarizzare con le dinamiche narrative, le scelte visive e le strutture sintattiche comunemente adottate nei prompt audiovisivi.
Di seguito si riporta un esempio rispettivamente di input e output facenti parte del dataset in questione. Al prompt *Technological Innovations in Healthcare* è associata la risposta *Use a combination of wide and close-up shots to depict cutting-edge healthcare technologies. Incorporate interviews with experts and detailed visuals to provide an educational documentary experience.* Questa e diverse altre coppie di input e output dovrebbero essere in grado di allenare il modello, abituandolo a individuare le sequenze cinematografiche adatte ad ogni situazione richiesta.
Tuttavia, seppur utile, questa prima base dati si è dimostrata insufficiente rispetto al grado di precisione richiesto da un progetto audiovisivo interamente prodotto con IA. L'obiettivo iniziale della pipeline produttiva non era semplicemente quello di generare testi tecnici, ma quello di creare prompt operativi e ben strutturati, contenenti tutte le componenti essenziali per la generazione audiovisiva. In particolare, si è osservato che un prompt efficace deve includere con chiarezza almeno cinque cate-

gorie: ambientazione, personaggi coinvolti, stile visivo, scelte registiche e obiettivo espressivo dell'inquadratura.
Per questo motivo, si è deciso di affiancare al primo dataset un'altra raccolta di dati molto più mirata, costruita interamente a partire dal materiale utilizzato durante la produzione del cortometraggio audiovisivo. Nello specifico, tutti i prompt generati durante la fase creativa con l'aiuto di *ChatGPT*, poi verificati, raffinati e validati manualmente, sono stati raccolti e strutturati affinché risultassero la base di partenza per predisporre un dataset più coerente rispetto al contesto progettuale e più vicino allo stile e alla qualità richiesta, utilizzandolo poi come base per una seconda fase di *fine-tuning* del sistema.
Il vantaggio principale di questo secondo dataset risiede nella sua natura attinente alla realtà del progetto: tutti i prompt contenuti al suo interno sono stati pensati per partire da input concreti, cioè uno storyboard già definito, che è esattamente il tipo di input previsto dalla pipeline di produzione finale. In tal senso, non era più necessario che il modello imparasse a immaginare da zero le scelte registiche o stilistiche, ma era fondamentale che avesse le istruzioni per tradurre in modo preciso, completo e funzionale ciò che lo storyboard richiedeva in termini di espressione visiva.
Il processo potrebbe essere in futuro raffinato e migliorato, includendo un numero più elevato di prompt, provenienti da ambiti e prodotti audiovisivi differenti, per una maggiore conoscenza nell'addestramento del modello di intelligenza artificiale. Parte del dataset utilizzato è disponibile in appendice alla Sezione M.
Questa fase di creazione dei dataset, articolata su due livelli complementari, il primo più generico e introduttivo, il secondo più specifico e aderente alla produzione cinematografica effettiva, è stata fondamentale per porre delle basi solide al successivo processo di *fine-tuning*. La capacità del sistema addestrato di generare prompt precisi, articolati e in linea con le esigenze narrative è infatti direttamente collegata alla qualità e alla strategia adottata nella preparazione di questi dati di addestramento.

3.4.3 Fine-tuning del modello di IA

Per effettuare il *fine-tuning* del modello di intelligenza artificiale è stata impiegata la piattaforma *Google Colab*. Questa scelta è stata dettata dalla possibilità di sfruttare GPU performanti, come le *NVIDIA Tesla T4*, che risultano particolarmente adatte per compiti di addestramento come quello del progetto in esame, permettendo anche tempi di esecuzione accettabili. La prima fase del processo ha previsto l'installazione delle librerie necessarie per l'addestramento. Sono state utilizzate le librerie *transformers* e altri strumenti legati alla gestione dei modelli linguistici, come *unsloth*, una soluzione che consente di effettuare *fine-tuning* in maniera ottimizzata e leggera soprattutto in termini di risorse, rendendola ideale quindi per ambienti con capacità computazionale limitata.
Successivamente è stato avviato il *Fast Language Model*, strumento utile per caricare modelli pre-addestrati e accessibili a tutti, in particolare quelli basati sull'architettura *LLaMA*, introdotta da Touvron et al. [39] e sviluppata da *Meta*. A partire dal modello *llama 3.2* con 3 miliardi di parametri, si è proceduto con l'importazione dei dataset preparati nelle fasi precedenti. Questi sono stati convertiti in un formato compatibile per l'addestramento, organizzati sotto forma di coppie *user-assistant*: più precisamente l'input dell'utente è la parte di *user* e il corrispettivo output è l'*assistant*, ovvero il prompt testuale dettagliato che il sistema deve apprendere.

Successivamente, con i dati pronti all'utilizzo, è stata avviata la fase di training vera e propria. Attraverso l'uso delle librerie precedentemente installate, il modello ha iniziato ad apprendere dalle coppie di dati fornite. In questa fase è stato necessario monitorare diversi parametri come, ad esempio, il tasso di apprendimento, la lunghezza del contesto, il *batch size* e il numero di epoche. Più volte sono state eseguite modifiche per ottimizzare il processo e ottenere risultati qualitativamente migliori. Il parametro di riferimento, utilizzato per comprendere l'andamento dell'addestramento, è stato principalmente il *training loss*, ossia la misura dell'errore che il sistema commetteva nel riprodurre la risposta corretta.
Durante questa fase, sono stati effettuati test di inferenza a intervalli regolari per verificare in tempo reale l'efficacia dell'addestramento. Questo procedimento ha permesso di comprendere quanto il modello stesse imparando effettivamente dalle coppie di dati e se fosse necessario proseguire o modificare i parametri dell'allenamento. Questa fase ha avuto un ruolo importante anche per evitare fenomeni di *overfitting*, ovvero l'apprendimento eccessivo dal dataset, con conseguente perdita di generalizzazione. Una volta raggiunto un livello di performance soddisfacente, il sistema addestrato è stato esportato nel formato *GGUF*, pensato per l'ottimizzazione e l'esecuzione in locale. È stata quindi scaricata l'intera cartella contenente i pesi e la configurazione del modello direttamente dalla piattaforma di *Google Colab*, dove l'addestramento è stato eseguito.
Una volta installato in locale, il modello allenato è completamente funzionante senza bisogno di internet o altri supporti, diventando uno strumento ancora più efficace per i progetti audiovisivi con pipeline legata all'IA. È importante sottolineare che l'intero processo di addestramento è stato eseguito sia sul primo dataset, più generico, che sul secondo dataset, costruito a partire da prompt reali utilizzati nella produzione del cortometraggio. Tuttavia, l'attenzione principale è stata rivolta al secondo, più completo, coerente e specificamente progettato per rispondere alle esigenze effettive della pipeline creativa. Il sistema addestrato, basato su dati reali, è stato successivamente testato in vari scenari, per avere un confronto netto con il livello raggiunto dai prompt in precedenza.
Il codice integrale utilizzato per la fase di *fine-tuning* è disponibile in appendice, alla Sezione N.

3.4.4 Utilizzo all'interno della pipeline

Al fine di semplificare e automatizzare ulteriormente la pipeline di produzione, si è voluto sfruttare il modello di intelligenza artificiale addestrato e installato localmente per rendere più efficiente l'intero processo di generazione dei prompt per la produzione video. Per farlo è stato sviluppato uno script in linguaggio *Python*, riportato di seguito. Questo codice ha lo scopo di automatizzare completamente la sequenza di operazioni necessarie per interrogare il sistema più volte, in maniera ordinata e continua, senza rendere obbligatorio l'intervento manuale per salvare le risposte generate per ogni prompt. Questo script può essere personalizzato in maniera semplice, rendendo possibile l'organizzazione più strutturata del file di output con i prompt generati; è sufficiente modificare alcune righe inserendo un controllo che, per ogni input contenente la scritta determinante l'inizio di una nuova scena, preveda la memorizzazione della stessa sequenza sul file di output, rendendo anche quest'ultimo ordinato.

Di seguito si riporta lo script utilizzato:

```
import subprocess
import time

input_file = "input.txt"
output_file = "output.txt"
model_name = "video_model_v2"

def chiamata_ollama(prompt):
    result = subprocess.run(
        ["ollama", "run", model_name],
        input=prompt.encode('utf-8'),
        stdout=subprocess.PIPE,
        stderr=subprocess.PIPE
    )
    return result.stdout.decode('utf-8').strip()

def main():
    with open(input_file, 'r', encoding='utf-8') as infile:
        lines = [line.strip() for line in infile if line.strip()]

    if not lines:
        print("Nessun prompt trovato nel file.")
        return

    with open(output_file, 'w', encoding='utf-8') as outfile:
        for idx, prompt in enumerate(lines, start=1):
            print(f"Eseguo richiesta {idx}...")
            response = chiamata_ollama(prompt)
            outfile.write(f"{response}\n\n")

    print("Tutte le risposte sono state salvate correttamente")

if __name__ == "__main__":
    main()
```

La logica alla base del codice è relativamente semplice: viene precedentemente predisposto un file di input testuale contenente, riga per riga, una serie di descrizioni sintetiche relative a tutte le inquadrature da generare. Questo file può contenere, ad esempio, le descrizioni già presenti nello storyboard sviluppato durante le fasi precedenti del progetto, oppure prompt più grezzi, scritti manualmente, che delineano in modo generale la volontà di rappresentazione visiva di ciascuna clip. Lo script legge poi ciclicamente ogni riga del file di input, trattandola come un'istruzione da trasmettere al modello addestrato. Per ogni descrizione letta, viene effettuata una chiamata all'IA attraverso una funzione dedicata che si occupa di avviare un sottoprocesso in grado di interrogare l'istanza locale del sistema allenato. Il prompt grezzo viene così fornito in input al modello e il sistema attende la generazione della relativa risposta. Una volta ricevuta, viene salvata in modo ordinato all'interno del

file di output, garantendo una corrispondenza diretta tra ciascun input e la rispettiva generazione. Il ciclo continua finché tutto il documento viene elaborato. Il sistema può essere ulteriormente personalizzato per scrivere, ad esempio, un'intestazione ogni volta che cambia il numero della scena, oppure per separare gruppi di prompt in base ai differenti blocchi narrativi del progetto audiovisivo. Queste personalizzazioni possono variare da produzione a produzione, in base alle esigenze specifiche del team creativo e possono essere facilmente implementate nel codice riportato sopra.
L'adozione di questo sistema automatizzato ha permesso di risolvere una delle criticità operative riscontrate durante la fase di produzione del progetto, dovuta alla generazione dei prompt gestita manualmente tramite successive richieste rivolte all'intelligenza artificiale. In quel contesto, era necessario effettuare la richiesta, attendere la risposta, verificarne la correttezza per il successivo trasferimento nel documento finale. Inoltre, la lettura dello storyboard da parte del modello non era sempre affidabile e necessitava di ripetuti controlli da parte dell'autore umano. Con l'introduzione del sistema addestrato ed eseguito in locale e grazie allo script automatizzato, tutto questo processo è stato reso molto più rapido, fluido e sistematico. La generazione dei prompt può infatti essere avviata attraverso un solo comando, lasciando che sia il sistema a occuparsi di tutte le fasi, dalla lettura all'output finale. Questo cambiamento porta a un alleggerimento del carico operativo e a una maggiore efficienza nell'intera pipeline creativa. Rimane comunque fondamentale, anche in un sistema altamente automatizzato, il ruolo del controllo umano. La verifica da parte dell'autore resta cruciale per garantire che i prompt generati siano coerenti, appropriati e adatti al contesto della scena. In alcuni casi, il modello potrebbe produrre descrizioni troppo generiche, eccessivamente complesse o lievemente incoerenti rispetto all'intento originario. Per questo motivo, una revisione e rifinitura del lavoro prodotto rappresenta una fase imprescindibile all'interno del flusso di lavoro.

3.4.5 Risultati del modello addestrato

I risultati prodotti dal sistema allenato sono molto interessanti in quanto rispettano pienamente la struttura di prompt composta da più sezioni, rendendo più efficace la produzione di clip video. Allo stesso tempo però, il modello ha mantenuto una certa libertà, non allineandosi completamente con i prompt inseriti nel dataset che avrebbero potuto influenzare in maniera troppo evidente quanto generato dal modello allenato.
Di seguito si riporta un confronto tra il prompt creato da *ChatGPT* e quanto generato dal sistema addestrato e utilizzato in locale, a partire dallo stesso input, ovvero il testo estratto dallo storyboard. Si nota una buona aderenza tra i due, soprattutto per quanto riguarda la struttura del prompt ma nonostante questo le descrizioni non sono uguali, frutto del fatto che durante l'addestramento è stato evitato l'*overfitting*, fenomeno che avrebbe portato ad un'eccessiva somiglianza tra le due risposte.

Aerial view of the city in the afternoon, shot with a drone. Extreme Wide Shot, Drone - Slow Descend, None, None, Warm atmosphere, natural afternoon light, long shadows to add depth.

Prompt 3.18: Input utilizzato per generare i due prompt definitivi visualizzabili in Output 3.13 e Output 3.14. In particolare, questa è la descrizione presente nello storyboard dello Shot 1.1.

This visual style is ultra-realistic and cinematic, aiming to show the world not just from its perspective, but also as if it were alive — a breathing, growing entity with a story of its own.

Main Goal / Message: To establish that this city is more than just a physical place — it has memories, emotions, and characters.

Setting / Location: Aerial view of the city in the afternoon. Shot with a drone. Extreme Wide Shot. The camera slowly descends through the air as it moves toward the city below — low enough to feel the movement of air around it.

Characters: None visible from this height. The entire landscape unfolds like a map — a rich tapestry of natural light, architectural forms, and movement.

Mood and Music: Warm atmosphere, natural afternoon light, long shadows to add depth. A soft ambient sound begins — distant wind, water flowing in the distance. No music. This scene should be almost silent.

Visual Style: Ultra-realistic and cinematic. Extreme Wide Shot with a slow descending camera movement that captures the entire cityscape below. The drone flies low, allowing for a sense of motion through the air. Light is golden, natural. Shadows are long and expressive — trees lean in, windows glow warmly from inside. Everything feels grounded and alive.

Output 3.13: Output generato dal modello addestrato. Da notare la somiglianza nella struttura con quanto prodotto da *ChatGPT*, mostrato in Output 3.14

This video must be ultra-realistic and cinematic in tone and execution. The scene should look like it was captured on a high-end digital cinema camera, with soft, natural lighting and carefully composed visuals.

Main Goal / Message: Introduce the setting — a Northern Italian coastal city — from a high, detached viewpoint, revealing the layered beauty and movement of the urban environment.

Setting / Location: Aerial view of a city in Northern Italy, located on the coast, during the afternoon. The drone reveals a patchwork of terracotta rooftops, narrow medieval alleys, modern buildings, and small harbors. The sea is visible in the distance, shimmering under the sunlight, softly lapping against the port structures. The streets below are populated by people and vehicles, though distant and miniature.

Characters: No visible characters in this shot — the cityscape is the sole protagonist.

Mood and Music: Warm and introspective. City sounds are muted and atmospheric: distant hum of traffic, seagulls, a breeze from the sea. A sparse, slow ambient score may be added (piano or synth pads).

Visual Style: Hyper-realistic and cinematic. Extreme Wide Shot captured with a drone in slow descending motion. The afternoon sun casts long, angled shadows over the buildings. The lighting is golden and natural. The framing emphasizes the contrast between the scale of the city and the quiet, solitary tone of the film's beginning.

Output 3.14: Output generato da *ChatGPT*. Il confronto con quanto prodotto dal modello addestrato, mostrato in Output 3.13, è un'ulteriore conferma dell'ottimo livello raggiunto dopo l'allenamento.

Il modello addestrato è stato poi testato inserendo in input dei comandi grezzi e poco articolati, per studiare la sua capacità nel renderli efficaci secondo la migliore struttura per la generazione di video. Di seguito sono riportati degli esempi, completamente slegati dal progetto audiovisivo in analisi, indicando prima i prompt grezzi e poi gli output strutturati generati dal modello.

Video of a young man walking down a street, while singing a song.

Video of a large castle between high moutains fully covered from the snow.

Prompt 3.19: Prompt testuali poco strutturati inseriti per indagare l'abilità del modello addestrato nel produrre prompt specifici e utili alla realizzazione di clip video mediante IA.

This video must be created in ultra-realistic style, showing the everyday life of ordinary people in a Northern Italian coastal city.
Main Goal / Message: To portray the emotional and physical recovery of the characters as they rebuild their lives.
Setting / Location: Same street from the beginning, but now with more people moving through it. The atmosphere is renewed — it's warmer and brighter than before.
Characters: Young man (mid-20s), walking down the street, singing softly. His voice may be interrupted by other sounds or voices in the background—traffic, conversations, a distant train whistle.
Mood and Music: Hopeful but fragile. The young man sings with an open heart but his voice is hesitant at times. As he walks, his melody begins to fade into ambient city noise — a soft echo of what once was lost.
Visual Style: Hyper-realistic and cinematic. Extreme Close-Up of the singer as he walks down the street, singing softly. His face, posture, and movements are natural — like he's sharing a private moment with the world but also leaving it behind. The background is composed by many visual elements that reinforce its urban feel — traffic, windows reflecting light or shadows, people walking or stopping to look at him, distant sounds like traffic lights changing, footsteps in alleys, the train whistle fading away into silence.

This video must be ultra-realistic and cinematic, focusing on the stillness and grandeur of the natural landscape surrounding the castle. The shot should amplify the sense of isolation and beauty in the scene.
Main Goal / Message: To portray the castle as part of a vast winter world that is quiet and peaceful, emphasizing its connection to the environment.
Setting / Location: High mountain peaks fully covered by snow, distant trees, sky covered with mist.
Characters: No characters appear—only the landscape. The castle remains still at center frame, surrounded by open space.
Mood and Music: Contemplative. Sound design focuses on natural wind, distant water flow if present, and snowfall in low tones. No music or sound effects are added.
Visual Style: Ultra-realistic and cinematic. Extreme Wide Shot with a slow dolly-in as the camera approaches from a distance. The landscape is open and still—no movement. The castle is at center frame. Snow covers everything except its walls. In the background, mountains stretch high into the sky covered in uniform snow. Light is soft and pale.

Output 3.15: Prompt dettagliati generati a partire dagli input forniti e mostrati in Prompt 3.19.

Capitolo 4

Risultati e conclusioni

4.1 Valutazione dell'efficacia della pipeline con IA

4.1.1 Confronto con i metodi di produzione tradizionali

Il confronto tra i metodi tradizionali di produzione audiovisiva e la pipeline adottata in questo progetto, basata sull'impiego estensivo dell'intelligenza artificiale, consente di mettere in luce differenze sostanziali non solo dal punto di vista operativo, ma anche sul piano culturale e progettuale. La produzione tradizionale richiede un articolato sistema di fasi interdipendenti, tra pre-produzione, riprese e post-produzione, che comportano l'impiego di numerose risorse umane e logistiche. Basti pensare, ad esempio, alla necessità di effettuare accurati sopralluoghi per la selezione delle ambientazioni da usare, alla selezione e gestione di attori tramite casting, alla pianificazione delle riprese con troupe specializzate o alla realizzazione di complesse scenografie. Ogni elemento richiede tempo, competenze specialistiche e un investimento economico non indifferente: una pipeline basata su IA, invece, permette di risparmiare tempo e costi, come viene anche riportato in uno studio di Zhang et al. [16].
La pipeline sviluppata attraverso l'intelligenza artificiale riduce in maniera drastica la complessità strutturale delle fasi produttive. Ambientazioni, attori, movimenti di macchina e persino voiceover sono generati a partire da input testuali, riducendo al minimo l'intervento manuale e annullando del tutto la necessità di una logistica fisica. Questa trasformazione porta con sé vantaggi evidenti in termini di rapidità, accessibilità e scalabilità. Come osservato nello studio di Li [40], l'adozione di strumenti generativi permette una forma di prototipazione narrativa pressoché immediata, in cui l'ideazione e la realizzazione possono avvenire in una finestra temporale estremamente ridotta rispetto ai metodi tradizionali.
Uno degli aspetti più evidenti riguarda la possibilità di effettuare un'iterazione rapida: mentre nel cinema tradizionale ogni modifica al girato implica costi e tempi elevati (nuove riprese, riorganizzazione delle risorse, montaggio aggiuntivo), con l'IA è possibile rigenerare una clip, un'immagine o un effetto audio in pochi secondi. Questo ha aperto scenari radicalmente nuovi per l'esplorazione creativa, permettendo una sperimentazione più libera e meno vincolata da costi e infrastrutture. Tuttavia, come evidenziato in uno studio sulla convergenza tra IA e produzione video [41], questa facilità di accesso può comportare anche una perdita di controllo stilistico e una forte standardizzazione delle scelte visive, fenomeno dovuto al fatto che le IA

generative tendono a produrre risultati secondo pattern statistici ricorrenti, appresi dai data set di addestramento.
Inoltre, mentre l'elaborazione manuale si fonda su un approccio fortemente collaborativo e centrato sulle capacità, ma soprattutto sulla sensibilità umana, la pipeline guidata dall'intelligenza artificiale tende a spostare la centralità del processo creativo verso l'autore singolo, capace di orchestrare il tutto tramite comandi scritti, trasformandosi in un *regista-moderatore*, sensazione condivisa da diverse persone intervistate nel lavoro di Inie et al. [42]. Ciò ha implicazioni importanti anche sul piano del linguaggio audiovisivo: se il cinema tradizionale si costruisce sul set, nel dialogo tra regia, attori, fotografia e scenografia, il video generato da IA prende forma attraverso l'interpretazione dei prompt da parte del modello utilizzato. La scrittura testuale sostituisce in larga parte la regia visiva e la grammatica filmica viene parzialmente mediata da codici impliciti nell'architettura dei modelli. Anche durante questo progetto è stato visto, infatti, che utilizzare una struttura del prompt ben definita e standardizzata permette di ottenere i risultati migliori: la chiave per realizzare il progetto è individuare la giusta combinazione affinché il modello possa lavorare al meglio all'interno della rete. Questo comporta ovviamente una forte uniformazione dei prodotti audiovisivi generati.
Un altro punto di confronto significativo riguarda la gestione delle emozioni e dell'empatia. Nei metodi tradizionali, la direzione degli attori, l'interpretazione vocale e l'uso della colonna sonora sono strumenti fondamentali per suscitare emozioni complesse. Nella pipeline basata sull'IA, questo compito è delegato a sistemi che simulano espressioni emotive sulla base di metriche sintetiche, con risultati ancora in parte limitati. Un chiaro esempio di questo problema è emerso nella conversione vocale del voiceover: l'empatia veicolata dalle voci sintetiche è ancora lontana da rappresentare in maniera fedele quelle umana.
Infine, va sottolineato che la pipeline adottata, pur offrendo numerosi vantaggi in termini di accessibilità, velocità e sperimentazione, richiede comunque un alto livello di supervisione umana. La figura dell'autore infatti non scompare ma si trasforma, diventa guida del processo, selezione critica degli output, responsabile dell'uniformità estetica e narrativa.

4.1.2 Punti di forza e limiti del processo

L'adozione dell'intelligenza artificiale all'interno del processo creativo ha messo in luce una serie articolata di punti di forza ma anche, inevitabilmente, di limiti strutturali e funzionali.
Tra i principali vantaggi riscontrati, vi è senza dubbio la notevole accelerazione dei tempi di produzione. A differenza di una pipeline tradizionale che richiede pianificazione dettagliata, a partire dalla ricognizione delle ambientazioni fino ad arrivare alle prove tecniche, il processo guidato dall'IA permette di generare immagini, scenari e figure umane direttamente da input testuali, riducendo drasticamente la necessità di mezzi logistici e risorse fisiche. L'eliminazione di barriere operative e l'automazione delle fasi iniziali ha offerto un vantaggio immediato in termini di rapidità e accessibilità, rendendo possibile esplorare molteplici alternative in tempi estremamente ridotti.
Un ulteriore punto di forza risiede nella capacità dell'intelligenza artificiale di essere un valido aiuto anche da un punto di vista creativo. Le idee generate in fase di

brainstorming hanno infatti certamente condizionato l'autore umano nell'impostazione generale del prodotto audiovisivo. In particolare, alcuni suggerimenti forniti dall'IA hanno portato sicuramente a valutare varie possibilità creative che normalmente sarebbero state suggerite da figure umane. Si può certamente discutere sulla qualità delle idee suggerite che ricadono ovviamente sempre nei limiti e nelle potenzialità del modello utilizzato, a differenza della creatività umana che non ha da questo punto di vista alcun limite. In particolari situazioni, avere comunque la possibilità di ottenere suggerimenti e poterli sfruttare per arricchire la propria idea è sicuramente un aspetto positivo.
Un altro aspetto senza dubbio positivo portato da questa tipologia di pipeline riguarda l'immediatezza: la sperimentazione è istantanea, anche attivando un dialogo costante tra intenzione umana e interpretazione algoritmica. La possibilità di rigenerare rapidamente contenuti e di raffinare in modo iterativo le produzioni ha rappresentato un'opportunità concreta per perfezionare l'idea progettuale in tempo reale, evitando rigidità tipiche dei processi più tradizionali. Questo modello di produzione flessibile è stato definito in letteratura come *adaptive co-creation* [43], indicando un paradigma in cui la creatività non è lineare ma esplorativa, basata su cicli continui di generazione, valutazione e rifinitura.
Tuttavia, a questi punti di forza si affiancano limitazioni altrettanto evidenti. In primo luogo, la qualità e la precisione del contenuto generato risultano strettamente legati alla chiarezza e alla costruzione del prompt. L'intera pipeline si è fondata su un processo descrittivo testuale, dove ogni parola assumeva un peso decisivo. Questo ha richiesto una costante attività di revisione, riscrittura e ottimizzazione del linguaggio, dimostrando come l'IA non può sostituire l'attività autoriale, ma ne modifica radicalmente le modalità. Come evidenziato da McCormack et al. [44], l'uso dell'IA nella produzione artistica impone un ruolo sempre più ibrido all'autore, che da creatore diventa anche curatore e supervisore. Come più volte già specificato, durante la realizzazione del progetto, questo aspetto ha rappresentato la base delle interazioni con i sistemi di intelligenza artificiale generativa, rendendo centrali le parti di supervisione, correzione e affinamento dei risultati prodotti.
Un altro aspetto da analizzare è legato alla gestione del controllo creativo. L'intelligenza artificiale, pur facilitando la produzione, impone una mediazione continua tra l'intenzione autoriale umana e la capacità interpretativa del modello. Nelle varie fasi del progetto, l'output restituito non corrispondeva sempre alle aspettative, rendendo necessario iniziare nuovamente il processo o intervenire con correzioni più o meno complesse. Questo è in parte dovuto al fatto che il ragionamento umano e i relativi prompt inseriti non sono e non saranno mai allineati completamente con l'IA e la naturale conseguenza è ottenere risultati inaspettati oppure non completamente coerenti con l'idea iniziale. In ogni caso questo aspetto ha rallentato in alcuni casi la pipeline, evidenziando che la rapidità dell'IA si accompagna inevitabilmente a un margine di imprevedibilità. Inoltre, come discusso da Goktas [45], la mancanza di trasparenza nei processi che si svolgono all'interno dei modelli generativi ostacola il pieno controllo autoriale, rendendo difficile prevedere esattamente quale sarà il risultato visivo del prompt fornito.
Dal punto di vista della fase di montaggio e post-produzione invece, è emerso che, nonostante il supporto tecnologico, molte operazioni sono al momento ancora prettamente manuali e tradizionali. Il montaggio, la sincronizzazione, la regolazione cromatica e la rifinitura finale hanno richiesto il costante intervento umano. Ad

oggi infatti, non esistono soluzioni di intelligenza artificiale affidabili in grado di automatizzare, in modo soddisfacente, l'intero processo di montaggio, soprattutto in progetti che richiedono senso del ritmo narrativo e cura del dettaglio. Sebbene siano emersi negli ultimi anni alcuni software e strumenti promettenti, spesso incentrati sull'editing automatico basato su pattern musicali, riconoscimento di parole chiave o tagli intelligenti, si è preferito non adottare alcuna di queste soluzioni per il presente progetto. La qualità, la precisione e il controllo richiesti non potevano essere garantiti da queste tecnologie, che risultano ancora sperimentali o limitate per quanto riguarda le funzionalità avanzate. Per altri utilizzi rappresentano invece delle ottime alleate: la possibilità, ad esempio, di rimuovere in maniera automatica tutti i silenzi da una traccia video, può essere sfruttata per effettuare il montaggio di un'intervista in maniera più rapida ed efficiente.
Altri aspetti positivi riguardano l'accessibilità di queste tecnologie. L'utilizzo dell'IA consente potenzialmente di ridurre i costi di produzione, rendendo accessibili pratiche cinematografiche e artistiche anche a chi non dispone di grandi budget o mezzi produttivi. Questo rappresenta un'opportunità democratica significativa, limitata ovviamente dalle capacità tecniche ed espressive di questi strumenti. Tuttavia, l'accesso alle piattaforme di intelligenza artificiale generativa è spesso limitato per ragioni economiche (piani a pagamento, accessi riservati), soprattutto nel caso dei migliori modelli in circolazione.
Un altro aspetto da considerare è la reale ed effettiva semplicità di utilizzo di queste tecnologie: se da una parte basta impartire semplici comandi per ottenere dei risultati, bisogna ricordare che, per ottenere dei risultati simili, almeno in parte, a quelli professionali, servono conoscenze tecniche non trascurabili, sia a livello cinematografico che registico, rendendo questo scenario accessibile solo a una parte di autori indipendenti.
In conclusione, l'esperienza progettuale ha evidenziato come la pipeline con intelligenza artificiale rappresenti un modello produttivo alternativo, efficace e innovativo, ma non privo di limiti. Essa non sostituisce sicuramente il processo creativo umano, bensì lo ristruttura, imponendo nuove competenze linguistiche, interpretative e gestionali. I vantaggi in termini di accessibilità, velocità e flessibilità sono innegabili, ma devono essere bilanciati con una consapevolezza dei vincoli stilistici, delle limitazioni espressive e della necessità di un intervento umano costante nella supervisione e rifinitura del prodotto finale.

4.1.3 Problemi di coerenza visiva e narrativa

Uno degli aspetti più critici emersi durante la realizzazione del progetto è stato il mantenimento di una coerenza visiva e narrativa costante all'interno dell'intero cortometraggio. Questo obiettivo ha richiesto un notevole impegno durante tutte le fasi della produzione, a partire dalla definizione dei prompt testuali fino all'organizzazione delle scene. In teoria ci si aspetterebbe che i modelli generativi avanzati, come quelli utilizzati per realizzare il prodotto audiovisivo, dovrebbero essere in grado di restituire risultati stilisticamente coerenti all'interno di uno stesso contesto narrativo. In pratica però, la riproduzione di elementi visivi omogenei nel tempo presenta ancora forti criticità strutturali. Numerose clip prodotte hanno evidenziato variazioni anche significative in dettagli che per il pubblico assumono un'importanza marginale ma, analizzati in un'ottica di valutazione critica del processo, rappresen-

tano sintomi concreti di instabilità nella pipeline. Ad esempio, lo stesso personaggio, pur mantenendo tratti simili da una scena all'altra, mostrava spesso modifiche non intenzionali nell'acconciatura, nell'abbigliamento o nei dettagli del volto, come la dimensione o il colore degli occhi. Allo stesso modo, ambienti generati per rappresentare il medesimo luogo in momenti diversi della narrazione si distinguevano per cromie lievemente alterate, per differenze nella disposizione architettonica o per la luce ambientale, nonostante l'utilizzo di prompt identici o comunque strutturati nello stesso modo. Queste discrepanze, pur non compromettendo la leggibilità della narrazione agli occhi dello spettatore, risultano particolarmente evidenti se si analizza il materiale con l'approccio analitico dell'indagine condotta in ambito accademico sulla produzione audiovisiva tramite IA.
Si riportano di seguito due esempi in particolare, che ben rappresentano i problemi evidenziati. In Figura 4.1, si può notare come, in due clip differenti ma appartenenti alla stessa scena, la protagonista che prima non indossa alcun orecchino, successivamente, nella seconda inquadratura presenta il gioiello al lobo dell'orecchio. Questo potrebbe essere giustificato con il semplice fatto che le due clip appartengono a giornate differenti, ma sono molteplici gli esempi analoghi dove invece la consequenzialità temporale è immediata, rendendo dunque ingiustificabile la variazione stilistica. Pensando a un utilizzo generico dell'IA, è impensabile che, ad ogni generazione, sia necessario andare ad analizzare minuziosamente tutti questi dettagli, eseguendo nuovamente i prompt oppure andando a modificare ogni volta la clip in questione utilizzando, ad esempio, la funzione *Remix* di *Sora*. Questo approccio allungherebbe a dismisura le tempistiche di produzione dei contenuti, andando quindi a perdere di fatto uno dei principali vantaggi che questa tecnologia offre.

Figura 4.1: In alto, frame di un'inquadratura in cui la protagonista non indossa alcun orecchino. La seconda immagine mostra invece un frame dove il gioiello è presente. Le due inquadrature appartengono però alla stessa scena.

Nel secondo esempio riportato in Figura 4.2, invece, è possibile notare come la coerenza nella gestione delle luci non sia sempre perfetta: un attimo prima l'artista cammina in piazza dove è presente un'illuminazione tipica del tardo pomeriggio e qualche clip dopo, il dettaglio della scultura presenta un cielo palesemente appartenente a un'altra fascia oraria. Sicuramente il lavoro di rifinitura fatto in post-produzione ha aiutato a rendere meno accentuata questa differenza, grazie all'applicazione di una leggera correzione ai colori della seconda clip, ma resta comunque il fatto che quanto prodotto dall'IA non rispecchia a pieno quanto richiesto.

Figura 4.2: In alto, l'illuminazione della scena è correttamente quella del tardo pomeriggio. In basso, invece, il cielo mostra chiaramente una luce completamente differente. Le due inquadrature sono però temporalmente in sequenza.

In particolare, questo aspetto evidenzia un'ulteriore difficoltà, ovvero il mantenimento di una forte coerenza narrativa. Infatti, il modello utilizzato non dispone di una memoria narrativa tra una generazione e l'altra: ciascuna clip viene trattata singolarmente. La conseguenza è che gli sviluppi narrativi devono essere costruiti esternamente dall'autore umano che ha la responsabilità di creare il racconto in modo coerente, utilizzando i prompt in maniera adeguata e intervenendo manualmente dove l'IA non è in grado di garantire una progressione logica. Questo limite è già ben noto nei contesti accademici: la creatività generativa artificiale soffre di un forte deficit strutturale nel mantenimento del contesto narrativo a lungo termine, poiché i modelli non sono progettati per elaborare continuità semantiche, ma sono strutturati solamente per rispondere localmente, ottimizzando il prompt fornito. Lo studio di Zhou et al. [46] si concentra proprio sul miglioramento della coerenza visiva nei contenuti generativi a lungo termine, introducendo il modello *StoryDiffusion*. Questo approccio utilizza meccanismi di *self-attention* condivisa per stabilizzare elementi ricorrenti come volti, posture e ambienti. Sebbene i risultati sperimentali mostrino un chiaro progresso tecnico rispetto ai modelli tradizionali, gli autori non affrontano direttamente il tema della coerenza narrativa o dell'intenzionalità autoriale. Ciò suggerisce che, pur migliorando la stabilità formale delle sequenze, resta aperta la questione di come strutturare un racconto che abbia un senso narrativo piena-

mente comprensibile e coerente, funzione che al momento continua a richiedere un intervento umano diretto.
Tutti i problemi finora analizzati non sono circoscritti solamente al presente progetto ma rappresentano una sfida strutturale già analizzata da tempo. Diversi studi evidenziano come la consistenza temporale e la continuità stilistica nelle sequenze generate da modelli testuali per la sintesi video, costituiscano tuttora uno dei principali limiti tecnici. L'assenza di meccanismi interni per il tracciamento di identità visive persistenti nel tempo, come ad esempio volti, oggetti o ambienti, porta i modelli a rigenerare in parte ogni nuova inquadratura, compromettendo la coerenza tra i diversi video. Similmente, Unterthiner et al. [47] sottolineano che l'instabilità visiva rappresenta uno dei principali fattori che limitano l'adozione professionale delle soluzioni di generazione video automatizzata. Questo è sicuramente un problema ma, molto probabilmente, è solo un limite tecnico e in futuro verranno sicuramente sviluppati sistemi in grado di riconoscere i personaggi e gli ambienti, mantenendoli coerenti all'interno dell'arco narrativo di un prodotto audiovisivo. Lo studio di Zhao et al. [48] evidenzia come i modelli attuali, pur generando sequenze visivamente interessanti, falliscano nel mantenere un'alta coerenza narrativa e la stabilità dei personaggi su una scala temporale lunga. Per colmare questa lacuna, gli autori propongono *MovieDreamer*, un'architettura gerarchica che combina pianificazione semantica e generazione visiva. Tuttavia, anche con tale innovazione, l'IA continua a operare senza una reale intenzione narrativa, confermando la necessità di un intervento umano per garantire direzione e senso al racconto audiovisivo. Al momento esistono già degli strumenti che permettono di replicare i soggetti in luoghi o con vestiti differenti, prendendo come riferimento un'immagine fornita dall'utente e rappresentante la persona da mantenere coerente. È solamente una questione di tempo necessario per l'evoluzione della tecnologia che poi potrà essere applicata anche all'ambito video, determinando così un notevole incremento delle capacità dei sistemi generativi.

4.1.4 Affrontare i problemi di coerenza

Nel contesto del progetto audiovisivo sviluppato, i problemi legati all'incoerenza visiva e narrativa generata dall'intelligenza artificiale sono stati affrontati attraverso una serie di strategie tecniche e operative, elaborate in corso d'opera e orientate alla costruzione di un racconto il più possibile continuo e credibile.
Il primo intervento ha riguardato l'organizzazione meticolosa del materiale prodotto: ogni scena generata tramite la piattaforma *Sora* è stata salvata e classificata in cartelle separate, una per ciascuna scena del cortometraggio. Questa suddivisione ha rappresentato un punto fondamentale per una supervisione chiara della coerenza interna, permettendo una rapida comparazione tra le clip candidate e una valutazione accurata di eventuali discrepanze visive o stilistiche. Aprendo le singole cartelle, infatti, le clip di una stessa scena si trovano una di fianco all'altra, permettendo un confronto mirato e notevolmente più efficace rispetto alla ricerca delle stesse clip all'interno della cronologia completa di generazione dei video. Passando il cursore sopra i diversi video, questi vengono riprodotti, permettendo quindi una rapida visione in serie di tutti gli elementi della stessa scena.

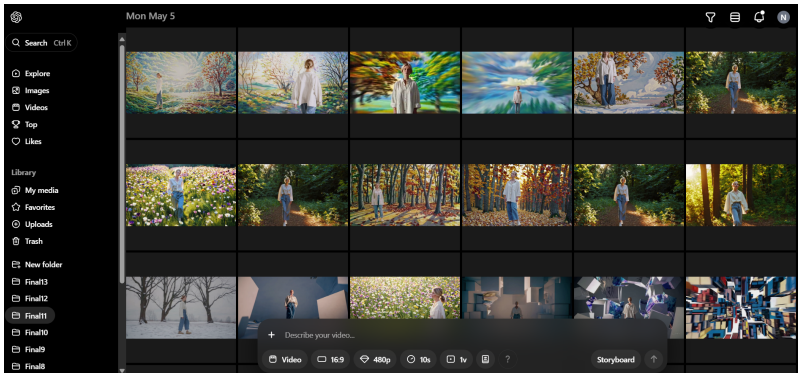


Figura 4.3: Immagine raffigurante una schermata della piattaforma *Sora*. In particolare, è qui mostrata la cartella contenente le clip scelte per la costruzione della scena 11.

Una tecnica non particolarmente ricercata ma tra le più ricorrenti ha riguardato la rigenerazione selettiva dei contenuti. In presenza di difformità significative, come ad esempio variazioni non volute nella fisionomia del personaggio principale, mutamenti significativi negli sfondi oppure alterazioni ambientali ingiustificate, si è proceduto a modificare leggermente il prompt fornito, tentando di forzare il modello a mantenere costanti determinati elementi. È stato riscontrato che inserire riferimenti puntuali e ripetuti, come *lo stesso abito della scena precedente* o *identico paesaggio urbano con la stessa angolazione*, non aiutava a rafforzare la memoria visiva all'interno delle generazioni; descrivere in maniera meticolosa e dettagliata l'elemento da modificare invece, produceva risultati spesso soddisfacenti. Se in seguito a più iterazioni le modifiche si dimostravano inefficaci, si procedeva con una nuova generazione della clip, accettando il rischio di ottenere una scena diversa ma coerente in termini narrativi e stilistici.
La funzione *Remix* fornita da *Sora* è stata utilizzata per modificare alcune clip visivamente promettenti ma con elementi incongruenti. Selezionando la clip in questione, si indicavano testualmente le modifiche desiderate, procedendo con una descrizione il più possibile particolareggiata. Tuttavia, anche in questo caso, la mancanza di un completo controllo del modello e la natura probabilistica dell'IA hanno richiesto un'attenta e costante selezione dei risultati. Le clip ritenute insoddisfacenti venivano eliminate a favore di versioni più coerenti sul piano visivo.

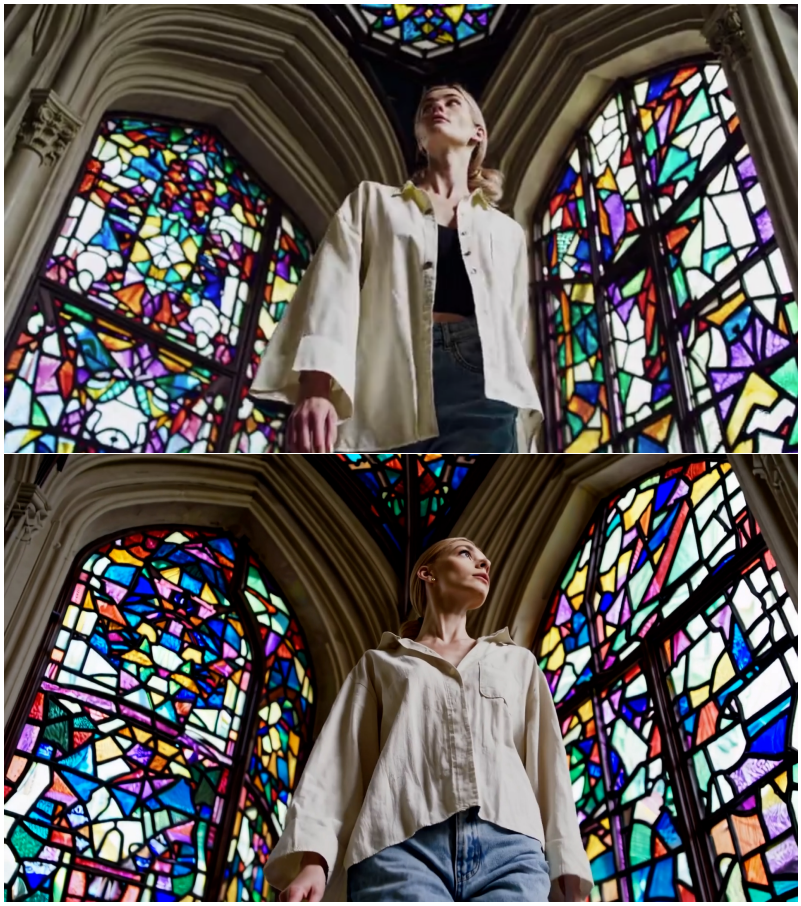


Figura 4.4: Esempio di modifica dovuta a un particolare non coerente: nella clip in alto, generata inizialmente, la camicia dell'artista è aperta, in contrasto con quanto rappresentato in altre clip. Si è quindi richiesto, tramite la funzione *Remix* e una descrizione dettagliata dell'elemento da modificare, di generare un nuovo video, ovvero quello presentato in basso, in linea con quanto mostrato nelle altre inquadrature.

In aggiunta, la tecnica di transizione tramite la funzione *Blend* si è rivelata una risorsa fondamentale per collegare due scene visivamente contrastanti, rendendo il passaggio narrativo più fluido e meno percepibile. Utilizzando diverse curvature di interpolazione, si è cercato di ottenere passaggi morfologici graduali delle immagini. Questo metodo si è dimostrato utile soprattutto in quelle sequenze che, pur avendo una continuità tematica, presentavano bruschi salti cromatici o formali. La tecnica è stata usata inoltre per generare clip di transizione particolarmente intriganti, che hanno saputo ben delineare e sfruttare elementi della scena in uscita per collegarsi a quella in entrata.
Anche la fase di montaggio ha giocato un ruolo cruciale nella gestione delle discontinuità: la selezione delle clip migliori e l'ordine di utilizzo scelto sono stati effettuati considerando il ritmo, la posizione della camera e l'illuminazione delle varie inquadrature. In alcuni casi, dove non era possibile generare un raccordo visivo accettabile, si è fatto ricorso a stacchi netti o pause a nero che, in modo narrativamente giustificato, hanno introdotto momenti di riflessione. Allo stesso tempo, sono stati utilizzati accorgimenti nella fase di correzione colore per uniformare, seppur marginalmente, le tonalità cromatiche, rendendo meno evidenti le discrepanze

tra scene contigue, senza mai però forzare la resa estetica originaria delle immagini generate.
Nel complesso, il mantenimento della coerenza all'interno di una pipeline basata sull'intelligenza artificiale ha richiesto un equilibrio costante tra rigore tecnico e flessibilità interpretativa. Nessuna delle tecnologie disponibili ha fornito una soluzione completamente automatizzata al problema della continuità visiva e narrativa. Tuttavia, attraverso una combinazione di strategie di rigenerazione, revisione e correzione cromatica, è stato possibile ottenere un prodotto in generale coerente, espressivo e narrativamente solido.

4.1.5 Problemi di generazione con IA

Una criticità rilevante riguarda l'uniformità estetica e le limitazioni creative intrinseche al modello. Le immagini generate, pur visivamente efficaci, tendevano a riprodurre soluzioni stilistiche ricorrenti, legate ai data set di addestramento. Questa tendenza alla standardizzazione, notata nella fase di generazione video con *Sora* e negli output vocali di *Eleven Labs*, ha richiesto un attento lavoro di selezione per evitare scelte visive troppo prevedibili. La letteratura scientifica, come nel caso del lavoro di Seshadri et al. [49], sottolinea come i modelli generativi spesso riflettano bias stilistici presenti nei dati di addestramento, limitando la varietà formale e simbolica dei contenuti. Un chiaro esempio di questo problema si è presentato nelle fasi iniziali del progetto, quando ancora non era stato definito con accuratezza il prompt per generare in maniera sempre uguale e coerente l'artista protagonista del cortometraggio. In quel caso, durante la progettazione delle immagini dello storyboard, più volte è capitato di nominare l'artista senza specificarne il sesso o l'età e questo spesso portava l'IA a generare figure umane maschili e anziane: questo aspetto mette in luce chiaramente il forte problema dei bias presenti in tutti i modelli di IA.
Un altro limite è dovuto alla scarsa empatia ed espressività dei modelli di trasformazione vocale utilizzati per il voiceover. Tale ostacolo è riconducibile principalmente al grado di sviluppo ancora limitato dei modelli vocali, in particolare per la lingua italiana che non dispone delle stesse capacità espressive e sfumature garantite invece dai modelli anglofoni, più ricchi in termini di addestramento e varietà di scelta. Anche questi ultimi presentano però delle lacune in termini emotivi, condizione che ha portato l'autore umano a scegliere di utilizzare la lingua italiana. Questo è però un forte limite che influenza ampiamente l'impatto emotivo del prodotto audiovisivo finale, rendendolo meno espressivo, soprattutto nei momenti più introspettivi e personali.
Un'ulteriore criticità emersa durante la fase di generazione riguarda l'incapacità, da parte del modello di intelligenza artificiale generativa, di gestire in modo coerente e tecnicamente corretto alcune situazioni complesse richieste dalla narrazione. In particolare, sono stati riscontrati limiti evidenti nella simulazione di movimenti articolati o nel rispetto delle regole fisiche all'interno di determinate sequenze. Sebbene in molte parti del cortometraggio, grazie all'ambientazione onirica e a tratti immaginaria del viaggio artistico, tali incoerenze fossero giustificabili o addirittura funzionali alla buona riuscita dell'opera, in altri casi queste mancanze si sono rivelate dei veri e propri impedimenti narrativi. In tal senso, un esempio significativo si è verificato nella fase conclusiva del viaggio, laddove la sceneggiatura prevedeva il ritorno dell'artista nel proprio studio. La sequenza avrebbe dovuto mostrare l'aper-

tura della porta, l'ingresso nel laboratorio e l'avvio della fase conclusiva del racconto. Nonostante i numerosi tentativi e le continue variazioni testuali nei prompt, il modello non è stato in grado di restituire nemmeno una clip coerente e tecnicamente soddisfacente. Gli output risultavano affetti da distorsioni nell'animazione, difetti nei dettagli architettonici, incongruenze nel movimento dell'artista o della porta e compenetrazioni non naturali tra il corpo e gli oggetti con cui interagiva. Di fronte a questa impossibilità tecnica, si è optato per una soluzione narrativa alternativa. È stata inserita una dissolvenza a nero più estesa rispetto a quelle normalmente utilizzate all'interno del progetto, segnando simbolicamente il passaggio dal mondo immaginario delle opere a quello reale. A tale scelta visiva si è associata una progettazione acustica più marcata. L'utilizzo di effetti sonori specifici, come il rumore della chiave nella serratura, il cigolio della porta e il suono dei passi all'interno dello studio, ha permesso allo spettatore di immaginare l'azione non visibile. Questo espediente, che rafforza il distacco tra la fine del viaggio e il ritorno al mondo reale, è stato sicuramente funzionale al ritmo e alla struttura della narrazione nonostante rappresenti una chiara limitazione nella generazione di video con intelligenza artificiale, non permettendo di sviluppare quanto studiato sul piano visivo e richiedendo un intervento creativo compensativo.

Infine, un'ulteriore difficoltà generativa emersa ha riguardato l'impossibilità di ottenere, attraverso i modelli di intelligenza artificiale generativa, uno stile visivo coerente con le specifiche correnti pittoriche richieste. L'idea progettuale prevedeva che le scene centrali del viaggio onirico, in particolare quelle comprese tra la sesta e l'undicesima, rappresentassero ciascuna un linguaggio artistico differente, in linea con le sperimentazioni pittoriche dell'artista protagonista. Questo approccio avrebbe permesso non solo una forte varietà estetica ma anche una continuità simbolica con il percorso interiore della protagonista. Nella pratica, i modelli utilizzati non sono riusciti a tradurre efficacemente l'estetica dei movimenti artistici specifici in un linguaggio cinematografico credibile e coerente. Le immagini prodotte, sebbene visivamente valide, non mostravano con chiarezza i tratti distintivi delle correnti pittoriche indicate nei prompt. A volte vi erano elementi pittorici accennati ma non coerenti con l'ambientazione o il personaggio; altre volte la scena risultava semplicemente generica, senza tenere conto del linguaggio pittorico richiesto. Di fronte a queste difficoltà, si è scelto in più occasioni di abbandonare l'intento iniziale di rappresentare lo stile pittorico, mantenendo invece uno stile realistico, più facilmente ottenibile ma soprattutto visivamente stabile. Questa alternanza ha portato a una struttura narrativa visivamente disomogenea, caratterizzata da un continuo passaggio tra scene stilizzate e altre più realistiche. Sebbene da un punto di vista narrativo ciò abbia generato un'interessante dinamica di andata e ritorno tra il mondo interiore dell'artista e la realtà, questa soluzione si è imposta più come necessità tecnica che come scelta compositiva. Per rendere coerente tale andamento, si è intervenuti a livello descrittivo, sfruttando il voiceover dell'artista come strumento compensativo: attraverso le parole, la protagonista riflette spesso sulla possibilità che il viaggio stia per concludersi (quando di fatto si ritrova in ambientazioni realistiche), salvo poi ritrovarsi in una nuova opera. Questa scelta ha quindi permesso di giustificare quella che, da un punto di vista tecnico, è stata una limitazione evidente dell'IA.

L'insieme delle problematiche riscontrate durante la produzione ha messo in evidenza uno dei limiti strutturali più rilevanti delle tecnologie di intelligenza artificiale applicate alla generazione video: la difficoltà nel guidare l'output verso uno stile visi-

vo specifico e coerente, soprattutto quando si tratta di riferimenti estetici complessi. Tale complicazione tecnica ha reso necessaria l'introduzione di soluzioni narrative alternative quali, ad esempio, la specifica di alcuni dettagli all'interno del voiceover della protagonista.
Analogamente, in altre parti del progetto, l'intelligenza artificiale ha mostrato importanti carenze nella gestione di sequenze complesse, come ad esempio nella fase di transizione tra il viaggio onirico e il rientro nello studio. In questi casi, la risposta del modello è stata spesso instabile, costringendo a intervenire con modifiche strutturali nel montaggio o con soluzioni sonore e narrative che sopperissero alla mancanza visiva. Sebbene tali strategie abbiano consentito di mantenere la coerenza e l'integrità del racconto, trasformando i limiti in risorse stilistiche, è evidente che in altri contesti creativi, meno aperti all'astrazione o alla sospensione simbolica, questi stessi vincoli potrebbero rappresentare un ostacolo concreto alla libertà progettuale, compromettendo la realizzabilità dell'idea originaria.

4.1.6 Relazione tra prompt e output

Le odierne IA generative (come i modelli *text-to-image* o *text-to-video*) hanno compiuto progressi enormi, generando contenuti visivi complessi a partire da semplici istruzioni testuali. Ciò ha portato a una profonda riflessione sul rapporto tra i prompt forniti in input e l'output prodotto. In generale, i sistemi di IA seguono alla lettera le istruzioni fornite, senza comprenderne il contesto o l'intento implicito: operano in modo statistico, associando parole e immagini a partire dai dati di addestramento. Questo può dar luogo a risultati visivi anche molto coerenti, ma privi della sfumatura voluta dall'autore umano.
Ad esempio, se il nostro intento narrativo è suscitare meraviglia o empatia, non basta dirlo nel prompt; l'IA enfatizzerà aspetti letterali e ovvi della richiesta, ignorando sfumature emotive non esplicitate. Ne risulta un output che contiene tutti gli elementi richiesti in maniera didascalica, senza la gerarchia di importanza o il sottotesto che un artista umano saprebbe inserire. L'autore umano è quindi costretto a iterare e ragionare con la macchina, modificando ripetutamente il prompt per far avvicinare il prodotto all'idea iniziale.
In questo processo di *prompt engineering*, ogni termine usato porta con sé un intero campo semantico, ma anche visivo, appreso dal modello utilizzato. Questa forte dipendenza dal linguaggio fa sì che lo stile e il contenuto nell'IA non possano essere spesso scindibili: basta una parola di troppo o un termine leggermente fuori luogo perché l'output restituisca inaspettatamente estetiche o soggetti indesiderati. Come riportato in un lavoro di Oppenlaender [50], esistono delle espressioni lessicali che, pur non facendo parte della descrizione oggettiva del soggetto, influenzano in modo diretto i parametri visivi come lo stile, l'illuminazione, la composizione e la palette cromatica. L'interazione fra elementi di linguaggio e componenti visuali diventa quindi di grande rilievo: il contenuto e la forma dell'immagine risultano difficilmente scindibili poiché modellati simultaneamente da ciascuna scelta linguistica. In particolare, lo studio evidenzia come l'aggiunta, la rimozione o la sostituzione di un singolo modificatore possa produrre alterazioni sostanziali nel risultato finale, rendendo il processo di prompt engineering una pratica altamente sensibile. Questo elemento porta le interfacce generative ad operare in maniera complessa e non sempre

prevedibile a variazioni minime del linguaggio, dimostrando quanto la componente verbale sia anche una forma di controllo estetico e tecnico.
L'analisi condotta da Dumpala et al. [51] evidenzia una marcata sensibilità dei modelli generativi visivi alle variazioni lessicali nei prompt, anche quando il significato semantico rimane invariato. I risultati mostrano che modifiche minime nella formulazione dei prompt o nell'ordine delle indicazioni fornite possono produrre fluttuazioni significative nelle performance del modello. Gli autori dimostrano inoltre che né l'uniformità all'interno dello stesso modello né quella tra modelli differenti è garantita in presenza di parafrasi o riordinamenti.
L'autore umano deve dunque sviluppare una nuova alfabetizzazione non solo linguistica ma anche visiva, imparando a prevedere come la macchina tradurrà ogni suggerimento verbale in segnali visivi. Questa relazione complessa tra input umano e output algoritmico costituisce la base dell'estetica generata dall'IA e richiede un approccio consapevole e sperimentale da parte del creativo.

4.2 Valutazione dell'estetica generativa

4.2.1 Estetica e forma visuale

Le dinamiche dei video generati con IA presentano ulteriori sfide estetiche e tecniche rispetto alla generazione di semplici immagini o testi. I cortometraggi e le animazioni prodotte dall'IA mostrano spesso un'estetica a prima vista promettente ma spesso problematica: sequenze fluide, talora oniriche, che però possono risultare incoerenti o vacue sul piano emotivo e narrativo. Una caratteristica ricorrente è una sorta di surrealismo: le scene appaiono visivamente impeccabili, ricche di dettagli e transizioni fluide, tuttavia generano un senso di vuoto nello spettatore. L'intelligenza artificiale, pur mostrando risultati tecnicamente elevati, fatica a gestire con coerenza e sensibilità elementi complessi come l'anatomia del corpo umano in movimento o scenografie visivamente dense. Queste limitazioni, oltre a generare artefatti visivi difformi dalla realtà, contribuiscono a una resa estetica spesso fredda e priva di coinvolgimento emotivo. Come evidenziato da Kimanos e Yderberg [52], le animazioni in cui compaiono distorsioni anatomiche o ambienti visivamente affollati richiedono un attento intervento umano o, in certi casi, l'abbandono delle soluzioni di IA in favore di metodi più controllabili manualmente.
In particolare, per quanto rilevato durante la produzione del prodotto audiovisivo, diverse clip generate catturano l'attenzione con immagini molto suggestive, spesso al rallentatore, con colori vividi e immerse in scenari fantastici, ma sembrano prive di intenzione. Sono certamente interessanti sotto un profilo puramente estetico ma insoddisfacenti sul piano emotivo. Lo spettatore ha spesso l'impressione che il video voglia comunicare *qualcosa*, ma quel *qualcosa* non arriva mai, lasciando solo l'esperienza momentanea del flusso visivo. Questa assenza di un messaggio chiaro è legata principalmente al fatto che l'IA non ha una volontà narrativa; di fatto non capisce perché un'inquadratura dovrebbe seguire un'altra se non per associazioni di bassa correlazione e conseguentemente agisce in maniera vaga, cercando di massimizzare il risultato per la singola inquadratura, senza avere una direzione generale prestabilita da seguire.

Dal punto di vista tecnico, i limiti attuali influiscono anch'essi sull'estetica dei video generati. Le piattaforme generative faticano a mantenere la consistenza di personaggi e oggetti per più di pochi secondi. Spesso l'output è limitato a clip di massimo cinque o dieci secondi prima che la qualità diminuisca o che i soggetti si trasformino. Video di durata maggiore, portano a cambi repentini di inquadrature, più inconsistenza e, in generale, a risultati peggiori. Per creare un video complessivo più lungo, bisogna generare molte clip brevi, ma ciò introduce il rischio di ottenere sbalzi visivi notevoli, come ad esempio cambiamenti nell'aspetto di volti, corpi, scenari e oggetti. Tutto questo incide negativamente sulla continuità narrativa, creando confusione e disorientamento nello spettatore. Anche la fisica e i movimenti rappresentano un punto debole: le animazioni di una persona che interagisce con l'ambiente, spesso risultano innaturali o incoerenti, con oggetti che fluttuano, ombre e luci che cambiano senza una logica precisa (Bai et al. [53]), mostrando chiaramente l'assenza di una completa comprensione fisica del mondo da parte dell'IA (Yuan et al. [54]). Inoltre, dettagli come testi scritti tendono a deformarsi nel tempo. Ad esempio, un'insegna contenente lettere può inizialmente essere comprensibile per poi dissolversi in glifi senza senso dopo pochi fotogrammi. Questi difetti tecnici sono in realtà rivelatori della natura dell'IA e lo spettatore esperto riconosce immediatamente l'utilizzo di questi strumenti proprio da certe oscillazioni, azioni e imperfezioni dell'immagine in movimento. Tali peculiarità possono essere interpretate creativamente; c'è chi, ad esempio, vede in questi glitch un richiamo alle avanguardie del cinema surrealista o all'estetica della *glitch art* contemporanea, come confermato dallo studio di Knight et al. [55], dove si mostra come i glitch nei video generati con IA, spesso considerati errori, possano essere reinterpretati creativamente in ambito artistico. Queste anomalie visive diventano ottimi strumenti espressivi ma la mancanza di controllo sui modelli rende difficile riprodurli intenzionalmente.
In ogni caso, per un uso narrativo convincente nel cinema, queste limitazioni pongono ancora forti vincoli espressivi. Il rischio è che i video generati si attestino su un registro paragonabile a un *sogno lucido* perenne, certamente affascinanti come esperienze visive astratte, ma inadeguati laddove serva una struttura narrativa solida o un linguaggio filmico coerente. Nel cortometraggio generato con IA sono state affrontate tutte queste sfide; più volte si è dovuto scartare generazioni dove la trasformazione delle figure umane risultava troppo discontinua, nonostante l'ambientazione fosse di grande impatto. Questa attenta selezione ad opera dell'autore umano è stata fondamentale per stabilizzare l'esperienza estetica e darle un senso filmico riconoscibile nonostante l'origine algoritmica delle immagini.

4.2.2 Estetica dell'uncanny valley

L'adozione sempre più diffusa dell'intelligenza artificiale generativa nella produzione cinematografica ha reso inevitabile un confronto diretto con il concetto di *uncanny valley*, ovvero quella soglia psicologica oltre la quale un'entità artificiale, pur avvicinandosi all'aspetto umano, provoca disagio anziché empatia. Questo fenomeno, inizialmente ipotizzato da Masahiro Mori nel 1970 [56], assume oggi una rilevanza cruciale all'interno dei flussi produttivi automatizzati, specialmente quando si affronta la generazione di volti, espressioni e movimenti corporei nei personaggi sintetici.

Nel contesto cinematografico e in particolare nella realizzazione di prodotti audiovisivi interamente prodotti con pipeline basata sull'IA, si osserva un confronto costante tra la verosimiglianza tecnica e la credibilità percettiva. Anche il cortometraggio prodotto, nel quale ambientazioni, attori e voiceover sono generati a partire da input testuali ha introdotto nuove sfide estetiche e cognitive. Tra queste, la più rilevante è proprio la gestione dell'asimmetria percettiva generata da output visivi e uditivi simili a quelli umani. La letteratura scientifica più recente documenta chiaramente questa criticità. Nel lavoro di Zhang et al. [16], attraverso un'analisi dei film prodotti nel contesto del *MIT AI Film Hack*, si evidenzia che il 100% delle opere ha integrato l'IA per la generazione dei video, ma la quasi totalità degli artisti ha evitato lo stile fotorealistico, preferendo approcci astratti o nel campo dell'animazione per aggirare le problematiche emotive generate dall'imperfezione di questi modelli. Le cause sono da ricondurre principalmente alla persistente inadeguatezza dell'IA generativa nel mantenere alta la coerenza dinamica e il realismo espressivo. I dati quantitativi raccolti indicano che la generazione di movimenti corporei coerenti viene considerata il fattore più importante per la qualità filmica (*6,62* su *7*), ma gli strumenti attuali ottengono un punteggio solamente pari a *4,55*, evidenziando un divario ancora sostanziale. Analogamente, la coerenza nei volti generati e il controllo della camera ottengono punteggi di desiderabilità superiori alla valutazione tecnica effettiva.

Entrando più nello specifico, il lavoro di Liao et al. [57] ha individuato le caratteristiche più deboli nella generazione di personaggi con sembianze umane: sincronizzazione labiale, microespressioni facciali e continuità tra angolazioni diverse della stessa scena. Tali imperfezioni rappresentano precisamente ciò che, secondo le neuroscienze cognitive [58], attiva l'effetto perturbante: la mente umana, dotata di un'alta sensibilità all'espressività sociale, tende a rilevare come disturbanti anche minime discrepanze nei volti artificiali, soprattutto quando si collocano in un punto intermedio di realismo. In ambito filmico, questo aspetto si traduce in una compromissione dell'identificazione empatica con il personaggio. Alcuni progetti hanno deliberatamente trasformato questa limitazione in linguaggio, adottando un'interpolazione a bassa frequenza per ridurre il contrasto tra aspettativa visiva e risposta cognitiva.

Tuttavia, nei casi in cui si persegua un'estetica più realistica, come nel cortometraggio prodotto, l'effetto *uncanny* rischia di diventare più evidente, specie nei momenti in cui il volto digitale assume ruoli drammatici, introspettivi o dialogici. Il voiceover sintetico, sebbene vantaggioso in termini di accessibilità e tempistiche produttive, può aggravare il problema se mal sincronizzato con le movenze labiali. In questo caso specifico però, si è scelto deliberatamente di non far mai parlare l'artista protagonista, riproducendo solamente quelli che sono i suoi pensieri in ogni fase del viaggio.

L'adozione di una pipeline basata su IA comporta dunque una responsabilità, imponendo una riflessione accurata sulle soglie percettive dell'umano, laddove ogni deviazione dalla plausibilità espressiva può trasformarsi in un disturbo emozionale, portando l'osservatore a un immediato distacco.

In conclusione, l'*uncanny valley* non rappresenta solo un ostacolo tecnico, ma anche una questione semiotica ed estetica, che chiama in causa il rapporto tra sintesi algoritmica e autenticità percepita.

4.2.3 Lo stile estetico dell'IA

Un aspetto centrale emerso sia prima che durante la produzione delle clip ha portato all'identificazione dello stile estetico dell'IA generativa. Ogni modello ha un'impronta stilistica differente, derivante principalmente dai dati di addestramento.
Si comincia a parlare di estetica proprio perché gli utenti hanno notato nel tempo alcuni pattern ripetitivi: combinazioni di colori, effetti e pose stereotipate, frutto dei data set su cui queste IA sono state allenate e che rendono alcuni prodotti immediatamente identificabili come artificiali.
La sfida per l'autore è dunque duplice: da un lato saper dialogare con questo stile di default, comprendendone il linguaggio, dall'altro personalizzare e arricchire l'estetica generata, introducendo variazioni e sperimentazioni di vario tipo. In pratica ciò si traduce in un ragionamento costruttivo con la macchina: si provano combinazioni di prompt insolite o si interviene con aggiustamenti minori per spingere l'output fuori dai sentieri più classici previsti dallo stile dell'IA.
Questa operazione di lettura e riscrittura del risultato è di grande rilevanza in quanto prevede l'osservazione attenta del prodotto, decodificando i riferimenti estetici e formulando in seguito nuovi input per generare variazioni in un ciclo iterativo. Il linguaggio verbale gioca un ruolo chiave perché la descrizione dettagliata del risultato voluto è l'unico mezzo per comunicare con l'algoritmo: padroneggiare sfumature linguistiche (sinonimi, registri, riferimenti culturali) permette di influire sullo stile visivo con sorprendente precisione.
Non a caso alcuni studiosi, come Serbanescu [59], parlano di *IA partecipativa*, immaginando sistemi progettati per cooperare con gli artisti e potenziarne il flusso di lavoro creativo. L'autrice propone l'utilizzo di un grafo della conoscenza basato su sette categorie narrative per supportare la creazione di personaggi complessi e coerenti. L'IA agisce come partner creativo, ampliando lo spazio concettuale e offrendo suggerimenti che stimolano nuove idee.
In prospettiva, man mano che tali strumenti diventeranno più raffinati, potremo vedere emergere nuovi linguaggi audiovisivi, nati in virtù di questa sinergia: forme narrative ibride in cui l'imprevedibilità algoritmica viene incanalata dall'autore umano in strutture significative del tutto originali. Già ora, l'esperienza maturata con la produzione del cortometraggio *La Tela Oscura*, suggerisce che l'estetica dell'IA non deve essere subita passivamente ma può essere plasmata tramite un uso creativo del linguaggio e una costante reinterpretazione critica e artistica dei risultati, facendo sempre attenzione di evitare il fenomeno dell'*uncanny valley*.
Di fronte a queste potenzialità, è essenziale adottare un'analisi critica dell'estetica generata dall'IA, sia nel discorso teorico sia nella pratica artistica. Un autore consapevole non può limitarsi a usare l'IA come una *black box* neutrale ma deve piuttosto esplicitare la comprensione di come lo strumento opera e valutarne criticamente i risultati prodotti. Affinché questa dimensione emerga, è necessaria un'impostazione critica da parte del creatore. Per avere un impatto duraturo, gli artisti che lavorano con l'IA devono mantenere sempre uno sguardo esaminatore sui dispositivi che utilizzano.
Non a caso, nel lavoro portato avanti, si è analizzato il cortometraggio non solo come prodotto visivo ma anche come processo, mettendo in discussione, in primo luogo, ogni output generato, rapportando poi il lavoro svolto con la teoria dei media digitali.

Questa riflessione ha permesso di riconoscere nell'output dell'IA non un semplice esercizio tecnico, ma un testo audiovisivo da leggere criticamente, individuando riferimenti, stereotipi iconografici replicati dalla macchina e lacune espressive, colmate in parte successivamente con il montaggio. Proprio l'atto di analizzare il risultato estetico indica che il ricercatore ha piena padronanza del medium veicolato dall'IA e non subisce passivamente il fascino della tecnologia.
Questa riflessione arricchisce l'opera stessa, inserendola nel più ampio discorso dei *media studies* e dell'estetica filmica contemporanea. Solamente attraverso una simile analisi critica si può infatti affermare di avere davvero il controllo creativo sull'IA, trasformando uno strumento avanzato ma statistico in un mezzo espressivo consapevole e al servizio di una visione artistica.

4.2.4 Estetica del cortometraggio generato

L'estetica complessiva del video generato riflette in modo diretto i limiti, le possibilità e le peculiarità proprie di una pipeline fondata sull'intelligenza artificiale generativa. Il risultato finale si caratterizza per un equilibrio formale che alterna immagini altamente suggestive, dense di atmosfera e cariche di contenuto simbolico, a segmenti visivi più imperfetti o disomogenei, nei quali si avverte con maggiore chiarezza la natura automatizzata del processo produttivo.
La mancanza della possibilità di un controllo puntuale su alcuni aspetti fondamentali dell'immagine, quali ad esempio l'uniformità cromatica, la coerenza spaziale o la gestione completa delle luci, ha determinato sicuramente la presenza di discontinuità visive importanti, assieme ad alcune variazioni involontarie nello stile e nella resa prospettica. Queste imperfezioni, anziché essere eliminate o nascoste del tutto, sono state integrate nel cortometraggio come se il contrasto tra realismo e astrazione, tra coerenza formale e rottura percettiva, fosse parte integrante dell'estetica dell'opera, trasformando un vincolo tecnico in una risorsa espressiva.
Va inoltre osservato che l'estetica si riconosce per una specifica impronta ibrida, a metà tra fotorealismo simulato e arte digitale sintetica. Questo stile peculiare, ormai riconoscibile e quasi autonomo, è frutto dell'addestramento del modello su data set eterogenei ma non curati per specifici riferimenti artistici o cinematografici. Di conseguenza, il cortometraggio finisce per assumere un'identità visiva propria che non coincide né con un'estetica umana tradizionale né con una completamente astratta, una sorta di *estetica transitoria* dove l'autorialità è condivisa e spesso contesa tra autore umano e algoritmo.
In sintesi, l'estetica del video non può essere valutata secondo i consueti canoni cinematografici di controllo totale sulla composizione dell'inquadratura, ma piuttosto secondo parametri nuovi, che tengano conto del linguaggio emergente delle intelligenze artificiali. Il lavoro svolto ha mostrato come, anche in presenza di limitazioni tecnologiche, sia possibile costruire un'identità visiva coerente e poetica, sfruttando in modo consapevole il potenziale e i vincoli dell'IA come elementi costruttivi, anziché vederli come ostacoli da nascondere.
Si analizzano ora i principali aspetti che hanno contribuito a definire l'estetica complessiva del prodotto audiovisivo.
Uno degli elementi più complessi, emersi durante la realizzazione del cortometraggio, ha riguardato la coerenza visiva complessiva, con particolare attenzione alla gestione dell'illuminazione tra le diverse clip. Nonostante l'impegno nella redazione dei

prompt per mantenere un'illuminazione uniforme in ogni scena, si sono riscontrate variazioni percettibili nella qualità e nell'intensità della luce, in particolare nell'incidenza sul volto della protagonista. In alcune sequenze, ad esempio, i raggi di luce colpiscono l'artista con angolazioni e temperature differenti anche a distanza di pochi secondi narrativi. Questo tipo di discontinuità, sebbene spesso impercettibile al pubblico in generale, può generare nello spettatore più attento una sensazione di instabilità visiva, compromettendo parzialmente la continuità percettiva del racconto. Si tratta di un limite strutturale dei modelli generativi attuali che non garantiscono il controllo preciso della sorgente luminosa o delle ombre prodotte dagli oggetti in scena.

Figura 4.5: In questi due frame, tratti da due inquadrature consecutive, si può notare chiaramente l'incoerenza della luce: la direzione è diametralmente opposta.

La gestione dello spazio scenico, in particolare la rappresentazione dello studio dell'artista, ha presentato alcune criticità sul piano dell'uniformità. Pur avendo predisposto nei prompt descrizioni dettagliate e reiterate dello stesso ambiente, le inquadrature generate mostrano talvolta ambienti differenti, con oggetti e arredamenti che variano sensibilmente da una clip all'altra. Se da un lato ciò potrebbe essere narrativamente giustificato con l'esplorazione di angoli diversi di uno spazio ampio, dall'altro può generare confusione sul piano percettivo. La difficoltà sta nel fatto che, nel contesto del cortometraggio, lo studio è spesso descritto come un ambiente unico e sempre uguale, una costanza che risulta parzialmente compromessa da variazioni inattese nella resa visiva delle sue componenti. Tuttavia il montaggio lineare

delle clip e la struttura narrativa centrata sull'introspezione dell'artista, consentono di accettare queste variazioni come eccezioni gradite. In tal senso, il linguaggio audiovisivo tende a minimizzare l'incoerenza spaziale attraverso l'uso del voiceover che nomina più volte lo studio, riducendo così l'impatto negativo sull'esperienza complessiva dello spettatore.

Figura 4.6: In questi frame, tratti da due inquadrature presenti in scene differenti del cortometraggio, si possono notare chiaramente le enormi differenze nella rappresentazione dello studio. È evidente la difficoltà di rappresentare coerentemente lo stesso spazio.

Un aspetto particolarmente rilevante dell'estetica generale del cortometraggio riguarda la resa dell'ultra-realismo, obiettivo perseguito in modo esplicito attraverso la dichiarazione, all'inizio di ciascun prompt, di voler ottenere un video con questo stile. Il risultato ottenuto si è rivelato complessivamente positivo, ma con differenze sostanziali a seconda della tipologia di scena. In particolare, le clip che raffigurano paesaggi naturali hanno dimostrato un livello di qualità visiva elevato. Queste sequenze, spesso caratterizzate da inquadrature ampie, hanno una luce ambientale coerente, la palette cromatica risulta equilibrata e i dettagli ben definiti. Si può affermare che, in questi casi, l'intelligenza artificiale sia riuscita a produrre una realtà paesaggistica che si avvicina molto alle aspettative visive tipiche di un prodotto cinematografico ad alta definizione.

Figura 4.7: Paesaggio naturale generato dall'intelligenza artificiale con ottimi risultati.

L'obiettivo dell'ultra-realismo può dirsi quindi parzialmente raggiunto. Se da un lato l'intelligenza artificiale è in grado di produrre ambienti naturali e scene panoramiche di grande impatto visivo, dall'altro incontra ancora difficoltà sostanziali nella riproduzione ad esempio di ambienti urbani complessi. Tali limiti sono stati affrontati nel progetto attraverso una selezione accurata delle clip e con l'adozione di scelte narrative che, in alcuni casi, hanno giustificato o coperto le mancanze visive. Nonostante ciò, resta evidente che l'aderenza ai codici del realismo cinematografico costituisce ancora uno spazio critico da colmare nei progetti audiovisivi realizzati tramite modelli generativi.
Uno degli aspetti più delicati nella valutazione estetica del cortometraggio riguarda la gestione degli oggetti presenti in scena e la loro interazione con la protagonista. Le clip che prevedevano azioni specifiche da parte del personaggio si sono rivelate particolarmente complesse da generare in modo coerente e credibile. L'intelligenza artificiale, pur mostrando un buon livello di definizione visiva nei singoli frame, incontra ancora notevoli difficoltà nel coordinare i movimenti del corpo umano con quelli di oggetti, soprattutto quando l'interazione richiede precisione fisica o reazioni proporzionali alla forza e al contatto. Tentativi di far eseguire all'artista gesti semplici come chiudere una porta, afferrare un pennello o posizionare una tela su un tavolo, hanno generato frequentemente artefatti visivi, compenetrazioni tra corpi solidi o anomalie fisiche. Questi problemi rappresentano una forte limitazione per qualsiasi progetto audiovisivo che aspiri a una rappresentazione realistica delle interazioni ambientali.

Figura 4.8: Una particolare difficoltà generativa è rappresentata dall'interazione tra gli oggetti e la protagonista stessa. Se tenuti fermi, come in questo caso, si ottengono risultati accettabili.

Al di là delle difficoltà specifiche legate all'interazione, la simulazione della fisica si è rivelata generalmente soddisfacente. La resa dei fluidi ha raggiunto una qualità accettabile in più occasioni. Ad esempio, quando compare una statua, l'acqua che scende lungo la pietra è discretamente verosimile, anche se non completamente aderente ai parametri fisici reali. Considerando la natura simbolica e onirica di gran parte della narrazione, tali piccole deviazioni dalla fisica del mondo reale sono state accettate come elementi funzionali all'estetica generale del progetto. Inoltre, anche gli effetti secondari legati alla gravità e all'inerzia si sono rivelati nella maggior parte dei casi ben integrati, contribuendo alla credibilità dell'immagine. In qualche occasione è possibile notare una leggera problematica relativa alla fisica del tessuto che compone la camicia della protagonista, dettaglio trascurabile e non percepito dallo spettatore. Nonostante questo, il comportamento del corpo della protagonista presenta a tratti rigidità evidenti, soprattutto nei movimenti degli arti superiori e nelle sequenze in cui cammina. La camminata, in particolare, è uno degli elementi più problematici. In alcune clip è possibile osservare la ripetizione irregolare della gamba che avanza o una sovrapposizione innaturale delle fasi motorie, rompendo così la fluidità del passo e generando un effetto visivo poco credibile. Anche i movimenti delle braccia possono talvolta apparire meccanici o al contrario fluidi, troppo lenti, privi sicuramente della naturale oscillazione e flessibilità tipiche del corpo umano. Queste anomalie, pur non invalidando l'efficacia narrativa della sequenza, possono creare un senso di disorientamento, soprattutto quando il personaggio è posto in primo piano e l'azione è centrale nella composizione dell'inquadratura. Le limitazioni sopra descritte hanno richiesto da parte dell'autore un lavoro aggiuntivo in fase di montaggio, con l'obiettivo di minimizzare l'impatto degli errori. Ad esempio, la scelta di posizionare parti della camminata incoerenti prossime al taglio con la successiva clip, ha permesso di mascherare la problematica in quanto, iniziando una nuova inquadratura, lo spettatore non ha il tempo di concentrarsi sul precedente errore. L'interazione con gli oggetti e la simulazione dei movimenti restano comunque uno dei limiti più evidenti delle attuali tecnologie di generazione video e rappresentano un'area critica su cui è ancora necessario intervenire con stratagemmi in post-produzione.

Figura 4.9: Esempio di problema nella rappresentazione naturale della camminata. Si può vedere chiaramente come la protagonista effettua due volte di fila il passo con la gamba destra, riportando quindi un movimento completamente innaturale. L'azione, se vista rapidamente, non crea grosse problematiche.

L'analisi della qualità visiva delle immagini generate dall'intelligenza artificiale all'interno del cortometraggio evidenzia un equilibrio complessivamente positivo tra resa tecnica e coerenza estetica. Le dominanti cromatiche impiegate nelle diverse sequenze si attestano prevalentemente su toni caldi che contribuiscono a costruire un'atmosfera emotivamente avvolgente e coerente con l'interiorità del personaggio principale. In numerose scene, questo trattamento cromatico si è rivelato partico-

larmente efficace nel veicolare sensazioni intime e personali, rafforzando il legame tra lo spettatore e la protagonista.

Figura 4.10: Frame particolarmente riuscito nel rappresentare un momento di interiorità del personaggio. La luce calda è certamente di grande aiuto per raggiungere l'obiettivo.

Nel segmento centrale del film, caratterizzato dal viaggio onirico della protagonista attraverso le sue opere artistiche, la gestione della luce assume una caratteristica meno realistica e più simbolica. In queste sequenze, la variazione continua delle fonti e delle intensità luminose non è percepita come incoerente, bensì come parte integrante della narrazione. In un contesto immaginario e sospeso, dove spazio e tempo non seguono logiche lineari, anche il linguaggio visivo può permettersi una maggiore libertà espressiva. In tal senso il cambiamento della luce tra una scena e l'altra è stato in grado di marcare il passaggio tra opere, stili e stati d'animo differenti.
Uno degli aspetti più sorprendenti e qualitativamente apprezzabili emersi nel corso della produzione del cortometraggio riguarda la gestione della profondità di campo. Le clip che includono dettagli ravvicinati, come ad esempio il volto durante i momenti di riflessione, hanno mostrato una qualità compositiva particolarmente elevata. L'intelligenza artificiale ha dimostrato una notevole padronanza nel simulare un'adeguata sfocatura dello sfondo, concentrando la messa a fuoco sul soggetto in primo piano, secondo logiche estetiche perfettamente coerenti con i principi della cinematografia tradizionale. Questa capacità ha consentito di conferire alle sequenze un forte impatto, valorizzando i dettagli significativi dell'azione, catturando così l'attenzione dello spettatore. In queste situazioni, l'algoritmo ha restituito immagini convincenti, ben illuminate e dotate di una consistenza realistica delle texture della pelle e degli elementi circostanti.

Figura 4.11: Inquadratura ravvicinata dell'artista protagonista. Da notare l'ottima gestione della profondità di campo.

La protagonista del cortometraggio è stata spesso inquadrata in situazioni in cui la profondità visiva assumeva un ruolo fondamentale e la resa dell'intelligenza artificiale si è dimostrata all'altezza, in particolare per la messa a fuoco e la nitidezza degli occhi e dei capelli. Questi elementi, quando correttamente generati, contribuiscono in modo determinante alla credibilità visiva e al coinvolgimento emotivo dello spettatore. Quindi, se da un lato alcune scene hanno mostrato espressività limitata, la qualità delle inquadrature, l'equilibrio tra il soggetto e lo sfondo e la precisione della messa a fuoco hanno compensato in parte tale mancanza, offrendo comunque un'immagine visivamente coinvolgente.
È opportuno sottolineare che il buon esito delle inquadrature ravvicinate non è da considerarsi scontato nel contesto della generazione tramite intelligenza artificiale. La gestione della profondità di campo, infatti, rappresenta una delle sfide più complesse per i modelli generativi, poiché implica una comprensione tridimensionale della scena e delle sue relazioni spaziali interne. Nel presente progetto, invece, tali inquadrature si sono spesso rivelate tra le più riuscite e coerenti, a dimostrazione del fatto che la tecnologia attuale è già in grado di supportare, se adeguatamente guidata, una resa visiva sofisticata. Questo punto di forza suggerisce che un utilizzo più strategico dei primi piani e dei dettagli potrebbe rappresentare una delle chiavi per compensare altri limiti ancora presenti nei modelli, garantendo al tempo stesso una resa narrativa e visiva di elevata qualità.
I movimenti di camera generati dall'intelligenza artificiale risultano in generale soddisfacenti, soprattutto quando la protagonista viene seguita lungo percorsi dinamici, come avviene in diverse camminate o attraversamenti ambientali. In tali casi, la camera si muove con una fluidità accettabile, mantenendo l'inquadratura centrata sul personaggio e garantendo un buon senso di continuità. Si rileva però una prevalenza di sequenze statiche o semi-statiche, che appaiono come semplici punti di osservazione fissi, simili a prospettive in prima persona o inquadrature neutre. Questa tendenza, osservata in particolare nelle clip generate in assenza di istruzioni specifiche, evidenzia un uso ancora parziale delle potenzialità cinematiche del modello. Una precisazione va fatta per i movimenti di camera più complessi come ad esempio carrellate, panning laterali o cambi di zoom progressivi che risultano difficili da ottenere senza una costruzione estremamente dettagliata del prompt. Il modello

tende infatti in generale a semplificare l'azione visiva, ricorrendo a soluzioni più conservative. Questo aspetto ha limitato in parte l'articolazione registica di alcune sequenze che, in una produzione tradizionale, avrebbero potuto beneficiare di maggiore varietà e dinamismo visivo.

Figura 4.12: Movimento di camera complesso, generato positivamente dall'intelligenza artificiale. La sequenza dei frame parte in alto a sinistra per poi proseguire in senso orario.

Un elemento significativo da considerare nel processo di valutazione estetica del cortometraggio riguarda la gestione dei movimenti incoerenti e dei glitch grafici all'interno delle clip generate. L'autore umano ha scelto fin da subito di scartare in maniera sistematica tutte le clip che presentavano imperfezioni evidenti. Questo ha consentito di ottenere un prodotto finale sostanzialmente privo di distorsioni percettibili o errori strutturali, evitando così la necessità di interventi correttivi a posteriori. Non sono stati rilevati oggetti fluttuanti, compenetrazioni evidenti tra corpo e sfondo o comportamenti fisicamente incoerenti nella resa ambientale, problemi comuni dei contenuti generati dall'intelligenza artificiale. Va comunque considerato che una parte consistente del cortometraggio si sviluppa in un contesto dichiaratamente onirico e immaginario, in cui la sospensione delle regole logiche dello spazio e del tempo consente una maggiore flessibilità visiva. In altri termini, laddove anche fossero presenti lievi incoerenze nella rappresentazione, queste risulterebbero facilmente integrabili nella narrazione, grazie al carattere sognante e simbolico del viaggio dell'artista all'interno delle sue opere.

Figura 4.13: Particolare fluttuante incontrato dall'artista durante il viaggio esplorativo.

Un limite stilistico rilevante che ha influenzato negativamente l'estetica inizialmente immaginata, riguarda l'incapacità dell'intelligenza artificiale di rappresentare in modo riconoscibile diversi stili pittorici. L'intento originario del progetto prevedeva che, dalla scena 6 alla scena 11, la protagonista attraversasse una sequenza di ambienti ispirati ciascuno a una corrente artistica differente. Per ogni scena era stato previsto uno stile pittorico ben definito, identificato precedentemente grazie a una ricerca accurata. Nonostante questo, nella fase successiva di generazione video, tale scelta non è stata mantenuta. Le clip prodotte presentavano risultati poco coerenti con gli stili pittorici indicati, tendendo a ricadere frequentemente in un'estetica generica e realistica, indipendentemente dalle indicazioni fornite. Questa difficoltà ha reso impossibile strutturare visivamente un percorso narrativo che riflettesse davvero la varietà degli stili artistici utilizzati dall'artista. La mancata riproduzione fedele di tratti distintivi come texture, pennellate o palette cromatiche proprie di ogni corrente pittorica ha vanificato il tentativo di costruire un viaggio stilisticamente variegato. Nonostante alcuni esperimenti abbiano prodotto clip visivamente suggestive, la maggior parte dei risultati si è rivelata troppo simile, rendendo difficile distinguere i passaggi tra una scena e l'altra basandosi solamente sul linguaggio estetico adottato.

Figura 4.14: Frame appartenenti al viaggio della protagonista. Ogni scena doveva essere rappresentata in uno stile pittorico differente. I limiti dell'intelligenza artificiale hanno portato ad adottare una soluzione ibrida, giustificata tramite il voiceover. In particolare, in alto sono presenti due esempi raffiguranti stili pittorici differenti mentre in basso lo stile è prettamente realistico, vista l'impossibilità di rappresentare altre correnti pittoriche.

Questa problematica ha evidenziato quanto, in uno stato dell'arte ancora in fase di evoluzione, la possibilità di intervenire tempestivamente sul processo creativo sia cruciale per la riuscita complessiva dell'opera. La realizzazione del cortometraggio ha potuto proseguire grazie a una riconfigurazione prettamente narrativa, in cui il fallimento tecnico è stato riconvertito grazie a una soluzione descrittiva. Il passaggio tra gli stili è stato sostituito da un'alternanza tra scene visionarie e scene più reali, giustificata a livello narrativo attraverso il voiceover dell'artista che commenta in maniera riflessiva la natura del proprio viaggio, pensando di essere al termine quando attraversa paesaggi più realistici e tornando invece a una dimensione onirica quando arriva in un ambiente pittorico.
Nonostante queste criticità, è possibile affermare che la qualità complessiva delle immagini generate si sia mantenuta sempre su un livello soddisfacente. Alcune clip, come il dettaglio floreale, si sono distinte per la precisione visiva. Ciò conferma che, se ben guidata e correttamente raffinata, la generazione video tramite intelligenza artificiale può produrre risultati di notevole impatto estetico, capaci di sostenere anche progetti di narrazione complessa come quello analizzato.

Figura 4.15: Frame appartenente a una clip particolarmente riuscita. I dettagli e la profondità di campo sono stati generati in maniera ottima.

Un aspetto particolarmente rilevante emerso durante il lavoro, riguarda la continua negoziazione tra l'estetica proposta dal modello generativo e la visione autoriale umana. Questo processo, lungi dall'essere lineare o prevedibile, si è rivelato il vero motore creativo del progetto: l'autore umano ha dovuto interpretare, correggere, guidare e in alcuni casi accogliere l'estetica prodotta dall'intelligenza artificiale, integrandola con le proprie scelte narrative. Il bilancio complessivo evidenzia come, pur tra limiti tecnici e necessarie ridefinizioni progettuali, l'esperienza creativa condotta si sia rivelata significativa, sia come esercizio di sperimentazione stilistica che come dimostrazione della capacità dell'autore umano di adattarsi e modellare l'opera, nonostante le restrizioni imposte dagli strumenti tecnologici.
Dal punto di vista stilistico, il cortometraggio conserva tratti riconoscibili dell'estetica dei contenuti generati con IA: movimenti macchinosi, posture rigide, inquadrature tipiche della logica generativa. Tuttavia, tali elementi non costituiscono necessariamente un difetto e possono essere visti come un nuovo codice visivo emergente, distintivo della produzione mediata dall'intelligenza artificiale, che però uniformato per tutti, risulta monotono.

4.2.5 Il carattere didascalico del cortometraggio generato

L'intelligenza artificiale, quando impiegata nella costruzione narrativa di prodotti audiovisivi, tende naturalmente a privilegiare una comunicazione visiva immediata, basata su strutture di linguaggio già codificate. Questo comportamento è legato alla logica combinatoria e probabilistica del suo funzionamento. I modelli, per costruire immagini coerenti, attingono da un archivio di soluzioni apprese attraverso data set, molti dei quali sono pieni di convenzioni simboliche e archetipi culturali. Ne risulta una predilezione per l'elemento didascalico, ovvero per quelle immagini che traducono in forma esplicita un'emozione, uno stato mentale o un passaggio esistenziale, riducendo la polisemia a favore dell'immediatezza comunicativa.
Tale meccanismo è evidente fin dalla prima sequenza del cortometraggio prodotto, quando la vista aerea pomeridiana della città, serve da codice visivo per introdurre la protagonista all'interno di una dimensione urbana solenne e rarefatta, carica di potenzialità narrative. Questa visione incarna un codice narrativo ben riconoscibile,

quello dell'artista nella metropoli, assieme alla propria arte ma in conflitto con il passato.
Un altro esempio è l'utilizzo ricorrente della contrapposizione fra città e natura, presente nei flashback che mostrano la giovane artista immersa spensieratamente in paesaggi montani, in netto contrasto con le inquadrature urbane più riflessive. Questa dicotomia, benché coerente con il percorso narrativo, rientra in uno schema ormai ricorrente in molte produzioni; la natura è vista come luogo autentico, la città come uno spazio di crisi esistenziale. Questo comportamento dell'intelligenza artificiale nel produrre narrazioni visive non è casuale, ma è riconosciuto anche dalla letteratura scientifica. Una ricerca di Wei et al. [5] sottolinea come l'IA generativa, pur eccellendo nell'arricchire scene e garantire coerenza estetica, tende a privilegiare schemi narrativi altamente formalizzati e convenzionali. All'interno del loro modello collaborativo di storytelling visivo, gli autori notano come l'intelligenza artificiale assuma il ruolo di amplificatore semantico, ossia un agente che arricchisce i dettagli ma fatica a gestire ambiguità o a suggerire significati multipli.
Un altro limite significativo dell'intelligenza artificiale generativa risiede nella modalità con cui veicola i significati all'interno della narrazione, anch'essa altamente didascalica. I modelli di IA, infatti, mostrano una notevole difficoltà strutturale nel costruire quei livelli di allusione che sono propri del linguaggio audiovisivo più maturo e prettamente umano. Manca in altre parole la capacità di suggerire contenuti attraverso la composizione, la ricorrenza simbolica, il montaggio parallelo o la relazione di elementi provenienti da scene differenti che possano emergere a posteriori nella mente dello spettatore come interpretazione autonoma. Al contrario, l'IA tende a esplicitare ogni passaggio narrativo con una linearità marcata, spesso traducendo le sequenze in maniera eccessivamente diretta e spiegata, quasi a voler ridurre ogni possibile ambiguità o fraintendimento. Questo approccio, se da un lato garantisce una maggiore accessibilità e immediatezza nella comprensione, dall'altro compromette la possibilità di un coinvolgimento interpretativo più profondo da parte dello spettatore, riducendo il prodotto audiovisivo a una semplice trasmissione unidirezionale di contenuti. La dimensione implicita, il non detto, l'intervallo tra le immagini restano al momento difficilmente implementabili dai sistemi generativi, ancora troppo vincolati a logiche predittive e alla restituzione coerente e immediata di un messaggio esplicito.
In conclusione, la grammatica visiva dell'intelligenza artificiale nel progetto, mostra una sorprendente coerenza formale e una raffinata padronanza estetica. Tuttavia, proprio questa coerenza rischia di trasformarsi in rigidità. L'uso massiccio di elementi didascalici riduce lo spettro interpretativo, riconducendo ogni immagine a un solo significato possibile. La capacità dell'IA di generare contenuti di alto livello è innegabile, ma la sua capacità di suggerire, lasciare spazi aperti, produrre elementi che non spieghino completamente tutto, resta ancora lontana.

4.2.6 L'emotività nel cortometraggio generato

Un aspetto importante da valutare riguarda l'espressività del volto umano e, più in generale, la gamma emozionale della protagonista all'interno del cortometraggio. Le sequenze che pongono in primo piano il viso dell'artista mostrano una certa staticità, le espressioni tendono a mantenersi uniformi, con variazioni minime tra una clip e l'altra. L'occhio, la bocca e i muscoli facciali non mostrano movimenti progres-

sivi, risultando spesso rigidi o bloccati in posizioni neutre. Questa caratteristica è particolarmente evidente nelle scene in cui la protagonista dovrebbe mostrare stati d'animo profondi, come momenti di incertezza, commozione o ispirazione creativa. In queste circostanze, la mancata articolazione espressiva riduce l'impatto emotivo della scena e spezza in parte l'immedesimazione dello spettatore.

Figura 4.16: Frame rappresentante una clip in cui l'espressività della protagonista è praticamente nulla.

Va tuttavia riconosciuto che non tutte le sequenze sono ugualmente carenti da questo punto di vista. In alcune clip (ad esempio nella scena 2 in cui l'artista sorride mentre ricorda la sua prima mostra) l'IA è riuscita a generare un volto con una mimica che ben si adatta al contesto narrativo. È verosimile ipotizzare che future evoluzioni tecnologiche, permetteranno di superare il limite, introducendo una maggiore ricchezza espressiva e una più fine gestione del linguaggio corporeo.
Un altro elemento fondamentale nell'equilibrio tra linguaggio visivo e narrazione è stato il ruolo del voiceover. La voce narrante, che accompagna e scandisce ogni fase del racconto, funge da guida immersiva, offrendo allo spettatore continui punti di riferimento interpretativi. Il voiceover è capace di chiarire l'azione, suggerire stati emotivi e delineare la trasformazione della protagonista. Purtroppo la componente vocale presenta ancora margini di miglioramento sul piano espressivo. Il modello vocale adottato per generare la voce in italiano, trasformata artificialmente a partire dalla voce dell'autore umano, non raggiunge ancora i livelli di profondità e naturalezza che sarebbero auspicabili per una narrazione empatica e coinvolgente. In particolare, si registra una certa monotonia e una limitata capacità di modulazione emotiva, soprattutto nei momenti più intensi o riflessivi del racconto. Il voiceover è stato calibrato in funzione di ogni scena, garantendo una completa coerenza tra immagini e storia. Nonostante i limiti ancora evidenti della componente vocale, la scelta di accompagnare ogni sequenza con il pensiero dell'artista ha rafforzato la fruibilità e l'unitarietà narrativa del progetto.

4.2.7 Linguaggio e intenzionalità dell'IA come autore

L'impiego dell'IA solleva spesso interrogativi sulla natura della creatività e dell'originalità nell'arte. È noto che i modelli attuali vengono allenati su miliardi di prodotti visivi creati dall'uomo e successivamente adattano quanto appreso; va ricordato però

che anche l'artista umano tradizionalmente crea rielaborando e inserendo influenze pregresse.
La differenza cruciale risiede infatti nella dimensione critica e interpretativa: un contenuto generato dall'IA è di per sé un artefatto grezzo, che diventa portatore di significato solamente attraverso l'interpretazione e il contesto. I modelli sono in grado di eseguire dettagliatamente quanto richiesto ma la qualità descrittiva può far percepire le opere generate tramite IA intellettualmente vuote, prive di significato, nonostante siano talvolta visivamente affascinanti.
In termini critici l'output generato soddisfa la richiesta con precisione ma non possiede quel significato aggiunto che nasce dall'intenzione e dalla sensibilità umana. Ne consegue che la creatività dell'IA va intesa in senso relazionale: la macchina offre materiale grezzo, spesso con stereotipi visivi o con un gusto collettivo derivato dai data set utilizzati in fase di allenamento e spetta all'umano reinterpretarlo. È proprio in questi termini che si può confermare l'assenza di intenzionalità da parte dell'IA nella creazione dei contenuti. In uno studio condotto da Hall e Schofield [60], anche quando non viene comunicata l'origine dell'opera, gli spettatori tendono ad attribuire un valore più elevato alle opere create da esseri umani. Come osservano gli autori, nonostante l'arte generata dall'IA possa risultare esteticamente comparabile, continua a possedere un valore percepito inferiore, segno di un bias intrinseco che privilegia il processo creativo e il tocco umano, elementi che l'IA fatica a replicare. Questo suggerisce che non è la forma dell'opera a determinare il suo valore ma la percezione di una presenza intenzionale e sensibile, presente dietro il gesto artistico. L'IA dunque, non viene rifiutata per i risultati che produce, ma per la mancanza di una profondità interpretativa che solo l'intervento umano può dare.
L'arte generativa diventa così un terreno di co-creazione in cui l'IA suggerisce una serie di possibilità e l'occhio umano seleziona, integra e apre nuove strade espressive. Solo attraverso questo intervento critico l'opera algoritmica può elevarsi da esercizio meramente didascalico a oggetto culturale significativo. L'atto creativo risiede tanto nell'uso dell'IA quanto nella lettura umana dei suoi risultati, in un ciclo continuo di proposta e valutazione. Solamente forti di questa consapevolezza è possibile comprendere la vera forza dell'IA, definendola come uno strumento complesso, capace spesso di produrre ottimi contenuti ma il cui output va sempre interrogato e raffinato dall'ingegno umano.
Questa riflessione trova conferma anche nei risultati dello studio di Rezwana et al. [61], secondo cui il vero potenziale dell'IA generativa si manifesta solo quando il suo output è riletto criticamente dall'umano, all'interno di un processo collaborativo orientato a valori condivisi. Gli autori sottolineano come l'IA, pur partecipando attivamente alla produzione creativa, non possa essere considerata autonoma né intenzionale e concludono che è strettamente necessario approfondire al più presto le questioni etiche e il punto di vista degli utenti riguardo la progettazione di sistemi di IA co-creativa equi e incentrati sull'uomo.

4.2.8 Conclusione

L'IA offre possibilità visive senza precedenti, combinando in maniera fluida forme, stili diversi e mettendo contemporaneamente a disposizione dell'artista un enorme quantitativo di contenuti. Questa realtà in espansione, come è successo per innovazioni tecnologiche nel corso della storia, stimola interrogativi fondamentali: cosa de-

finisce la creatività? Dove risiede il significato artistico quando l'opera è co-prodotta da una macchina?
L'IA porta con sé però i propri limiti e bias, che si riflettono direttamente nell'estetica dei contenuti generati: tende a reiterare schemi esistenti, privilegia soluzioni visive statisticamente sicure, talvolta stereotipate e manca di quella visione selettiva guidata dalla sensibilità e dall'intento che contraddistingue l'artista umano.
L'IA non rappresenta né la fine dell'arte né una radicale rivoluzione creativa, bensì una continuazione tecnologica di realtà già presenti nella *new media art* e nelle avanguardie del passato. In altre parole, si inserisce in un contesto preparato da decenni di sperimentazioni digitali, innovazioni e tentativi falliti, portando a una forte rielaborazione culturale che in parte era già stata esplorata dall'arte contemporanea. Spetta a noi utilizzatori orientare l'IA verso risultati che non siano meri esercizi di stile, bensì contributi originali per il panorama artistico. Come osserva Gerner [62], l'estetica dell'intelligenza artificiale può essere intesa come una nuova fonte di innovazione computazionale e algoritmica, capace di ispirare arte e riflessione sociale. Di fronte all'indeterminatezza visiva e alla coerenza simulata delle immagini neurali, diventa cruciale interrogarsi sui presupposti estetici e filosofici che guidano la produzione generativa. In quest'ottica, l'estetica dell'IA richiede un'indagine critica, capace di aprire la scatola nera dell'algoritmo per esplorare i limiti e le sfide che pone alla sensibilità artistica contemporanea. Comprendere i principi di base dell'IA nella cultura odierna è importante per essere alfabetizzati digitalmente ed è un passo essenziale per chi si definisce creativo. Solo consapevoli di ciò, sarà possibile sfruttare criticamente il potenziale estetico dell'intelligenza artificiale, evitando sia l'entusiasmo ingenuo sia il rifiuto pregiudiziale, collocando le nuove opere generate come continuazione della storia dell'arte.

4.3 Impatto del video

4.3.1 Feedback e reazioni del pubblico

L'analisi dei dati raccolti attraverso il questionario rivolto al pubblico al termine della proiezione del prodotto audiovisivo, ha permesso di verificare l'impatto complessivo del cortometraggio realizzato con l'intelligenza artificiale. Le risposte, seppur diverse, hanno evidenziato alcune tendenze chiare, che permettono una riflessione utile per comprendere l'efficacia complessiva del lavoro svolto e per valutare criticamente le potenzialità e i limiti dell'IA nei processi creativi.
Nonostante non siano presenti dati relativi all'età, sesso e nazionalità del pubblico presente in sala durante l'evento, si può generalmente affermare che, da un punto di vista anagrafico, gli spettatori erano equamente distribuiti tra le diverse fasce d'età. Anche la percentuale tra pubblico femminile e maschile è bilanciata, con una netta predominanza di persone di nazionalità italiana. Hanno risposto alle domande 68 persone e successivamente i dati sono stati puliti prima di procedere con l'analisi.
Un primo approfondimento riguarda la capacità del prodotto di veicolare emozioni. Il pubblico afferma di aver provato maggiormente *confusione*, *nostalgia*, *tristezza* ma anche *curiosità* e *noia*. L'impatto emotivo, pur essendo presente, non ha sviluppato un coinvolgimento profondo tra gli spettatori. Questo aspetto fa emergere un chiaro limite nella dimensione espressiva e nella capacità di stimolare stati d'animo

complessi da parte dell'IA, suggerendo che i modelli generativi impiegati siano in grado, al momento, di eseguire meccanicamente quanto richiesto dall'autore umano, senza produrre però risultati esaustivi per quanto riguarda la creazione di atmosfere emotive complesse. Rimane comunque interessante l'impatto su una parte del pubblico che ha dimostrato curiosità verso il prodotto audiovisivo proiettato, avendo probabilmente un'inclinazione più marcata a esplorare e accogliere questa nuova tecnologia.
Rimanendo sempre nello stesso ambito, alla domanda se fosse stato percepito un cambiamento nello stato d'animo dopo la visione del cortometraggio, il 20% del pubblico ha risposto in maniera negativa, segno che per una parte degli spettatori il prodotto audiovisivo non è stato in grado di veicolare particolari emozioni. Altre persone, invece, hanno affermato di essere *ispirate*, *stimolate*, oppure di essere entrate in uno stato di maggiore *riflessione*. Altri spettatori ancora hanno evidenziato di sentirsi maggiormente *sfiduciati* (alcuni verso il futuro) o *ansiosi*.
L'alto punteggio ottenuto alla domanda presente in Figura 4.17, denota che l'estetica dei video generati con IA è facilmente riconoscibile e quindi in questo senso devono ancora essere compiuti dei miglioramenti. Non è questo però l'unico fattore a influenzare la risposta; anche l'audio imperfetto, come riportato in alcune risposte, ha contribuito sicuramente ad aumentare questa percezione, oltre allo stile realistico che si è scelto di adottare.

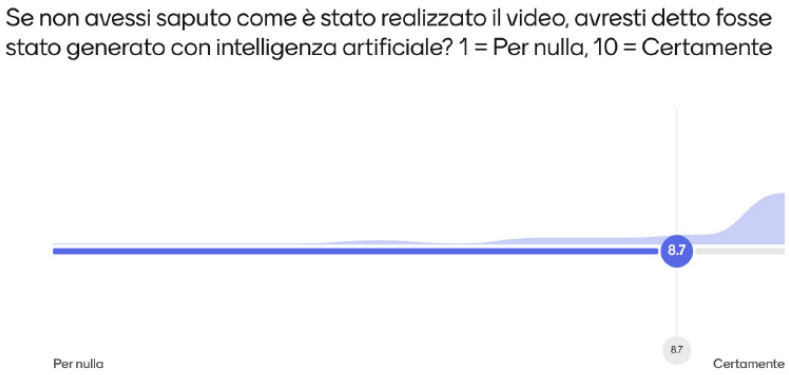


Figura 4.17: Il risultato di questa domanda fa capire quanto sia ancora facilmente riconoscibile un video generato con intelligenza artificiale.

Un aspetto emerso frequentemente durante l'analisi delle risposte, specie alla domanda sull'impatto complessivo del video, riguarda la percezione della lunghezza del prodotto audiovisivo finale. Sia in questa che in altre domande, infatti, il cortometraggio è stato definito *lungo*, indicando un certo livello di affaticamento e noia nella visione del contenuto e, quindi, una difficoltà a mantenere l'attenzione per tutta la durata dell'opera. Questo tipo di reazione va interpretato in modo articolato: non è il semplice dato oggettivo sulla durata a costituire il problema, ma va analizzata piuttosto la qualità del coinvolgimento narrativo e visivo che il prodotto è riuscito a suscitare.
Da un punto di vista tecnico infatti, un cortometraggio è un qualsiasi prodotto audiovisivo della durata inferiore ai 52 minuti, come riportato dalle linee guida del *Ministero della Cultura* [63]. Normalmente però, si usa produrre contenuti che non eccedano i 30 minuti, con la maggior parte dei cortometraggi che tendono ad avere

una durata attorno ai 15 minuti. Oggettivamente il prodotto audiovisivo del progetto (15 minuti totali, inclusi titoli di coda) rientra tranquillamente nelle tempistiche comuni a quel tipo di lavoro e la lunghezza percepita, appare più come una conseguenza indiretta di altri fattori. In particolare, alcune sezioni del cortometraggio possono risultare poco coinvolgenti, con una narrazione che tende in alcuni momenti a risultare ripetitiva o eccessivamente didascalica, come in alcune sequenze del viaggio onirico che non trasmettono nulla di concreto da un punto di vista narrativo. Un altro fattore influente è la progressione narrativa lineare e priva di sorprese, caratteristica che ha probabilmente inciso sulla sensazione di durata eccessiva, rendendo il flusso visivo meno dinamico e dunque più difficile da seguire. Un'altra possibile causa di questa reazione risiede nel fatto che il pubblico non è sempre riuscito a cogliere pienamente il contesto di sperimentazione che ha guidato la realizzazione dell'opera. Il cortometraggio è infatti nato come un prodotto di ricerca, pensato anche come laboratorio narrativo in cui esplorare le possibilità dell'intelligenza artificiale nelle varie fasi produttive. In questo senso, l'intenzione non era necessariamente quella di generare un prodotto perfettamente bilanciato in termini di ritmo e impatto emotivo, quanto piuttosto quella di testare diverse soluzioni estetiche e tecniche all'interno di un unico contenitore.
Non tutte le reazioni sono state negative. Alcuni spettatori hanno comunque definito il video *carino* e *interessante*, riconoscendo lo sforzo di costruzione e la particolarità della proposta. La presenza di commenti che evidenziano una sorta di *straniamento* conferma che il prodotto audiovisivo ha comunque lasciato un segno, suscitando interrogativi e riflessioni.

Figura 4.18: Grafico generato dalle risposte alla domanda *Con una parola, come descriveresti l'impatto complessivo del video?*

Alla domanda se il personaggio principale sia verosimile o meno, il pubblico si è espresso con un voto medio di 4,4 su 10. Il risultato è sicuramente notevole, visto che la coerenza e continuità nella rappresentazione di figure umane è certamente uno dei punti dove i modelli di intelligenza artificiale faticano di più. Un punteggio così contenuto evidenzia una distanza percettiva ancora marcata tra l'intento rappresentativo e l'effettiva resa visiva. Il personaggio, pur costruito con attenzione in fase

di prompt engineering, ha sicuramente sofferto di variazioni minime ma altamente percepibili nei tratti fisionomici, nelle espressioni o nell'integrazione con l'ambiente visivo circostante. Tali discrepanze, anche se sottili, tendono a essere notate dallo spettatore e incidono negativamente sul senso di realismo e credibilità del personaggio. Inoltre, va considerato che la natura sintetica della voce può aver contribuito a rafforzare una sensazione di artificialità del personaggio, rendendo più difficile per il pubblico stabilire un coinvolgimento empatico autentico.
Infine, è stato chiesto al pubblico quale fosse il messaggio suggerito dal cortometraggio. Si è riscontrato, in linea generale, un ampio interesse a questa domanda, con vari spunti proposti dagli spettatori. Diverse persone hanno individuato la *speranza* come messaggio principale, altri hanno sottolineato il fatto che il prodotto audiovisivo porta a considerare la *riflessione* come una fase importante della vita. Altre persone ancora hanno individuato il *trovare le proprie radici* e il *lottare per un mondo migliore* come messaggio principale e circa il 10% del pubblico ha commentato che l'ultima frase del cortometraggio rappresenta bene il significato del video: *tornare indietro, a volte, è necessario per andare avanti.* Infine, una considerazione va fatta sulla parte di pubblico che non ha trovato un particolare messaggio nella visione del prodotto audiovisivo, a dimostrazione che il contenuto del video poteva essere veicolato certamente in modo migliore. In generale però, questa domanda ha certamente sottolineato che un messaggio è veicolabile in maniera netta anche attraverso l'uso dell'IA.
I dati raccolti ma non presentati completamente sono disponibili per la consultazione alla Sezione L dell'appendice.

4.3.2 Analisi dell'efficacia comunicativa del cortometraggio

Uno degli aspetti maggiormente emersi dall'analisi qualitativa dei dati raccolti riguarda la difficoltà, da parte del prodotto audiovisivo, di trasmettere emozioni in modo autentico. L'impressione generale restituita dal pubblico è stata quella di un contenuto visivamente anche ben costruito, ma con una componente emotiva piuttosto piatta o comunque non sufficientemente articolata per innescare un coinvolgimento profondo. Gli aspetti da migliorare riguardano infatti la durata e la scarsa emotività del cortometraggio, segno che da un punto di vista tecnico qualitativo le riprese non hanno influenzato negativamente la fruizione del prodotto. Il limite emotivo sembra essere legato, almeno in parte, all'attuale incapacità dei modelli di restituire quelle sfumature affettive, ambiguità emotive e piccole variazioni espressive che costituiscono l'essenza dell'empatia narrativa. Se da un lato l'IA è riuscita a produrre contenuti anche coesi tra di loro, dall'altro ha mostrato un'insufficienza strutturale nel veicolare tensioni e contrasti emotivi che tradizionalmente caratterizzano i momenti di maggior intensità nelle opere audiovisive.
Un elemento positivo risiede invece nell'efficacia comunicativa del messaggio veicolato, in quanto buona parte del pubblico è riuscita ad individuarlo chiaramente, dimostrando che, da questo punto di vista, il prodotto audiovisivo è riuscito nel suo intento. C'è da considerare il fatto che il messaggio finale è più volte riportato in maniera esplicita, sia dalla protagonista che nei titoli di coda. Più volte inoltre, lungo l'intera durata del cortometraggio, vengono riportati esplicitamente gli elementi chiave, utili allo spettatore per comprendere al meglio quanto l'autore vuole esprimere. La didascalicità portata avanti dai modelli di IA, infatti, esplicita il tutto,

rendendo maggiormente efficace la diffusione del messaggio finale, probabilmente eccessivamente semplificato dall'IA, visto che una parte del pubblico, seppur ridotta, l'ha trovato banale. L'IA, non riuscendo a gestire con efficacia il non detto, ha optato per un registro fortemente esplicito, didascalico, in cui le informazioni narrative venivano dichiarate in modo diretto, senza offrire la possibilità di una lettura sottotraccia. Questo approccio, che ha garantito una chiarezza comunicativa, ha ridotto drasticamente l'impatto evocativo dell'opera, escludendo le forme di partecipazione cognitiva ed emotiva che rendono maggiormente fruibile un'opera.
Una parte significativa delle criticità emerse potrebbe essere attribuita a un elemento certamente non trascurabile, dato dal forte disallineamento tra il pubblico presente al momento della proiezione del cortometraggio e quello originariamente coinvolto nella fase preliminare di raccolta dei dati. I primi concerti svolti a fine 2024, avevano visto la partecipazione di spettatori attivamente coinvolti nel processo creativo del progetto, attraverso le preferenze, emozioni e sensazioni espresse, raccolte e successivamente analizzate per dare vita al cortometraggio. Tuttavia, al momento della proiezione finale del cortometraggio, il pubblico presente era, per la maggior parte, diverso da quello della prima fase esperienziale. Questo potrebbe aver influenzato la sensibilità degli spettatori, in quanto il cortometraggio proposto non era parte della loro esperienza ma frutto di riflessioni altrui. Questa distanza, anche solo in termini di contesto e aspettative, ha probabilmente influito negativamente, favorendo un giudizio di estraneità o di eccessiva freddezza emotiva del prodotto. D'altra parte, chi ha fornito inizialmente i dati, sicuramente ha avuto maggior piacere vedendoli trasformati in immagini, creando così una connessione tra aspettative e cortometraggio. Inoltre, è sicuramente stimolante ed emozionante per coloro che in precedenza avevano partecipato ai concerti, ricercare all'interno del prodotto audiovisivo le preferenze precedentemente espresse, catturando così la loro attenzione e coinvolgimento durante la proiezione.
In conclusione, l'IA può supportare la costruzione di strutture narrative ma non può ancora eguagliare la capacità umana di intuire e gestire piccoli aspetti interpretativi e tensioni emozionali complesse. Un utilizzo consapevole e umanamente guidato dello strumento può portare ad ottenere dei buoni risultati, evitando una comunicazione appiattita, eccessivamente didascalica ed emotivamente poco efficace.

4.3.3 Corrispondenza tra intenti progettuali e risultati

L'intero progetto è nato dalla volontà di esplorare in modo operativo e concreto le potenzialità attuali dell'intelligenza artificiale all'interno della produzione audiovisiva, con un approccio altamente sperimentale e orientato a verificare praticamente i risultati ottenuti. Uno dei primi obiettivi è stato quello di comprendere il panorama di strumenti effettivamente disponibili per affrontare ogni fase della realizzazione di un cortometraggio: dalla scrittura alla generazione di video, fino all'audio e alla post-produzione. Questa fase preliminare è stata condotta tramite un'indagine, supportata dalla consultazione di benchmark accademici aggiornati e, successivamente, attraverso una sperimentazione pratica e diretta degli strumenti valutati positivamente. Sono stati testati numerosi modelli di IA, interrogati tutti con prompt identici, per valutare non solo la qualità del risultato ma anche fattori quali rapidità, stabilità della piattaforma e costo. La varietà di soluzioni disponibili si è rivelata molto ampia; alcuni strumenti, come *Sora* per la generazione video o *Suno*

per la colonna sonora, hanno mostrato un'elevata affidabilità, mentre altri, seppur promettenti, si sono rivelati limitati nella personalizzazione o non sostenibili sul piano tecnico, come le implementazioni locali di modelli, troppo esigenti in termini di risorse richieste, nonostante l'alta precisione.
Una volta individuate le piattaforme operative, è stato necessario strutturare un processo produttivo coerente che permettesse di passare da un'idea narrativa iniziale a un prodotto audiovisivo finito. Questo processo ha dato vita a una pipeline precisa che ha subito più volte delle modifiche per essere ottimizzata nel produrre concretamente quanto richiesto in ciascuna fase. Inizialmente la pre-produzione è stata gestita in maniera ibrida con l'utilizzo di *ChatGPT*, che ha portato poi alla scrittura di un soggetto, sviluppato successivamente in una sceneggiatura dettagliata. Su questa base sono stati progettati uno storyboard visivo e conseguentemente tutti i prompt adatti a generare le varie clip, per concludere infine con la produzione delle sequenze video vere e proprie, attraverso l'utilizzo della piattaforma *Sora*. Durante lo stadio di post-produzione, l'utilizzo dell'IA è stato secondario in quanto si è preferito non fare affidamento su troppi strumenti di IA, visto il loro funzionamento ancora non perfetto in questa fase.
In generale, il rapporto tra autore umano e macchina è stato fondamentale nello svolgimento del lavoro e la progressione dell'attività ha mostrato nel tempo come l'ordine e la chiarezza metodologica siano cruciali quando si opera con intelligenze artificiali generative che non possono ancora adattarsi, con la stessa fluidità dell'intuizione umana, ai cambiamenti in corso d'opera. Ogni piccola modifica richiedeva infatti prompt precisi per essere compresa dai modelli, come capitato più volte durante la generazione della storia principale, dove cambiare una parte della narrazione significava spiegare nel dettaglio tutte le conseguenze che la modifica avrebbe portato.
In linea generale, la pipeline utilizzata si è dimostrata efficace, portando a produrre il prodotto finale senza particolari problemi, soprattutto tenendo conto che è stata impiegata una sola risorsa umana per l'intero progetto. In particolare, è importante sottolineare il fatto che la creazione dei documenti di pre-produzione e la successiva generazione dei singoli contenuti video sono una diretta conseguenza del buon lavoro svolto nella prima parte del progetto, fondamentale per la buona riuscita di tutte le fasi successive. Preparare nel dettaglio, sempre sfruttando l'intelligenza artificiale, tutti i testi scritti necessari, si è rivelata essere una scelta che ha facilitato enormemente le successi fasi della produzione.
Un altro tema centrale, già emerso in realtà durante l'analisi del panorama audiovisivo, riguarda la coerenza narrativa e stilistica dei prodotti generati con IA. I modelli, specialmente quelli dedicati alla generazione di contenuti visivi, non garantiscono sempre un allineamento semantico o estetico tra le varie clip prodotte. Questo problema è stato riscontrato poi in diverse fasi della produzione, quando si è verificato che immagini teoricamente correlate, appartenenti ad esempio alla stessa scena, presentavano variazioni importanti nell'illuminazione, nei colori o addirittura nella fisionomia del soggetto principale rappresentato. È stato dunque svolto un lavoro dedicato per permettere all'IA di avere tutte le informazioni necessarie a mantenere una continuità narrativa plausibile, creando prompt precisi e specifici in modo tale da facilitare la generazione costante degli stessi ambienti e personaggi. Successivamente, è stato necessario intervenire manualmente per selezionare le clip migliori, ponendo una particolare attenzione alla successione visiva e scegliendo solamente i

video più coerenti tra loro. Questa parte della ricerca si è rivelata indispensabile per compensare l'incapacità dei modelli generativi attuali a mantenere coerenza tra output successivi ed ha rappresentato un'importante fase dell'intero lavoro. In parte si è riuscito a rendere marginali gli errori generativi anche se, all'occhio attento e critico, le differenze non sfuggono.
Dal punto di vista tecnico, è infine emersa con chiarezza l'esistenza di limiti strutturali nelle attuali tecnologie. Nonostante i modelli impiegati siano stati in grado di generare contenuti visivamente credibili, si è rilevata una notevole difficoltà nella gestione dei registri emotivi più complessi, nel costruire sottotesti e nell'articolare un messaggio in forma indiretta o simbolica. Il risultato, in molti casi, proponeva una comunicazione eccessivamente esplicita e didascalica, che ha reso il messaggio più semplice da cogliere ma meno interessante. Anche la ricezione da parte del pubblico ha confermato questo aspetto: molti spettatori hanno avvertito una certa distanza emotiva dal racconto, percependolo più come dimostrazione tecnica che come esperienza narrativa.
Nel complesso, il progetto ha permesso di rispondere a tutti i principali interrogativi presenti ad oggi sul funzionamento dell'intelligenza artificiale nella produzione audiovisiva. La ricerca condotta ha permesso di individuare le piattaforme più adatte per ciascuna fase, di definire una pipeline coerente e funzionale, di verificare empiricamente le difficoltà nella costruzione di un'estetica e una narrazione coesa e di osservare con attenzione i limiti derivanti dall'utilizzo dell'intelligenza artificiale in campo audiovisivo. Le soluzioni messe in campo, sebbene non perfette, hanno comunque dimostrato che un approccio ibrido e guidato consente di ottenere un risultato efficace, confermando il valore della sperimentazione come pratica centrale nella ricerca artistica contemporanea.

4.3.4 Riflessione critica sull'impatto del cortometraggio

La presentazione del cortometraggio ha evidenziato come un prodotto audiovisivo costruito attraverso strumenti di intelligenza artificiale possa risultare meno accessibile a un pubblico non specializzato e non preparato al senso della sperimentazione. La tecnologia e, conseguentemente, il prodotto finale, se non contestualizzato, rischia infatti di essere assimilato a un'estetica pari a quella dell'arte contemporanea, facendo emergere ancora di più la netta distanza che spesso contraddistingue i prodotti creati tramite IA. Lo spettatore non è in grado di comprendere tutti i limiti tecnici che sono attualmente presenti, sospendendo in parte il suo giudizio e apprezzando piuttosto lo sforzo di ricerca e sperimentazione che si vuole portare avanti. L'obiettivo principale del progetto, infatti, non è mai stato quello di raggiungere una perfezione estetica o narrativa tipica delle grandi produzioni cinematografiche, bensì quello di indagare, attraverso la pratica, i limiti e le potenzialità reali dell'intelligenza artificiale nella produzione audiovisiva. Ogni fase, pertanto, dalla scrittura alla generazione visiva, è stata concepita come una ricerca, in cui gli errori e le imperfezioni sono diventati parte integrante dell'analisi. Ad esempio, la decisione di utilizzare modelli T2V, pur presentando delle chiare limitazioni nella resa del movimento o nella continuità temporale, è stata una scelta strategica per verificare fino a che punto tali strumenti potessero sostenere una narrazione complessa mantenendo allo stesso tempo un'alta coerenza. Lo stesso vale per la voce sintetica della protagonista, modificata tramite clonazione, dove non si puntava a un realismo perfetto,

ma si voleva capire quale tipo di compromesso fosse accettabile e come tale voce potesse essere interpretata dal pubblico.
In ogni fase della produzione, non si è cercato a tutti i costi di nascondere i limiti dell'IA, ma la ricerca si è concentrata nel trovare la metodologia migliore per evitare o rendere meno evidenti le problematiche, anche per riflettere criticamente sul loro impatto percettivo ed emotivo. Questo approccio sperimentale permette di raccogliere dati, osservazioni e suggestioni non finalizzate alla realizzazione di un prodotto finale di successo, al pari degli standard dell'industria cinematografica, ma piuttosto alla costruzione di un sapere operativo utile per futuri progetti e per comprendere meglio il ruolo che l'IA può assumere nei processi creativi. La logica non è quindi quella della ricerca della perfezione, ma dell'esplorazione, dove c'è spazio anche per il fallimento. Questo si è potuto osservare chiaramente nel momento in cui parte del pubblico ha espresso difficoltà nel decifrare la natura e le emozioni veicolate dall'opera, probabilmente a causa di una narrazione che non ne svelava in alcun modo le reali intenzioni e le specifiche tecniche di riferimento. In questo senso, una forma più breve e diretta, maggiormente vicina a un semplice video, avrebbe potuto facilitare la ricezione generale del prodotto audiovisivo, portando a una comprensione più immediata a discapito del contenuto sperimentale, in realtà fondamentale per lo sviluppo del lavoro di tesi.
È evidente che opere nate da processi come quelli resi possibili dall'intelligenza artificiale, necessitino di una doppia chiave di lettura: da un lato la fruizione più classica, visiva ed emozionale, dall'altro la restituzione di una metodologia creativa in continua evoluzione e in piena sperimentazione, che risulti però comprensibile. Se viene meno quest'ultima, il rischio è che lo spettatore colga il prodotto audiovisivo come qualcosa di estremamente artificiale, senza andare a considerare più ampiamente il lavoro svolto, andandolo a classificare come incapace di avvicinarsi alle consuete dinamiche della produzione artistica umana. D'altra parte è utile analizzare il cortometraggio anche da un punto di vista estetico ed emozionale, parametri centrali in ambito audiovisivo, utili per comprendere nel dettaglio come il lavoro prodotto si colloca rispetto a flussi più tradizionali. Proprio per questi motivi, la mediazione comunicativa assume un ruolo centrale: momenti introduttivi specifici, approfondimenti con l'autore in seguito alla visione o anche semplici materiali esplicativi potrebbero colmare il divario tra l'intenzione autoriale e l'interpretazione del pubblico. Certamente non risulta semplice organizzare queste attività, sopratutto nel contesto in cui si è svolta la presentazione del cortometraggio ma, in linea generale, una maggiore attenzione a questi elementi avrebbe probabilmente favorito una ricezione del prodotto più adeguata, soprattutto in un contesto in cui lo scarto tra la forma visiva e il processo creativo è così altamente marcato.
Va ricordata, infine, la natura stessa dell'output generato: trattandosi di video, il prodotto tende a essere considerato come opera compiuta, da valutare secondo criteri estetici ed emozionali classici. Tuttavia, la funzione principale del progetto non era quella di proporre un contenuto da inserire nel normale panorama della distribuzione audiovisiva, bensì di offrire un terreno di riflessione sulle possibilità e i limiti della collaborazione tra uomo e macchina. In quest'ottica, il cortometraggio non deve essere considerato fine a se stesso, ma un punto di partenza per ragionare sulle possibilità future della creatività mediata da IA.
Nel corso dell'intero processo di sviluppo del cortometraggio il rapporto tra autore umano e intelligenza artificiale ha costituito il fulcro metodologico e di indagine

del progetto. Questo dialogo continuo ha prodotto risultati positivi ma ha saputo evidenziare anche alcune problematiche, permettendo un'analisi concreta dei meccanismi di co-creazione. Tra gli aspetti più riusciti si evidenzia la capacità dell'IA di generare rapidamente molteplici versioni di uno stesso input, come accaduto nella fase di stesura della sceneggiatura, dove i suggerimenti forniti dai modelli hanno permesso all'autore umano di esplorare varianti narrative differenti in tempi contenuti, sempre mantenendo coerenza con il lavoro precedentemente svolto. L'IA, infatti, è sempre stata in grado di leggere e consultare i documenti di pre-produzione prodotti durante le prime fasi, consentendo uno scambio di idee coerenti con quanto deciso e scritto in precedenza. Questi aspetti hanno facilitato un processo creativo iterativo e altamente produttivo, rendendo l'IA una valida alternativa agli sviluppi classici di un prodotto audiovisivo, pur presentando le dovute limitazioni. Analogamente, l'impiego dell'IA nella fase di produzione, ha consentito di ottenere risultati stilisticamente coerenti, accelerando una fase solitamente lunga e complessa, che richiede un ampio spettro di conoscenze differenti, ottenibili solamente con l'aiuto di specialisti del settore: operatori di camera, addetti alle luci, costumisti e molti altri ruoli dell'industria cinematografica sono stati integrati in un unico strumento, che ha dialogato con l'autore umano per generare il prodotto audiovisivo finale. Tutti questi elementi dimostrano come l'interazione uomo-macchina possa realmente ampliare le possibilità espressive e organizzative in contesti produttivi, soprattutto nel caso di risorse insufficienti, permettendo un controllo efficace pur restando all'interno di un contesto altamente automatizzato.
Sono ovviamente presenti però dei limiti strutturali nella relazione con l'intelligenza artificiale, in particolare, l'impossibilità da parte dell'IA di operare e produrre una sintesi concettuale complessa. Nella costruzione narrativa, ad esempio, l'IA ha fortemente privilegiato strutture logiche lineari, spesso didascaliche, rendendo difficile introdurre elementi aperti all'interpretazione o sottotesti non espliciti. Allo stesso modo, la generazione visiva e quella sonora hanno mostrato rigidità nella gestione del tono e delle sfumature emotive, faticando a produrre soluzioni visive realmente sofisticate o capaci di adattarsi fluidamente al contesto narrativo. Un altro aspetto critico riguarda la gestione delle modifiche; l'IA, in molti casi, non è in grado di intervenire puntualmente sugli elementi già generati, rendendo di conseguenza necessari numerosi tentativi per ottenere minime variazioni. La richiesta di piccole modifiche risulta difficilmente applicabile, con l'IA che non è in grado di apportare cambiamenti senza stravolgere completamente quanto precedentemente prodotto. Questo problema rallenta il flusso di lavoro, riducendo la precisione del controllo autoriale sul prodotto audiovisivo finale. Infine, la qualità dei contenuti generati non è sempre di alto livello, come dimostrato dalla necessità di scartare a priori quasi la metà delle clip generate. Mentre alcuni output erano pienamente soddisfacenti, altri richiedevano revisioni sostanziali o non risultavano affatto utilizzabili, rallentando la produzione del cortometraggio.
Tutte queste considerazioni portano ad affermare con certezza che il ruolo autoriale umano è ancora imprescindibile, sia come filtro critico e analitico, sia come motore della creatività del progetto.
Con il progredire dell'esperienza operativa e l'affinamento delle capacità tecniche, il dialogo tra uomo e macchina potrà evolvere verso una forma di collaborazione sempre più integrata, delineando in maniera sempre più netta un nuovo paradigma di progettazione audiovisiva co-autoriale. In tale prospettiva, l'intelligenza artifi-

ciale non va considerata come uno strumento passivo o un esecutore tecnico, ma deve essere usata come un interlocutore progettuale, capace di incidere sulla buona riuscita del prodotto finale, lavorando sul linguaggio, sui tempi e sulle logiche della narrazione, seppur con dei limiti. Affinché questa interazione possa generare risultati significativi, è necessario che i prodotti di tali processi vengano inseriti entro un quadro che ne chiarisca tanto il metodo quanto il senso. Ogni progetto sperimentale vive infatti di un equilibrio delicato tra l'autonomia del risultato finale e la consapevolezza del percorso che lo ha generato: se da un lato le opere devono poter vivere di vita propria, confrontandosi con altre produzioni artistiche in modo indipendente, dall'altro è fondamentale fornire chiavi interpretative che ne esplicitino la genesi e le intenzioni. Solo così si può evitare una lettura superficiale e trasformare l'esperienza di visione in un momento critico, capace di valutare il ruolo dell'umano, dell'IA e della loro possibile coesistenza nei processi creativi del presente ma sopratutto in quelli del futuro.

4.3.5 Differenze tra aspettativa e ricezione reale

Nel corso della progettazione e realizzazione del lavoro, le aspettative iniziali in merito alla ricezione da parte del pubblico erano caratterizzate da un equilibrio ideale: si immaginava una divisione piuttosto simmetrica tra chi, magari per età o distanza dal contesto tecnologico, avrebbe trovato il prodotto difficile da comprendere e chi invece, dotato di una maggiore familiarità con le nuove tecnologie, ne avrebbe apprezzato la portata sperimentale. Si pensava quindi che una parte del pubblico potesse manifestare perplessità o distacco, mentre un'altra fosse maggiormente entusiasta e coinvolta. Tuttavia, la risposta effettiva è risultata più complessa e meno equilibrata del previsto, tendendo più verso la perplessità che verso l'entusiasmo, a dimostrazione del fatto che il contesto in cui si inserisce un prodotto audiovisivo è determinante per la sua comprensione e accettazione.
Molti dei presenti al concerto dove è stato presentato il cortometraggio non avevano partecipato alle fasi iniziali del progetto, di raccolta dati e impressioni, da cui il lavoro audiovisivo ha preso forma. Questo ha certamente generato una distanza maggiore: chi aveva fornito i dati avrebbe probabilmente sentito il cortometraggio come parte di un percorso condiviso, mentre il pubblico in sala, per la maggior parte estraneo al processo, si è trovato di fronte a un prodotto che non sentiva proprio. In un contesto in cui il contenuto viene generato a partire da tracce emotive e da elementi raccolti, la relazione con chi ha fornito quei dati è centrale nella restituzione del prodotto audiovisivo in quanto l'assenza di tale connessione potrebbe influire sul risultato finale. Va detto però che il cortometraggio deve poter esser fruito anche indipendentemente dal progetto da cui è nato e da questo punto di vista il pubblico in sala rappresenta fedelmente questo tipo di spettatore.
Un altro aspetto importante da considerare è la difficoltà nel confrontare un'opera generata dall'IA con un prodotto cinematografico tradizionale. Questo confronto è di base sbagliato, in quanto ad oggi ancora non possiamo affermare che i modelli siano allo stesso livello delle produzione cinematografica classica. I parametri di valutazione, infatti, non possono essere gli stessi, così come non lo sono le condizioni produttive e le finalità. L'intento del progetto non era quello di produrre un cortometraggio secondo i canoni classici (almeno non completamente), ma di esplorare le possibilità offerte da una nuova modalità di creazione audiovisiva, analizzandone i

punti di forza e di debolezza. Alcuni tra i presenti al concerto hanno compreso questo approccio e in diversi hanno espresso l'idea che tale progetto avrebbe avuto maggiore efficacia e sarebbe stato meglio compreso se accompagnato da una presentazione più approfondita del contesto e degli obiettivi. In effetti, una contestualizzazione più precisa avrebbe probabilmente aiutato il pubblico a orientarsi meglio nella visione e a interpretare correttamente il cortometraggio.
L'esperienza legata alla realizzazione e presentazione del prodotto audiovisivo si è rivelata, nel complesso, profondamente significativa e ricca di stimoli, soprattutto in termini di riflessione sull'attuale rapporto tra la produzione artistica e l'intelligenza artificiale. Inizialmente si poteva pensare di sfruttare questa ricerca come una semplice dimostrazione tecnica ma si è invece voluto trasformarla fin da subito in un vero e proprio esperimento critico, capace di generare interrogativi, confronti e prese di posizione chiare. L'interesse personale che ha animato ogni fase del lavoro, dalla sperimentazione dei vari strumenti di IA alla costruzione narrativa, è nato proprio dal desiderio di esplorare questa nuova tecnologia e di confrontarsi con i limiti, anche tecnici, dei nuovi modelli generativi.
L'incontro diretto con il pubblico ha offerto spunti di grande rilevanza. Diverse persone, anche appartenenti a fasce d'età meno vicine al linguaggio digitale, si sono dimostrate curiose e desiderose di approfondire, fermandosi al termine dello spettacolo per porre domande specifiche sul lavoro svolto. Questo dimostra che non è tanto la generazione o il background culturale a determinare la capacità di accoglienza, quanto la disponibilità al confronto e all'ascolto. In questo senso, il dialogo successivo alla proiezione è diventato parte integrante del progetto stesso, contribuendo a definire l'importanza non solo come prodotto artistico ma anche come oggetto di riflessione collettiva.
Molte persone hanno dimostrato e manifestato curiosità e apertura; alcuni spettatori hanno posto domande tecniche e metodologiche, cercando di capire quali fossero state le maggiori difficoltà incontrate e complimentandosi per il lavoro svolto. Altri si sono interessati invece agli strumenti utilizzati, dimostrando in primo luogo una buona conoscenza della tecnologia ma manifestando allo stesso tempo il desiderio di comprendere più a fondo i meccanismi della generazione automatica. Questo ha generato un confronto stimolante, in cui è stato possibile chiarire che si trattava di un progetto sperimentale, pensato non per offrire un prodotto perfetto ma per testare i limiti e le potenzialità della relazione tra essere umano e macchina. È stato molto interessante ed è sicuramente importante continuare a dialogare in questo senso perché certamente l'IA non è una tecnologia che interessa a tutti ed è un argomento che divide molto, ma è fondamentale capire meglio, tutti assieme, in che direzione andare e tutto quello che ci aspetta in futuro, consci del fatto che ci sono molte persone interessate e disposte a dialogare per comprendere tutto ciò che ruota attorno all'intelligenza artificiale.

4.4 Considerazioni e possibili sviluppi

4.4.1 Miglioramenti possibili della pipeline con IA

Alla luce dell'esperienza svolta è possibile delineare alcune proposte concrete per migliorare la pipeline produttiva basata su intelligenza artificiale, considerando sia

i punti di forza che i limiti emersi nel corso del progetto. Tra gli aspetti più critici si evidenzia innanzitutto la gestione della coerenza visiva e narrativa: sebbene la generazione dei video sia stata rapida ed efficace, permangono discrepanze nella resa dei personaggi e delle ambientazioni tra una scena e l'altra, come ad esempio differenze nei tratti somatici o variazioni cromatiche negli ambienti. Un possibile sviluppo futuro potrebbe consistere nell'implementazione di moduli dedicati ad aumentare l'aderenza narrativa ed estetica, attraverso sistemi di referenza visiva interna, capaci di mantenere meglio un'alta coerenza su più clip senza dover rigenerare manualmente le scene. Alcuni modelli di IA stanno già lavorando su queste funzioni, con le prime versioni ancora non ottimizzate però per produrre sequenze video complesse e articolate. In futuro sarà certamente possibile lavorare su questi elementi per ottenere un prodotto audiovisivo finale migliore.
Un altro aspetto sicuramente da migliorare riguarda il processo di generazione del voiceover che necessita di ulteriore ottimizzazione, specialmente per quanto riguarda la lingua italiana. Gli attuali modelli, seppur utili, non riescono a trasmettere efficacemente le emozioni presenti negli audio, compromettendo la potenza espressiva di certe sequenze. Una proposta concreta potrebbe essere lo sviluppo o l'utilizzo di modelli TTS multilingue addestrati su dataset emotivamente marcati e bilanciati, in grado di adattarsi dinamicamente al contesto narrativo. Diverse piattaforme ci stanno già lavorando, migliorando la qualità di giorno in giorno. Un'attenta ricerca e sperimentazione tra i diversi modelli disponibili sarà necessaria in futuro per trovare l'opzione migliore. Alternativamente, si potrebbe provare ad adottare una soluzione ibrida, basata su campionamenti reali espressivi, combinati con la flessibilità dell'IA, che potrebbe a questo punto prendere spunto da quanto campionato realmente per produrre un audio artificiale più adatto al prodotto audiovisivo generato.
L'aspetto che sicuramente potrà essere maggiormente migliorato all'interno della pipeline di produzione riguarda il montaggio e, in generale, la fase di post-produzione, attualmente ancora quasi esclusivamente manuale. Tra le varie cose, si potrebbe sperimentare l'integrazione di modelli di IA combinati con strumenti di video editing automatico che possano suggerire montaggi coerenti sulla base della sceneggiatura fornita. Gli strumenti attuali provano a replicare qualcosa di simile, fallendo però nella maggior parte delle situazioni: la tecnologia infatti non è ancora pronta per questa tipologia di scelte, anche perché dovrebbe imparare a valutare attentamente i risultati prodotti e solo in un secondo momento, tramite un'analisi dettagliata, capire come editare l'intero materiale visivo fornito. Andrebbe inoltre potenziata la fase riguardante la correzione e gradazione del colore, in quanto con l'avanzare delle nuove tecnologie, nasceranno nuovi modelli di IA addestrati su dataset di riferimento in ambito cinematografico, che porteranno di conseguenza a una maggiore facilità nel garantire armonia cromatica tra le varie clip generate, mantenendo una forte coerenza all'interno dell'intera durata del video.
Un altro possibile sviluppo futuro riguarda l'interoperabilità tra gli strumenti utilizzati. Attualmente, la pipeline si basa su differenti piattaforme, scelte accuratamente in base alle migliori caratteristiche presentate da ognuna. Uno sviluppo estremamente utile sarebbe quello di creare un sistema unificato di generazione, capace di gestire in modo centralizzato i vari moduli della pipeline di produzione e la loro relativa revisione, garantendo maggiore uniformità e riducendo gli errori dovuti al passaggio tra differenti fasi produttive. Uno strumento utile in questo senso potrebbe essere quello relativo alla generazione per *nodi*, dove tutta la pipeline produttiva

può essere strutturata attraverso la connessione di più nodi all'interno di una rete, ciascuno dei quali genera in particolare un aspetto del prodotto audiovisivo finale, tenendo conto degli input indicati, che possono essere a loro volta output generati da altri nodi oppure immagini, testi e indicazioni fornite dall'autore umano. Si lascia di seguito un esempio di progetto sviluppato con una piattaforma che permette di lavorare con la generazione per *nodi*.

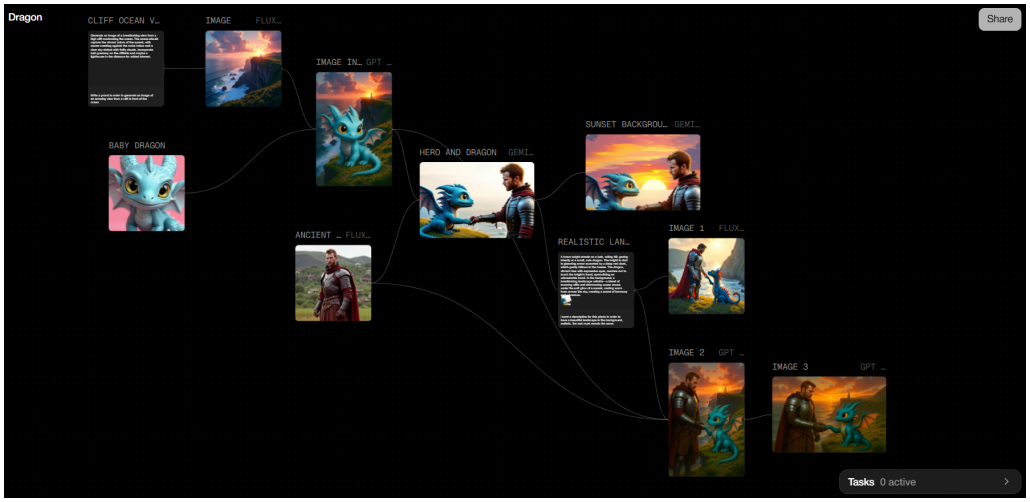


Figura 4.19: Un interessante sviluppo è rappresentato dalla generazione attraverso nodi, dove ogni singolo elemento della rete può essere generato usando modelli di IA differenti, utilizzando input provenienti da altri nodi. In figura è rappresentato un progetto sviluppato in fase di ricerca sulla piattaforma *Flora Fauna AI*.

Infine, rimane centrale il ruolo dell'autore umano a tutela dell'intenzionalità, della coerenza narrativa e della qualità estetica. Con l'evoluzione dei modelli multimodali e la crescente accessibilità di strumenti avanzati, è auspicabile lo sviluppo di interfacce sempre più co-autoriali, capaci di proporre alternative creative piuttosto che semplici esecuzioni di prompt e favorendo un'interazione più dialettica tra autore e macchina.
In conclusione, la pipeline sviluppata si è rivelata solida e ha portato a un risultato finale soddisfacente, pienamente replicabile anche in futuro per altri progetti. Ci sono ovviamente ampi margini di ottimizzazione, che vanno di pari passo con l'evoluzione della tecnologia negli anni. Gli spunti emersi da questa esperienza rappresentano dei punti di partenza per ridefinire e potenziare un modello produttivo ibrido sempre più raffinato, preciso e capace di dialogare con l'evoluzione tecnologica e culturale del linguaggio audiovisivo contemporaneo.

4.4.2 Potenzialità dell'IA per la produzione audiovisiva

L'impiego dell'intelligenza artificiale nella produzione audiovisiva mette in luce una serie di potenzialità che stanno già modificando in parte le pipeline di produzione, i ruoli professionali e i linguaggi estetici del settore audiovisivo. Una tra le più importanti è la possibilità di accelerare e semplificare molte fasi del processo produttivo, rendendo meno complicate operazioni complesse e permettendo una maggiore rapidità di visualizzazione del prodotto audiovisivo. Ad esempio, strumenti T2V come

Sora permettono di generare clip visive partendo da semplici descrizioni testuali, offrendo una base visiva utile sia in fase di pre-produzione che in quella di produzione. Se da un punto di vista tecnico si riescono a generare video, è chiaramente possibile generare anche immagini per definire personaggi, ambientazioni e farsi un'idea del prodotto audiovisivo già in fase iniziale.
In futuro saranno sempre più gli strumenti come *Runway* o *Kling* a permettere di gestire dinamiche di camera, illuminazione e ritmo narrativo attraverso interfacce intuitive, rendendo la regia un'interazione diretta con il linguaggio naturale. Ma non ci si limita al visivo; in ambito sonoro infatti, strumenti come *Suno* consentono la composizione di colonne sonore coerenti con lo stile narrativo, mentre piattaforme come *ElevenLabs* permettono di generare voiceover credibili in diverse lingue, contribuendo a una maggiore semplicità di distribuzione dei contenuti. Questo aspetto è importante in quanto permetterà, grazie all'IA, di condividere con un pubblico molto più ampio i contenuti prodotti, rendendo più accessibile a tutti la fruizione di prodotti audiovisivi. Un altro aspetto chiave è la capacità dell'IA di supportare la fase ideativa: i modelli generativi possono aiutare nella scrittura di sceneggiature, formulando logline e sinossi, non solo completando frasi ma anche suggerendo strutture narrative alternative. Da questo punto di vista, sebbene l'IA possa offrire molti spunti interessanti, c'è il rischio di incorrere in una uniformazione generale dei contenuti. Visto il funzionamento tecnico dei modelli, l'intelligenza artificiale non è in grado di generare contenuti e idee nuove, ma parte sempre da una riformulazione dei dati di addestramento che non contengono niente altro che dati prodotti nella storia dall'essere umano. La naturale conseguenza è quindi quella di riprodurre contenuti simili a quanto proposto in passato, senza dare un contributo innovativo al panorama cinematografico.
Un'altra grande potenzialità dell'applicazione dell'IA all'interno dei flussi di lavoro è la possibilità di effettuare iterazioni rapide, modificando elementi specifici su richiesta dell'autore e facilitando un processo creativo flessibile. Questo approccio può favorire in particolare la sperimentazione, l'ibridazione dei generi e la costruzione di estetiche non convenzionali, in maniera semplice ed economica: la possibilità di produrre rapidamente diverse versioni dello stesso contenuto audiovisivo potrebbe portare rapidamente ad escludere oppure scegliere nuove sperimentazioni che altrimenti, per motivi economici o di tempo, non si potrebbero esplorare.
In prospettiva, queste potenzialità lasciano intravedere un futuro in cui l'IA non sostituisce ma amplia la capacità espressiva dell'essere umano, trasformando il linguaggio audiovisivo in un ambiente sempre più accessibile e interattivo. Ciò che rende particolarmente interessante questa fase storica è proprio la grande quantità di sperimentazione e ricerca che l'IA consente di effettuare. Non sono presenti infatti modelli generativi perfetti né tanto meno si riescono ad ottenere output di livello professionale. È possibile però esplorare i diversi strumenti da un punto di vista tecnico e creativo, andando ad individuare e perfezionare delle pipeline operative che integrano parzialmente l'IA al loro interno. Ad esempio, ora è possibile generare in pochi minuti diverse varianti di uno stesso video, confrontarle, selezionare quella più adatta e continuare la produzione. Questo processo, se accompagnato da una visione autoriale solida, può tradursi in un perfezionamento della pipeline di produzione, fino a raggiungere una forma altamente sofisticata.
Allo stesso tempo, l'accesso sicuramente più ampio agli strumenti di produzione porta con sé nuove possibilità anche per artisti indipendenti e sperimentatori, che

possono finalmente dimenticarsi dei vincoli produttivi tradizionali e portare avanti progetti personali, utilizzando strumenti facilmente accessibili. Questa democratizzazione degli strumenti creativi consente di sviluppare linguaggi alternativi e narrazioni completamente innovative, svincolate dai canoni industriali e commerciali che spesso determinano scelte stilistiche e contenutistiche. La possibilità di lavorare in completa autonomia, combinando modelli testuali, visivi e sonori, offre alti margini di libertà espressiva, in cui anche gli errori di generazione possono diventare a volte parte integrante della grammatica creativa.
Questa situazione apre scenari di grande interesse per la ricerca artistica e accademica, dove l'interazione tra macchina e autore umano può generare nuove forme ibride, capaci di interrogare il presente, esplorando nuove forme di co-creazione artistica.

4.4.3 Implicazioni etiche e artistiche dell'uso di IA nel cinema

Dal punto di vista estetico e creativo, l'IA consente di esplorare modalità completamente nuove per la narrazione e la messa in scena, aprendo a linguaggi visivi fino a poco tempo fa irrealizzabili. L'utilizzo di modelli generativi, ad esempio, permette di costruire mondi immaginari attraverso semplici descrizioni testuali e, allo stesso tempo, sistemi di clonazione vocale consentono di creare dialoghi con voci sintetiche fedelmente modellate su voci esistenti o create da zero. Queste tecnologie rendono possibile un cinema sperimentale, abbattendo le barriere economiche e tecniche che limitano l'accesso alla produzione audiovisiva a molteplici artisti. Rimane comunque da sottolineare il fatto che l'utilizzo di questo nuovo approccio richiede in ogni caso alte competenze tecniche e un costante monitoraggio da parte dell'autore umano.
Le nuove alternative offerte dall'IA comportano anche implicazioni etiche significative. Una delle questioni più urgenti riguarda la tutela del lavoro creativo del passato: la capacità dell'IA di generare immagini, voci o testi passa attraverso modelli che sono stati addestrati su dati prodotti da autori umani, ponendo seri interrogativi sul diritto d'autore, sull'attribuzione e sulla responsabilità creativa nell'utilizzo dell'IA. Un esempio emblematico è quello di attori la cui immagine vocale o corporea viene riprodotta artificialmente senza consenso o di registi il cui stile viene replicato in altre opere senza autorizzazione. Questo apre uno scenario delicato, in cui l'identità artistica rischia di essere svuotata del suo significato, trasformandola solamente in un dato manipolabile. Un altro nodo critico riguarda la trasparenza nei processi creativi: il pubblico ha infatti il diritto di sapere se un'opera è stata generata con l'ausilio dell'IA e in che misura. In assenza di norme chiare a livello internazionale, si rischia di alimentare un clima di diffidenza, in cui il confine tra autentico e artificiale diventa sempre meno evidente. Chiaramente questo non significa dover limitare l'utilizzo dell'IA ma al contrario serve rafforzare una cultura più ampia e aperta, che sappia distinguere tra quella che è una sperimentazione consapevole e ciò che invece vuole essere un utilizzo disinvolto delle tecnologie. Per questi e altri motivi è auspicabile la definizione di linee guida condivise che tutelino i diritti degli autori, proteggano l'integrità delle opere ma, allo stesso tempo, lascino spazio alla sperimentazione contemporanea. Occorre dunque costruire un sistema di norme che non blocchi l'innovazione, ma che permetta di svilupparla con attenzione e rispetto del passato.
In conclusione, l'intelligenza artificiale sarà inevitabilmente uno degli strumenti centrali nella produzione audiovisiva del futuro. Per questa ragione, è fondamentale ini-

ziare fin da ora un processo di ricerca, sperimentazione, confronto e studio, in modo tale da essere pronti alle sfide che l'evoluzione tecnologica continuerà a proporci.

Appendice

A Domande poste al pubblico

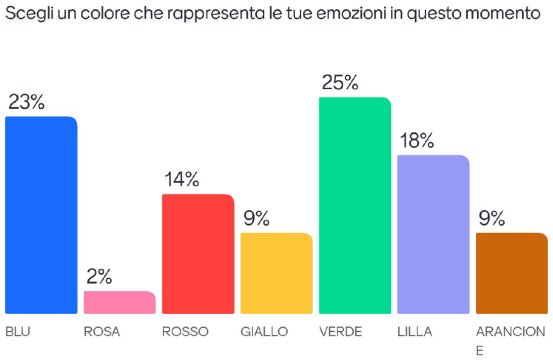
Di seguito le domande poste durante l'evento *Una passeggiata creativa.*

1. *Scegli un colore che rappresenta le tue emozioni in questo momento.*
2. *Nuvola di emozioni 1 - Scrivi una, due o tre parole per descrivere le emozioni e le sensazioni che l'ascolto ha stimolato.*
3. *PENSIERI E IMMAGINAZIONE - Quali pensieri sono emersi durante l'ascolto. Quali immagini hai visualizzato?*
4. *RICORDI - Quale esperienza di vita passata ti ha ricordato l'ascolto? Cosa hai visualizzato?*
5. *Nuvola di emozioni 2 - Scrivi una o più parole per descrivere cosa senti fisicamente in termini di benessere o malessere.*
6. *DIFFERENZE - Hai notato, prima e dopo l'ascolto, delle differenze nel tuo umore e stato d'animo?*

Di seguito, invece, le domande poste durante l'evento *Song Writing: intelligenza naturale generativa.*

1. *Scrivi 1, 2 o 3 parole per descrivere emozioni, immagini e pensieri che la poesia ti ha suscitato.*
2. *Il tema del VIAGGIO che cosa ti riporta alla memoria, quali ricordi, desideri, film, canzoni, persone? Scrivi una o più frasi.*
3. *Condividi la prima parola che ti viene in mente su come ti senti ora.*
4. *Se dovessi rappresentare la canzone con una stagione, quale sarebbe?*
5. *Se la canzone fosse un viaggio, verso quale destinazione ti porterebbe?*
6. *Quale COLORE assoceresti al tema trattato oggi?*
7. *Ti piacerebbe ascoltare la canzone nel video creato con Intelligenza Artificiale o preferiresti sentire altro (e cosa)?*
8. *Il tuo contributo servirà per creare un video con l'Intelligenza Artificiale, quale stile di VIDEO ti piacerebbe vedere?*

B Grafici dei dati raccolti



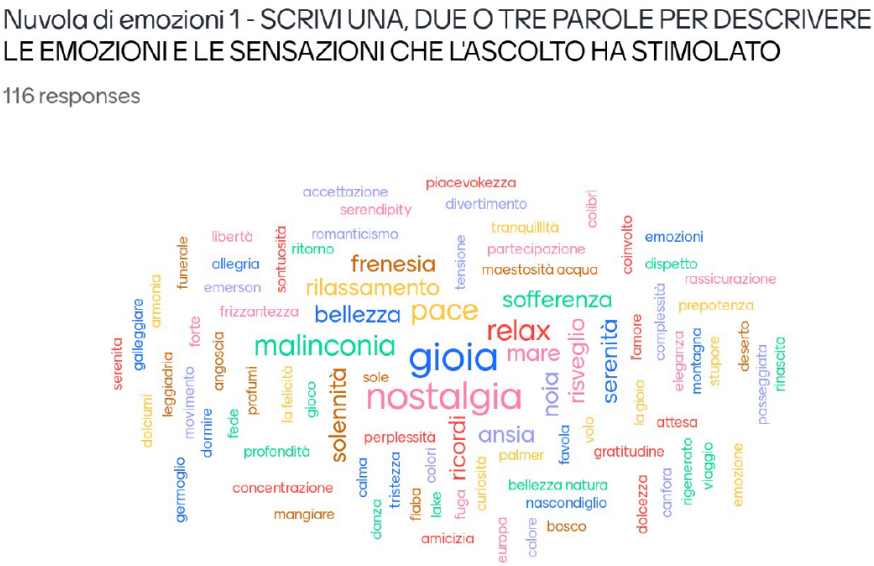


Figura 1: Grafici relativi alle domande poste durante l'evento *Una passeggiata creativa*.

Figura 2: Prima parte di grafici relativi alle domande poste durante l'evento *Song Writing: intelligenza naturale generativa.*

Figura 3: Seconda parte di grafici relativi alle domande poste durante l'evento *Song Writing: intelligenza naturale generativa.*

C Dati elaborati

Dalla pagina seguente sono mostrate tutte le tabelle contenenti i dati organizzati e rielaborati per essere resi facilmente disponibili anche ai modelli di intelligenza artificiale.

Alla fine del progetto verrà creato un video con l'Intelligenza Artificiale, che prenderà ispirazione anche dal tuo contributo e che verrà proiettato durante i concerti di musica contemporanea previsti per giugno 2025. Quale stile di VIDEO ti piacerebbe vedere?
Documentario, anche musicale o astratto
Genere non molto movimentato con risvolti naturalistico
.
Video con immagini reali non astratte
Tutto
Animato
Docureale
Un cortometraggio che racconti le emozioni e la loro evoluzione attraverso riferimenti a fatti/opere realist
Video che simula un'esperienza 3D, con tante immagini che si susseguono.
Moderno
Fantascienza
Narrativo
elementi astratti
Esplosione
Moderno

Scegli un colore che rappresenta le tue emozioni in questo momento	
BLU	23%
ROSA	2%
ROSSO	14%
GIALLO	9%
VERDE	25%
LILLA	18%
ARANCIONE	9%

Se dovessi rappresentare la canzone con una stagione, quale sarebbe?	
ESTATE	11
PRIMAVERA	4
AUTUNNO	3
INVERNO	1

Quale COLORE assoceresti al tema trattato oggi?	
Parola	**Conteggio**
Blu	10
Azzurro	9
Turchese	4
Bianco	4
Giallo	4
Arancio	3
Celeste	3
Verde	3
Rosa	3
Verde corallino	2
Arancione	2
Ghiaccio	1

Analisi	
Positivo	14
Negativo	2
Altro	2

Scrivi una, due o tre parole per descrivere le emozioni e le sensazioni che l'ascolto ha stimolato	
Parola	**Conteggio**
Gioia	5
Nostalgia	5
Bellezza	3
Malinconia	3
Pace	3
Relax	3
Ansia	2
Frenesia	2
Mare	2
Noia	2
Ricordi	2
Rilassamento	2
Risveglio	2
Serenità	2
Solennità	2
Accettazione	1
Acqua	1
Allegria	1
Amicizia	1
Amore	1
Angoscia	1
Armonia	1
Attesa	1
Bosco	1
Calma	1
Calore	1
Canfora	1
Colibrì	1
Colori	1
Complessità	1
Concentrazione	1
Curiosità	1
Danza	1
Deserto	1
Dispetto	1
Dolcezza	1
Dolciumi	1
Dormire	1
Emerson	1
Emozione	1
Emozioni	1
Europa	1
Fede	1
Favola	1
Felicità	1
Fiaba	1
Forte	1
Frizzantezza	1
Funerale	1
Fuga	1
Galleggiare	1
Germoglio	1
Gioco	1
Gratitudine	1
Lago	1
Leggiadria	1
Libertà	1
Maestosità	1
Mangiare	1
Montagna	1
Movimento	1
Nascondiglio	1
Natura	1
Palmer	1
Partecipazione	1
Passeggiata	1
Perplessità	1
Piacevolezza	1
Prepotenza	1
Profondità	1
Profumi	1
Rassicurazione	1
Rinascita	1
Ritorno	1
Romanticismo	1
Serendipity	1
Sole	1
Sofferenza	1
Sontuosità	1
Stupore	1
Tensione	1
Tranquillità	1
Tristezza	1
Viaggio	1
Volo	1
Volo	1
Viaggio	1

Quali pensieri sono emersi durante l'ascolto. Quali immagini hai visualizzato?
Foresta
Infanzia
Gruppo di danza nella mia giovane esperienza universitaria
Mongolfiera
Cascata Uccellini Burattino
Paesaggio naturale
La strega e il suo rumoroso corteo
Sincronia in processione
Trincea di una guerra poco prima dell'attacco
Viaggio
C’era una volta in un castello…
Fende il mare la nave che mi porta via
Una passeggiata in montagna
Un bosco
Carovana di cammelli nel deserto
Cartoni animati
Una passeggiata attraverso i suoni e le sensazioni inserite nei diversi quadri
Momenti belli in situazioni tristi, ero davanti alla mia finestra mentre nevicava, d’inverno, ed ero bambinc
La corteccia di un albero buio
Vita in una foresta in diversi momenti del tempo
Un amore passato, intenso ma triste, che non tornerà mai più
Cielo
Danza su un tappeto orientale
Suoni della naturaCinguettio e battito di ali in un prato fiorito
Passeggiata in un bosco fantastico e poi per le colline toscane
Luogo fiabesco
Russia
Il nonno sulla sua sedia sa di canfora e tabacco
Dame e cavalieri in balli antichi, foglie che volteggiano e parlano tra loro
Una foresta di betulle con ruscello e animali
mendicando denaro per strada e l'indifferenza dei passanti che voltano la schiena steppa russa, danzatori dai costumi tradizionali dai colori accesi danzano sulla neve
Dialogo tra figlio e mamma che lo rincuora. Cerimonia religiosa importante. Annuncio primaverile import
Concerto in luogo aperto
Estate
Natura
Passeggiata in montagna,spiccare un volo, viaggio in città
Gnomo cattivoPasseggiata nel boscoStrega cattivaCarro nel fango
Ogni quadro ha fatto emergere una scena diversa dall'altra tra castelli medievali, navi in partenza, giudi
Una strada acciottolata di paese
Una passeggiata in montagna, tra ruscelli e prati.
Gnomi diversi tra loro, con vestiti dai colori vivi, che scendono da un pendio sbucando da un bosco dai
Ho immaginato di entrare in una casa piena di stanze. Ogni stanza ha una musica diversa.
Una piazza, a Budapest, alla fine di un viaggio solitario. E una salita, in mezzo al bosco, in compagnia
Un fratello dispettoso
Cielo e mare
Storia/Film narrante sofferenza, resistenza, rinascita e fioritura

La toeletta delle piccole vanità
Sentiero
Immagine di danze , movimenti armoniosi , paesaggio fiabesco . In alcuni momenti sensazione di fede , preghiera e conforto .
La dolcezza di una rinascita
Carrellata di sensazioni
Galleggiare su un materassino in mezzo al mare
Un'anatra, il suo stagno, uccelli felici che volano al tramonto di un giorno autunnale
Persone che corrono in una foresta
La linea
La musica è un linguaggio universale
Inquietudine nel bel mezzo di una tormenta
Qualcuno si aggira nello stagno
Pianoforte
Rivoli d'acqua
La Corazzata Potemkin
Un sottobosco in cui filtra luce
Vecchio lunapark pieno di luci
Bosco dai colori autunnali
Un torrone alle mandorle (o forse è cianuro?)
Biancaneve nel bosco Pantera rosa intenta a seguire qualcuno Scena di un corteo funebre dove vengono inquadrati una serie di personaggi che hanno avuto a che
Bosco
La leggerezza che vorrei sentire in mezzo al traffico cittadino
Il oane venduto tutti i giorni
Tempi passati dove la cortesia ed il vezzo la fanno da padrone
Città di notte piove strade lucide guizzare di topi vita notturna dei vicoli La mucca ha indossato il suo paio di stivali più belli per recarsi al mercato, ma non riesce a salire le
L'uccello che si canta sull 'alberoMi sento la tranquillità
Un pipistrello si apre nel petto
Ricordi di tempi passati remoti e tempi non molto lontani. Sentimento di dolcezza e appagamento.
Alba
Liberiamo i colombi!
Una nube di farina si alza dal sacco battuto
Orto
Creare contatto
Ape che si risveglia all'interno di un fiore che si schiude Uccelli tropicali grandi foglie verdi Un quadro di Rousseau fiori grandi felci sottobosco
Il fiume che si trova da qualche parte in montagna
Pensieri di serenità, immagini di fiabe
Funerale
L'ultimo brano mi ha ricondotto al mio luogo d'infanzia era una cascina con tanti animali da cortile e mi s
Un germoglio che si fa stada nel terreno fino a vedere la luce
Ricordi di momenti sereni insieme ad amici
Lungomare

Quale esperienza di vita passata ti ha ricordato l'ascolto? Cosa hai visualizzato?
Infanzia
Infanzia
Anni di studio della musica
Me da giovane
Uno scoiattolo corre e salta tra i rami Sembra sempre che possa cascare, i rami si piegano ma lui ha sempre successo
Estate
La prepotenza della tirannia ed il cano di libertà
Lo spettacolo di danza classica all'Alfieri
Il furgoncino dei brigidini alla fiera
L'infanzia. La morte di mio padre
Infanzia
Serate in collegio
Gli anni 70 dei Emerson Lake &Palmer
La vita spensierata da bambina
Quando mi sono perso in montagna da piccolo .
InfanziaPasseggiate nella natura
Balli a corte
La neve che cade davanti alla finestra
Ascolto musicale in età giovanile
Esperienza da musicista in orchestra
Conservatorio
Vacanza e spensieratezza
Infanzia
Sentiero di Santiago
Scuola materna, giocando con altri bambini
Tempo sospeso
GiovinezzaIl prossimo futuro che prevedo in relax
Ho ricordato alcune musiche di film che narravani ks ca cia alla volpe
Nessun ricordo
Giochi con amici nei prati durante i pic nic
L'ascolto di fiabe la sera prima di andare a dormire
I cartoni animati guardati con affianco la nonna
Quando ero giovane e guardavo i cartoni della Disney con i miei figli piccoli Ricordi di questa estate e dei giochi con mio nipote
Giochi per strada
Bambine danzanti su un tappeto multicolore incrociano sguardi, esperienze, camminano insieme verso i
Mi ricordo quando sono piccola una volta sono andata con la mia famiglia a visitare la mia nonna
Scolastica, spensieratezza
Una casa in mezzo al nulla, un uomo assente, la solitudine e poi… la scelta, la fuga, la rinascita
Quando tornerò?
L'infanzia e le nevicate in montagna
Colonna sonora di cartoni dell' infanzia
Ascolto musicale in età giovanile
La narrazione di una fiaba
Altri concerti da camera con amici
Un pomeriggio in un caffè storico di Torino con una delle persone che mi ha amato di più in tutta la mia

L'adolescenza con i suoi alti e bassi
Adolescenza :esperienza di ascolto musica dal vivo
Ho visualizzato una sala museo di scienze naturali. La musica mi ha fatto sentire a casa, vicino alla radio accesa.
Le estati dell'infanzia
Io che gioco nelle strade del mio paese mentre una fitta nebbia mi nasconde agli amici chi altrui e mi fa
Cielo .
Passeggiata in un viale alberato
L’odore buono delle poltrone di velluto dei teatri, la sera
Cielo e mare
Le scorribande giovanili
Camminate lungo lago in età adolescente
Mia nonna, le sue mani, il suo profumo d'amore
Sette cieli .
Estate vacanza
Creatività : dalla musica ai sensi
Il mio lavoro in teatro, dietro le quinte
Vita presente per periodo con meno impegno
Ho visualizzato prati in fiore e uccellini in volo
Quando sono andata con la mia famiglia a nuotare al mare ero contenta
Una passeggiata

Il tuo contributo servirà per creare un video con l’Intelligenza Artificiale, quale stile di VIDEO ti piacerebbe vedere?
Animazione
Illustrazione
Cartone animato
Natura
Ironico
Romantico
Video con colori chiari/pastello
Pubblicita'
Animazione con elementi astratti
Docufilm
Natura
Fantastico/ astratto
Sperimentale, astratto, teatro fisico, cyberg
Paesaggi, natura, persone, sguardi
Contemporaneo
Animazione
Comunque recitato non documentario
Animato
Divertente
Musicale
Paesaggistico con colori neon

Scrivi una o più parole per descrivere cosa senti fisicamente in termini di benessere o malessere.	
Parola	**Conteggio**
Serenità	6
Rilassamento	5
Relax	4
Caldo	3
Leggerezza	3
Pace	3
Sorriso	3
Ansia	2
Benessere	2
Emozione	2
Felice	2
Pacatezza	2
Piacevolezza	2
Rilassata	2
Solennità	2
Spensieratezza	2
Trasportata dal suono	2
Agitazione	1
Allegria	1
Appagamento	1
Calma	1
Chiaro	1
Connessione	1
Corpo teso	1
Crativita'	1
Disagio	1
Disarmonia	1
Distrazione	1
Fastidio uditivo	1
Felicità	1
Formicolii	1
Galleggiamento	1
Gioia	1
Gioia infantile	1
Guscio	1
Ispirazione	1
Inquietudine	1
Leggera	1
Liberato	1
Libero	1
Lucido	1
Mente rilassata	1
Molto attento	1
Molto teso	1
Nutrimento	1
Pelle d'oca	1
Presenza	1
Respiro profondo	1
Sfarfallio	1
Sonno	1
Sonnolenza	1
Sospensione	1
Tepore	1
Tensione	1
Time-out	1
Troppe emozioni	1
Trasparente	1
Tranquillità	1
Vibrazioni intense	1
Voglia di riscatto	1
Vulnerabilità	1
Vibrazioni intense	1
Voglia di riscatto	1

Hai notato, prima e dopo l'ascolto, delle differenze nel tuo stato d'animo?
No
Sono più serena
AffannoSerenità
Più ottimismo
Meno in tensione dopo l'ascolto
Più tranquilla
Stanchezza e desiderio di andare dormire
Più rilassato
Più presente
Più allegra
Più rilassato
Più sereno alla fine del concerto
No, è rimasto invariato
Tensione, caldo, ansia prima. Rilassamento, leggerezza, curiosità dopo.
Più sereno
Meno ansioso
Passato dallo stress alla serenità: attiva
Si è accesa una vena di creatività
Più rilassata
La musica mi ha aiutato a rilassarmi
Si, prima ero agitata ed ora sono rilassata
Prima curiosità Dopo : tanti ricordi
Più calma
Si, sensazione di pace
Più rilassato, più divertito, più affamato
Prima umore neutro (aspettativa) poi pesantezza.
Prima dell ascolto ero curiosa di sapere cosa sarebbe stata l esperienza. Dopo sono stata contenta di aver partecipato
Curiosità, benessere
Senso di rilassatezza e piacevolezza con stimoli sensoriali interessanti e stimolanti
Più sorridente
Stima per i musicisti!!
Maggiore consapevolezza
Rispetto a prima, maggiore apertura verso l'esterno
Più energico su come affrontare prossime situazioni nella vita
Più contento
Più centrata su di me
Mi sento più serena, e' stato poacevole
Più rilassato
Inizialmente tesa poi mi sono rilassata e mi sono fatta trasportare a immagini di vita regale
Energia fisica recuperata.Umore alleggerito.
Prima e durante irrequieta, poi rilassata
Attivazione di sensazioni poco frequenti
Si la mente si é concentrata sul suono
Prima di l'ascolto ho sentito rumori e dopo ho sentito la felicità e la tranquillità e la fantasia
Più rilassata

Sono arrivata trafelata e agitata per il traffico automobilistico trovato, ma dopo l'ascolto ho ritrovato un s
La musica classica dal vivo mi rilassa e migliora il mio umore
No
Nessuna differenza, stesso malessere di prima
Mi ha riconnesso con le mie sensazioni corporee mettendo distanza con le ansie della giornata
La felicità della concentrazione
Più tranquilla
Più serena
Il mio umore é sereno
No

Ti piacerebbe ascoltare la canzone nel video creato con Intelligenza Artificiale o preferiresti sentire altro (e cosa)?
Altre cose
Assolutamente la canzone!
La canzone!
Certo!
No
La canzone
La canzone!!!
La canzone
Quello che volete, basta che sia una musica ricca di ARMONIA
Si
Canzone creata con AI
Si, sono curioso
Sì. (Ma non posso sapere se è idoneo, perché dovrei vedere il video e valutare di conseguenza)
La canzone! :)
La canzone
Mi piacerebbe ascoltare ka canzone creata con IA
Certo
Certamente

Scrivi una, due, tre parole per descrivere emozioni, immagini e pensieri che la poesia ti ha	
Parola	**Conteggio**
Libertà	7
Viaggio	6
Leggerezza	4
Aria	3
Solitudine	3
Mare	2
Pace	2
Riflessione	2
Serenità	2
Speranza	2
Tranquillità	2
Vita	2
Acqua	1
Alberi	1
Alda Merini	1
Anima	1
Armonia	1
Bellezza	1
Calma	1
Country road	1
Desiderio	1
Divenire	1
Dolcemente	1
Fantasia	1
Fine	1
Fluire	1
Forza	1
Gioia	1
Interesse	1
Lasciare andare	1
Luce	1
Melanconia	1
Natura	1
Orizzonte	1
Precarietà	1
Quiete	1
Raggi del sole	1
Ricordo	1
Rimembranza	1
Senso di calma	1
Sensorialita	1
Sogno	1
Tempo presente	1
Un bel viaggio	1
Vaghezza	1

Condividi la prima parola che ti viene in mente su come ti senti ora.	
Parola	**Conteggio**
Gioia	3
Condivisione	2
Rilassato	2
Soddisfatto	2
Azzurro	1
Allegra	1
Allegria	1
Amore	1
Ballo matto	1
Bellezza	1
Bene	1
Cantautrice	1
Contenta	1
Condivisione	1
Creatività	1
Divertito	1
Emozionante	1
Emozionata	1
Emozionato	1
Euforia	1
Felice	1
Felicità	1
Giovane	1
Le parole diventano musi	1
Libertà	1
Mare	1
Naufragio	1
Partecipazione	1
Risata	1
Risate	1
Sorpresa	1
Sereno	1
Spensierato	1
Vita	1

Il tema del VIAGGIO che cosa ti riporta alla memoria, quali ricordi, desideri, film, canzoni, persone
La bellezza della scoperta
"I veri viaggiatori sono coloro che partono per partire" (Baudelaire)
Viaggio perchè sono libero, per esplorare e conoscere nuove persone, luoghi e culture.
l'Il non ricordo
L'ultimo giorno eravamo solo il mio zaino e io. Più di 14 ore di cammino tra lacrime, dolore, prati e cielo.
Il finestrino del treno si colora dei colori del paesaggio, colori che fuggono e corrono come me.
I miei viaggi in treno verso il Sud, cercavo soprattutto me stesso.
Viaggio, fuori e dentro di sè. Verso qualcosa, per scappare, per ritrovarsi, per tornare. Viaggiare è sempre trasformarsi.
The power of love
Le canzoni cantate con gli amici in macchina
Lontananza, speranza, futuro e impatto di nuove realtà anche belle
Viaggiare in solitaria in auto, con musica e pensieri che, si rincorrono e si intrecciano, un viaggio nel via
L'oceano che si infrange sulle rocce L'alba nel deserto La scoperta ed i profumi
Mantello, binario, abbraccio
Lunghe tratte in macchina in attesa di arrivare a destinazione...
Le pizze sulla spiaggia
Che il viaggio finisca per poterne intraprendere uno nuovo
"Sarai invincibile" disse la vita, "ma prima ti renderò forte" ♥
Bambine giocano nel camper, danzano, abbracciano il mondo dolcemente, curiose di vedere mari, esplo
Nasce il viaggio quando muore il sole, come il mare sale
I tempi di un'estate gioiosa e calorosa, familiare.
Un lungo viale, alberi alberi e in fondo un pallido sole. Odori, odori sconosciuti, puzza, profumo? Non so
Into The wild
La scoperta di nuove realtà, nuove persone e nuove culture
La mia bicicletta e il sole e il vento del Marocco.
Paesaggi che ti aprono il cuore e senti la tua gioia
Il viaggio può essere un percorso di formazione, di scoperta di sé oltre che del mondo. Ha luci e ombre,
Staccarsi dalla routine quotidiana per scoprire nuove situazioni, persone, colori, profumi. Il tempo assum
Lago di Atitlan, il sorriso antico dei Maya, un cielo di vulcani.
Camminando nel mattino della mia seconda vita trovo il tempo che è passato su una strada che ho perc
Appunti disordinati di viaggio... 2006 Cuba ...e la più buona aragosta. E birra della vita...
Ero pronto per viaggiare ma mi sono perso, non ricordo la destinazione. Lista fatta , giacca , sacco a pelo , torcia. Fari nella notte. Zaino in spalle, bussola in mano, ma dove andare?
Viaggiare con una playlist degli Afterhours
2020...primo giorno di lock down... Gatto che sale nel mio camper e non va più via ..
Dove cielo tocca mare...
La nave parte , una terra si allontana, una terra si avvicina
Camminare piano, gustare il mondo

Elicriso,mirto,granito. Sardegna.
Gatto il gatto...
Viaggio per sentirmi libera e leggera ma, soprattutto, per ritrovare me stessa e lasciare indietro le preoc
Per la sola ragione del viaggio,viaggiare. Faber -cit
C'è chi ci va in nave, chi invece dentro un'auto A volte sopra un camper ed altre in bicicletta
Avanti avanti avanti si può spingere di più!
Continui la tua strada con lo spirito di un bambino per raggiungere la tua meta
Bambino in riva al mare la coda tira al gatto. Ti ho visto camminare, ma l'ombra si allontana
Cammino per ritrovarmi, appunti disordinati. Il vento c'è l'ho contro, ma non mi fa un torto. Non so se sto scappando, oppure sto tornando. Continuo la mia strada, la meta è annullata.
Nel sole della sera Ti ho visto camminare Ti voglio accompagnare Ma l'ombra tua scompare

Se la canzone fosse un viaggio, verso quale destinazione ti porterebbe?	
Parola	**Conteggio**
Mare	5
Maratea	2
Antartide	1
Barbados	1
Bari	1
Busan	1
Calabria	1
Calanques rocce rosse	1
Cancun	1
Capo verde	1
Creta	1
Felicità	1
Isola	1
Lago	1
Libertà	1
Oceano	1
Pace	1
Paradiso	1
Prato	1
Sardegna	1
Senerità	1
Sicilia	1
Tenerife	1
Una scogliera	1
Ventotene	1

D Logline e sinossi

Una giovane artista, ispirata dalla vitalità della città, si perde tra murales pulsanti di vita e opere che sembrano respirare, per poi lasciarsi chiamare dalla vastità incontaminata della natura, intraprendendo un viaggio che fonde l'arte urbana con il richiamo selvaggio dell'ambiente.

Una giovane artista di 35 anni, riconosciuta per le sue opere disseminate in tutta la città, sente un'inquietudine profonda nonostante il successo. La città la opprime, la sua arte sembra incastrata in un mondo artificiale che non le appartiene più. Ripensa alla sua infanzia trascorsa tra le montagne, immersa nella natura incontaminata. Il dilemma è chiaro: rimanere in città, dove il suo lavoro trova spazio e pubblico, oppure tornare nella natura, dove sente di appartenere davvero?

E Soggetto cinematografico

Nel laboratorio di città, un'artista di trentacinque anni si muove tra le sue tele e i suoi schizzi sparsi ovunque. Il disordine dello studio riflette il suo stato d'animo, irrequieto e combattuto. Da tempo avverte una crescente insoddisfazione, una sensazione di distacco dalla vita. Lavora, espone le sue opere, riceve apprezzamenti, ma le manca qualcosa.

Davanti a lei, una tela incompiuta cattura la sua attenzione. È un paesaggio non ancora definito, privo del suo elemento essenziale. Lo osserva a lungo, quasi ipnotizzata, come a cercare lì la risposta ai suoi dubbi esistenziali. Sa che quel quadro è importante, ma ancora non capisce il perché.

Sopraffatta dai pensieri, decide di uscire alla ricerca delle sue tracce in città. Alcuni suoi murales colorano i muri grigi, le sue sculture popolano le piazze, le sue installazioni interagiscono con il paesaggio urbano. Sono la sua firma, il suo segno in città. Arriva davanti a un murale, dipinto su una parete di mattoni. L'opera raffigura "una strada acciottolata di paese". Il disegno è fermo, immobile. L'artista si sofferma, lo osserva attentamente, poi riprende a camminare. Mentre cammina, qualcosa di straordinario accade: le opere iniziano a prendere vita.

Poco più avanti, nella piazza centrale, si erge una scultura moderna ispirata a "il fiume che si trova da qualche parte in montagna". L'acqua scolpita sembra quasi reale. L'artista si avvicina e, improvvisamente, la scultura prende vita: i solchi rigidi si sciolgono in un flusso d'acqua vero, che scorre e riflette la luce dei lampioni. L'artista si blocca, sorpresa. Per un istante resta in silenzio, poi si allontana, con un senso di meraviglia crescente.

In fondo a un vicolo, su un muro largo e consunto, c'è un'altra sua opera. È un murale che rappresenta "un bosco fantastico con creature misteriose". Ma mentre lo guarda, qualcosa cambia. Dai contorni del disegno iniziano a emergere radici reali, poi rami, poi alberi veri che si innalzano fino a coprire parte della strada. Il cemento si crepa, la città svanisce. L'artista si ritrova immersa nella sua creazione.

Attorno a lei, dai rami iniziano a sbocciare "fiori grandi e felci sottobosco", intrecciandosi al terreno che diventa improvvisamente liquido. Il suolo sotto i suoi piedi si scioglie in "rivoli d'acqua", e senza opporre resistenza, l'artista viene trasportata dalla corrente. Ora è in un paesaggio montano, circondata da ruscelli e neve. Raccoglie un po' di neve tra le mani, la osserva scivolare tra le dita, ma mentre cade, cambia forma: la neve si trasforma in "cielo e mare". L'acqua si espande intorno a lei, le onde si increspano dolcemente, il paesaggio muta ancora.

Cammina lungo la riva del mare, ma la sabbia cambia consistenza sotto i suoi passi. Le sue impronte affondano e, poco alla volta, la sabbia si trasforma in "spighe di grano", che ondeggiano leggere. L'artista allunga la mano e raccoglie una spiga. Tra le sue dita, si irrigidisce e si trasforma in una scultura, una struttura leggera che rappresenta "una danza tra foglie che volteggiano nel vento". Le foglie scolpite iniziano a distaccarsi e a fluttuare nell'aria. L'artista le segue con lo sguardo: si librano sempre più in alto fino a fondersi nel cielo, diventando parte di "un cielo pieno di mongolfiere".

Si lascia trasportare in alto, mentre le mongolfiere diventano sempre più grandi. Una di esse inizia a dissolversi, rivelando al suo interno un nuovo scenario: "un castello immerso in un paesaggio fiabesco". L'artista avanza verso l'ingresso, ma non appena supera il portone, il castello si deforma, le pareti mutano colore, le finestre si fondono, l'intero spazio cambia. L'architettura si trasforma in "un sogno astratto con forme geometriche che si muovono".

Le forme si sovrappongono e si distorcono, poi si ricompongono fino a delineare i profili di una città. L'artista si ritrova ora in una "città di notte, piove, con strade lucide", i neon pulsano, i riflessi dell'acqua brillano sui marciapiedi. Si guarda intorno, ma tutto inizia a dissolversi. Le immagini si sovrappongono, come se il tempo scorresse a una velocità irreale. Il paesaggio cambia ancora e l'artista si accorge di camminare attraverso "un quadro in cui si alternano le stagioni". Ogni passo la porta avanti in un ciclo continuo: inverno, primavera, estate, autunno.

Poi, all'improvviso, tutto svanisce.

Dopo questo viaggio onirico, improvvisamente si ritrova nella sua bottega, come se nulla fosse accaduto. Ma dentro di lei tutto è cambiato. Si volta e osserva di nuovo la tela incompiuta. Ora capisce. Quel paesaggio sfocato, quelle linee ancora incerte rappresentano la sua casa d'infanzia, il paesino di montagna dove è cresciuta. Senza esitazione, inizia a raccogliere le sue cose: scatoloni, tele, strumenti, tutto viene imballato. La decisione è presa: lascerà la città. Non perché rinnega la sua arte, ma perché ha capito che per continuare a creare deve ritrovare le sue radici.

L'ultima scena la mostra in un paesaggio completamente diverso: le montagne, il cielo limpido, la natura che la circonda. È seduta all'aperto, la tela davanti. Osserva per un attimo il paesaggio reale, poi abbassa lo sguardo sul quadro e comincia a dipingere. Con un ultimo colpo di pennello, l'opera incompiuta che ha scatenato tutto è finalmente completa e l'irrequietezza svanisce per sempre.

F Sceneggiatura

FADE IN:

EXT. CITTÀ - POMERIGGIO
Una VISTA AEREA della città, ripresa da un drone. I raggi del sole filtrano tra i palazzi, creando un contrasto di luci e ombre. Le strade pulsano di vita, la gente si muove tra le piazze e i vicoli.

EXT. STUDIO D'ARTE - POMERIGGIO
Una via affollata. L'insegna dello studio spicca tra le altre. Attraverso la vetrata si intravedono tele, pennelli e sculture incomplete. L'ARTISTA, 35 anni, assorta, osserva il traffico dalla finestra.

INT. STUDIO D'ARTE - POMERIGGIO
La stanza è un caos creativo. Tele sparse, schizzi su fogli accatastati. La luce naturale taglia l'aria con un bagliore polveroso.

Primo piano su una TARGA DI RICONOSCIMENTO appesa al muro: "Premio Giovani Artisti".

L'ARTISTA fruga nervosamente tra i suoi lavori, come cercando qualcosa. Il respiro è irregolare. Si blocca davanti a una TELA INCOMPIUTA.

Si ferma. La osserva. Qualcosa dentro di lei è irrisolto.

EXT. STRADE DELLA CITTÀ - POMERIGGIO
L'ARTISTA cammina senza una direzione precisa. I suoi occhi scivolano sulle opere urbane che la circondano: murales, sculture, installazioni. Ogni cosa sembra parlare al suo inconscio.

Si ferma di fronte a un MURALE raffigurante una strada di paese. La sua mano sfiora la pittura.

FLASHBACK: Una BAMBINA corre per le strade di un paese di montagna.

EXT. PIAZZA DELLA CITTÀ - POMERIGGIO
L'ARTISTA arriva in piazza, osserva una scultura, che improvvisamente prende vita, iniziando a far scorrere acqua. L'ARTISTA si guarda attorno sorpresa, ma nessuno sembra notare nulla di strano.

FLASHBACK: Una BAMBINA fissa un quadro per la prima volta, incantata.

EXT. VICOLI - POMERIGGIO
Si inoltra nei vicoli. Davanti a lei, un altro murale: un bosco incantato.
Primo piano sulle linee del dipinto che iniziano a muoversi.
Le radici dipinte si estendono dalla parete. La città svanisce.

EXT. BOSCO MAGICO - POMERIGGIO
L'ARTISTA si ritrova immersa in un bosco surreale. Alberi colossali, fiori pulsanti di luce, il suolo sembra liquido. Il terreno sotto i suoi piedi inizia a scivolare.

EXT. ACQUA SOTTERRANEA
Un fiume nascosto la trascina dolcemente lungo il suo corso.

EXT. PAESAGGIO MONTANO - GIORNO
La corrente la lascia su un prato innevato. Si inginocchia, raccoglie la neve tra le mani. La neve si trasforma in acqua.

EXT. SPIAGGIA - TRAMONTO
L'acqua si espande fino a diventare un oceano. La riva si trasforma in un campo di grano dorato che ondeggia nel vento.

Primo piano su una spiga tra le dita dell'ARTISTA. Si irrigidisce, trasformandosi in una SCULTURA DI FOGLIE.
Le foglie scolpite si staccano, levitando verso il cielo.

EXT. CIELO - GIORNO
Le foglie diventano MONGOLFIERE e si allontanano nel cielo. Nel mentre L'ARTISTA osserva la scena da una collina. Spunta un castello incantano in lontananza.

EXT. CASTELLO FIABESCO - GIORNO
Un castello appare tra le montagne. L'ARTISTA si avvicina, incuriosita, fino a entrarci.

INT. CASTELLO - TRASFORMAZIONE
Le pareti mutano forma e colore. Le finestre si dissolvono in un vortice geometrico. La realtà si decompone.

EXT. CITTÀ NOTTURNA - PIOGGIA
Dalle geometrie emerge una metropoli notturna. Neon lampeggiano, l'acqua piovana riflette luci distorte.

EXT. QUADRO DELLE STAGIONI
L'ARTISTA cammina all'interno di un dipinto che cambia stagioni a ogni passo. Inverno. Primavera. Estate. Autunno.
Poi tutto svanisce.

INT. STUDIO D'ARTE - TRAMONTO
L'ARTISTA si ritrova nel suo studio. Il tempo sembra non essere passato, ma dentro di lei tutto è cambiato.

Primo piano sulla TELA INCOMPIUTA. Ora capisce chiaramente: è la sua casa d'infanzia.

L'ARTISTA inizia a raccogliere le sue cose. Ha deciso. Deve partire.

EXT. PAESAGGIO MONTANO - GIORNO
Ora è immersa nella natura. Cielo limpido. Montagne silenziose.
L'ARTISTA dipinge. L'ULTIMA PENNELLATA chiude il cerchio.

FADE OUT.

G Inquadrature in inglese

1.1, "Aerial view of the city in the afternoon, shot with a drone.", Extreme Wide Shot, Drone - Slow Descend, None, None, "Warm atmosphere, natural afternoon light, long shadows to add depth."
1.2, "High-angle shot of a square with a central fountain, people moving around.", Wide Shot, Static, Passersby, None, "Urban dynamism, contrast between architectural elements and the flow of people."
1.3, "Exterior shot of the young artist's studio, a lively street with people walking.", Wide Shot, Slow Tracking, Passersby, None, "The studio appears as the main point of interest, embedded in the urban context."
1.4, "We move closer to the studio, showing the shop window and entrance more clearly.", Medium Wide Shot, Slow Zoom-In, None, None, "Details on the window with street reflections."
1.5, "Lateral dolly shot inside the studio, showing paintings, sculptures, and artistic tools.", Wide Shot, Lateral Dolly, None, Natural light illuminating the space., "The studio is chaotic but rich with works in various stages of creation."
1.6, "Close-up of a recognition plaque awarded to the artist.", Close-Up, Static, None, Soft lighting to highlight it., "Suggests the artist's past success."
1.7, "The lateral dolly continues until we see the artist from behind, frantically rummaging through papers.", Medium Shot, Lateral Dolly, Artist, None, "The frantic behavior reveals her restless state of mind."
1.8, "We see the artist's workspace as she frantically searches through papers.", Close-Up, Static, Artist, None, "Rapid hand movements, nervous and irregular gestures."
1.9, "The artist turns around and moves toward another corner of the studio.", Medium Full Shot, Tracking Shot, Artist, None, "Tense expression, she seems to be searching for something important."
1.10, "She finds a newspaper article about herself and reads it for a moment.", Close-Up, Static, Artist, None, "The article suggests her past and the recognition she received."
1.11, "The artist sets down the article and moves toward some of her displayed works.", Medium Shot, Slow Push-In, Artist, None, "She approaches the works with a mix of nostalgia and reflection."
1.12, "She stops in front of an unfinished painting that we have yet to see.", Close-Up, Static, Artist, None, "Moment of hesitation, expressing her creative block."
1.13, "Angle change: now we see the unfinished painting with the artist in the foreground.", Medium Shot, Static, Artist, Soft lighting emphasizing the canvas., "The painting now becomes the center of the scene."
1.14, "The artist slowly turns, keeping her gaze on the canvas before looking away.", Medium Shot, Slow Push-In, Artist, None, "A meaningful gesture, signaling the start of a new thought."
1.15, "The artist walks through the studio, the camera precedes her in a steadicam shot.", Medium Full Shot, Steadicam - Tracking, Artist, None, "The movement suggests transition and change."
1.16, "The artist, seen from afar, exits the studio and closes the door behind her.", Wide Shot, Static, Artist, None, "A final look at the studio, leaving the past behind."

1.17, "We return to the unfinished canvas, now alone in the studio.", Medium Shot, Slow Dolly-In, None, Soft and enveloping light., "The painting remains as a symbol of her inner struggle."

2.1, "The artist walks through the city, framed from the front, lost in thought.", Medium Full Shot, Tracking Shot, Artist, None, "Reflective atmosphere, the artist does not interact with her surroundings."
2.2, "Flashback: Aerial view of a small town, with a young girl running through the hills.", Extreme Wide Shot, Drone - Slow Descend, Young girl (young artist), None, "The natural environment is bright, contrasting with the present-day city."
2.3, "The artist turns a corner and begins to observe urban artworks around her.", Medium Shot, Static, Artist, None, "Focused expression, she seems to start connecting with her past."
2.4, "Flashback: The young artist as a child, looking at a canvas for the first time, mesmerized.", Close-Up, Slow Push-In, Young girl, Warm light emphasizing the moment., "The image of the canvas hints at the child's artistic future."
2.5, "The artist walks until she finds herself in front of a mural and turns to look at it.", Medium Wide Shot, Static, Artist, None, "The mural is large and detailed, her first point of contact with her artistic memory."
2.6, "Flashback: Teenage artist leaving her small town with a suitcase, heading to the city.", Medium Shot, Tracking Shot, Teenage artist, Slight blur around the edges for a nostalgic effect., "Determined expression, but with a hint of melancholy."
2.7, "The artist approaches the mural, touching it gently with her hand.", Close-Up, Static, Artist, Subtle lighting variation emphasizing the touch., "A moment of connection between past and present."
2.8, "Flashback: The teenage artist receives her first recognition in the city.", Medium Shot, Slow Push-In, Teenage artist, Light vignette to emphasize the memory., "Joy is evident, but her gaze carries a hint of nostalgia."
2.9, "The artist turns away from the mural and resumes walking.", Medium Shot, Tracking Shot, Artist, None, "A visual transition between memory and the present."
2.10, "Flashback: Young adult artist inaugurating her first exhibition.", Medium Wide Shot, Slow Push-In, Young artist, Warm lighting, celebratory atmosphere., "The space is full of people, but her gaze seems distant."
2.11, "The artist looks at the mural one last time, then continues walking.", Medium Full Shot, Tracking Shot, Artist, None, "Closing the cycle of memories, ready to move forward."

3.1, "The artist enters the central square, observing a modern sculpture.", Wide Shot, Tracking Shot, Artist, Passersby, None, "A wide view of the square, another person walking in the background."
3.2, "The artist approaches the sculpture, intrigued, while the shot remains focused on her.", Medium Shot, Slow Push-In, Artist, None, "Her gaze is concentrated, the shot highlights her reaction."
3.3, "Close-up of the sculpture: the carved grooves start to vibrate slightly.", Detail Shot, Slow Zoom-In, None, Subtle vibration effect on the stone., "The carved details seem to move imperceptibly."

3.4, "Suddenly, the sculpted water begins to flow slowly.", Close-Up, Static, None, Fluid animation of the water coming to life., "Contrast between the stone and the movement of the water."
3.5, "Medium shot of the artist freezing as she observes the water flowing.", Medium Shot, Static, Artist, Light reflections on the moving water., "Expression of astonishment, the sound of water dominates the scene."
3.6, "The artist slowly turns to one side, looking for other people.", Medium Shot, Static, Artist, None, "A slight head movement as she scans her surroundings."
3.7, "Artist's POV: she sees people walking in the square, unaware of the situation.", POV Shot, Handheld, Passersby, None, "From her point of view, everything seems normal to others."
3.8, "The artist turns in the other direction, still looking for confirmation.", Medium Shot, Static, Artist, None, "Still doubtful, she searches for someone who might have seen what happened."
3.9, "Artist's POV: she sees more people continuing to walk, noticing nothing.", POV Shot, Handheld, Passersby, None, "To them, the sculpture still appears immobile."
3.10, "Medium shot of the artist with a confused expression, unable to understand what happened.", Medium Shot, Static, Artist, None, "A puzzled look, slight head movement as she reflects."
3.11, "Medium full shot: the artist turns and slowly walks away from the square.", Medium Full Shot, Tracking Shot, Artist, None, "A surreal atmosphere as she leaves the sculpture behind."

4.1, "Lateral dolly following the artist as she walks through the alleyways, stopping in front of the mural.", Wide Shot, Lateral Dolly, Artist, None, "Narrow alleyways, a more intimate and isolated atmosphere."
4.2, "The artist observes the mural, framed from a low angle.", Medium Shot (Low Angle), Static, Artist, None, "The angle emphasizes the size and visual impact of the mural."
4.3, "Close-up of the lines in the painting as they begin to move.", Detail Shot, Macro Shot, None, Subtle animation effect on the lines., "The movement is faint, almost imperceptible."
4.4, "Medium shot of the artist looking down at her feet, noticing something strange.", Medium Shot, Static, Artist, None, "The framing builds suspense, urging the audience to discover what's happening below."
4.5, "Close-up of the ground cracking, with roots beginning to emerge.", Detail Shot, Slow Push-In, None, Cracking effect and root growth., "The ground slowly splits, revealing the transformation taking place."
4.6, "Wide shot with tilt-up: the painted trees start to come to life and rise.", Wide Shot, Tilt-Up, None, Effect of trees growing., "The setting becomes surreal, the mural expands and transforms into something alive."
4.7, "Full shot: the ground beneath the artist's feet continues to break apart.", Full Shot, Static, Artist, Cracking effect on the ground beneath her feet., "The change is now evident, the terrain is transforming under her."
4.8, "Close-up of the artist looking around as the world transforms.", Close-Up, Handheld, Artist, Visual distortion effects to emphasize the transformation., "Change is happening all around her, creating a sense of disorientation."

4.9, "Extreme close-up of the artist with an expression of pure surprise.", Extreme Close-Up, Static, Artist, None, "Detail on the artist's emotions, the shock of what she is experiencing."
4.10, "Medium full shot: the artist turns around, taking in the new environment.", Medium Full Shot, 360° Camera Pan, Artist, The environment is completely transformed., "She looks around, trying to understand what has happened."
4.11, "Wide shot: an overview of the new environment she finds herself in.", Wide Shot, Static, Artist, "Completely transformed landscape, fantastical elements.", "The setting has evolved into a magical, natural world."

5.1, "The artist finds herself in a magical forest, with a very wide shot to emphasize the setting.", Extreme Wide Shot, Slow Zoom-In, Artist, Soft lighting effects and dreamlike atmosphere., "A wide frame that fully immerses the viewer in the scene."
5.2, "The camera begins to move in slowly, revealing details of the vegetation.", Wide Shot, Slow Dolly-In, Artist, "Light filtering through the trees, magical effect.", "Leaves gently moving, the environment starts to feel alive."
5.3, "The frame tightens further, revealing the shades of colors and plants.", Medium Wide Shot, Slow Push-In, None, The colors seem to pulse slightly., "Gradual transition into a closer look."
5.4, "Medium shot of the artist observing the scene, fascinated.", Medium Shot, Static, Artist, None, "The frame sets up for the arrival of a key detail."
5.5, "Close-up of a flower blooming, with a slow and fluid movement.", Detail Shot, Macro Shot - Slow Push-In, None, Animation of the flower's growth., "Complete focus on the moving flower, vibrant and saturated colors."
5.6, "The camera slowly widens, showing more flowers blooming.", Medium Shot, Slow Zoom-Out, None, Visual transition from detail to landscape., "The environment gradually expands, creating a sense of immersion."
5.7, "The artist notices that something is changing.", Medium Close-Up, Static, Artist, Light beginning to reflect off her skin., "A first reaction is visible on her face."
5.8, "Tilt-down: the artist looks down at her feet as water begins to spread.", Medium Shot, Slow Tilt-Down, Artist, Distorted reflection in the liquid., "The scene shifts to a more focused moment, emphasizing transformation."
5.9, "The water starts to expand rapidly across the landscape.", Wide Shot, Slow Zoom-Out, None, The water grows and moves slowly., "The environment changes quickly."
5.10, "The camera continues pulling back, showing the water filling the space.", Wide Shot, Slow Zoom-Out, Artist, The water continues to expand., "An intermediate transition to help the audience adapt to the large final shot."
5.11, "Extreme wide shot: the river now dominates the scene, carrying the artist away.", Extreme Wide Shot, Dolly Follow, Artist, Movement of the water gently carrying the artist., "The river becomes the focus of the shot, the scene is dynamic and fluid."

6.1, "The artist finds herself in a mountainous landscape, surrounded by snow and streams.", Extreme Wide Shot, Slow Zoom-In, Artist, "Cold and surreal atmosphere, light mist.", "A wide view emphasizing isolation and the vastness of the landscape."

6.2, "The artist bends down and picks up some snow in her hands.", Medium Shot, Static, Artist, Details of the snow's texture., "An intimate shot, the artist connects with the new environment."
6.3, "Close-up of the snow in the artist's hands as it starts to transform.", Close-Up, Slow Push-In, None, Transition effect from snow to water., "Focus on the metamorphosis of matter."
6.4, "The snow slips through her fingers, transforming into liquid water.", Extreme Close-Up, Macro Shot, None, Animation of the transition from snow to water., "Detailed effect on the melting process."
6.5, "The water continues to expand, reflecting the sky.", Medium Wide Shot, Slow Zoom-Out, Artist, The surface of the water becomes a mirror for the sky., "A surreal transition between natural elements."
6.6, "The artist watches the water spreading around her, amazed.", Medium Close-Up, Static, Artist, Reflected light from the transformed landscape., "Expression of wonder and discovery."
6.7, "The surrounding environment slowly dissolves, merging with the horizon.", Extreme Wide Shot, Slow Fade-Out, Artist, "Dissolution of the environment, the boundary between sky and sea vanishes.", "A poetic closure of the scene, transitioning into a new world."

7.1, "The artist walks along the shoreline, immersed in the golden light of sunset.", Wide Shot, Tracking Shot, Artist, "Golden sunset light, reflections in the water.", "A warm, enveloping atmosphere, the sound of waves in the background."
7.2, "Close-up of the sand under the artist's feet as it starts to change texture.", Close-Up, Handheld, None, Transition of sand into wheat stalks., "A gradual visual transformation, the texture of the sand shifting."
7.3, "Close-up of the wheat stalks emerging from the sand.", Close-Up, Slow Push-In, None, Natural growth effect of the stalks., "The stalks sprout slowly, with a delicate motion."
7.4, "The wheat stalks sway in the wind, the sea has now disappeared.", Medium Wide Shot, Slow Dolly-In, None, Natural movement of the stalks in the wind., "The scene has completely transformed into a wheat field."
7.5, "The artist picks up a wheat stalk and observes it curiously.", Medium Shot, Static, Artist, None, "An expression of curiosity, the wind gently brushing the scene."
7.6, "The stalk in her hands begins to stiffen, transforming into a small sculpture.", Close-Up, Macro Shot, None, Effect of metamorphosis of the stalk., "The object slowly changes shape between the artist's fingers."
7.7, "Extreme close-up of the artist, surprised by what is happening in her hands.", Extreme Close-Up, Static, Artist, Subtle illumination on her face., "Detail on the artist's emotions, a mix of disbelief and wonder."
7.8, "The sculpture represents a dance of leaves swirling in the wind.", Medium Shot, Static, None, Subtle motion effect on the sculpted leaves., "Contrast between the solid material and the lightness of the wind is emphasized."
7.9, "The sculpted leaves detach from the sculpture and start floating into the sky, transforming into real leaves.", Wide Shot, Tilt-Up, None, Dissolution effect from sculpted leaves to real ones., "A visual transition setting up for the next scene."

8.1, "The leaves begin to transform into hot air balloons as they rise into the sky.", Wide Shot, Slow Tilt-Up, None, Visual transition from leaves → hot air balloons., "The transformation is gradual and harmonious, with a soft fading effect."
8.2, "The hot air balloons drift into the sky, not yet revealing the artist.", Extreme Wide Shot, Slow Zoom-Out, None, Vibrant colors contrasting with the clouds., "A fairy-tale panorama, harmonious visual composition."
8.3, "As the balloons ascend, the frame begins to reveal the artist from below.", Medium Wide Shot, Slow Tilt-Down, Artist, Soft light on the artist's face., "The movement is delicate, emphasizing her observation of the balloons."
8.4, "The artist watches the hot air balloons float in the sky, her expression captivated.", Medium Close-Up, Static, Artist, Light reflections in her eyes., "A contemplative gaze, a moment of connection with the dreamy landscape."
8.5, "As the balloons drift away, a fairy-tale castle emerges in the background.", Wide Shot, Slow Push-In, Artist, "Soft transition, the castle gradually becomes visible.", "A fog or fade-in effect, the castle appears slowly."
8.6, "The artist slowly walks toward the castle, immersed in the dreamlike atmosphere.", Medium Full Shot, Tracking Shot, Artist, Ethereal lighting on the architectural elements., "A moment of transition, her journey continues."

9.1, "The artist slowly approaches the entrance of the castle.", Medium Full Shot, Tracking Shot, Artist, "Ethereal atmosphere, diffused light.", "The castle appears imposing yet welcoming, the journey progresses."
9.2, "The artist crosses the castle's threshold, but something begins to change.", Medium Shot, Slow Push-In, Artist, "Slight visual distortion, first signs of environmental alterations.", "A moment of transition between reality and surrealism."
9.3, "The castle walls begin to shift in shape and color.", Wide Shot, Slow Zoom-Out, None, Effects of architectural metamorphosis., "The environment becomes increasingly unstable and surreal."
9.4, "The castle windows merge into a single moving surface.", Close-Up, Handheld, None, Fluid animation of surfaces warping., "The entire space is in constant mutation, dreamlike atmosphere."
9.5, "The entire setting changes, dissolving into moving geometric shapes.", Extreme Wide Shot, Slow Tilt-Up, None, Fading of architecture into abstract elements., "The castle no longer exists, now it's a flow of shapes and colors."

10.1, "The geometries start moving, rotating and transforming around the artist.", Wide Shot, Slow Pan, Artist, Animated movement of geometric shapes., "A fluid and harmonious motion, the artist is suspended among abstract figures."
10.2, "The artist floats among the moving geometries, observing their transformation.", Medium Shot, Tracking Shot, Artist, "Reduced gravity effect, delicate movements.", "The shapes seem to respond to her presence."
10.3, "The geometries begin to reassemble, forming the skyline of a nighttime city under the rain.", Wide Shot, Slow Zoom-Out, None, Transition from geometric forms to a cityscape., "A fluid reconstruction effect, neon lights and urban reflections emerge."
10.4, "The wet streets reflect neon lights, creating a surreal atmosphere.", Medium Wide Shot, Slow Zoom-In, None, Dynamic light effects on the water surface., "Saturated colors, cyberpunk atmosphere."

10.5, "The artist walks slowly through the nighttime city, observing the surroundings.", Medium Shot, Handheld, Artist, "Raindrops on surfaces, depth effect.", "A contemplative moment, with a slowed-down rhythm."

11.1, "The artist finds herself inside a painting divided into four sections, each representing a season.", Wide Shot, Slow Zoom-In, Artist, Transition effect between seasons., "The artist stands at the center of the painting, displaying the passage of time."
11.2, "Lateral dolly as the artist walks through winter, surrounded by snow.", Full Body Shot, Lateral Dolly, Artist, "Snow falling gently, cold atmosphere.", "Her footsteps leave light imprints in the snow, the sky is gray."
11.3, "Lateral dolly as the artist crosses into spring, flowers blooming around her.", Full Body Shot, Lateral Dolly, Artist, "Petals dancing in the air, vivid colors.", "The atmosphere warms up, the gentle wind moves the green grass."
11.4, "Lateral dolly as the artist walks through summer, bathed in golden light.", Full Body Shot, Lateral Dolly, Artist, "Visible heat in the air, wind through the trees.", "The sun is high in the sky, the shadows are short."
11.5, "Lateral dolly as the artist walks through autumn, surrounded by falling golden leaves.", Full Body Shot, Lateral Dolly, Artist, Leaves detaching from the trees, dancing in the wind., "Soft light, the sky has warm, orange tones."
11.6, "Suddenly, everything begins to dissolve, the painting loses its definition.", Extreme Wide Shot, Slow Fade-Out, Artist, The seasons overlap and disappear., "The landscape loses its consistency, setting up the next scene."

12.1, "The artist finds herself back in her studio, in total silence.", Wide Shot, Static, Artist, Warm sunset light., "The contrast between the dreamlike journey and reality is striking."
12.2, "The artist looks around, realizing she has returned.", Medium Shot, Slow Push-In, Artist, "Subtle quiet effect, slight reflection of light on the canvases.", "A thoughtful expression, but without confusion."
12.3, "She turns and observes the unfinished painting.", Over-the-shoulder Shot, Static, Artist, "The blurred effect on the canvas now has meaning.", "A symbolic framing, the painting represents the artist's past."
12.4, "The artist approaches the canvas with a newfound sense of awareness.", Close-Up, Slow Push-In, Artist, Subtle light variation to emphasize the moment., "A determined expression, the doubt has vanished."
12.5, "Detail of the artist's hand gently touching the canvas, almost affectionately.", Extreme Close-Up, Static, Artist, Soft lighting to highlight the emotional weight of the gesture., "She caresses the painting as if finally understanding its meaning."
12.6, "Close-up on the artist: from a serious, almost sad face, she begins to smile.", Close-Up, Static, Artist, Warm lighting effect accompanying her change in expression., "A moment of realization, now everything makes sense."
12.7, "She starts gathering her belongings, packing canvases and tools.", Medium Wide Shot, Tracking Shot, Artist, Small dust particles floating in the sunset light., "The gesture is fluid and determined, without hesitation."
12.8, "Close-up of her hands putting brushes and small objects into a box.", Close-Up, Static, Artist, Subtle reflections of sunset light on the tools., "Delicate movements, but with purpose."

12.9, "Lateral dolly as the artist moves through the studio, gathering the last items.", Full Body Shot, Lateral Dolly, Artist, Warm light following her movements., "The progressive emptying of the studio is visible."
12.10, "Medium shot of the artist carefully sealing a box.", Medium Shot, Static, Artist, "Soft light on the box, symbolizing a new beginning.", "A symbolic closure."
12.11, "Wide shot of the studio now filled with packed boxes, with the paintings removed from the walls.", Wide Shot, Slow Zoom-Out, None, "A slight dust effect in the air, illuminated by the sunset.", "The studio has completely changed, marking the start of a new chapter in her life."

13.1, "A new setting: imposing mountains, a clear sky, and untouched nature.", Extreme Wide Shot, Slow Pan, None, "Intense natural light, vibrant colors of nature.", "The atmosphere is serene, in contrast to the city she left behind."
13.2, "The artist is sitting outdoors, with the unfinished painting in front of her.", Medium Wide Shot, Static, Artist, A gentle breeze moving the grass around her., "The composition is balanced, with the artist in harmony with the landscape."
13.3, "The artist gazes at the landscape, taking a deep breath.", Close-Up, Slow Push-In, Artist, Natural light highlighting her calm face., "Expression of peace and inner realization."
13.4, "She lowers her gaze to the canvas, still unfinished.", Over-the-shoulder Shot, Static, Artist, "Visible texture of the canvas, the painting is close to completion.", "The connection between the real landscape and the painting becomes evident."
13.5, "Close-up of the brush as it touches the canvas, the painting comes to life.", Close-Up, Macro Shot, Artist, "Smooth brush movement, the color blends perfectly.", "An intimate moment of creation."
13.6, "Frontal medium shot of the artist as she continues painting, her gaze intense.", Medium Shot, Static, Artist, Natural light illuminating her face and the canvas., "A focused expression, steady hands adding details."
13.7, "Full body shot from behind: the artist seated, deeply immersed in completing her work.", Full Body Shot, Lateral Dolly, Artist, Gentle movements of the wind around her., "The scene captures the artist fully engaged in her craft, a moment of deep connection."
13.8, "The painting is complete, seen from the perspective of the canvas with a slight zoom out.", Medium Wide Shot, Slow Zoom-Out, None, "Defined colors, the painting is now finished.", "The artwork reflects the surrounding landscape, symbolizing the completion of her journey."
13.9, "The artist looks at her finished work, smiling with satisfaction and emotion.", Close-Up, Static, Artist, "Eyes shining with emotion, warm light enhancing the moment.", "A moment of realization and inner peace, the film closes on this image."
13.10, "Extreme wide shot of the mountain, with the artist visible in the distance, almost tiny.", Extreme Wide Shot, Static, Artist, "Soft, warm lighting, an atmosphere of peace.", "The vastness of nature compared to the artist, a significant visual closure."

H Storyboard

Di seguito sono riportate le tavole dello storyboard completo.

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.1
Azione: Vista aerea della città nel pomeriggio ripresa con un drone
Focale: Grandangolo (16-35mm)
Movimento Macchina: Extreme Wide Shot Drone - Slow Descend
Personaggi: Nessuno
VFX: Nessuno
Indicazioni Aggiuntive: Atmosfera calda luci naturali del pomeriggio

Shot: 1.2
Azione: Ripresa dall'alto di una piazza con una fonte
Focale: Grandangolo (16-35mm)
Movimento Macchina: Wide Shot
Personaggi: Statico
VFX: Passanti
Indicazioni Aggiuntive: Dinamismo urbano contrasto tra elementi architettonici

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.3
Azione: Ripresa esterna del laboratorio della giovane artista
Focale: Standard (50mm)
Movimento Macchina: Wide Shot
Personaggi: Slow Tracking
VFX: Passanti
Indicazioni Aggiuntive: L'atelier appare come punto di incontro della creatività

Shot: 1.4
Azione: Ci avviciniamo all'atelier mostrando meglio la vetrina e l ingresso
Focale: Normale (35-50mm)
Movimento Macchina: Medium Wide Shot
Personaggi: Slow Zoom In
VFX: Nessuno
Indicazioni Aggiuntive: Dettagli sulla vetrina con riflessi delicati

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.5
Azione: Carrello laterale all'interno dell'atelier mostrando tele
Focale: Normale (35-50mm)
Movimento Macchina: Sculture e strumenti artistici
Personaggi: Wide Shot
VFX: Lateral Dolly
Indicazioni Aggiuntive: Luce naturale che illumina lo spazio

Shot: 1.6
Azione: Dettaglio su una targa di riconoscimento dell'artista
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Nessuno
VFX: Leggera illuminazione per evidenziarlo
Indicazioni Aggiuntive: Suggerisce il successo passato dell'artista

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.7
Azione: Il dolly laterale riprende fino a mostrare una parte significativa del laboratorio
Focale: Standard (50mm)
Movimento Macchina: Slow Tracking
Personaggi: Nessuno
VFX: Nessuno
Indicazioni Aggiuntive: Nessuna indicazione aggiuntiva

Shot: 1.8
Azione: Vediamo il laboratorio dell'artista mentre cerca qualcosa
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: L'artista osserva intensamente l'ambiente circostante

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

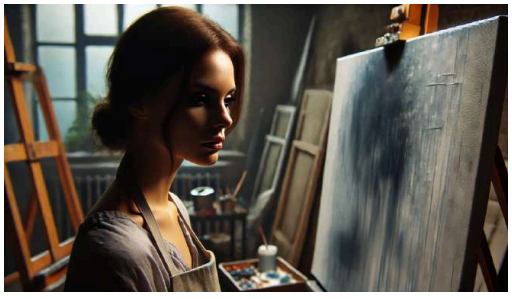
Shot: 1.9
Azione: L'artista si rivolta e si sposta verso un nuovo angolo dello spazio
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Trova un articolo di giornale su di lei e lo legge attentamente

Shot: 1.10
Azione: L'artista nota dettagli nell'articolo che suggeriscono il suo passato
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: L'articolo suggerisce il suo passato e il riconoscimento del suo talento

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.11
Azione: L'artista posa l'articolo e si dirige verso alcune delle opere esposte
Focale: Medium Shot
Movimento Macchina: Slow Push-In
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Si avvicina alle opere con un misto di nostalgia e riflessione



Shot: 1.12
Azione: L'artista si ferma davanti a un dipinto incompiuto
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Momento di esitazione che esprime il blocco creativo

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.13
Azione: Cambio d'angolazione: ora il dipinto incompiuto appare in primo piano con l'artista sullo sfondo
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Illuminazione soffusa
Indicazioni Aggiuntive: Il dipinto diventa il fulcro della scena

Shot: 1.14
Azione: L'artista si gira lentamente, mantenendo lo sguardo fisso sulla tela prima di distoglierlo
Focale: Medium Shot
Movimento Macchina: Slow Push-In
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Gesto significativo che annuncia l'inizio di un nuovo pensiero

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.15
Azione: L'artista attraversa lo studio, seguita dalla camera in una ripresa steadycam
Focale: Medium Full Shot
Movimento Macchina: Steadycam Tracking
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Il movimento suggerisce una transizione e un cambiamento

Shot: 1.16
Azione: Vista da lontano, l'artista esce dallo studio e chiude la porta alle sue spalle
Focale: Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Uno sguardo finale allo studio che simboleggia il distacco dal passato

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 1.17
Azione: Si ritorna sul dipinto incompiuto, ora lasciato solo nello studio
Focale: Medium Shot
Movimento Macchina: Slow Dolly-In
Personaggi: Nessuno
VFX: Illuminazione soffusa e avvolgente
Indicazioni Aggiuntive: Il dipinto resta come simbolo della sua lotta interiore

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.1
Azione: L'artista cammina per la città, inquadrata frontalmente e persa nei suoi pensieri
Focale: Medium Full Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Atmosfera riflessiva in cui l'artista non interagisce con l'ambiente circostante

Shot: 2.2
Azione: Flashback: vista aerea di un piccolo paese, con una giovane ragazza che corre tra le colline
Focale: Extreme Wide Shot
Movimento Macchina: Drone Slow Descend
Personaggi: Giovane ragazza (futura artista)
VFX: Nessuno
Indicazioni Aggiuntive: L'ambiente naturale appare luminoso, in netto contrasto con la città attuale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.3
Azione: L'artista gira ad un angolo e inizia a osservare le opere d'arte urbane intorno a lei
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Espressione concentrata che suggerisce l'inizio di una riconnessione con il suo passato

Shot: 2.4
Azione: Flashback: la giovane artista da bambina osserva per la prima volta un dipinto, rapita
Focale: Close Up
Movimento Macchina: Slow Push-In
Personaggi: Bambina
VFX: Luce calda che enfatizza il momento
Indicazioni Aggiuntive: L'immagine del dipinto preannuncia il futuro artistico della bambina

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.5
Azione: L'artista cammina finché non si trova davanti a un murale e si ferma per osservarlo
Focale: Medium Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Il murale, ampio e dettagliato, rappresenta il primo contatto con i suoi ricordi artistici

Shot: 2.6
Azione: Flashback: l'adolescente artista lascia il suo piccolo paese con una valigia, diretta verso la città
Focale: Medium Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista adolescente
VFX: Leggera sfocatura ai bordi per un effetto nostalgico
Indicazioni Aggiuntive: Espressione determinata, ma con una punta di malinconia

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.7
Azione: L'artista si avvicina al murale e lo tocca delicatamente con la mano
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Variazione luminosa sottile che enfatizza il gesto
Indicazioni Aggiuntive: Momento di connessione tra il passato e il presente

Shot: 2.8
Azione: Flashback: l'adolescente artista riceve il suo primo riconoscimento in città
Focale: Medium Shot
Movimento Macchina: Slow Push-In
Personaggi: Artista adolescente
VFX: Leggera vignettatura per accentuare il ricordo
Indicazioni Aggiuntive: L'espressione di gioia è evidente, accompagnata da un tocco di nostalgia

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.9
Azione: L'artista si allontana dal murale e riprende a camminare
Focale: Medium Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Transizione visiva tra il ricordo e il presente

Shot: 2.10
Azione: Flashback: la giovane artista da adulta inaugura la sua prima mostra
Focale: Medium Wide Shot
Movimento Macchina: Slow Push-In
Personaggi: Giovane artista
VFX: Illuminazione calda, atmosfera di festa
Indicazioni Aggiuntive: L'ambiente è affollato, ma il suo sguardo appare distante

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 2.11
Azione: L'artista osserva il murale per l'ultima volta e poi riprende a camminare
Focale: Medium Full Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Chiusura del ciclo dei ricordi, pronta a proseguire

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.1
Azione: L'artista entra nella piazza centrale e osserva una scultura moderna
Focale: Wide Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista e passanti
VFX: Nessuno
Indicazioni Aggiuntive: Ampia visuale della piazza con una presenza in sottofondo

Shot: 3.2
Azione: L'artista si avvicina alla scultura, incuriosita, mentre l'inquadratura resta focalizzata su di lei
Focale: Medium Shot
Movimento Macchina: Slow Push-In
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Il suo sguardo concentrato evidenzia la reazione all'opera

TITLE L’ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.3
Azione: Close-up sulla scultura, in cui i solchi scolpiti iniziano a vibrare leggermente
Focale: Detail Shot
Movimento Macchina: Slow Zoom-In
Personaggi: Nessuno
VFX: Effetto di vibrazione sottile sulla pietra
Indicazioni Aggiuntive: I dettagli scolpiti appaiono quasi in movimento

Shot: 3.4
Azione: Improvvisamente, l'acqua scolpita inizia a scorrere lentamente sulla superficie
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Nessuno
VFX: Animazione fluida che trasforma l'acqua in movimento
Indicazioni Aggiuntive: Forte contrasto tra la staticità della pietra e il fluire dell'acqua

TITLE L’ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.5
Azione: Inquadratura media dell'artista che rimane immobile osservando il flusso d'acqua
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Riflessi di luce sull'acqua in movimento
Indicazioni Aggiuntive: Espressione di stupore, dominata dal suono dell'acqua

Shot: 3.6
Azione: L'artista si gira lentamente verso un lato, alla ricerca di altre persone
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Un leggero movimento del capo mentre scruta l'ambiente

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.7
Azione: Dal suo punto di vista, l'artista osserva le persone che camminano nella piazza, ignare della situazione
Focale: POV Shot
Movimento Macchina: Handheld
Personaggi: Passanti
VFX: Nessuno
Indicazioni Aggiuntive: Dal suo punto di vista, per gli altri tutto appare normale

Shot: 3.8
Azione: L'artista si gira nella direzione opposta, continuando a cercare conferme
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Ancora dubbiosa, cerca qualcuno che possa aver notato l'accaduto

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.9
Azione: Dal suo punto di vista, l'artista vede altre persone continuare a camminare, ignare
Focale: POV Shot
Movimento Macchina: Handheld
Personaggi: Passanti
VFX: Nessuno
Indicazioni Aggiuntive: Per gli altri, la scultura appare ancora immobile

Shot: 3.10
Azione: Inquadratura dell'artista con un'espressione confusa, incapace di comprendere cosa sia accaduto
Focale: Close-up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Sguardo perplesso e un leggero movimento del capo, in segno di riflessione

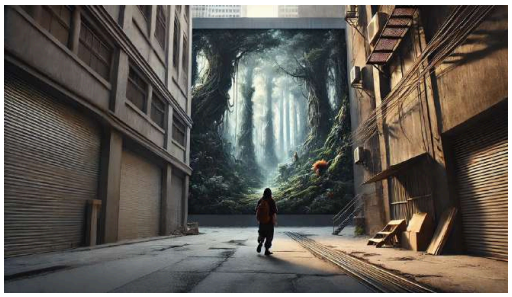
TITLE L’ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 3.11
Azione: Inquadratura a figura intera: l'artista si gira e cammina lentamente via dalla piazza
Focale: Medium Full Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Atmosfera surreale mentre lascia la scultura alle spalle

TITLE L’ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025



Shot: 4.1
Azione: L'artista cammina per i vicoli e si ferma davanti a un murale.
Focale: Wide Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Vicoli stretti e atmosfera intima

Shot: 4.2
Azione: L'artista osserva il murale da un'angolazione bassa.
Focale: Medium Shot (Low Angle)
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: L'angolazione enfatizza la grandezza e l'impatto visivo del murale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 4.3
Azione: Close-up delle linee del murale che iniziano a muoversi leggermente.
Focale: Detail Shot
Movimento Macchina: Slow Zoom-In
Personaggi: Nessuno
VFX: Effetto di vibrazione sottile
Indicazioni Aggiuntive: Il movimento è quasi impercettibile, creando un'atmosfera surreale

Shot: 4.4
Azione: L'artista guarda in basso, notando qualcosa di strano sotto i suoi piedi.
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: L'inquadratura crea suspense invitando a scoprire il dettaglio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 4.5
Azione: Close-up del terreno che si incrina con radici che emergono lentamente.
Focale: Detail Shot
Movimento Macchina: Slow Push-In
Personaggi: Nessuno
VFX: Effetto di crepatura e crescita delle radici
Indicazioni Aggiuntive: La trasformazione del terreno è evidente e simbolica

Shot: 4.6
Azione: In un wide shot con tilt-up, gli alberi dipinti sul murale iniziano a prendere vita e a sollevarsi.
Focale: Wide Shot
Movimento Macchina: Tilt-Up
Personaggi: Nessuno
VFX: Effetto di crescita degli alberi
Indicazioni Aggiuntive: Il murale si espande, trasformandosi in un ambiente vivo e surreale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 4.7
Azione: Full shot: il terreno sotto i piedi dell'artista continua a fratturarsi.
Focale: Full Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Effetto di incrinatura al suolo
Indicazioni Aggiuntive: La trasformazione del paesaggio diventa palpabile

Shot: 4.8
Azione: Close-up dell'artista che guarda intorno con effetti di distorsione visiva.
Focale: Close-Up
Movimento Macchina: Handheld
Personaggi: Artista
VFX: Effetto di distorsione per enfatizzare la trasformazione
Indicazioni Aggiuntive: Il senso di disorientamento è marcato

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 4.9
Azione: Extreme close-up dell'artista, espressione di puro stupore.
Focale: Extreme Close-Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Dettaglio che mostra il shock e l'emozione

Shot: 4.10
Azione: Inquadratura a figura intera: l'artista si gira e osserva l'ambiente trasformato.
Focale: Medium Full Shot
Movimento Macchina: 360° Camera Pan
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: La panoramica rivela il cambiamento radicale del paesaggio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 4.11
Azione: Wide shot: visione d'insieme del nuovo ambiente, surreale e magico.
Focale: Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Il paesaggio si è evoluto in un mondo fantastico e naturale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.1
Azione: L'artista si ritrova in un bosco magico con una ripresa molto ampia per enfatizzare l ambiente
Focale: Extreme Wide Shot
Movimento Macchina: Slow Zoom In
Personaggi: Artista
VFX: Effetti di luce soffusa e atmosfera onirica
Indicazioni Aggiuntive: Inquadratura aperta che immerge completamente nella scena

Shot: 5.2
Azione: La camera inizia ad avvicinarsi lentamente rivelando dettagli della vegetazione
Focale: Wide Shot
Movimento Macchina: Slow Dolly In
Personaggi: Artista
VFX: Luce filtrata tra gli alberi effetto magico
Indicazioni Aggiuntive: Si nota il movimento delle foglie e l ambiente inizia a sentirsi più vivo

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.3
Azione: L'inquadratura si stringe ancora rivelando le sfumature dei colori e delle piante
Focale: Medium Wide Shot
Movimento Macchina: Slow Push In
Personaggi: Nessuno
VFX: I colori sembrano pulsare leggermente
Indicazioni Aggiuntive: Transizione graduale verso il dettaglio

Shot: 5.4
Azione: Medium Shot dell'artista che osserva la scena affascinata
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: L'inquadratura prepara l'arrivo del dettaglio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.5
Azione: Dettaglio di un fiore che sboccia con un movimento lento e fluido
Focale: Detail Shot
Movimento Macchina: Macro Shot Slow Push In
Personaggi: Nessuno
VFX: Animazione della crescita dei fiori Focus totale sul fiore in movimento colori vibranti e saturi
Indicazioni Aggiuntive: Nessuna indicazione aggiuntiva

Shot: 5.6
Azione: La camera si allarga lentamente mostrando più fiori che sbocciano
Focale: Medium Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Transizione visiva dal dettaglio al paesaggio L ambiente si amplia gradualmente creando una sensazione di immersione
Indicazioni Aggiuntive: Nessuna indicazione aggiuntiva

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.7
Azione: L'artista nota che qualcosa sta cambiando
Focale: Medium Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce che inizia a riflettersi sulla pelle
Indicazioni Aggiuntive: Si vede una prima reazione sul volto dell'artista

Shot: 5.8
Azione: Tilt Down l'artista guarda sotto i suoi piedi mentre l'acqua si espande
Focale: Medium Shot
Movimento Macchina: Slow Tilt Down
Personaggi: Artista
VFX: Riflessione deformata nel liquido
Indicazioni Aggiuntive: La scena passa a un focus più ristretto enfatizzando la trasformazione

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.9
Azione: L'acqua inizia a espandersi velocemente nel paesaggio
Focale: Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: L acqua cresce e si muove lentamente
Indicazioni Aggiuntive: L'ambiente cambia rapidamente

Shot: 5.10
Azione: La camera continua ad allontanarsi mostrando l acqua che cresce e riempie lo spazio
Focale: Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Artista
VFX: L acqua continua ad espandersi lentamente
Indicazioni Aggiuntive: Un passaggio intermedio che aiuta lo spettatore ad adattarsi alla grande inquadratura finale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 5.11
Azione: Extreme Wide Shot il fiume ora domina la scena e la corrente trasporta l artista via
Focale: Extreme Wide Shot
Movimento Macchina: Dolly Follow
Personaggi: Artista
VFX: Movimento dell'acqua che trasporta dolcemente l artista
Indicazioni Aggiuntive: Il fiume diventa il centro dell'inquadratura la scena è dinamica e fluida

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 6.1
Azione: L artista si ritrova in un paesaggio montano circondata da neve e ruscelli
Focale: Extreme Wide Shot
Movimento Macchina: Slow Zoom In
Personaggi: Artista
VFX: Atmosfera fredda e surreale nebbia leggera
Indicazioni Aggiuntive: Ampia visuale per enfatizzare l isolamento e la vastità del paesaggio

Shot: 6.2
Azione: L artista si china e raccoglie della neve tra le mani
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Dettagli sulla texture della neve
Indicazioni Aggiuntive: Inquadratura intima l artista entra in contatto con il nuovo ambiente

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 6.3
Azione: Close Up della neve nelle mani dell'artista mentre inizia a trasformarsi
Focale: Close Up
Movimento Macchina: Slow Push In
Personaggi: Nessuno
VFX: Effetto di transizione dalla neve all'acqua
Indicazioni Aggiuntive: Focus sulla metamorfosi della materia

Shot: 6.4
Azione: La neve scivola tra le dita e si trasforma in acqua liquida
Focale: Extreme Close Up
Movimento Macchina: Macro Shot
Personaggi: Nessuno
VFX: Animazione della transizione da neve ad acqua
Indicazioni Aggiuntive: Effetto dettagliato sulla fusione della materia

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 6.5
Azione: L'acqua continua ad espandersi riflettendo il cielo
Focale: Medium Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Artista
VFX: Superficie dell acqua che diventa specchio del cielo
Indicazioni Aggiuntive: Transizione surreale tra elementi naturali

Shot: 6.6
Azione: L'artista osserva l'acqua che si espande intorno a lei stupita
Focale: Medium Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Illuminazione riflessa dal nuovo paesaggio
Indicazioni Aggiuntive: Espressione di meraviglia e scoperta

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 6.7
Azione: L'ambiente circostante si dissolve lentamente fondendosi con l orizzonte
Focale: Extreme Wide Shot
Movimento Macchina: Slow Fade Out
Personaggi: Artista
VFX: Dissolvenza dell'ambiente confine tra cielo e mare si annulla
Indicazioni Aggiuntive: Chiusura poetica della scena passaggio a un nuovo mondo

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 7.1
Azione: L'artista cammina lungo la riva immersa nella luce del tramonto
Focale: Wide Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Luce dorata del tramonto riflessa nell'acqua
Indicazioni Aggiuntive: Atmosfera calda avvolgente suono delle onde di sottofondo

Shot: 7.2
Azione: Dettaglio sulla sabbia sotto i piedi dell'artista mentre inizia a cambiare consistenza
Focale: Close Up
Movimento Macchina: Handheld
Personaggi: Nessuno
VFX: Transizione della sabbia in spighe di grano
Indicazioni Aggiuntive: Graduale trasformazione visiva texture della sabbia che muta

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 7.3
Azione: Close Up sulle spighe di grano che iniziano a emergere dalla sabbia
Focale: Close Up
Movimento Macchina: Slow Push In
Personaggi: Nessuno
VFX: Effetto di crescita naturale delle spighe
Indicazioni Aggiuntive: Le spighe spuntano lentamente con un movimento delicato

Shot: 7.4
Azione: Le spighe di grano ondeggiano nel vento mentre il mare è ormai sparito
Focale: Medium Wide Shot
Movimento Macchina: Slow Dolly In
Personaggi: Nessuno
VFX: Movimento naturale delle spighe al vento
Indicazioni Aggiuntive: La scena si trasforma completamente in un campo di grano

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 7.5
Azione: L'artista raccoglie una spiga e la osserva incuriosita
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Nessuno
Indicazioni Aggiuntive: Espressione di curiosità mentre il vento accarezza la scena

Shot: 7.6
Azione: La spiga nelle sue mani inizia a irrigidirsi trasformandosi in una piccola scultura
Focale: Close Up
Movimento Macchina: Macro Shot
Personaggi: Nessuno
VFX: Effetto di metamorfosi della spiga
Indicazioni Aggiuntive: L oggetto cambia lentamente forma tra le dita dell'artista

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 7.7
Azione: Extreme Close Up sull'artista sorpresa da ciò che sta accadendo tra le sue mani
Focale: Extreme Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Illuminazione sottile sul volto
Indicazioni Aggiuntive: Dettaglio sulle emozioni dell'artista incredulità e meraviglia

Shot: 7.8
Azione: La scultura rappresenta una danza di foglie che volteggiano nel vento
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Nessuno
VFX: Effetto di leggero movimento delle foglie scolpite
Indicazioni Aggiuntive: Il contrasto tra materiale solido e leggerezza del vento è enfatizzato

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 7.9
Azione: Le foglie scolpite si staccano dalla scultura e iniziano a fluttuare nel cielo trasformandosi in vere foglie
Focale: Wide Shot
Movimento Macchina: Tilt Up
Personaggi: Nessuno
VFX: Effetto di dissolvenza delle foglie scolpite in foglie reali
Indicazioni Aggiuntive: Transizione visiva che prepara alla scena successiva

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 8.1
Azione: Le foglie iniziano a trasformarsi in mongolfiere mentre salgono nel cielo
Focale: Wide Shot
Movimento Macchina: Slow Tilt Up
Personaggi: Nessuno
VFX: Transizione visiva in cui le foglie diventano mongolfiere
Indicazioni Aggiuntive: La trasformazione è graduale armoniosa con effetto di dissolvenza

Shot: 8.2
Azione: Le mongolfiere si allontanano nel cielo senza ancora mostrare l artista
Focale: Extreme Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Colori vividi in contrasto con le nuvole
Indicazioni Aggiuntive: Panorama fiabesco composizione visiva armoniosa

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 8.3
Azione: Mentre le mongolfiere salgono l'inquadratura inizia a rivelare l'artista dal basso
Focale: Medium Wide Shot
Movimento Macchina: Slow Tilt Down
Personaggi: Artista
VFX: Luce soffusa sul volto dell'artista
Indicazioni Aggiuntive: Il movimento è delicato enfatizzando la sua osservazione delle mongolfiere

Shot: 8.4
Azione: L'artista osserva le mongolfiere fluttuare nel cielo con espressione affascinata
Focale: Medium Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Effetto di riflessi di luce negli occhi
Indicazioni Aggiuntive: Sguardo contemplativo momento di connessione con il paesaggio onirico

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 8.5
Azione: Mentre le mongolfiere si allontanano sullo sfondo emerge un castello fiabesco
Focale: Wide Shot
Movimento Macchina: Slow Push In
Personaggi: Artista
VFX: Transizione dolce il castello diventa lentamente visibile
Indicazioni Aggiuntive: Effetto di nebbia o dissolvenza il castello appare gradualmente

Shot: 8.6
Azione: L'artista si incammina lentamente verso il castello immersa nell'atmosfera fiabesca
Focale: Medium Full Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Luce eterea sugli elementi architettonici
Indicazioni Aggiuntive: Momento di transizione il viaggio continua

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 9.1
Azione: L'artista avanza lentamente verso l ingresso del castello
Focale: Medium Full Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Atmosfera eterea luce diffusa
Indicazioni Aggiuntive: Il castello appare imponente ma accogliente il viaggio continua

Shot: 9.2
Azione: L'artista varca la soglia del castello ma qualcosa inizia a cambiare
Focale: Medium Shot
Movimento Macchina: Slow Push In
Personaggi: Artista
VFX: Leggera distorsione visiva prime alterazioni dell ambiente
Indicazioni Aggiuntive: Momento di transizione tra realtà e surreale

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 9.3
Azione: Le pareti del castello iniziano a mutare forma e colore
Focale: Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Effetti di metamorfosi dell'architettura
Indicazioni Aggiuntive: L'ambiente diventa sempre più instabile e surreale

Shot: 9.4
Azione: Le finestre del castello si fondono in un unica superficie in movimento
Focale: Close Up
Movimento Macchina: Handheld
Personaggi: Nessuno
VFX: Animazione fluida delle superfici che si deformano
Indicazioni Aggiuntive: L'ambiente è in costante mutazione atmosfera onirica

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 10.1
Azione: Le geometrie iniziano a muoversi ruotando e trasformandosi attorno all'artista
Focale: Wide Shot
Movimento Macchina: Slow Pan
Personaggi: Artista
VFX: Animazione delle forme geometriche
Indicazioni Aggiuntive: Movimento fluido e armonioso l artista è sospesa tra le figure astratte

Shot: 10.2
Azione: L'artista fluttua tra le geometrie in movimento osservando il loro mutamento
Focale: Medium Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Effetto di gravità ridotta movimenti delicati
Indicazioni Aggiuntive: Le forme sembrano rispondere alla sua presenza

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 10.3
Azione: Le geometrie iniziano a ricomporsi formando il profilo di una città notturna sotto la pioggia
Focale: Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Transizione da forme geometriche a skyline cittadino
Indicazioni Aggiuntive: Effetto di ricostruzione fluida neon e luci urbane emergono

Shot: 10.4
Azione: Le strade bagnate riflettono i neon creando un atmosfera surreale
Focale: Medium Wide Shot
Movimento Macchina: Slow Zoom In
Personaggi: Nessuno
VFX: Effetto di luci dinamiche sulla superficie dell'acqua
Indicazioni Aggiuntive: Colori saturi atmosfera cyberpunk

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 10.5
Azione: L'artista cammina lentamente attraverso la città notturna osservando l ambiente
Focale: Medium Shot
Movimento Macchina: Handheld
Personaggi: Artista
VFX: Gocce di pioggia sulle superfici effetto di profondità
Indicazioni Aggiuntive: Momento contemplativo ritmo rallentato

Shot: 10.6
Azione: Close Up sull'artista mentre lo sfondo cambia dietro di lei
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Effetto di dissolvenza temporale dietro di lei
Indicazioni Aggiuntive: L'ambiente cambia mentre cammina

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 10.7
Azione: Medium Shot che completa la transizione e il cambio di scenario
Focale: Medium Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Artista
VFX: Chiusura visiva della trasformazione
Indicazioni Aggiuntive: Il paesaggio si stabilizza nella sua nuova forma

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 11.1
Azione: L'artista si ritrova in un quadro diviso in quattro parti ognuna rappresentante una stagione
Focale: Wide Shot
Movimento Macchina: Slow Zoom In
Personaggi: Artista
VFX: Effetto di transizione tra le stagioni
Indicazioni Aggiuntive: L'artista è al centro del quadro che mostra il passaggio tra le stagioni

Shot: 11.2
Azione: Carrello laterale mentre l'artista cammina attraverso l inverno circondata dalla neve
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Neve che cade dolcemente atmosfera fredda
Indicazioni Aggiuntive: I suoi passi lasciano impronte leggere sulla neve il cielo è grigio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 11.3
Azione: Carrello laterale mentre l'artista attraversa la primavera i fiori sbocciano attorno a lei
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Petali di fiori che danzano nell'aria colori vividi
Indicazioni Aggiuntive: L'atmosfera si scalda il vento leggero muove l erba verde

Shot: 11.4
Azione: Carrello laterale mentre l artista cammina nell'estate avvolta da luce dorata
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Effetto di calore visibile nell'aria vento tra gli alberi
Indicazioni Aggiuntive: Il sole è alto nel cielo le ombre sono corte

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 11.5
Azione: Carrello laterale mentre l artista cammina nell'autunno tra foglie dorate che cadono
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Foglie che si staccano dagli alberi e danzano nel vento
Indicazioni Aggiuntive: La luce è morbida il cielo ha toni caldi e aranciati

Shot: 11.6
Azione: Tutto inizia a dissolversi improvvisamente il quadro perde definizione
Focale: Extreme Wide Shot
Movimento Macchina: Slow Fade Out
Personaggi: Artista
VFX: Le stagioni si sovrappongono e svaniscono
Indicazioni Aggiuntive: Il paesaggio perde consistenza preparando alla prossima scena

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 12.1
Azione: L'artista si ritrova nel suo studio in un silenzio totale
Focale: Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce calda del tramonto
Indicazioni Aggiuntive: Il contrasto tra il viaggio onirico e la realtà è forte

Shot: 12.2
Azione: L'artista guarda attorno a sé realizzando che è tornata
Focale: Medium Shot
Movimento Macchina: Slow Push In
Personaggi: Artista
VFX: Effetto di quiete lieve riflesso della luce sulle tele
Indicazioni Aggiuntive: Espressione pensierosa ma senza confusione

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 12.3
Azione: Si volta e osserva la tela incompiuta
Focale: Over the Shoulder Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: L effetto sfumato della tela ora ha un significato
Indicazioni Aggiuntive: Inquadratura simbolica il quadro rappresenta il passato dell'artista

Shot: 12.4
Azione: L'artista si avvicina alla tela con un nuovo senso di consapevolezza
Focale: Close Up
Movimento Macchina: Slow Push In
Personaggi: Artista
VFX: Leggera variazione della luce per enfatizzare il momento
Indicazioni Aggiuntive: Espressione decisa il dubbio è svanito

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 12.5
Azione: Dettaglio della mano dell'artista che accarezza la tela quasi con affetto
Focale: Extreme Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce soffusa per enfatizzare l emotività del gesto
Indicazioni Aggiuntive: L artista accarezza il dipinto come se comprendesse finalmente il suo significato

Shot: 12.6
Azione: Close Up sull'artista da un viso serio e quasi triste che inizia a sorridere
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Effetto di luce calda che accompagna il cambiamento di espressione
Indicazioni Aggiuntive: Momento di realizzazione ora tutto ha senso

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

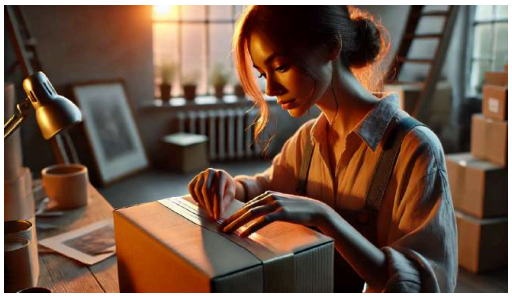
Shot: 12.7
Azione: Inizia a raccogliere le sue cose impacchettando tele e strumenti
Focale: Medium Wide Shot
Movimento Macchina: Tracking Shot
Personaggi: Artista
VFX: Piccole particelle di polvere nella luce del tramonto
Indicazioni Aggiuntive: Il gesto è fluido e determinato senza esitazione

Shot: 12.8
Azione: Dettaglio delle mani che ripongono pennelli e piccoli oggetti in una scatola
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Leggeri riflessi della luce del tramonto sugli strumenti
Indicazioni Aggiuntive: Movimenti delicati ma con decisione

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 12.9
Azione: Carrello laterale mentre l'artista cammina nello studio raccogliendo le ultime cose
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Effetto di luce calda che accompagna i suoi movimenti
Indicazioni Aggiuntive: Si vede il progressivo svuotamento dello studio



Shot: 12.10
Azione: Medium Shot dell'artista mentre chiude con cura uno scatolone
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce soffusa sulla scatola segno di un nuovo inizio
Indicazioni Aggiuntive: Momento di chiusura simbolica

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 12.11
Azione: Wide Shot dello studio ormai riempito di scatoloni le opere rimosse dalle pareti
Focale: Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Leggero effetto di polvere nell'aria illuminata dal tramonto
Indicazioni Aggiuntive: Lo studio è cambiato completamente segno del nuovo capitolo della sua vita

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 13.1
Azione: Un nuovo ambiente montagne imponenti cielo limpido e natura incontaminata
Focale: Extreme Wide Shot
Movimento Macchina: Slow Pan
Personaggi: Nessuno
VFX: Luce naturale intensa colori vibranti della natura
Indicazioni Aggiuntive: L'atmosfera è serena contrasto con la città lasciata alle spalle

Shot: 13.2
Azione: L'artista è seduta all'aperto con la tela incompiuta di fronte a sé
Focale: Medium Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Leggera brezza che muove l'erba intorno a lei
Indicazioni Aggiuntive: La composizione è equilibrata con l artista in armonia con il paesaggio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 13.3
Azione: L'artista osserva il paesaggio inspirando profondamente
Focale: Close Up
Movimento Macchina: Slow Push In
Personaggi: Artista
VFX: Luce naturale evidenzia il suo volto sereno
Indicazioni Aggiuntive: Espressione di calma e realizzazione interiore

Shot: 13.4
Azione: Abbassa lo sguardo sul quadro ancora incompiuto
Focale: Over the Shoulder Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Texture della tela visibile il dipinto è vicino alla sua conclusione
Indicazioni Aggiuntive: Il legame tra il paesaggio reale e il quadro si fa evidente

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 13.5
Azione: Close Up del pennello che sfiora la tela la pittura prende forma
Focale: Close Up
Movimento Macchina: Macro Shot
Personaggi: Artista
VFX: Movimento fluido del pennello il colore si fonde perfettamente
Indicazioni Aggiuntive: Momento intimo e concentrato sulla creazione

Shot: 13.6
Azione: Medium Shot frontale dell'artista mentre continua a dipingere con sguardo intenso
Focale: Medium Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce naturale che illumina il volto e la tela
Indicazioni Aggiuntive: Espressione concentrata mano ferma mentre aggiunge dettagli

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 13.7
Azione: Full Body Shot da dietro l'artista seduta intenta a completare l opera
Focale: Full Body Shot
Movimento Macchina: Lateral Dolly
Personaggi: Artista
VFX: Leggeri movimenti del vento attorno a lei
Indicazioni Aggiuntive: La scena mostra l'artista immersa nella pittura un momento di connessione profonda

Shot: 13.8
Azione: L'opera è completa vista dall'inquadratura della tela con un leggero zoom out
Focale: Medium Wide Shot
Movimento Macchina: Slow Zoom Out
Personaggi: Nessuno
VFX: Colori definiti il dipinto ora è terminato
Indicazioni Aggiuntive: Il quadro riflette il paesaggio attorno simbolo della chiusura del cerchio

TITLE L'ARTISTA E LA CITTA (PROVVISORIO)

NAME MATTIA MORGUT YEAR & SUBJECT 2025

Shot: 13.9
Azione: L'artista osserva la sua opera completata con un sorriso soddisfatto e commosso
Focale: Close Up
Movimento Macchina: Statico
Personaggi: Artista
VFX: Occhi brillanti di commozione luce calda che esalta l emozione
Indicazioni Aggiuntive: Un momento di realizzazione e pace interiore il film si chiude su questa immagine

Shot: 13.10
Azione: Extreme Wide Shot della montagna con l'artista visibile in lontananza quasi minuscola
Focale: Extreme Wide Shot
Movimento Macchina: Statico
Personaggi: Artista
VFX: Luce calda e morbida atmosfera di pace
Indicazioni Aggiuntive: L'immensa natura rispetto all'artista chiusura visiva significativa

I Tabelle con associazioni dei dati

Dato dal documento	Come è stato ritrovato
"Città di notte piove strade lucide guizzare di topi vita notturna dei vicoli"	Questa descrizione dettagliata è riscontrabile nelle sequenze della città notturna, dove neon pulsanti e riflessi sulle strade bagnate creano un'atmosfera misteriosa e onirica, riflettendo il contrasto tra l'urbanità e l'esperienza interiore dell'artista.
"Elementi astratti"	La dissolvenza del castello in forme geometriche astratte e la trasformazione dei murales in ambienti astratti creano un linguaggio visivo che unisce realtà e fantasia, enfatizzando il tema della metamorfosi interiore.
"La bellezza della scoperta"	L'idea che ogni nuovo quadro o scena riveli un aspetto inedito del mondo si traduce in sequenze graduali e rivelatrici, in cui i dettagli naturali emergono come metafora della scoperta di sé e del mondo circostante.
"I veri viaggiatori sono coloro che partono per partire"	Questa citazione, ispirata a Baudelaire, incarna lo spirito dell'avventura e il desiderio di trasformazione, riflettendo il percorso dell'artista che, intraprendendo il viaggio, si apre a nuove possibilità e a una rinnovata identità.

Tabella 1: Associazioni principali tra le citazioni presenti nel documento contenente i dati e i loro corrispondenti utilizzi narrativi nel cortometraggio. Sono riportate tra virgolette le esatte parole presenti nel documento di progetto.

Dato dal documento	Come è stato ritrovato
"Neve"	Elemento ricorrente che, trasformandosi in acqua, simboleggia rinascita e fluidità.
"Scultura"	La scultura urbana diventa dinamica, fondendo arte e natura in un movimento fluido.
"Città di notte"	La città prende vita attraverso forme astratte e neon, evocando un'atmosfera surreale.
"Nostalgia"	Di fronte alla tela incompiuta, la protagonista evoca, con tono malinconico, il ricordo della sua casa d'infanzia e del paesino di montagna, trasmettendo una profonda sensazione di nostalgia e desiderio di ritorno alle origini.
"Passeggiata in montagna, tra ruscelli e prati"	La narrazione visiva traduce questa immagine in sequenze in cui l'artista percorre sentieri naturali, immergendosi nel paesaggio e riconnettendosi con le proprie radici e il fluire della vita.
"Documentario, anche musicale o astratto"	Le preferenze espresse per un video ibrido si rispecchiano nell'approccio narrativo: il cortometraggio unisce riprese reali, elementi documentaristici e tocchi di astrazione, creando una sinfonia visiva e musicale in grado di coinvolgere emotivamente lo spettatore.
"Video che simula un'esperienza 3D, con tante immagini che si susseguono"	Le tecniche cinematografiche – tracking shot, dolly e zoom – generano un effetto immersivo e dinamico, simulando un'esperienza tridimensionale che trasporta lo spettatore in un viaggio visivo fluido tra realtà e animazione.
"Docureale"	L'approccio docureale emerge dalle riprese autentiche dello studio, dai flashback e dai momenti riflessivi, che combinano la realtà quotidiana con una narrazione emotiva intensa, raccontando il percorso interiore dell'artista.
"Moderno e Fantascienza"	La fusione di tecniche cinematografiche moderne con elementi surreali – come il castello che si dissolve in geometrie astratte – crea un linguaggio visivo innovativo, capace di unire il realismo con la fantasia futuristica.

Tabella 2: Associazioni principali tra le citazioni presenti nel documento contenente i dati e i loro corrispondenti utilizzi narrativi nel cortometraggio. Sono riportate tra virgolette le esatte parole presenti nel documento di progetto.

Dato dal documento	Come è stato ritrovato
"Viaggio"	L'artista attraversa ambientazioni diverse (dallo studio in città al bosco e infine alla montagna), mostrando sia un percorso fisico che un'evoluzione interiore.
"Gioia"	Compare nel momento in cui l'artista, completato il quadro, sorride con commozione, segnando il superamento della crisi creativa.
"Serenità"	Si avverte nel volto rilassato della protagonista quando osserva la tela finita immersa nel paesaggio naturale.
"Speranza"	È espressa dal desiderio di tornare alle origini come nuovo inizio, trasformando il ritorno in un atto di rinnovata fiducia.
"Infanzia"	Richiamata attraverso il ricordo della casa del paesino di montagna, che emerge dalle linee sfocate presenti sulla tela incompiuta.
"Natura"	Presente nei boschi, nei ruscelli e nel paesaggio montano, che riflettono e accompagnano il cambiamento interiore dell'artista.
"Solitudine"	Si percepisce nelle scene in cui l'artista cammina da sola tra vicoli e piazze, evidenziando l'isolamento emotivo come spinta alla trasformazione.
"Rinascita"	Simbolizzata dalla metamorfosi della neve in acqua, che poi si fonde con cielo e mare, indicando un processo di rinnovamento profondo.
"Mare"	Presente nella scena in cui la sabbia si trasforma in spighe di grano, collegando il fluire emotivo dell'artista con l'immagine del mare e del cambiamento.
"Un amore passato"	Raccontato attraverso murales e ricordi, con immagini che collegano la protagonista alle sue radici familiari e culturali.
"Sogno"	Catturato dalle sequenze oniriche in cui il castello si dissolve in forme geometriche e l'artista fluttua in un mondo irreale.
"Tempo presente"	Sottolinea il momento in cui l'artista torna nello studio e comprende che il 'qui e ora' è ciò che conta.
"Fuga"	Intesa come allontanamento dalla città per ritrovare sé stessa, reso tramite la sequenza in cui l'artista esce dallo studio verso la natura.
"Danza"	Raffigurata dal movimento coreografato delle foglie scolpite e delle spighe di grano che si muovono in sincronia, evocando fluidità.

Tabella 3: Altre associazioni, più inerenti alla sfera emotiva, rilevate tra dati raccolti e narrazione del prodotto audiovisivo. Sono riportate tra virgolette le esatte parole presenti nel documento di progetto.

J Prompt generati

Di seguito vengono riportati solamente i prompt relativi alla prima inquadratura di ogni scena. Il file completo è disponibile al seguente link.

Link al file completo.

```
SCENA 1

Prompt - Shot 1.1
This video must be ultra-realistic and cinematic in tone and
execution. The scene should look like it was captured on a high-end
digital cinema camera, with soft, natural lighting and carefully
composed visuals.

Main Goal / Message:
 Introduce the setting-a Northern Italian coastal city-from a high,
 detached viewpoint, revealing the layered beauty and movement of
 the urban environment.

Setting / Location:
 Aerial view of a city in Northern Italy, located on the coast,
 during the afternoon. The drone reveals a patchwork of terracotta
 rooftops, narrow medieval alleys, modern buildings, and small
 harbors. The sea is visible in the distance, shimmering under the
 sunlight, softly lapping against the port structures. The streets
 below are populated by people and vehicles, though distant and
 miniature.

Characters:
 No visible characters in this shot - the cityscape is the sole
 protagonist.

Mood and Music:
 Warm and introspective. City sounds are muted and atmospheric:
 distant hum of traffic, seagulls, a breeze from the sea. A sparse,
 slow ambient score may be added (piano or synth pads).

Visual Style:
 Hyper-realistic and cinematic.
 Extreme Wide Shot captured with a drone in slow descending motion.
 The afternoon sun casts long, angled shadows over the buildings.
 The lighting is golden and natural. The framing emphasizes the
 contrast between the scale of the city and the quiet, solitary tone
 of the film's beginning.
```

SCENA 2

Prompt - Shot 2.1
This video must be ultra-realistic and cinematic, opening the second scene in continuity with the emotional atmosphere. We are in a Northern Italian coastal city, transitioning from the quiet of the studio to the complexity of the outside world.

Main Goal / Message:
To show the artist walking through the city with a reflective, isolated demeanor. She is physically present, but emotionally distant from her surroundings.

Setting / Location:
Urban streets of a Northern Italian city by the sea. The neighborhood is vibrant and layered: worn facades, modern details, small shops, and glimpses of the coastline between alleyways. The light is soft, late afternoon, with shadows stretching across cobblestones.

Characters:
The artist walks slowly through the streets.
She is a woman in her mid-30s, with a slender build and fair skin. Her facial features are soft and natural, and her blonde, lightly wavy hair is styled in a relaxed low chignon. She wears an oversized off-white linen shirt with rolled-up sleeves, high-waisted faded blue jeans, and beige canvas sneakers. Subtle silver earrings and a thin ring may be visible. Her gaze is low, distant, immersed in thought.

Mood and Music:
Introspective. The sound design focuses on footsteps, distant voices, seagulls overhead, and the low murmur of the city. A minimal, ambient musical layer may accompany the moment-soft pads or distant textures.

Visual Style:
Hyper-realistic and cinematic.
Medium Full Shot with a tracking camera movement from the front. The artist walks directly toward the lens, surrounded by the city's soft chaos. The background blurs slightly, focusing on her facial expression, which carries a mix of melancholy and self-reflection. Light flickers across walls and glass surfaces, adding texture and movement to the background.

SCENA 3

Prompt - Shot 3.1
This video must be ultra-realistic and cinematic, set in a public square of a Northern Italian coastal city during the afternoon. The tone is observational and grounded in natural rhythm.

Main Goal / Message:
To present a new urban environment and mark the entry into a more open, contemplative moment. The artist approaches a sculpture that quietly commands the center of the space.

Setting / Location:
A central sqaure with modern and classical architectural elements: cobblestones, surrounding cafes, civic buildings. In the middle, a large modern sculpture rises-abstract but natural in form, carved from pale stone, perhaps suggesting the flow of water or wind.

Characters:
The artist, mid-30s, fair skin, slender, with soft natural features. Her blonde, lightly wavy hair is styled in a low, relaxed chignon. She wears an oversized off-white linen shirt with sleeves rolled up, high-waisted faded blue jeans, and beige canvas sneakers. As she enters the square, her gaze is focused on the sculpture ahead.

Mood and Music:
Open and curious. Background ambiance includes footsteps, the chatter of people in the distance, clinking glasses. No music-just space and light.

Visual Style:
Hyper-realistic and cinematic.
Wide Shot with a tracking camera movement following the artist as she walks into the sqaure. The framing includes passersby and the geometry of the space, slowly revealing the sculpture. Natural light enhances the stone texture and draws shadows across the pavement. The tone is observational-something is beginning, but nothing has shifted yet.

SCENA 4

Prompt - Shot 4.1
This video must be ultra-realistic and cinematic, set in the narrow alleyways of a Northern Italian coastal city. The light is softer now, filtered through high walls and hanging cables.

Main Goal / Message:
To show the artist entering a more intimate and mysterious area of the city. Her walk suggests introspection and quiet anticipation. The world begins to close in, becoming more internal.

Setting / Location:
A narrow stone alley, typical of northern seaside towns. The walls are aged with chipped paint, small balconies, wires overhead. It's quieter here-no cars, only footsteps and echoes.

Characters:
The artist, mid-30s, with a slender build, fair skin, and soft natural features. Her blonde, lightly wavy hair is styled in a relaxed low chignon. She wears an oversized off-white linen shirt with rolled sleeves, high-waisted faded blue jeans, and beige canvas sneakers. Her hands are in her pockets or gently brushing the stone as she walks. Her gaze is straight ahead, thoughtful.

Mood and Music:
Anticipatory. Ambient sounds dominate: her footsteps, birds on rooftops, distant voices bouncing between the walls. A subtle tonal bed may begin-warm, minimal.

Visual Style:
Hyper-realistic and cinematic.
Wide Shot with a lateral dolly movement, keeping the artist centered in frame as she walks. The camera follows from the side, gliding along the alley. Textures are rich-stone, brick, laundry swaying. Light shifts rhythmically as she passes under hanging structures or sunlit gaps.

SCENA 5

Prompt - Shot 5.1
This video must be ultra-realistic and cinematic. The artist, having crossed the threshold between city and nature, now finds herself fully immersed in a vibrant, expansive forest. The atmosphere is alive, but still.

Main Goal / Message:
To establish the vastness and dreamlike beauty of the magical forest. The artist is small against the grandeur of nature-but fully present within it.

Setting / Location:
An open forest clearing filled with towering trees, flowering plants, moss, and streams of light. The canopy opens above,

allowing sunbeams to fall like spotlights. The world feels enchanted, but still grounded in natural forms.

Characters:
The artist stands within this world. She is a woman in her mid-30s with fair skin, light freckles, and delicate natural features. Her blonde, lightly wavy hair is tied into a low, relaxed chignon. She wears an oversized off-white linen shirt with sleeves rolled up, high-waisted faded blue jeans, and beige canvas sneakers. She stands still, gazing outward, slowly breathing, overwhelmed by what surrounds her.

Mood and Music:
Awe and stillness. Nature sounds fill the soundscape-leaves rustling, distant birdsong, the occasional drop of water. A soft ambient musical tone begins: long, sustained, almost sacred.

Visual Style:
Hyper-realistic and cinematic.
Extreme Wide Shot with a slow zoom-in. The artist is a small figure amid towering trunks and soft vegetation. Light moves through the space naturally. Mist hovers in low pockets. Colors are vivid but realistic: green, golden, deep brown, mossy grey. The forest breathes-the viewer should feel drawn into its rhythm.

SCENA 6

Prompt - Shot 6.1
This video must be created in a painting style, where the entire scene is composed through delicate dots of color that merge from a distance, creating an optical shimmer of light and atmosphere, showing the arrival of the artist in a new, high-altitude landscape. The atmosphere is pure, surreal, and vast.

Main Goal / Message:
To establish a radical change in setting. The artist is now in a mountainous environment-isolated, open, and ethereal.

Setting / Location:
A high mountain plateau, surrounded by snow-covered peaks, scattered boulders, and shimmering streams. Low mist lingers near the ground. The water from the previous scene has dissipated, leaving behind a glistening, pristine terrain.

Characters:
The artist is present, mid-30s, fair skin, light freckles, and delicate features. Her blonde, lightly wavy hair is tied in a low,

relaxed chignon. She wears her iconic outfit: oversized off-white linen shirt with rolled sleeves, high-waisted faded blue jeans, and beige canvas sneakers. She stands still, adjusting her gaze, trying to orient herself.

Mood and Music:
Suspended stillness. A gentle wind is heard, along with soft crunches of snow or gravel underfoot. A faint, cold ambient tone may hover in the background.

Visual Style:
Hyper-realistic and cinematic.
Extreme Wide Shot with a slow zoom-in. The artist is a small figure in a vast, white-blue landscape. The light is diffuse and cold, but not hostile. Everything feels endless and quiet-like a sacred threshold.

SCENA 7

Prompt - Shot 7.1
This video must be rendered in a Fauvist painting style, with intense, expressive brushstrokes and bold colors that amplify the emotional energy of the natural environment, showing the artist walking along the shore of a surreal landscape. The atmosphere is golden and suspended in time.

Main Goal / Message:
To portray a moment of calm and reflection as the artist walks through a dreamlike beach environment. Her steps are soft, her gaze focused forward, bathed in light.

Setting / Location:
A transformed shoreline. The sea is calm and far in the distance. The sky is painted in the warm tones of sunset-amber, gold, soft pink. The sand has a delicate, almost velvety texture.

Characters:
The artist, mid-30s, fair skin, light freckles, soft features. Her blonde, lightly wavy hair is tied in a low relaxed chignon. She wears an oversized off-white linen shirt with sleeves rolled up, high-waisted faded blue jeans, and beige canvas sneakers. Her pace is slow, deliberate, contemplative.

Mood and Music:
Warm and introspective. The sound of soft waves, wind through empty space. A light piano motif or ambient pad may accompany the moment.

Visual Style:
Hyper-realistic and cinematic.
Wide Shot with a tracking camera movement, following the artist from the side as she walks. Her shadow stretches behind her. The beach extends endlessly, textured in fine, light sand. Every element -fabric, light, texture-must be tangible and glowing.

SCENA 8

Prompt - Shot 8.1
This video must be created in the style of Vincent van Gogh, with swirling brushwork, emotional intensity, and rich, vibrant textures that give the scene a dreamlike, poetic resonance, showing the final stage of the sculpture's metamorphosis: the floating leaves begin to transform into hot air balloons as they rise into the sky.

Main Goal / Message:
To portray a transformation of scale and symbol. The once-small leaves become large, elegant flying objects, suggesting a passage toward elevation and vision.

Setting / Location:
The sky above the wheat field. The horizon is still visible, bathed in the final golden glow of sunset.

Characters:
No human figures in this frame. The leaves are the protagonists.

Mood and Music:
Awe and ascent. The score begins to evolve-gentle strings, warm and layered, rising softly.

Visual Style:
Hyper-realistic and cinematic.
Wide Shot with a slow tilt-up, following the leaves as they slowly morph into hot air balloons mid-air. Their surface transforms from soft organic texture to structured balloon fabric, painted with natural tones and patterns. Light reflects off their contours as they begin to drift upward. Everything remains elegant and physical-no magical sparkles, just believable movement and texture.

SCENA 9

Prompt - Shot 9.1
This video must be designed in the visual language of Italian Futurism, using dynamic compositions, angular fragmentation, and

rhythmic repetition, capturing the artist as she slowly approaches the entrance of the castle. The structure appears majestic, almost sacred.

Main Goal / Message:
To convey the tension between awe and intimacy. The castle is monumental, but does not intimidate-it invites.

Setting / Location:
Exterior, late day. The fairy-tale castle glows with diffused light and an ethereal atmosphere. Its architecture is tall, intricate, with a curved gate and overgrown details.

Characters:
The artist, mid-30s, fair skin, freckles, delicate features, with blonde, softly wavy hair in a low chignon. She wears an oversized off-white linen shirt with rolled sleeves, high-waisted faded blue jeans, and beige canvas sneakers. She walks slowly, looking up at the entrance.

Mood and Music:
Subtle tension and reverence. A soft ambient pad, light wind brushing leaves or stone.

Visual Style:
Hyper-realistic and cinematic.
Medium Full Shot, tracking shot following the artist from behind as she approaches the entrance. The castle looms slowly into full view, its presence growing across the frame.

SCENA 10

Prompt - Shot 10.1
This video must be crafted in a Cubist visual style, where space and forms are deconstructed into overlapping geometric planes, showing the artist suspended inside an abstract dimension made of floating geometries. The entire space is in motion, responding to her presence.

Main Goal / Message:
To visually introduce a space that represents pure perception: harmonious, abstract, and completely detached from the real.

Setting / Location:
An undefined space filled with slowly rotating geometric shapes-spheres, arches, translucent cubes. Everything floats. Soft lighting and volumetric shadows.

Characters:
The artist, a woman in her mid-30s with fair skin, soft facial features, and blonde lightly wavy hair tied in a relaxed low chignon. She wears an oversized off-white linen shirt with rolled-up sleeves, high-waisted faded blue jeans, and beige canvas sneakers. Her body floats gently among the shapes. She looks around with a calm, observant expression.

Mood and Music:
Hypnotic and weightless. Soft ambient textures, deep layers with no rhythm-just flow.

Visual Style:
Ultra-realistic and cinematic.
Wide Shot, slow pan around the scene. The shapes move harmoniously. The artist appears small among them. The lighting is ambient and soft, with subtle reflections on the skin and clothes.

SCENA 11

Prompt - Shot 11.1
This video must be portrayed in an Impressionist painting style, capturing the fleeting light, color, and atmosphere of the seasons with soft edges and delicate, luminous brushwork, showing the artist standing in the center of a painting divided into four parts, each representing a different season.

Main Goal / Message:
To introduce the metaphorical painting-world of the four seasons. The artist is at the center, visually bridging all transitions.

Setting / Location:
A surreal pictorial space where the environment is divided into four sections-winter, spring, summer, autumn-coexisting and separated by invisible borders.

Characters:
The artist, a woman in her mid-30s with fair skin, soft facial features, and . Her blonde, lightly wavy hair is tied in a relaxed low chignon. She wears an oversized off-white linen shirt with rolled-up sleeves, high-waisted faded blue jeans, and beige canvas sneakers. She stands still in the center of the scene.

Mood and Music:
Quiet, majestic, suspended. Each quadrant has a unique ambient tone corresponding to its season.

Visual Style:
Ultra-realistic and cinematic.
Wide Shot with a slow zoom-in. The camera slowly approaches the artist. Behind her, seasons transition subtly. Colors, weather, and vegetation are rich and detailed. Visual effects ensure soft blending between segments of the painting.

SCENA 12

Prompt - Shot 12.1
This video must be ultra-realistic and cinematic, showing the artist back inside her studio, in total silence.

Main Goal / Message:
To contrast the dreamlike journey she has just experienced with the grounded, quiet return to reality.

Setting / Location:
The artist's studio at sunset. The space is filled with warm, golden light. No movement. No sound.

Characters:
The artist, a woman in her mid-30s with fair skin, light freckles, and soft natural features. Her blonde, lightly wavy hair is tied into a relaxed low chignon. She wears an oversized off-white linen shirt with rolled-up sleeves, high-waisted faded blue jeans, and beige canvas sneakers. She stands at the center of the studio, completely still.

Mood and Music:
Absolute silence. Only a faint ambient tone, like a held breath.

Visual Style:
Ultra-realistic and cinematic.
Wide Shot, static camera. The light bathes the entire space. The contrast between the immobility of the scene and the emotional weight is key.

SCENA 13

Prompt - Shot 13.1
This video must be ultra-realistic and cinematic, introducing a vast mountain landscape. The atmosphere is serene, untouched, and filled with vibrant natural light.

Main Goal / Message:
To mark the beginning of a new chapter. The artist is no longer confined to the past or her inner journey-she is immersed in nature, at peace.

Setting / Location:
An expansive mountain setting with a clear blue sky and untouched grassy terrain. No signs of human activity. Everything is vibrant and illuminated.

Characters:
No visible characters in this shot.

Mood and Music:
Open and peaceful. Only the sound of wind moving through grass, distant birds.

Visual Style:
Ultra-realistic and cinematic.
Extreme Wide Shot, slow panoramic movement. The landscape unfolds across the screen, revealing cliffs, ridges, valleys, and patches of wildflowers. Light reflects on each element with physical texture and clarity.

K Cortometraggio

Di seguito sono presenti il link e il codice QR, utilizzabili per visualizzare il cortometraggio nella sua forma finale.

Link al cortometraggio.

Figura 4: Codice QR per il cortometraggio.

L Feedback del pubblico

Di seguito vengono mostrati tutti i dati puliti e organizzati riguardanti il feedback fornito dal pubblico successivamente alla proiezione del cortometraggio.

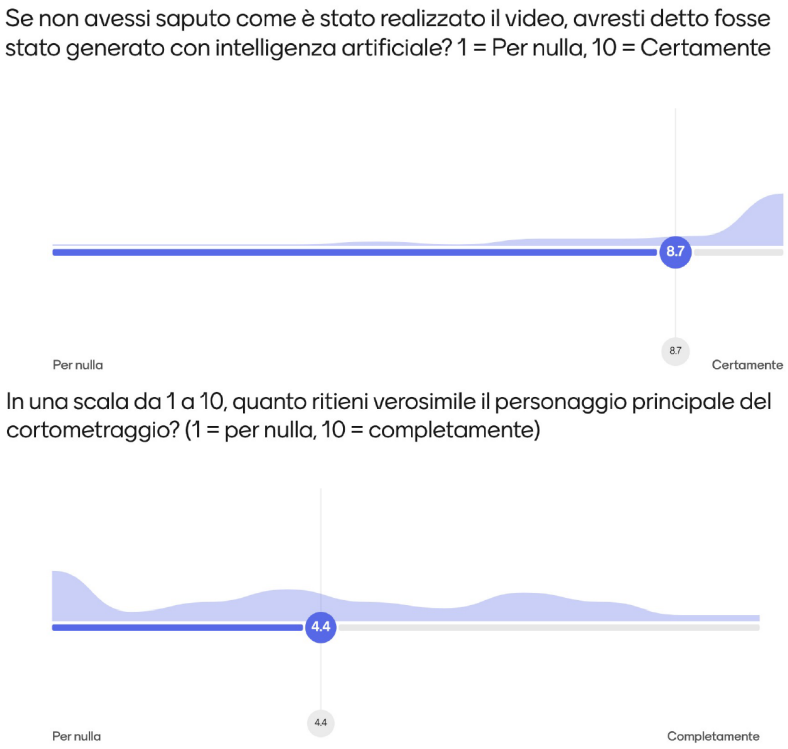


Figura 5: Grafici relativi a due domande poste al pubblico.

Che emozioni senti di aver provato durante la visione di questo video?

Emozione	Frequenza
confusione	6
curiosità	4
noia	4
smarrimento	3
incertezza	3
tristezza	3
nostalgia	3
malinconia	2
riflessione	2
nessuna emozione	2
disagio	1
pace	1
stabilità	1
straniamento	1
libertà	1
senza anima	1
meraviglia	1
introspezione	1
inquietudine	1
creatività	1
trip mentale	1
leggerezza	1
tranquillità	1
emozione	1
deconcerted	1
perplessità	1
euforia	1
irritazione	1
freddezza	1
perdizione	1
vuoto	1
spirale	1
felicità	1
dispersione	1
angoscia	1
sensibilità	1
squallore	1
dilemma	1
scoperta	1
paranoia	1
sorpresa	1
montagna	1
passato	1

Con una parola, come descriveresti l'impatto complessivo del video?

Parola	Frequenza
lungo	3
interessante	3
carino	2
banale	2
straniante	2
ripetitivo	2
malinconico	1
bello	1
carico	1
triste	1
freddo	1
originale	1
trip mentale	1
migliorabile	1
percezioni	1
noioso	1
split minded	1
vertiginoso	1
lenta	1
confusionario	1
intenso	1
nostalgico	1
artificiale	1
strano	1
ansiante	1
riflessivo	1
introspezione	1
disarmante	1
viaggio	1
consapevolezza	1
volenteroso	1
ispirante	1
mediocre	1
impattante	1
introspettivo	1
stimolante	1
scontato	1
deludente	1

Hai percepito un cambiamento nel tuo stato d'animo dopo la visione del video? Se sì, in che modo?

Risposta	Frequenza
si	12
no	11
angoscia	2
riflessività	1
rassicurazione	1
ansia	1
sfiducia nel futuro	1
incertezza sul futuro	1
la forza della subtesi	1
ho sentito la mancanza dell'anima	1
ho pensato: questo lavoro l'ho già fatto!	1
impazienza	1
sni	1
io torno ma al mare	1
coraggio	1
nostalgia	1
sfiducia	1
non tornero indietro questa volta	1
psichedelia	1
consapevolezza	1
un po' di noia	1
mi ha angosciata	1
malinconia	1
solo più noia!	1
pensierosa	1
noia	1

Quale messaggio suggerisce secondo te il cortometraggio?
Coraggio di scegliere
Pensiero
Approccio un po’ retorica
Riflessioni sulle proprie scelte
Trovare il proprio posto nel mondo
Sfruttare a proprio favore i cambiamenti
Saper apprezzare se stessi
Viaggio
Futuro
Impara l'arte e mettila da parte
Analisi interiore
Che devo tornare a casa
Riflessione
Futuro
Riconnettersi alle proprie radici
Riflettere sulle proprie scelte
Autoanalisi
Coraggio
Ricercare quello che veramente si vuole fare e dove si vuole stare, entrando in contatto con il proprio essere
Coraggio di cambiare. Il messaggio è ottimo, il cortometraggio no
Nostalgia
Trovare le proprie radici
Ritrovare le radici
Attenzione nei confronti delle proprie emozione e bisogni
Quello detto dalla frase finale
Bisogna uscire dai propri confini per ricominciare
Evolversi
Non aver paura del passato
Basta viaggi!
Non allontanarsi da sé come l'AI ci spinge a fare
L'arte non è per tutti
Apertissima
Capacità di mettersi in discussione
Speranza
Che non è mai troppo tardi per andare nella direzione della propria felicità
Un passo indietro può significare andare avanti
Introspezione, voglia di cambiare, lottare per un mondo migliore
Il ruolo dell'arte nella vita
Amore per l'ambiente e il proprio lavoro
Don’t be normal. Be unic and follow your creativity
Le scelte vanno pensate e fatte con coraggio
Speranza
Nulla
Banali riflessioni

M Dataset per fine-tuning

Si riporta di seguito solamente una parte dei dati utilizzati per il *fine-tuning* del modello. L'intero dataset è disponibile per la consultazione e l'eventuale download al seguente link.

Link al dataset completo.

User	**Prompt**
Aerial view of the city in the afternoon, shot with a drone.	This video must be ultra-realistic and cinematic in tone and execution...
High-angle shot of a square with a central fountain, people moving around.	This video must be ultra-realistic and cinematic, reflecting the poetry of real urban life in Northern Italy...
Exterior shot of the young artist's studio, a lively street with people walking.	This video must be ultra-realistic and cinematic, embedding the artist's world within the daily rhythm of a Northern Italian seaside city...
We move closer to the studio, showing the shop window and street outside.	This video must be ultra-realistic and cinematic, revealing the creative atmosphere through layers of detail...
Lateral dolly shot inside the studio, showing paintings and materials.	This video must be ultra-realistic and cinematic, capturing the tactile nature of the artist's work and environment...

Tabella 4: Esempio di alcune coppie di dati utilizzate per la fase di *fine-tuning* del modello. Nella prima colonna è presente una breve descrizione della clip, mentre nella seconda si riporta l'intero prompt generato in precedenza per realizzare il video. I prompt sono gli stessi presenti in Sezione J, qui troncati per una migliore resa visiva.

N Codice per fine-tuning

Il codice riportato di seguito è stato implementato in un *notebook* sulla piattaforma *Google Colab*, disponibile per consultazione ed eventualmente utilizzo al seguente link.

Link al *notebook* completo.

```
# Installazione delle librerie

!pip uninstall -y transformers accelerate
!pip install bitsandbytes
!pip install transformers
!pip install trl
!pip install accelerate
!pip install unsloth


# Inizializzazione del modello

import unsloth
from unsloth import FastLanguageModel
import torch

max_seq_length = 2048
dtype = None    # None for auto detection. Float16 for Tesla T4
load_in_4bit = False  # Use 4bit quantization to reduce memory usage.

model, tokenizer = FastLanguageModel.from_pretrained(
    model_name = "unsloth/Llama-3.2-3B-Instruct",
    max_seq_length = max_seq_length,
    dtype = dtype,
    load_in_4bit = load_in_4bit,
)

model = FastLanguageModel.get_peft_model(
    model,
    r = 16,
    target_modules = ["q_proj", "k_proj", "v_proj", "o_proj",
                      "gate_proj", "up_proj", "down_proj",],
    lora_alpha = 16,
    lora_dropout = 0,
    bias = "none",
    use_gradient_checkpointing = "unsloth",
    random_state = 3407,
    use_rslora = False,
    loftq_config = None,
)
```

```
from unsloth.chat_templates import get_chat_template

tokenizer = get_chat_template(
    tokenizer,
    chat_template = "llama-3.1",
)
FastLanguageModel.for_inference(model)

messages = [
    {"role": "user", "content": "scene of a young painting"},
]
inputs = tokenizer.apply_chat_template(
    messages,
    tokenize = True,
    add_generation_prompt = True,
    return_tensors = "pt",
).to("cuda")

from transformers import TextStreamer
text_streamer = TextStreamer(tokenizer, skip_prompt = True)
_ = model.generate(input_ids = inputs, streamer = text_streamer,
                   max_new_tokens = 128, use_cache = True,
                   temperature = 1.5, min_p = 0.1)


# Formattazione del testo per il formato di LLama3.2

from unsloth.chat_templates import get_chat_template

tokenizer = get_chat_template(
    tokenizer,
    chat_template = "llama-3.1",
)

import pandas as pd
df = pd.read_csv("/content/sora_prompts_parsed.csv",
        sep=";",
        encoding="utf-8-sig",
        engine="python",
        usecols=["User", "Prompt"]
      )

df = pd.read_csv(
    "/content/prompts_final2.csv",
    sep=",",
    encoding="utf-8-sig",
    engine="python",
```

```
    usecols=["prompt", "user"]
)

df[['prompt', 'user']] = df[['user', 'prompt']]
df.columns = ["User", "Prompt"]

df.head()

df = df.dropna()

from datasets import Dataset
df["conversations"] = df.apply(
    lambda x: [
        {"content": x["User"], "role": "user"},
        {"content": x["Prompt"], "role": "assistant"}
    ], axis=1
)

dataset = Dataset.from_pandas(df.drop(columns=["User", "Prompt"]))

def formatting_prompts_func(examples):
    convos = examples["conversations"]
    texts = [tokenizer.apply_chat_template(convo, tokenize = False,
                add_generation_prompt = False) for convo in convos]
    return { "text" : texts, }
pass

from unsloth.chat_templates import standardize_sharegpt
dataset = standardize_sharegpt(dataset)
dataset = dataset.map(formatting_prompts_func, batched = True,)


# Addestramento del modello

from trl import SFTTrainer
from transformers import TrainingArguments, DataCollatorForSeq2Seq
from unsloth import is_bfloat16_supported

trainer = SFTTrainer(
    model = model,
    tokenizer = tokenizer,
    train_dataset = dataset,
    dataset_text_field = "text",
    max_seq_length = max_seq_length,
    data_collator = DataCollatorForSeq2Seq(tokenizer = tokenizer),
    dataset_num_proc = 2,
    packing = False,        # Can make training 5x faster
    args = TrainingArguments(
```

```
        per_device_train_batch_size = 2,
        gradient_accumulation_steps = 4,
        warmup_steps = 5,
        # num_train_epochs = 1, # Set this for 1 full training run.
        max_steps = 15,
        learning_rate = 2e-4,
        fp16 = not is_bfloat16_supported(),
        bf16 = is_bfloat16_supported(),
        logging_steps = 1,
        optim = "adamw_8bit",
        weight_decay = 0.01,
        lr_scheduler_type = "linear",
        seed = 3407,
        output_dir = "outputs",
        report_to = "none",
    )
)

from unsloth.chat_templates import train_on_responses_only
trainer = train_on_responses_only(
    trainer,
    instruction_part =
        "<|start_header_id|>user<|end_header_id|>\n\n",
    response_part =
        "<|start_header_id|>assistant<|end_header_id|>\n\n",
)


# Show current memory stats

gpu_stats = torch.cuda.get_device_properties(0)
start_gpu_memory = round(
         torch.cuda.max_memory_reserved() / 1024 / 1024 / 1024, 3)
max_memory = round(gpu_stats.total_memory / 1024 / 1024 / 1024, 3)
print(f"GPU = {gpu_stats.name}. Max memory = {max_memory} GB.")
print(f"{start_gpu_memory} GB of memory reserved.")

trainer_stats = trainer.train()


# Inferenza con Streaming

FastLanguageModel.for_inference(model) # Enable 2x faster inference

messages = [
    {"role": "user", "content": "video of a tall guy jumping"},
]
```

```
inputs = tokenizer.apply_chat_template(
    messages,
    tokenize = True,
    add_generation_prompt = True,
    return_tensors = "pt",
).to("cuda")

from transformers import TextStreamer
text_streamer = TextStreamer(tokenizer, skip_prompt = True)
_ = model.generate(input_ids = inputs, streamer = text_streamer,
                   max_new_tokens = 128, use_cache = True,
                   temperature = 1.5, min_p = 0.1)


# Esporto in GGUF per Ollama

model.save_pretrained_gguf("/content/model_tmp", tokenizer,
                           quantization_method="f16")

# Save to q4_k_m GGUF
# model.save_pretrained_gguf("model", tokenizer,
                             quantization_method = "q4_k_m")

# Save to 8bit Q8_0
# model.save_pretrained_gguf("model", tokenizer,)


# Download the folder with Model

!zip -r /content/ModelfileFinal.zip /content/model_tmp
```

Bibliografia

[1] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[2] Oscar Sharp. Sunspring. Short film, 2016. Executive Producer: Walter Kortschak. Genres: Comedy, Science Fiction, Short, Experimental. Released on June 9, 2016.

[3] Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. Generating videos with scene dynamics. *Advances in neural information processing systems*, 29, 2016.

[4] Roberto Beccaria. LUCAS - Libera Unione di Cultura Arte e Spettacolo, 2002.

[5] Mengyi Wei, Chenjing Jiao, Chenyu Zuo, Lorenz Hurni, and Liqiu Meng. Constructing ai ethics narratives based on real-world data: Human-ai collaboration in data-driven visual storytelling. *arXiv preprint arXiv:2502.00637*, 2025.

[6] Andrea Rossini Pietro Lafiandra and Flavio Pizzorno. Miss polly had a dolly, 2023.

[7] Hooroo Jackson. Dreadclub: Vampire's verdict, 2024.

[8] 20th Century Studios. Morgan - ai edited trailer, 2016.

[9] Francesca Fini. Love, 2025.

[10] Cao Yiwen. What's next?, 2025.

[11] KNGMKRlabs. The first humans, 2024.

[12] Peter Luisi. The last screenwriter, 2024.

[13] Fabian Stelzer. Salt, 2024.

[14] NeoCinemaAi. The paperclip max – reimagined, 2024.

[15] Aaron McGurk and Chris Lapham. Ai p's & q's, 2024.

[16] Ruihan Zhang, Borou Yu, Jiajian Min, Yetong Xin, Zheng Wei, Juncheng Nemo Shi, Mingzhen Huang, Xianghao Kong, Nix Liu Xin, Shanshan Jiang, et al. Generative ai for film creation: A survey of recent advances. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 6267–6279, 2025.

[17] Samangi Wadinambiarachchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. The effects of generative ai on design fixation and divergent thinking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2024.

[18] Anil R Doshi and Oliver P Hauser. Generative ai enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290, 2024.

[19] Mi Zhou, Vibhanshu Abhishek, Timothy Derdenger, Jaymo Kim, and Kannan Srinivasan. Bias in generative ai. *arXiv preprint arXiv:2403.02726*, 2024.

[20] Abeer Ibrahim. Impact of using artificial intelligence in visual art performance: Artificial intelligence on the design industry. *Research Journal in Advanced Humanities*, 4:14, 07 2023.

[21] Glenn Marshall. The crow, 2022.

[22] Corridor Digital. Anime rock, paper, scissors, 2023.

[23] Lucasfilm. Star wars: Field guide, 2025.

[24] Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, Tobi Walsh, Armin Hamrah, Lapo Santarlasci, Julia Betts Lotufo, Alexandra Rome, Andrew Shi, and Sukrut Oak. The ai index 2025 annual report, April 2025.

[25] Aditi Singh. A survey of ai text-to-image and ai text-to-video generators. In *2023 4th International Conference on Artificial Intelligence, Robotics and Control (AIRC)*, pages 32–36. IEEE, 2023.

[26] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Mmlu: Massive multitask language understanding, 2020.

[27] Allen Institute for AI and University of Washington. Wildbench, 2024.

[28] Howard Yen, Tianyu Gao, Minmin Hou, Ke Ding, Daniel Fleischer, Peter Izsak, Moshe Wasserblat, and Danqi Chen. Helmet: How to evaluate long-context models effectively and thoroughly. In *The Thirteenth International Conference on Learning Representations*, 2025.

[29] Wei-Lin Chiang, Lianmin Zheng, Ying Sheng, Anastasios Nikolas Angelopoulos, Tianle Li, Dacheng Li, Banghua Zhu, Hao Zhang, Michael Jordan, Joseph E Gonzalez, et al. Chatbot arena: An open platform for evaluating llms by human preference. In *Forty-first International Conference on Machine Learning*, 2024.

[30] Ziqi Huang, Fan Zhang, Xiaojie Xu, Yinan He, Jiashuo Yu, Ziyue Dong, Qianli Ma, Nattapol Chanpaisit, Chenyang Si, Yuming Jiang, et al. Vbench++: Comprehensive and versatile benchmark suite for video generative models. *arXiv preprint arXiv:2411.13503*, 2024.

[31] Dian Zheng, Ziqi Huang, Hongbo Liu, Kai Zou, Yinan He, Fan Zhang, Yuanhan Zhang, Jingwen He, Wei-Shi Zheng, Yu Qiao, et al. Vbench-2.0: Advancing video generation benchmark suite for intrinsic faithfulness. *arXiv preprint arXiv:2503.21755*, 2025.

[32] Serkan Sulun, Paula Viana, and Matthew EP Davies. Video soundtrack generation by aligning emotions and temporal boundaries. *arXiv preprint arXiv:2502.10154*, 2025.

[33] Iwona Christop, Marek Kubis, et al. Cloneval: An open voice cloning benchmark. *arXiv preprint arXiv:2504.20581*, 2025.

[34] MOSAIC Project Consortium. MOSAIC: Multilingual open science for ai communication, 2024.

[35] Yixin Liu, Kai Zhang, Yuan Li, Zhiling Yan, Chujie Gao, Ruoxi Chen, Zhengqing Yuan, Yue Huang, Hanchi Sun, Jianfeng Gao, et al. Sora: A review on background, technology, limitations, and opportunities of large vision models. *arXiv preprint arXiv:2402.17177*, 2024.

[36] OpenAI. Overview - openai api platform documentation, 2025.

[37] Ryan Imgrund. Prompt engineer and elevator - custom gpt on chatgpt, 2024.

[38] ruvnet. 300+ cinematic sora video prompts, 2024.

[39] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

[40] Yan Li. Impact of artificial intelligence on creative digital content production. *Journal of Digital Art Engineering and Multimedia*, 6(2):121, 2019.

[41] Wilson Yan, Danijar Hafner, Stephen James, and Pieter Abbeel. Temporally consistent transformers for video generation. In *International Conference on Machine Learning*, pages 39062–39098. PMLR, 2023.

[42] Nanna Inie, Jeanette Falk, and Steve Tanimoto. Designing participatory ai: Creative professionals' worries and expectations about generative ai. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2023.

[43] Kaveh Abhari and Elizabeth J Davidson. Creative co-production: the adaption of an open innovation model in creative industries. In *Information systems and management in media and entertainment industries*, pages 119–130. Springer, 2017.

[44] Jon McCormack, Toby Gifford, and Patrick Hutchings. Autonomy, authenticity, authorship and intention in computer generated art. In *International conference on computational intelligence in music, sound, art and design (part of EvoStar)*, pages 35–50. Springer, 2019.

[45] Polat Goktas. Ethics, transparency, and explainability in generative ai decision-making systems: A comprehensive bibliometric study. *Journal of Decision Systems*, pages 1–29, 2024.

[46] Yupeng Zhou, Daquan Zhou, Ming-Ming Cheng, Jiashi Feng, and Qibin Hou. Storydiffusion: Consistent self-attention for long-range image and video generation. *Advances in Neural Information Processing Systems*, 37:110315–110340, 2024.

[47] Thomas Unterthiner, Sjoerd Van Steenkiste, Karol Kurach, Raphael Marinier, Marcin Michalski, and Sylvain Gelly. Towards accurate generative models of video: A new metric & challenges. *arXiv preprint arXiv:1812.01717*, 2018.

[48] Canyu Zhao, Mingyu Liu, Wen Wang, Weihua Chen, Fan Wang, Hao Chen, Bo Zhang, and Chunhua Shen. Moviedreamer: Hierarchical generation for coherent long visual sequence. *arXiv preprint arXiv:2407.16655*, 2024.

[49] Preethi Seshadri, Sameer Singh, and Yanai Elazar. The bias amplification paradox in text-to-image generation. *arXiv preprint arXiv:2308.00755*, 2023.

[50] Jonas Oppenlaender. A taxonomy of prompt modifiers for text-to-image generation. *Behaviour & Information Technology*, 43(15):3763–3776, 2024.

[51] Sri Harsha Dumpala, Aman Jaiswal, Chandramouli Sastry, Evangelos Milios, Sageev Oore, and Hassan Sajjad. Sensitivity of generative vlms to semantically and lexically altered prompts. *arXiv preprint arXiv:2410.13030*, 2024.

[52] Bahar Kimanos and Anton Yderberg. On human perception: Opinions on anomalies found in ai-generated animation, 2024.

[53] Chen Bai, Zeman Shao, Guoxiang Zhang, Di Liang, Jie Yang, Zhuorui Zhang, Yujian Guo, Chengzhang Zhong, Yiqiao Qiu, Zhendong Wang, et al. Anything in any scene: Photorealistic video object insertion. *arXiv preprint arXiv:2401.17509*, 2024.

[54] Ye Yuan, Jiaming Song, Umar Iqbal, Arash Vahdat, and Jan Kautz. Physdiff: Physics-guided human motion diffusion model. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 16010–16021, 2023.

[55] Jamal Knight, Andrew Johnston, and Adam Berry. Artistic control over the glitch in ai-generated motion capture. *arXiv preprint arXiv:2308.08576*, 2023.

[56] Masahiro Mori. The uncanny valley: the original essay by masahiro mori. *Ieee Spectrum*, 6(1):6, 1970.

[57] Tingting Liao, Hongwei Yi, Yuliang Xiu, Jiaxiang Tang, Yangyi Huang, Justus Thies, and Michael J Black. Tada! text to animatable digital avatars. In *2024 International Conference on 3D Vision (3DV)*, pages 1508–1519. IEEE, 2024.

[58] Jun'ichiro Seyama and Ruth S Nagayama. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence*, 16(4):337–351, 2007.

[59] Anca Serbanescu et al. Human-ai system co-creativity for building narrative worlds. *IASDR 2023: Life-Changing Design*, 2023.

[60] Johnathon Hall and Damian Schofield. The value of creativity: Human produced art vs. ai-generated art. *Art and Design Review*, 13(1):65–88, 2024.

[61] Jeba Rezwana and Mary Lou Maher. Identifying ethical issues in ai partners in human-ai co-creation. *arXiv preprint arXiv:2204.07644*, 2022.

[62] Alexander Gerner. Towards ai aesthetics. *Semeiosis - transdisciplinary journal of semiotics*, 11:1–30, 12 2023.

[63] Ministero della Cultura. Focus cortometraggi, 2024.