



**POLITECNICO
DI TORINO**

Dipartimento di Automatica e Informatica
Collegio di Ingegneria Informatica, del Cinema e Meccatronica
Corso di Laurea in Ingegneria del Cinema e dei Mezzi di Comunicazione

**Progettazione di esperienze VR assistite dall'intelligenza
artificiale: " Locked up"**

Relatore: Chiar.mo Prof. Riccardo ANTONINO

Tesi di Laurea di:
Dario SCARAFUGGI
Matr. 280007

Anno Accademico 2022 / 2023

Indice

Introduzione	I
Presentazione dell'argomento e del contesto	I
Domande ed obiettivi di ricerca	II
Struttura della tesi	III
Revisione della letteratura	1
Introduzione all'intelligenza artificiale	1
Storia dell'Intelligenza Artificiale.....	1
Concetti chiave dell'Intelligenza Artificiale.....	2
Tecniche e Algoritmi di Intelligenza Artificiale	3
Sfide e Prospettive Future dell'Intelligenza Artificiale	4
Reti neurali artificiali (Artificial Neural Network)	4
Reti neurali profonde (Deep Neural Network).....	7
Introduzione all'attuale workflow VR	7
Applicazioni teoriche dell'IA nella visualizzazione, modellazione 3D	10
Applicazioni teoriche dell'IA nel texturing.....	11
Stilizzazione e trasferimento di stile	14
Trasferimento neurale di stile:	14
Tecniche di trasferimento di stile e trasferimento neurale di stile	15
Sezione 1: Tecniche di trasferimento di stile.....	16
Sezione 2: Trasferimento neurale di stile	16
L'ambiente di sviluppo.....	18
Unity	18
Vantaggi di Unity:	18
Unreal Engine.....	20
Presentazioni delle principali IA valutate per il workflow	23
IA basate su immagini.....	23
DALL-E	23

Midjourney	26
Stable Diffusion.....	29
AI Testuali	34
ChatGPT.....	34
Metodologia	37
Descrizione del progetto "Locked Up”	37
Presentazione del lavoro svolto in precedenza.....	40
Processo di integrazione dell’IA nel workflow	45
Design dell’esperienza	45
Concettualizzazione Artistica	55
Modelling e Texturing	61
Fase di Generazione e Prototipazione del Gioco	66
Risultati.....	67
Conclusioni	76

Introduzione

Presentazione dell'argomento e del contesto

“Progettazione di esperienze in realtà virtuale assistite dall’intelligenza artificiale: locked Up”
sottende ad una domanda: quale sarà e quali potranno essere i cambiamenti al processo produttivo di un’esperienza VR?

Partendo dalla revisione di un progetto VR ancora in fase di sviluppo al quale ho preso parte per curare il texturing, del quale spiegherò in seguito, spinto dalle nuove tecnologie mi sono domandato come fosse possibile rendere i risultati attesi al meglio. Ho ricercato nel campo della generazione procedurale e dell’intelligenza artificiale nuove metodologie applicative che, già ad oggi, possono essere messe in pratica in un’esperienza VR.

Locked Up è nata da una collaborazione tra Robin Studio, il Politecnico di Torino (con le figure, come registi e ideatori, di Alessandro VISCONTI e Antonio MESSINA) e La Biennale di Venezia, lo scopo di quest’esperienza è quello di presentare e far immedesimare l’utente con il protagonista, un adolescente Hikikomori non binary di nome Robin, e del quale vestirà in prima persona i panni. Robin è quindi un giovane che a causa delle ansie sociali decide di rinchiudersi nelle sicure e accoglienti mura domestiche abbandonando la vita al di fuori di esse.

A partire quindi dalle specifiche delle texture e dei modelli low poli da realizzare, per i quali ho prima impostato un flusso di lavoro convenzionale poi procedurale, al fine di poter uniformare lo stile dell’esperienza a quello degli artboard. Essendomi sempre appassionato alla generazione di contenuti in maniera autonoma, avendo per questo corso di studi approfondito tutti i concetti alla base e le principali modalità operative della linea di produzione di questo tipo di esperienze, ho cercato di spingermi oltre, cercando, quando possibile, di automatizzare il processo creativo in modo che mi portasse al risultato atteso ma che potesse anche essere slegato ed implementato in ordine opposto al flusso di ragionamento che mi ci ha portato (quindi prima generare un concept e, seguendo tutta la catena, da questo giungere agli asset per il motore di rendering) facendomi stimolare dalle nuove frontiere tecnologiche in continua evoluzione legate all’Intelligenza Artificiale.

Domande ed obiettivi di ricerca

Come anticipato, la realizzazione delle texture e la cura dell'estetica della modellazione sono stati input che mi hanno portato a domandare se vi fosse un metodo più efficace e veloce per rendere il risultato atteso, ovverosia, quello di mantenere, integrare ed ampliare lo stile grafico dei modelli dell'esperienza.

Si andrà quindi a capire se sia applicabile, e in che maniera, una metodologia automatizzata con programmi attualmente in sviluppo e disponibili al pubblico sul mercato (quanto più in forma gratuita o a basso impatto economico) per giungere ad un prodotto che soddisfi le specifiche e valutandone gli esiti ottenuti.

Lo scopo di questa tesi è pertanto quello di tracciare un solco operativo e dimostrare come l'intelligenza artificiale possa snellire, velocizzare e migliorare il processo generativo e creativo per la realizzazione di esperienze immersive.

Il mio attuale lavoro e la mia passione in merito all'argomento della realtà virtuale mi hanno spinto a ricercare in questo campo al fine di rendere possibile la creazione di ambienti sempre più coinvolgenti e interessanti con l'ausilio dell'intelligenza artificiale.

Le intelligenze artificiali sono in continuo sviluppo, la stessa Chat GPT, nella stesura di questa tesi ha avuto uno scatto di performance non indifferente dalla versione 3.5 alla 4.0.

Pertanto, la trattazione bibliografica dalla quale reperire informazioni è relativa a metodi che da un mese all'altro possono rapidamente diventare inadeguati, per tale ragione lo studio nel settore è stato spinto dalla ricerca sul campo.

L'analisi stessa del workflow è stata dimensionata ai maggiori attori in campo, a scapito di alcune seppur valide alternative, ma di realtà minori, per la stabilità e il supporto che i colossi del settore possono fornire, in termini di assistenza, risultati e frequenza di aggiornamento.

La tesi vuole pertanto gettare delle fondamenta e fornire accorgimenti per impostare il modus operandi nel settore nonché, conscio della rapidità quasi esponenziale della crescita di queste AI, fornire importanti spunti per la realizzazione o ottimizzazione di processi. Lo scopo è quindi quello di poter orientare e guidare il flusso di lavoro in ottica creativa in un ambiente in costante crescita.

Struttura della tesi

Questa tesi è suddivisa in tre capitoli (escludendo questa sezione introduttiva): il primo è una revisione della letteratura in merito alla tesi e mostrerà il funzionamento a grandi linee degli strumenti usati, il secondo mostrerà il flusso per la generazione di contenuto a partire dallo studio svolto per la creazione di texture per il progetto, il terzo capitolo valuterà i risultati e trarrà le conclusioni del progetto ipotizzando altre applicazioni o futuri approfondimenti.

In coda segue un'appendice: per la bibliografia.

Per quanto riguarda la revisione della letteratura ho deciso di non entrare troppo nel merito del funzionamento delle reti neurali o dei processi di trasferimento di stile, come, in generale, di sviscerare nei minimi dettagli operativi il processo creativo di un'esperienza VR (che è stato una delle principali trattazioni del mio corso di studio) tuttavia questa sezione vuole, per chi si avvicina a queste tecnologie, presentarne la struttura ed il contesto al fine di completare e comprendere meglio le idee e gli strumenti utilizzati.

La metodologia non segue il processo deduttivo con il quale ho costruito questa tesi ma presenta, a partire dai miei risultati, il flusso di lavoro in ordine, assumendo che si parta dall'inizio del progetto, quando ancora si sta definendo il carattere dell'esperienza, suddiviso per fasi dal concept sino alla generazione degli asset in Unity, concludendo con alcune applicazioni e suggerimenti di sviluppo del codice sulle render engines.

La parte di risultati ha prodotto esiti maggiormente misurabili, avendo un possibile confronto della metodologia applicativa, come anticipato, in merito alla parte di texture ed estetica 3D, non avendo io lavorato alle altre parti in prima persona, posso solo mostrare quelli che sono i frutti, documentare il tempo speso e confrontare con il risultato ottenuto in precedenza.

Revisione della letteratura

Introduzione all'intelligenza artificiale

Di seguito andremo ad analizzare a grandi linee alcuni concetti base per comprendere alcune delle strutture su cui si fonda questa trattazione. Partendo da una definizione piuttosto consolidata si può dire che:

“L'intelligenza artificiale IA è un ramo dell'informatica che si occupa della creazione di sistemi e software in grado di eseguire compiti che normalmente richiedono l'intelligenza umana, questi compiti includono apprendimento, ragionamento, percezione, comprensione del linguaggio naturale e l'interazione con l'ambiente”.

“Russell, S. J., & Norvig, P. (2016). Artificial intelligence: A modern approach. Pearson.”

Questa definizione, piuttosto generale, racchiude però il senso profondo dell'intelligenza artificiale ovvero eseguire i compiti che normalmente richiedono l'intelligenza umana.

Sino ad oggi, soprattutto nel campo del linguaggio naturale, era difficile si potesse pensare che una macchina avesse possibilità di superare un test di Turing o che un artista o un fotografo potesse sentire il suo lavoro minacciato nella forma in cui lo conosciamo.

L'IA ha fatto progressi significativi negli ultimi anni e ha trovato e sta trovando sempre più applicazioni in diversi settori, spesso e volentieri anche molto complessi dal punto di vista intellettuale, tra cui medicina, finanza, ingegneria e intrattenimento.

In questa breve introduzione andremo ad esplorare la storia, i concetti e le diverse tecniche di intelligenza artificiale.

Storia dell'Intelligenza Artificiale

La ricerca sull'IA ha avuto inizio negli anni '50 del XX secolo, con l'obiettivo di creare macchine che potessero imitare l'intelligenza umana¹ più precisamente la

storia dell'intelligenza artificiale, intesa come sperimentazione scientifica, inizia con la conferenza Dartmouth Summer Research Project on Artificial Intelligence del 1956 (in seguito indicata come "Conferenza di Dartmouth"). Alan Turing, un pioniere dell'informatica, propose il test, che poi prese il suo nome, come un criterio per determinare se una macchina è in grado di pensare come un essere umano². Nel corso degli anni, la ricerca sull'IA ha attraversato diverse fasi, tra cui periodi di ottimismo e di scetticismo, ma ha continuato a progredire e a svilupparsi.

Basandosi sul paradigma dell'Intelligenza Artificiale Debole, che vedremo nella prossima sezione, a partire dagli anni Ottanta, sono state sviluppate le prime applicazioni di Intelligenza Artificiale in ambito industriale.

In particolare, la prima intelligenza artificiale applicata in ambito commerciale fu R1, sviluppata nel 1982 dall'azienda Digital Equipment per configurare gli ordini di nuovi computer. Quattro anni dopo l'azienda era in grado di risparmiare 40 milioni di dollari all'anno.

Nel 1997 l'intelligenza artificiale finì sui giornali di tutto il mondo dopo che il calcolatore Deep Blue fu in grado di vincere una partita a scacchi contro il campione del mondo in carica Garry Kasparov.

Successivamente, sistemi di IA, furono in grado di vincere contro un essere umano anche a Go e ad altri giochi.

Nel 2011 Sebastian Thrun and Peter Norvig organizzarono, con la Stanford University, un corso online denominato Artificial Intelligence. Con 160 mila iscritti da tutto il mondo fu il primo MOOC di grande successo e contribuì a far conoscere lo stato dell'arte dell'IA del 21° secolo.

Più recentemente, nel 2018, l'Università di Helsinki ha reso disponibile Elements of AI un corso interattivo, tradotto anche in italiano.

Nel 2020 inizia la ricerca avanzata di Open AI, l'anno successivo esce Dall-E, ma è il 2022 l'anno delle AI in cui escono praticamente tutti i tool usati nella tesi.

Concetti chiave dell'Intelligenza Artificiale

L'intelligenza artificiale può essere suddivisa in due categorie principali: l'IA debole e l'IA forte³. L'IA debole si riferisce a sistemi che sono progettati per eseguire compiti specifici, circoscritti e non possiedono una vera intelligenza o coscienza. Esempi di IA debole includono sistemi di riconoscimento vocale, motori di ricerca e assistenti virtuali. D'altro canto, l'IA forte, si riferisce a sistemi

che possiedono una vera intelligenza e coscienza simili a quelle degli esseri umani. Attualmente, l'IA forte rimane ancora un obiettivo teorico e non è stata ad oggi interamente realizzata.

Un altro concetto chiave nell'IA è l'apprendimento automatico o Machine Learning (ML), un sottoinsieme dell'IA che si concentra sull'abilità delle macchine di imparare da dati senza essere state esplicitamente programmate per lo specifico caso presentato facendo quindi assunzioni e associazioni⁴. L'apprendimento automatico ha portato a importanti progressi nell'IA, consentendo ai sistemi di migliorare le proprie prestazioni nel tempo attraverso gli errori ed i successi, quello che noi definiamo: l'esperienza.

Tecniche e Algoritmi di Intelligenza Artificiale

L'intelligenza artificiale utilizza una varietà di tecniche e algoritmi per risolvere problemi e imitare l'intelligenza umana. Alcune delle principali tecniche includono:

Reti neurali artificiali (Artificial Neural Network): Le ANN sono modelli computazionali ispirati al funzionamento del cervello umano e sono costituite da unità di elaborazione interconnesse chiamate neuroni artificiali⁵. Le reti neurali sono in grado di apprendere e adattarsi ai dati, consentendo loro di eseguire compiti come il riconoscimento di immagini, la traduzione automatica e il riconoscimento vocale.

Reti neurali profonde (Deep Neural Network): sono una classe avanzata di ANN, caratterizzate da molteplici strati nascosti, che permettono l'apprendimento di rappresentazioni complesse e astratte dei dati⁶.

(Queste due reti saranno trattate più approfonditamente nel successivo capitolo).

Algoritmi genetici: Gli algoritmi genetici sono ispirati alla teoria dell'evoluzione e utilizzano operatori genetici come la selezione, la mutazione e il crossover per cercare soluzioni ottimali a problemi complessi⁷. Questi algoritmi sono particolarmente utili nella risoluzione di problemi di ottimizzazione e di ricerca nello spazio delle soluzioni.

Sistemi esperti: I sistemi esperti sono programmi basati su conoscenza che imitano il ragionamento di un esperto umano in un dominio specifico, utilizzando regole e rappresentazioni simboliche⁸. Questi sistemi sono stati

utilizzati con successo in diverse applicazioni tra cui la diagnosi medica, la pianificazione finanziaria e il controllo di processi industriali.

Apprendimento per rinforzo (Reinforcement Learning RL):

L'apprendimento per rinforzo è un tipo di apprendimento automatico in cui un agente impara a compiere azioni in un ambiente per massimizzare una ricompensa cumulativa nel tempo⁹. Questo approccio è stato utilizzato con successo in una varietà di applicazioni, tra cui il controllo robotico, la gestione delle risorse e l'addestramento di agenti intelligenti in giochi e simulazioni.

Sfide e Prospettive Future dell'Intelligenza Artificiale

Nonostante i notevoli progressi nell'IA, esistono ancora diverse sfide e problemi aperti che devono essere affrontati. Tra questi vi sono, nonostante i grandi sviluppi in merito ad una profonda comprensione del linguaggio naturale, la visione artificiale, la robotica e l'intelligenza emotiva¹⁰. Inoltre, l'IA forte, che implica la creazione di macchine dotate di una vera intelligenza e coscienza, rimane, ad oggi, un obiettivo teorico.

Le prospettive future dell'intelligenza artificiale sono promettenti, con un crescente interesse e investimenti nel campo. Si prevede che l'IA avrà un impatto significativo su numerosi settori, tra cui la medicina, l'istruzione, il trasporto e la sicurezza¹¹. Tuttavia, è importante considerare anche i potenziali rischi e le implicazioni etiche associate all'IA, come la disoccupazione, la privacy (vedasi il recente *primato* italiano su Chat GPT) e la sicurezza¹².

In conclusione, l'intelligenza artificiale è un campo di ricerca in continua evoluzione che mira a creare sistemi e software in grado di eseguire compiti che richiedono l'intelligenza umana. La ricerca sull'IA ha portato a sviluppi significativi in diversi settori e ha introdotto una serie di tecniche e algoritmi innovativi. Nonostante le sfide e i problemi aperti, l'IA ha il potenziale per trasformare numerosi aspetti della vita quotidiana e del lavoro.

Reti neurali artificiali (Artificial Neural Network)

Nel nostro caso andremo ad interfacciarci con intelligenze artificiali piuttosto “evolute” basate sulle reti neurali profonde, pertanto, cercherò di fornire una breve introduzione in merito.

Come già anticipato, una rete neurale è un modello di intelligenza artificiale ispirato al funzionamento del cervello umano.

È composta da una serie di nodi (i neuroni artificiali) interconnessi tra di loro, tramite funzioni di connessione “sinapsi” che elaborano l'informazione in modo distribuito e parallelo. Ogni nodo riceve una serie di input (dalle altre unità della rete) che vengono pesati in base all'importanza e successivamente combinati per produrre un output.

Le reti neurali possono essere addestrate su un insieme di dati di esempio per apprendere una determinata funzione o per identificare pattern all'interno dei dati stessi. Questo processo di apprendimento viene effettuato tramite l'aggiustamento dei pesi delle connessioni tra i nodi le “sinapsi”, in modo da minimizzare una funzione di errore.

Le reti neurali possono essere utilizzate in una vasta gamma di applicazioni, tra cui la classificazione o generazione di immagini, il linguaggio naturale etc.

Per comprendere meglio il concetto alla base delle reti neurali possiamo fornire un esempio:

Prendiamo un problema di classificazione di immagini, dove vogliamo che la rete neurale classifichi un'immagine in una delle 10 categorie (ad esempio, "cane", "gatto", "auto", "bicicletta", ecc.).

il problema si compone di diversi problemi da risolvere: “capire cosa sto vedendo”, “fornire un responso” e “classificare”.

Per “vedere” dobbiamo convertire le informazioni in un formato accettato da codice, pertanto, il nostro primo passo sarà quello di convertire l'immagine.

Per fare ciò, in questo caso, ci servirà applicare varie fasi di filtraggio attraverso filtri detti “filtri convoluzionali”, un'operazione matematica utilizzata in una rete neurale per l'estrazione delle caratteristiche dell'immagine. Il filtro è una matrice di valori numerici (pesi) che viene applicata all'immagine di input durante la prima fase (fase di convoluzione).

Durante quest'operazione, il filtro scorre sull'immagine di input e moltiplica i valori di pixel dell'immagine in ciascuna posizione per i corrispondenti pesi del

filtro. Il risultato della moltiplicazione viene poi sommato e assegnato alla posizione centrale di una matrice di output, chiamata mappa di caratteristiche.

L'utilizzo di filtri convoluzionali consente alla rete neurale di rilevare ed estrarre automaticamente le caratteristiche salienti delle immagini come i bordi, i contorni, i pattern e le forme. La dimensione e la forma del filtro convoluzionale determinano la specificità delle caratteristiche estratte.

In sintesi, i filtri convoluzionali consentono di estrarre le caratteristiche dell'immagine in modo automatizzato.

Fatto ciò, con i nostri dati, passiamo al secondo strato (un livello di pooling), che riduce la dimensione dell'immagine estratta dal livello precedente.

Questi strati convoluzionali e di pooling si ripetono per diversi strati, ognuno dei quali estrae sempre caratteristiche più complesse dell'immagine.

Dopo l'ultimo strato convoluzionale, l'immagine viene "appiattita" in un vettore unidimensionale.

A questo punto dobbiamo "pensare" cioè passare questo codice in una rete di "neuroni", nodi, come detto in precedenza collegati tra loro con "sinapsi", funzioni pesate, per fare ciò passeremo attraverso uno o più strati completamente connessi (o "dense").

Uno strato dense, o completamente connesso, è uno strato di una rete neurale in cui ogni neurone è connesso ad ogni neurone del layer successivo. In questo tipo di strato, ogni input ottenuto a partire dal filtro ha un peso che viene moltiplicato dal valore di attivazione del neurone corrispondente e quindi sommato agli altri input pesati. Questa somma viene poi passata attraverso una funzione di attivazione, che produce l'output del neurone.

In questa fase dobbiamo "classificare" ossia giungere ad un responso, per fare ciò il valore ricadrà dentro i parametri previsti oppure fuori da essi, a fare queste valutazioni sarà lo stato di output che utilizzerà queste caratteristiche per fare la classificazione finale in una delle 10 categorie.

Ciò che rende particolare la rete neurale è il fatto che i pesi delle sinapsi, i valori di output delle funzioni e i parametri possono subire variazioni in base alle validazioni e all'apprendimento.

Lo scopo pertanto è quello di settare in modo efficace questi parametri per far sì che il flusso del responso cada nelle caratteristiche già "viste" o "simili" e possa quindi riconoscere l'oggetto.

Reti neurali profonde (Deep Neural Network)

Una versione più complessa di queste ANN, sono le DNN, (Deep Neural Network) reti neurali artificiali composte da molti strati di elaborazione dei dati. Grazie alla loro profondità queste reti sono in grado di apprendere rappresentazioni gerarchiche sempre più complesse dei dati di input.

Le DNN hanno avuto successo in molte aree, tra cui il riconoscimento delle immagini, la traduzione automatica e il riconoscimento del parlato. Uno dei motivi per cui le DNN hanno avuto successo è il loro utilizzo in combinazione con l'apprendimento supervisionato¹³. In questo tipo di apprendimento, la rete è addestrata su un set di dati di input etichettati, ad esempio immagini di gatti e cani, e le corrispondenti etichette di output.

Oltre all'apprendimento supervisionato, le DNN sono state utilizzate anche in altri tipi di apprendimento automatico, come l'apprendimento non supervisionato e l'apprendimento rinforzato (DRL). Nell'apprendimento non supervisionato la rete è addestrata su dati di input non etichettati, ad esempio immagini senza etichetta. La rete deve quindi trovare modelli e strutture nei dati di input senza la guida di etichette di output¹⁴.

Nell'apprendimento rinforzato, la rete apprende attraverso il feedback da un ambiente esterno, ad esempio un gioco. La rete è ricompensata quando fa scelte giuste e penalizzata quando fa scelte sbagliate. L'obiettivo è trovare una strategia di azione che massimizzi la ricompensa a lungo termine¹⁵.

Le DNN richiedono una grande quantità di dati e risorse computazionali per essere addestrate, ma i recenti progressi nella tecnologia hardware e software hanno reso possibile l'addestramento di reti ancora più profonde e complesse. Le DNN sono diventate uno strumento fondamentale e saranno queste quelle utilizzate per le generazioni di contenuti nei successivi paragrafi.

Introduzione all'attuale workflow VR

Creare un'esperienza di realtà virtuale attualmente è un'opera complessa, varia a seconda delle piattaforme, degli strumenti usati e dei dispositivi target.

Tuttavia, di seguito, andrò a trattare un flusso di lavoro che può essere applicato come canovaccio senza troppo legarsi a specifiche di sistema.

Ideazione e progettazione: Prima di iniziare a creare l'esperienza VR, è fondamentale pianificare e progettare accuratamente il concept, gli obiettivi e gli elementi interattivi dell'esperienza. Ciò include la definizione del pubblico di destinazione, la creazione di storyboard e la progettazione di layout e interfacce utente. Per avere un'idea chiara di come strutturare un'esperienza VR, si possono consultare le linee guida di design per la realtà virtuale, come quelle fornite da Google¹⁶ e Oculus¹⁷.

A questo punto si avrà bisogno di uno strumento o di un motore di gioco che supporti lo sviluppo VR. Alcuni dei più popolari e quelli di cui tratteremo sono Unity¹⁸ e Unreal Engine¹⁹. È quindi necessario scegliere il motore che si adatta meglio alle proprie esigenze e competenze. Ogni piattaforma offre documentazione e risorse specifiche per aiutare a iniziare con lo sviluppo VR²⁰
21.

Dopo la concettualizzazione si deve passare alla creazione di risorse 3D, essa può essere suddivisa nei seguenti passaggi:

Concept art e design: Si deve iniziare con la creazione di concept art e bozzetti per i modelli 3D, come personaggi, oggetti e ambienti. Questo aiuta a stabilire uno stile visivo coerente e a identificare le esigenze specifiche di modellazione, texturing e animazione

Modellazione 3D: Si devono creare modelli 3D utilizzando software di modellazione come Blender²², Maya²³ o 3ds Max²⁴. È corretto mantenere la geometria dei modelli il più semplice possibile per ottimizzare le prestazioni nella realtà virtuale, concentrandosi su un buon bilanciamento tra dettaglio e complessità.

UV mapping e texturing: Dopo aver completato la modellazione, si procede col creare una mappa UV per ogni modello. La mappa UV consente di applicare texture 2D ai modelli 3D in modo corretto e preciso. Successivamente, si devono creare le texture utilizzando software di pittura e fotoritocco come Adobe Photoshop²⁵, Substance Painter²⁶ o Quixel Mixer²⁷. Per la realtà virtuale, è importante solitamente utilizzare texture ad alta risoluzione e tecniche di

texturing avanzate, come il PBR (Physically Based Rendering), per ottenere un aspetto realistico e dettagliato.

Rigging e animazione: Se i modelli 3D richiedono animazione, è necessario creare uno scheletro (rig) e assegnare i pesi dei vertici ai giunti appropriati. Successivamente, si possono creare animazioni utilizzando tecniche come keyframing, motion capture o procedural animation. Per la realtà virtuale, è importante prestare attenzione all'ottimizzazione delle animazioni, riducendo il numero di keyframes e il peso dei file di animazione quando possibile.

Ottimizzazione delle risorse 3D: Prima di importare le risorse 3D nel motore di gioco, è opportuno verificare che siano ottimizzate per la realtà virtuale. Ciò può includere la riduzione della geometria, il riutilizzo di texture e l'utilizzo di tecniche di LOD (Level of Detail).

Importazione nel motore di gioco: a questa fase si dovranno importare le risorse 3D nel motore di gioco selezionato e iniziare a costruire la scena o l'ambiente virtuale. Posizionare gli oggetti, le luci e le telecamere e creare le interazioni e le animazioni necessarie. Ogni motore di gioco offre strumenti e processi specifici per l'importazione e l'integrazione delle risorse che vanno considerate a monte di questa fase per rendere questo processo quanto più snello possibile.

Implementazione di algoritmi VR: in seguito, si passa da quella che è un'esperienza desktop a quella che è una vera esperienza immersiva utilizzando gli SDK (Software Development Kits) forniti dai produttori di dispositivi VR, come Oculus²⁸, HTC Vive²⁹ o Valve Index³⁰, per integrare il supporto per i controller e le interazioni specifiche del dispositivo. La documentazione fornita dai produttori di dispositivi VR spiegherà come utilizzare gli SDK e come implementare correttamente le interazioni per garantire una buona esperienza utente consentendo di passare da una prima persona desktop a una renderizzazione a doppio “schermo” bi focale sul display target.

Test e review: l'ultimo passaggio è il test dell'esperienza VR su diversi dispositivi e verificare che le interazioni funzionino correttamente e che le prestazioni siano ottimali. Apportando modifiche e ottimizzazioni alle risorse,

alle interazioni e alla logica di gioco se necessario. Sarà opportuno utilizzare strumenti come il profilo di rendering (profiler) e le statistiche di prestazioni offerte dai motori di gioco per identificare eventuali colli di bottiglia e migliorare le prestazioni dell'applicazione. Inoltre, va tenuto presente che diversi dispositivi VR potrebbero richiedere diversi livelli di ottimizzazione, quindi, bisogna assicurarsi di testare l'esperienza su una varietà di hardware.

Infine, una volta completata l'esperienza VR, si deve prevedere una fase di pubblicazione ed eventuale distribuzione attraverso le piattaforme di distribuzione specifiche del dispositivo, come Oculus Store³¹ o SteamVR³². Ogni piattaforma ha requisiti e linee guida specifici per la pubblicazione e la distribuzione ³³³⁴³⁵.

Applicazioni teoriche dell'IA nella visualizzazione, modellazione 3D

Con l'intelligenza artificiale possiamo snellire alcuni dei passaggi indicati nella fase di concept, cuore generativo dell'esperienza VR, e farci supportare in vari compiti dalla stessa.

La bibliografia in merito subisce sempre rapidi aggiornamenti, pertanto, la mia trattazione conterrà perlopiù citazioni su articoli.

Uno dei contributi più grandi ed inediti che l'IA nel 2022 può offrire è la generazione di contenuti. L'IA può essere utilizzata per generare automaticamente modelli 3D, ambienti e texture, risparmiando tempo e risorse. Ad esempio, gli algoritmi di machine learning possono essere addestrati per creare modelli 3D di edifici o terreni basati su dati reali o parametri specifici.

Denoising e miglioramento delle immagini: L'IA può essere utilizzata per ridurre il rumore nelle immagini renderizzate e migliorare la qualità delle texture. Ad esempio, NVIDIA ha sviluppato un algoritmo di denoising basato sull'IA chiamato OptiX AI-Accelerated Denoiser, che migliora significativamente la qualità delle immagini renderizzate in tempo reale³⁶.

Ottimizzazione della geometria: L'IA può essere utilizzata per semplificare automaticamente la geometria dei modelli 3D, mantenendo al contempo una buona approssimazione dell'aspetto originale. Ciò può migliorare le prestazioni

e ridurre la quantità di memoria necessaria per visualizzare i modelli 3D, specialmente in applicazioni in tempo reale come videogiochi e realtà virtuale³⁷.

Riconoscimento e rigging automatico dei personaggi: L'IA può essere utilizzata per riconoscere automaticamente la struttura scheletrica dei modelli 3D dei personaggi e applicare automaticamente un rig adeguato³⁸. Ciò può accelerare il processo di rigging e garantire risultati più accurati e consistenti³⁹.

Generazione automatica di animazioni: L'IA può essere utilizzata per generare automaticamente animazioni realistiche per i personaggi e gli oggetti 3D⁴⁰. Ad esempio, gli algoritmi di machine learning possono essere addestrati per creare animazioni di camminata, corsa o salto basate su dati di motion capture o animazioni esistenti⁴¹.

Questi sono solo alcuni esempi delle applicazioni dell'IA nella visualizzazione, modellazione 3D. L'uso dell'intelligenza artificiale in questi ambiti sta diventando sempre più comune grazie ai miglioramenti nella potenza di calcolo e ai progressi nella ricerca sull'apprendimento automatico e profondo. Man mano che l'IA continua a evolversi è probabile che vedremo ulteriori innovazioni e applicazioni in questi settori che possono semplificare e migliorare il processo di creazione di contenuti 3D, la qualità delle immagini e delle animazioni prodotte.

Applicazioni teoriche dell'IA nel texturing

L'intelligenza artificiale (IA) ha avuto un impatto significativo sul texturing e sullo stile artistico nei settori della computer grafica, della modellazione 3D e della creazione di contenuti digitali. L'IA ha rivoluzionato il modo in cui gli artisti e gli sviluppatori creano texture, stili e design unici. Di seguito sono riportati alcuni degli impatti principali dell'IA sul texturing e sullo stile artistico:

Generazione automatica di texture: L'IA può essere utilizzata per generare automaticamente texture dettagliate e realistiche da un'immagine di input o da dati grezzi, riducendo il tempo e lo sforzo necessari per creare texture manualmente. Algoritmi come i GAN (Generative Adversarial Networks)

possono essere addestrati per produrre texture realistiche e coerenti per una vasta gamma di materiali e oggetti⁴²

Miglioramento delle texture: L'IA può essere utilizzata per migliorare la qualità delle texture esistenti, ad esempio aumentando la risoluzione delle immagini o ripristinando dettagli persi a causa della compressione. Tecniche come l'upscaling basato sull'apprendimento profondo e il super-resolution possono migliorare la qualità delle texture senza la necessità di ricreare manualmente l'immagine ad alta risoluzione⁴³.

Trasferimento di stile e stilizzazione: L'IA ha reso possibile il trasferimento di stile tra diverse immagini o modelli 3D, consentendo agli artisti di applicare lo stile di un'immagine di riferimento a un'altra immagine o a un modello 3D. Tecniche come il Neural Style Transfer utilizzano reti neurali convoluzionali (CNN) per analizzare lo stile di un'immagine di riferimento e applicarlo a un'altra immagine⁴⁴.

Creazione di stili artistici unici: L'IA può essere utilizzata per generare stili artistici unici basati su un insieme di dati di addestramento. Ad esempio, gli artisti possono addestrare un algoritmo di IA su un insieme di opere d'arte in uno stile specifico e utilizzare l'algoritmo per creare nuove opere d'arte in quel particolare stile. Questo consente agli artisti di esplorare nuove tecniche e stili che potrebbero non essere stati possibili senza l'IA⁴⁵.

Assistenza nella creazione di contenuti: Gli strumenti basati sull'IA possono aiutare gli artisti a creare contenuti più rapidamente e facilmente, fornendo suggerimenti e automazione per il processo di texturing e stilizzazione. Ad esempio, gli strumenti di IA possono suggerire automaticamente texture o materiali appropriati per un modello 3D in base al contesto o all'ambiente circostante.

Adattamento delle texture: Gli algoritmi di IA possono adattare automaticamente le texture agli oggetti e agli ambienti in tempo reale, consentendo un maggiore grado di realismo e coerenza nelle scene virtuali. Questo può essere particolarmente utile per applicazioni come i videogiochi, dove gli oggetti e gli ambienti devono essere rappresentati in modo realistico e coerente in diverse condizioni di illuminazione e di prospettiva[7].

Stilizzazione in tempo reale: L'IA ha reso possibile la stilizzazione in tempo reale di contenuti 3D e video, permettendo agli artisti e ai creatori di applicare stili artistici diversi in tempo reale durante la riproduzione o l'interazione con i contenuti. Questo può essere utilizzato in applicazioni come i videogiochi, la realtà virtuale e la realtà aumentata per creare esperienze visive uniche e coinvolgenti.

Collaborazione tra artisti e IA: L'intelligenza artificiale può essere utilizzata come uno strumento collaborativo tra artisti e macchine, dove l'IA può suggerire idee, stili e tecniche basate sull'input dell'artista e viceversa. Ciò consente una maggiore sperimentazione e innovazione nel campo dell'arte digitale e della computer grafica.

Personalizzazione dei contenuti: Gli algoritmi di IA possono essere utilizzati per personalizzare automaticamente le texture e gli stili in base alle preferenze degli utenti o ai dati demografici. Questo può portare a esperienze più coinvolgenti e personalizzate per gli utenti finali, sia che si tratti di applicazioni di intrattenimento, come i videogiochi, o di applicazioni professionali, come la formazione e la simulazione.

Stilizzazione e trasferimento di stile

La stilizzazione e il trasferimento di stile sono tecniche che utilizzano l'intelligenza artificiale per applicare lo stile visivo di un'immagine di riferimento a un'altra immagine o texture. Questi metodi sono basati su algoritmi di deep learning, in particolare sulle reti neurali convoluzionali Convolutional Neural Network (CNN), che sono state addestrate per estrarre e manipolare le caratteristiche stilistiche delle immagini.

Trasferimento neurale di stile:

Neural Style Transfer: Il trasferimento di stile neurale è una delle tecniche più popolari per la stilizzazione e il trasferimento di stile. Il metodo, introdotto da Gatys et al. nel 2015⁴⁶, utilizza una CNN pre-addestrata per estrarre le caratteristiche stilistiche e di contenuto da due immagini di input: un'immagine di riferimento per lo stile e un'immagine di riferimento per il contenuto. L'algoritmo combina poi queste caratteristiche per generare un'immagine di output che mantiene il contenuto dell'immagine di riferimento per il contenuto e lo stile dell'immagine di riferimento per lo stile.

Fast Style Transfer: Sebbene il trasferimento di stile neurale originale produca risultati di alta qualità, il processo di ottimizzazione può essere lento e computazionalmente costoso. Per affrontare questo problema, Johnson et al. hanno introdotto una tecnica chiamata "fast style transfer" nel 2016⁴⁷. Questo metodo addestra una CNN per eseguire il trasferimento di stile in una sola passata, riducendo notevolmente il tempo di elaborazione rispetto al trasferimento di stile neurale originale.

Arbitrary Style Transfer: L'arbitrary style transfer è una tecnica che consente di applicare lo stile di un'immagine a un'altra immagine in modo arbitrario, senza dover addestrare una rete specifica per ogni combinazione di stile e contenuto. Un esempio di questa tecnica è il metodo proposto da Huang e Belongie nel 2017⁴⁸, che utilizza un approccio chiamato "adaptive instance normalization" per combinare le caratteristiche stilistiche e di contenuto delle immagini.

Stilizzazione video: La stilizzazione e il trasferimento di stile possono anche essere applicati a video e animazioni, mantenendo al contempo la coerenza temporale tra i frame. Ad esempio, Ruder et al. hanno proposto un metodo per

il trasferimento di stile temporale coerente nel 2016⁴⁹, che estende il trasferimento di stile neurale originale per gestire sequenze video.

La stilizzazione e il trasferimento di stile basati sull'IA offrono nuove possibilità per la per la creazione artistica, il design e la generazione di contenuti. Essi possono essere utilizzati in una vasta gamma di applicazioni, tra cui:

Creazione di opere d'arte digitali: Gli artisti possono utilizzare la stilizzazione e il trasferimento di stile per creare opere d'arte uniche, combinando lo stile di artisti famosi o di altre immagini di riferimento con le proprie immagini e fotografie⁵⁰.

Modifica e miglioramento di texture per modelli 3D: La stilizzazione e il trasferimento di stile possono essere utilizzati per modificare e migliorare le texture utilizzate nei modelli 3D e nelle scene virtuali, ad esempio applicando uno stile particolare o creando variazioni di una texture esistente.

Stilizzazione di interfacce utente e applicazioni: I designer di interfacce utente e applicazioni possono utilizzare la stilizzazione e il trasferimento di stile per creare interfacce utente personalizzate e temi che si adattano alle preferenze stilistiche degli utenti⁵¹.

Pubblicità e marketing: Le aziende possono utilizzare la stilizzazione e il trasferimento di stile per creare materiali pubblicitari e promozionali accattivanti, ad esempio combinando lo stile del loro marchio con immagini di prodotti o scene⁵².

Animazione e film: La stilizzazione e il trasferimento di stile possono essere utilizzati per creare animazioni e film con uno stile visivo unico e distintivo, ad esempio combinando lo stile di diversi artisti o generi.

Man mano che la ricerca e lo sviluppo nell'IA e nel deep learning continuano a progredire, è probabile che vedremo ulteriori miglioramenti nelle tecniche di stilizzazione e trasferimento di stile, nonché nuove applicazioni e possibilità creative.

Tecniche di trasferimento di stile e trasferimento neurale di stile

Ecco alcuni esempi di tecniche di trasferimento di stile e trasferimento neurale di stile, divisi in due sezioni:

Sezione 1: Tecniche di trasferimento di stile

Image Analogies: L'approccio Image Analogies, introdotto da Hertzmann et al. nel 2001⁵³, è una delle prime tecniche di trasferimento di stile. Si basa sull'utilizzo di esempi di coppia di immagini, in cui una coppia di immagini mostra lo stile originale e lo stile target. La tecnica genera un'immagine stilizzata combinando la struttura dell'immagine di input con l'aspetto dell'immagine di stile, utilizzando una serie di filtri non lineari. Tuttavia, il metodo può essere lento e richiedere una grande quantità di memoria per elaborare immagini ad alta risoluzione.

Patch-based Style Transfer: L'approccio Patch-based Style Transfer, come quello proposto da Efros e Freeman nel 2001⁵⁴, consiste nell'utilizzo di piccoli segmenti o "patch" dell'immagine di stile per costruire un'immagine stilizzata. L'algoritmo cerca corrispondenze tra le patch dell'immagine di stile e le regioni dell'immagine di input e li combina per creare un'immagine di output. Questo metodo può produrre risultati visivamente interessanti, ma può anche essere lento e computazionalmente costoso.

Sezione 2: Trasferimento neurale di stile

Neural Style Transfer (NST): Il Neural Style Transfer di Gatys et al.⁵⁵ è una delle tecniche più popolari e influenti basate sull'apprendimento profondo per il trasferimento di stile. Si basa su reti neurali convoluzionali (CNN) pre-addestrate, come la VGG-19, per estrarre le caratteristiche stilistiche e di contenuto delle immagini. L'immagine stilizzata viene poi generata attraverso un processo di ottimizzazione che minimizza una funzione obiettivo composta da una combinazione di perdite di stile e contenuto.

Fast Style Transfer: Il Fast Style Transfer, proposto da Johnson et al. nel 2016⁵⁶, è una versione più veloce del Neural Style Transfer. Invece di utilizzare un processo di ottimizzazione per generare l'immagine stilizzata, il Fast Style Transfer addestra una CNN per eseguire il trasferimento di stile in una sola passata. Ciò riduce notevolmente il tempo di elaborazione rispetto al Neural Style Transfer originale.

Arbitrary Style Transfer: L'Arbitrary Style Transfer, come il metodo proposto da Huang e Belongie nel 2017⁵⁷, consente di applicare lo stile di un'immagine a un'altra immagine in modo arbitrario, senza dover addestrare una rete specifica

per ogni combinazione di stile e contenuto. Questo approccio utilizza una tecnica chiamata "adaptive instance normalization" per combinare le caratteristiche stilistiche e di contenuto delle immagini.

Stilizzazione video: Il Neural Style Transfer può essere esteso per stilizzare video e animazioni frame per frame, mantenendo la coerenza temporale tra i frame⁵⁸. Per esempio, Ruder et al. hanno proposto un metodo per il trasferimento di stile temporale coerente nei video⁵⁹. Questo approccio combina il Neural Style Transfer con un termine di regolarizzazione temporale che assicura la coerenza tra i frame adiacenti. Il risultato è un video stilizzato in cui lo stile viene applicato in modo uniforme e coerente lungo tutto il video.

Stilizzazione basata su semantica: Alcune tecniche di trasferimento di stile neurale tengono conto delle informazioni semantiche delle immagini per produrre risultati più accurati e dettagliati. Ad esempio, Chen et al. hanno presentato Deep Photo Style Transfer, che utilizza un termine di regolarizzazione semantica per preservare i dettagli dell'immagine di contenuto⁶⁰. Questo approccio è particolarmente utile per applicazioni come il miglioramento fotografico e la manipolazione di immagini.

Stilizzazione multi-stile: Tecniche come MUNIT (Multimodal Unsupervised Image-to-Image Translation) di Huang et al.⁶¹ e AdaIN (Adaptive Instance Normalization) di Ghiasi et al.⁶² consentono di combinare più stili in un'unica immagine. Questi metodi utilizzano un approccio di apprendimento non supervisionato per apprendere una varietà di stili e permettono di generare immagini con una combinazione arbitraria di stili.

In sintesi, le tecniche di trasferimento di stile e trasferimento neurale di stile offrono una vasta gamma di possibilità per la creazione artistica, il design, la generazione di contenuti e altre applicazioni. Man mano che la ricerca nel campo dell'apprendimento profondo e dell'intelligenza artificiale continua a progredire, è probabile che vedremo ulteriori miglioramenti e nuove tecniche emergenti nel campo del trasferimento di stile.

L'ambiente di sviluppo

Unity

Unity, è stato estensivamente trattato e usato durante tutto il mio corso di studi, quindi, per questo progetto, è stata la scelta più naturale, pertanto andrò a esplicitare cosa sia Unity e che potenzialità abbia ad oggi nel 2023.

Sviluppato da Unity Technologies, è un motore di gioco multiplatforma e un ambiente di sviluppo integrato (IDE) ampiamente utilizzato per la creazione di videogiochi, applicazioni interattive, simulazioni, visualizzazioni architettoniche e altre esperienze 3D per una varietà di piattaforme, come PC, console, dispositivi mobili e dispositivi VR e AR⁶³

Vantaggi di Unity:

Multiplatforma: Unity permette agli sviluppatori di creare e distribuire contenuti su diverse piattaforme, come PC, Mac, console, dispositivi mobili, VR e AR, utilizzando un'unica base di codice⁶⁴.

Ampia comunità e risorse didattiche: Unity vanta una vasta comunità di sviluppatori e numerose risorse formative, tra cui tutorial, forum, blog e corsi online, facilitando l'apprendimento e il supporto durante lo sviluppo⁶⁵.

Asset Store: Lo store di Unity offre risorse predefinite, come modelli 3D, animazioni, effetti sonori, script e altro, che possono essere acquistate e vendute, accelerando lo sviluppo e riducendo i costi di produzione ⁶⁶.

Programmazione visiva: Unity supporta la programmazione visiva tramite plug-in come Playmaker e Bolt, permettendo a designer e artisti di creare logiche di gioco e interazioni senza scrivere codice, facilitando la collaborazione tra membri del team con competenze diverse ⁶⁷.

Supporto nativo VR e AR: Unity offre un robusto supporto per lo sviluppo di applicazioni VR e AR, con integrazioni native per dispositivi popolari come Oculus Rift, HTC Vive, PlayStation VR e Microsoft HoloLens. Gli sviluppatori possono utilizzare le API e gli strumenti specifici di Unity per creare esperienze immersive ottimizzate per le piattaforme VR e AR⁶⁸.

Per quanto riguarda la VR, Unity offre diversi vantaggi:

Ottimizzazione delle prestazioni: Unity fornisce strumenti e funzionalità per ottimizzare le prestazioni delle applicazioni VR, come il rendering single-pass stereo, il rendering dinamico degli oggetti e l'occlusion culling⁶⁹, garantendo un'esperienza VR fluida e confortevole per gli utenti.

Interazioni VR: Unity semplifica lo sviluppo di interazioni VR con strumenti e framework come Unity XR Toolkit e SteamVR Plugin, che forniscono componenti predefiniti per gestire input dei controller, tracciamento della posizione, teletrasporto e altre interazioni comuni in VR⁷⁰.

Creazione di ambienti 3D: Unity è ottimo per la creazione di ambienti 3D e mondi virtuali immersivi, grazie alle sue potenti funzionalità di illuminazione, ombreggiatura e post-elaborazione⁷¹.

Supporto per l'audio spazializzato: Unity supporta l'audio spazializzato, essenziale per un'esperienza VR immersiva, consentendo agli sviluppatori di posizionare le sorgenti audio nello spazio 3D e gestire automaticamente l'audio spazializzato ⁷².

Collaborazione e prototipazione rapida: Unity offre strumenti per la collaborazione in tempo reale tra membri del team e la prototipazione rapida di idee e concetti. Ciò è particolarmente utile nello sviluppo di esperienze VR, poiché permette agli sviluppatori di iterare rapidamente e testare le interazioni e i design in un ambiente virtuale⁷³.

Compatibilità con l'intelligenza artificiale: Unity è compatibile con librerie e framework di intelligenza artificiale, come TensorFlow e ML-Agents, permettendo agli sviluppatori di integrare funzionalità di IA nelle loro applicazioni VR, come il riconoscimento vocale, la visione artificiale e l'apprendimento automatico ⁷⁴.

In sintesi, Unity è un'opzione popolare e potente per lo sviluppo di applicazioni VR grazie alle sue numerose funzionalità, al supporto multiplatforma, alla vasta comunità di sviluppatori e alle risorse di apprendimento disponibili. La sua flessibilità e i suoi strumenti nativi per VR consentono agli sviluppatori di creare esperienze virtuali immersive e ottimizzate per un'ampia varietà di dispositivi e piattaforme.

Unreal Engine

Nonostante Unity sia il fulcro come engine di questo progetto mi pare doveroso mettere per completezza di trattazione anche il principale competitor con le sue potenzialità all'interno di questa lista. Data la scalabilità di questo prodotto programmato in C++ con la possibilità però di integrare vari altri script e le nuove versioni uscite che sfruttano l'AI per ridurre in maniera autonoma alla complessità dei poligoni dell'ambiente di gioco non mi risulta difficile immaginare come esso potrebbe nel giro di poco tempo diventare leader nel settore.

Entrando nel merito il motore di gioco 3D e l'ambiente di sviluppo integrato (IDE) Unreal Engine, sviluppato da Epic Games, è famoso per la sua grafica di alta qualità e viene impiegato per creare videogiochi, simulazioni, visualizzazioni architettoniche e applicazioni di realtà virtuale (VR) e realtà aumentata (AR) su varie piattaforme, come PC, console di gioco, dispositivi mobili e dispositivi VR e AR⁷⁵

I vantaggi di Unreal Engine includono:

Grafica di qualità superiore: Grazie a funzioni avanzate come il ray tracing in tempo reale, l'illuminazione globale dinamica e sistemi di effetti visivi complessi, Unreal Engine è noto per la sua capacità di produrre grafiche di alta qualità e realismo⁷⁶

Multiplatforma: Come Unity, Unreal Engine consente agli sviluppatori di creare e distribuire facilmente contenuti su diverse piattaforme, tra cui PC, Mac, console di gioco, dispositivi mobili, dispositivi VR e AR, con un unico codice sorgente⁷⁷

Blueprint Visual Scripting: Unreal Engine offre il sistema Blueprint Visual Scripting, che permette agli sviluppatori di creare logiche di gioco e interazioni senza scrivere codice, facilitando la prototipazione rapida e la collaborazione tra membri del team con diverse competenze⁷⁸.

Asset Store e risorse gratuite: Con l'Unreal Marketplace, gli sviluppatori possono accedere a risorse predefinite e pronte all'uso, come modelli 3D, animazioni, effetti sonori e script. Epic Games offre periodicamente risorse

gratuite e di alta qualità tramite il programma Unreal Engine Free Monthly Content⁷⁹

Supporto nativo per VR e AR: Unreal Engine offre un solido supporto per lo sviluppo di applicazioni VR e AR, con integrazioni native per dispositivi popolari come Oculus Rift, HTC Vive, PlayStation VR e Microsoft HoloLens. Gli sviluppatori possono utilizzare le API e gli strumenti specifici di Unreal Engine per creare esperienze immersive e ottimizzate per le piattaforme VR e AR⁸⁰.

In termini di VR, Unreal Engine presenta numerosi vantaggi:

Ottimizzazione delle prestazioni: Unreal Engine fornisce strumenti e funzionalità per migliorare le prestazioni delle applicazioni VR, come il rendering instanced stereo, il forward rendering e l'occlusion culling⁸¹. Queste funzionalità sono cruciali per garantire un'esperienza VR fluida e confortevole agli utenti.

Interazioni VR avanzate: Unreal Engine offre strumenti e componenti predefiniti per gestire l'input dei controller VR, il tracciamento della posizione, le interazioni fisiche e altre interazioni comuni in VR, come il teletrasporto. Gli sviluppatori possono utilizzare il sistema di animazione e di ragdoll avanzato di Unreal per creare interazioni VR più realistici e dinamici⁸².

Creazione di ambienti 3D realistici: Grazie alla sua eccellente capacità di creare grafica di alta qualità, Unreal Engine è ideale per la creazione di ambienti 3D realistici e mondi virtuali immersivi. Gli sviluppatori possono sfruttare le potenti funzionalità di illuminazione, ombreggiatura e post-elaborazione di Unreal Engine per creare scenari visivamente accattivanti e coinvolgenti⁸³.

Audio spazializzato: Unreal Engine supporta l'audio spazializzato, che è fondamentale per creare un'esperienza VR immersiva. Gli sviluppatori possono posizionare le sorgenti audio nello spazio 3D e Unreal Engine gestirà automaticamente l'audio spazializzato, consentendo agli utenti di percepire la direzione e la distanza delle sorgenti audio all'interno dell'ambiente virtuale⁸⁴.

Collaborazione e prototipazione rapida: Unreal Engine offre strumenti per la collaborazione in tempo reale tra membri del team e la prototipazione rapida di idee e concetti. Questo è particolarmente utile nello sviluppo di esperienze

VR, in quanto permette agli sviluppatori di iterare rapidamente e testare le interazioni e i design in un ambiente virtuale⁸⁵.

Compatibilità con l'intelligenza artificiale: Unreal Engine supporta librerie e framework di intelligenza artificiale come TensorFlow e NVIDIA Deep Learning SDK, consentendo agli sviluppatori di integrare funzionalità di IA nelle loro applicazioni VR, come il riconoscimento vocale, la visione artificiale e l'apprendimento automatico⁸⁶.

In conclusione, Unreal Engine rappresenta un'eccellente opzione per lo sviluppo di applicazioni VR di alta qualità grazie alle sue avanzate funzionalità grafiche, al supporto multipiattaforma, ai potenti strumenti di interazione VR e all'integrazione con tecnologie di intelligenza artificiale. La sua flessibilità e i suoi strumenti nativi per VR consentono agli sviluppatori di creare esperienze virtuali immersive e ottimizzate per una vasta gamma di dispositivi e piattaforme.

Presentazioni delle principali IA valutate per il workflow

IA basate su immagini

Nel campo della generazione di immagini, che ci serviranno per alcuni degli scopi sopracitati, abbiamo tre principali attori, dei quali andrò a presentare in breve il funzionamento, i vantaggi e soprattutto i costi relativi all'uso. Tutti e tre hanno in comune alcune caratteristiche: sono tutti quanti reti neurali a apprendimento profondo (DNN) e tutti quanti hanno all'interno nel loro core un modulo di riconoscimento del testo.

DALL-E

È come ordine di rilascio il più “vecchio” quindi lo presenterò per primo. Sebbene, poi, non sia stato molto usato in questa tesi è piuttosto efficace e si è stabilizzato molto nella versione DALL-E 2.

A seguito della pubblicazione dei modelli NLP GPT e GPT-2, OpenAI decide di applicare l'architettura dei transformer alla generazione delle immagini. A tal proposito, nel giugno 2020 pubblica ImageGPT, il quale utilizzando lo stesso codice di GPT-2 produce immagini coerenti.

Il 5 gennaio 2021, OpenAI presenta sul suo blog il modello DALL-E, capace di generare immagini a partire da una didascalia che l'utente inserisce. DALL-E mostra di comprendere fedelmente le istruzioni della descrizione testuale rendendosi capace di apprendere concetti spaziali e temporali.

Malgrado le capacità di DALL-E siano state ampiamente mostrate nel blog di OpenAI e nelle relative pubblicazioni, il codice del modello non è mai stato rilasciato interamente, fatto che ha favorito il rilascio di numerosi modelli testo-immagine: Cogview, DALL-E Mini⁸⁷, ruDALL-E⁸⁸.

Questa sezione andrà a descrivere il funzionamento di DALL·E, un portale online basato sull'intelligenza artificiale che permette di generare immagini a partire da semplici descrizioni testuali. DALL·E è un algoritmo di intelligenza artificiale sviluppato da OpenAI, è in grado di apprendere concetti di spazio e tempo e di generare immagini realistiche basandosi sui comandi testuali forniti

dall'utente. Tra le varie funzioni offerte, è possibile apportare modifiche al contesto delle immagini generate o caricare immagini sulla piattaforma, creando diverse varianti ispirate all'immagine originale.

Per quanto riguarda i costi, al momento della stesura, DALL·E può essere testato gratuitamente nella sua versione beta con un sistema di crediti. Si ricevono 50 crediti all'iscrizione e 15 crediti gratuiti ogni mese successivo. Ogni credito permette di effettuare una richiesta originale (prompt) a DALL·E, che restituisce quattro immagini, oppure una richiesta di modifica/variazione, che restituisce tre immagini. Se si necessita di ulteriori crediti, è possibile acquistarli con prezzi a partire da 115 crediti/15 dollari (equivalenti a 460 immagini generate)⁸⁹.

Per utilizzare DALL·E, è necessario accedere alla home page del servizio, completare la registrazione e iniziare a generare immagini tramite descrizioni testuali. Dopo la registrazione, si deve inserire il proprio nome e cognome, specificare il numero di telefono e convalidare il codice a 6 cifre ricevuto tramite SMS. Una volta completato questo processo, è possibile iniziare a creare immagini con DALL·E.

Per generare un'immagine, si deve inserire una parola o frase nel campo di testo situato al centro della pagina e fare clic sul pulsante "Generate". L'algoritmo di DALL·E è in grado di comprendere anche la lingua italiana, ma per maggiore accuratezza si consiglia l'utilizzo dell'inglese. Dopo aver generato l'immagine desiderata, è possibile creare ulteriori varianti o utilizzare strumenti di modifica per ritoccare la foto.

DALL·E offre anche la possibilità di caricare una foto per modificarla o generarne delle varianti automatiche. Tuttavia, è importante non caricare foto di persone senza il loro consenso. Se non si ha un'idea precisa dell'immagine da creare, è possibile utilizzare la funzione "Surprise me" per generare immagini suggerite dall'algoritmo stesso, rendendo l'esperienza ancora più interessante e stimolante.

In sintesi, DALL·E rappresenta un potente strumento basato sull'intelligenza artificiale che permette di generare immagini a partire da semplici descrizioni testuali. Grazie alle sue funzionalità avanzate, è possibile creare immagini originali, modificarle e generare varianti in modo automatico. Sebbene sollevi alcune questioni etiche, l'IA come DALL·E dimostra come la tecnologia possa essere utilizzata per scopi creativi e innovativi.



An expressive oil painting of a basketball player dunking, depicted as an explosion of a nebula

Per utilizzare al meglio questo strumento, si consiglia di familiarizzare con le varie funzioni offerte e di sperimentare con diverse descrizioni testuali, tenendo presente che, per ottenere risultati più accurati, come già accennato è preferibile utilizzare l'inglese. Inoltre, è importante essere consapevoli dei limiti etici e delle responsabilità che derivano dall'utilizzo di immagini generate artificialmente, soprattutto quando si tratta di rappresentare persone o situazioni reali.

Nota importante, al momento della trattazione, anche la versione più avanzata uscita in questo tardo 2022, DALL-E 2 è disponibile in una beta gratuita con generazioni limitate e fruibile senza costi aggiuntivi.

Midjourney

Midjourney⁹⁰ è il tool che, come vedremo in seguito, ho usato di più per generare immagini di concept art con discreti risultati

Midjourney è un algoritmo di intelligenza artificiale testo-immagine, capace di generare immagini di alta qualità basate su istruzioni testuali. Esso produce immagini caratterizzate da colori complementari, dettagli nitidi, un uso artistico di luce e ombra, simmetria e prospettiva.

Utilizzando Midjourney, si possono creare immagini originali partendo da un input testuale. L'algoritmo eccelle nella creazione di ambientazioni in grado di imitare specifiche tendenze stilistiche o pittori particolari. La generazione di un'immagine richiede circa 50 secondi, a seconda della versione e della modalità richiesta, permettendo di osservare la progressiva creazione dell'opera d'arte, dalle linee e colori sfocati fino alla miniatura ad alta definizione.

Per quanto riguarda i costi, Midjourney offre un periodo di prova gratuito basato sul tempo di calcolo richiesto al server. Il sistema funziona tramite GPU time, ovvero i minuti necessari per generare un'immagine. Durante il periodo di prova gratuita, sono disponibili 25 minuti di GPU time, equivalenti a circa 25 richieste di creazione di immagini, incluse eventuali modifiche, ingrandimenti o conversioni ad alta definizione.

Al termine del periodo di prova, è possibile sottoscrivere uno dei seguenti piani di abbonamento offerti da Midjourney:

Basic Plan — al costo di 8 dollari/mese con 200 minuti al mese di GPU time e la possibilità di generare le immagini in modalità privata.

Standard Plan — al costo di 24 dollari/mese con 15 ore al mese di GPU time, la modalità “relax” per eseguire un numero illimitato di lavori con tempi di attesa leggermente più lunghi e la possibilità di generare le immagini in modalità privata.

Pro Plan — al costo di 48 dollari/mese con 30 ore al mese di GPU time, la modalità “relax” e la possibilità di generare le immagini in modalità privata.

L'abbonamento può essere annullato in qualsiasi momento senza alcun costo, e si potrà continuare a usufruire dei vantaggi offerti fino alla scadenza.

In merito ai diritti d'autore, secondo i termini e le condizioni del servizio, tutte le immagini prodotte con gli account Free Trial sono distribuite con licenza Creative Commons e possono essere utilizzate liberamente da altri. Con gli abbonamenti a pagamento, è possibile creare opere d'arte in modalità privata per poterle rielaborare personalmente e non in dominio pubblico. Tuttavia, la normativa statunitense, all'avanguardia in merito a queste tecnologie, impedisce di reclamare il copyright di opere create dall'intelligenza artificiale. Pertanto, le immagini generate con Midjourney possono essere utilizzate per usi commerciali senza poterne reclamare il possesso.

Midjourney non funziona tramite il proprio sito web, ma viene eseguito interamente su Discord. All'interno di un canale pubblico sul server, si può inserire il proprio comando (prompt) e, successivamente, l'algoritmo restituisce l'immagine richiesta. Pertanto, è necessario disporre di un account Discord che ricordiamo essere completamente gratuito.

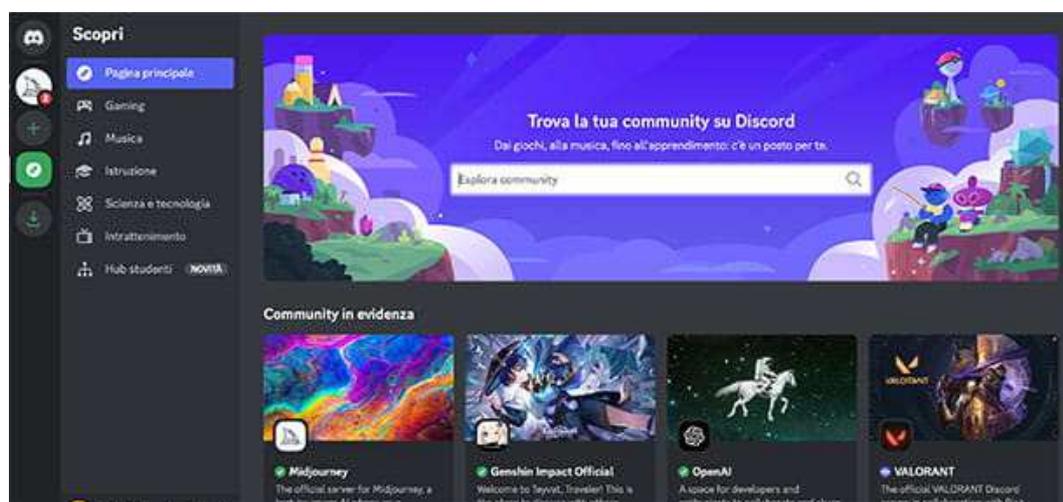
Per accedere al server di Midjourney su Discord, è possibile cliccare su "Hai già un invito? Unisciti a un server" e inserire il link di invito discord.gg/midjourney, premendo il tasto Unisciti. In alternativa, è possibile accedere tramite il sito ufficiale di Midjourney e cliccare sul bottone Sign In in home page, ma è comunque necessario disporre delle credenziali Discord.

Una volta all'interno del server di Midjourney, si può accedere a uno dei canali elencati nel pannello a sinistra sotto la voce "Newcomer Rooms" (ad esempio, newbies-120). All'interno del pannello principale, si noteranno le immagini generate dagli altri utenti. Per utilizzare Midjourney, è necessario digitare il comando `/imagine` seguito dalla parola o frase desiderata, premere Invio sulla tastiera del computer e attendere circa 50 secondi per la generazione dell'immagine.

L'algoritmo di Midjourney comprende molte lingue, tra cui l'italiano, ma per una maggiore accuratezza, anche in questo caso, è preferibile utilizzare termini in

inglese. Una volta generata l'immagine, sarà visualizzata come nuovo messaggio. Si possono poi utilizzare ulteriori comandi per ottenere un ingrandimento dell'immagine, generare una nuova immagine basata sullo stile e composizione dell'originale, o richiedere quattro nuove immagini dallo stesso prompt.

Per scaricare le immagini, è sufficiente cliccare sul file generato, premere Apri sul browser per visualizzare il risultato a grandezza naturale e salvare l'immagine sul proprio dispositivo secondo le modalità previste dal browser.



Home page di discord con in alto a sinistra il server di midjourney

Oltre al comando `"/image"`, si possono utilizzare i seguenti comandi:

`/info`: per avere informazioni sul proprio profilo e verificare i minuti di servizio disponibili.

`/settings`: per visualizzare e modificare i parametri di base dell'intelligenza artificiale.

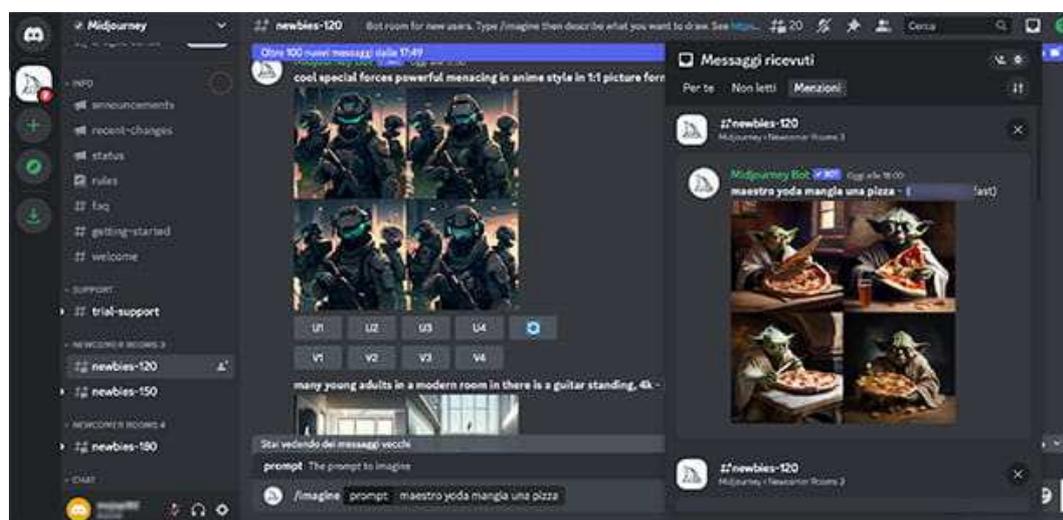
`/help`: per accedere alle informazioni utili e alla guida di Midjourney.

Dopo aver familiarizzato con l'utilizzo di Midjourney su Discord, si può sfruttare al massimo il potenziale dell'intelligenza artificiale per generare immagini di alta qualità. Per ottenere risultati migliori, è importante formulare il prompt in modo chiaro e descrittivo, specificando eventuali dettagli che si desiderano includere nell'immagine.

Se si desidera sperimentare con diverse varianti di uno stesso tema, è possibile utilizzare il comando `"/imagine"` seguito da diverse parole o frasi correlate al soggetto di interesse. In questo modo, si può confrontare e valutare i risultati ottenuti per scegliere l'immagine che meglio soddisfa le proprie esigenze.

È importante ricordare che durante il periodo di prova gratuita, tutte le immagini prodotte con gli account Free Trial sono distribuite con licenza Creative Commons e possono essere utilizzate liberamente da altri. Tuttavia, con gli abbonamenti a pagamento, è possibile creare opere d'arte in modalità privata, consentendo una maggiore protezione dei propri lavori.

In conclusione, Midjourney è uno strumento potente e versatile per generare immagini basate su testo tramite intelligenza artificiale. Seguendo le istruzioni fornite, si può sperimentare con diverse tematiche e stili per creare opere d'arte uniche e originali.



esempio di inserimento di prompt

Stable Diffusion

Stable diffusion è il tool che, unito al plug in in Blender, da i risultati più efficaci e performanti in termini di generazione di textures real time, pertanto, vale la pena approfondire questa IA più a fondo delle precedenti.

Inoltre la trattazione in merito è più approfondita grazie al fatto che il codice è completamente open source.

Anche Stable Diffusion è un modello di apprendimento automatico profondo (DNN), attualmente con codice open source e pesi del modello di pubblico dominio, pubblicato nel 2022, utilizzato principalmente per generare immagini dettagliate a partire da descrizioni di testo, è un modello di diffusione latente, una variante di rete neurale generativa profonda sviluppata dal gruppo CompVis alla LMU di Monaco. Il modello è stato rilasciato da una collaborazione tra Stability AI, CompVis LMU e Runway con il supporto di EleutherAI e LAION⁹¹.

A differenza degli altri modelli ha bisogno, per funzionare, di un hardware (che sebbene sia abbastanza facile da reperire nel 2022) consti di una GPU modesta con almeno 10 GB di VRAM abbandonando così il modello cloud based delle precedenti due AI.

Stable Diffusion utilizza una variante del modello di diffusione (DM), chiamato modello di diffusione latente (LDM)⁹². Introdotto già nel 2015, i modelli di diffusione vengono addestrati con l'obiettivo di rimuovere le successive applicazioni del rumore gaussiano sulle immagini di addestramento che possono essere pensate come una sequenza di autocodificatori per la riduzione del rumore.

È costituito da tre parti: l'autoencoder variazionale (VAE), U-Net e un codificatore di testo opzionale. Il codificatore VAE comprime l'immagine dallo spazio dei pixel a uno spazio latente di dimensioni inferiori, acquisendo un significato semantico più fondamentale dell'immagine come spiegato in precedenza con l'esempio della rete convoluzionale.

Il blocco U-Net, composto da una dorsale ResNet, pulisce poi il segnale in uscita dalla diffusione diretta all'indietro per ottenere una rappresentazione latente. Infine, il decoder VAE genera l'immagine finale riconvertendo la rappresentazione nello spazio dei pixel.

A seconda di cosa si trovi come confronto in questa fase di riduzione del rumore, ad esempio, se si ha del testo o un'immagine essa può essere condizionata, in modo flessibile.

I dati di condizionamento codificati sono esposti a U-Net (riduzione rumore) tramite un meccanismo di attenzione incrociata. Per il condizionamento del testo, il codificatore di testo CLIP ViT-L/14 fisso e pre-addestrato viene

utilizzato per trasformare i prompt di testo in uno spazio di incorporamento. I ricercatori indicano che il vantaggio degli LDM sono una maggiore efficienza computazionale per la formazione e la generazione⁹³⁹⁴.

Dati di addestramento

Stable Diffusion è stato addestrato su coppie di immagini e didascalie tratte da LAION-5B, un set di dati pubblicamente disponibile derivato dai dati di Common Crawl prelevati dal Web, in cui 5 miliardi di coppie immagine-testo sono state classificate in base alla lingua, filtrate in set di dati separati per risoluzione, una probabilità prevista di contenere una filigrana e un punteggio "estetico" previsto (ad es. qualità visiva soggettiva)⁹⁵.

Limitazioni

Stable Diffusion presenta problemi di degrado e imprecisioni in determinati scenari. Poiché il modello è stato addestrato su un set di dati costituito da immagini con risoluzione 512×512 , la qualità delle immagini generate peggiora notevolmente quando le specifiche dell'utente si discostano dalle immagini con risoluzione "prevista" 512×512 . Il modello inoltre a causa di una scarsità di immagini dettagliate non è sufficientemente addestrato per comprendere gli arti e i volti umani o di animali a causa della mancanza di caratteristiche rappresentative nel database. Richiedere al modello di generare immagini di questo tipo può confonderlo. La generazione di arti di animali, ad esempio, ha un tasso di fallimento osservato del 25% quando si cerca di generare l'immagine di un cavallo⁹⁶.

Un altro problema conosciuto anche ai creatori di Stable Diffusion è che sussiste un bias algoritmico poiché il modello è stato addestrato principalmente su immagini con descrizioni in inglese⁹⁷. Di conseguenza, le immagini generate rafforzano i pregiudizi sociali e provengono da una prospettiva occidentale.

Inoltre modello come per i due precedenti fornisce risultati più accurati per i prompt scritti in inglese rispetto a quelli scritti in altre lingue con culture occidentali "bianche" che spesso sono la rappresentazione su cui è stato addestrato⁹⁸.

Capacità

Il modello supporta la capacità di generare nuove immagini da zero attraverso l'uso di un prompt di testo che descrive gli elementi da includere o omettere. Come prompt può ricevere anche in input delle immagini esistenti per incorporare nuovi elementi descritti da testo (ad es. sintesi guidata di immagini). Inoltre, il modello consente anche di alterare parzialmente le immagini esistenti tramite inpainting e outpainting, se utilizzato con un'interfaccia utente appropriata che supporta tali funzionalità, di cui esistono numerose diverse implementazioni open source⁹⁹.

I requisiti per l'esecuzione del modello come già accennato sono 10 GB o più VRAM, tuttavia gli utenti con meno VRAM possono scegliere di caricare i pesi con precisione float16 anziché float32 predefinito per compensare le prestazioni del modello con un utilizzo VRAM inferiore.

Generazione da testo a immagine



Esempio di campionamento con selezione per esclusione o inserimento

Lo script di campionamento da testo a immagine all'interno di Stable Diffusion, noto come "txt2img", utilizza un prompt di testo oltre a parametri di opzione assortiti che coprono i tipi di campionamento, le dimensioni dell'immagine di output e i valori "seed" che influiscono sull'immagine di output. Gli utenti possono scegliere di randomizzare questo valore per esplorare diversi output generati o utilizzare lo stesso per ottenere il medesimo output di un'immagine generata in precedenza.

Il risultato sarà quindi un'immagine basata sull'interpretazione del modello del prompt.

Le immagini così generate sono contrassegnate con una filigrana digitale invisibile per consentire agli utenti di identificare un'immagine generata da Stable Diffusion, sebbene questa filigrana perda la sua efficacia se l'immagine

viene ridimensionata o ruotata¹⁰⁰, un timido primo passo verso la protezione da contraffazione e deep-fake.

Un'altra opzione configurabile è il valore della scala di guida che consente all'utente di regolare il grado di aderenza dell'immagine di output al prompt. I casi d'uso più sperimentali possono optare per un valore di scala inferiore, mentre i casi d'uso che mirano a risultati più specifici possono utilizzare un valore più alto.

Modifica dell'immagine

Stable Diffusion include anche un altro script di campionamento, "img2img", che utilizza una stringa di testo, un percorso a un'immagine esistente e un valore di intensità compreso tra 0,0 e 1,0. Lo script genera una nuova immagine basata sull'immagine originale che presenta anche elementi forniti all'interno della stringa di testo. Il valore dell'intensità indica la quantità di rumore aggiunta all'immagine di output. Un valore di intensità più elevato produce una maggiore variazione all'interno dell'immagine, ma può produrre un'immagine che non è semanticamente coerente con la stringa fornita.

Lo stesso processo può essere utile anche per l'upscaling dell'immagine, in cui viene aumentata la risoluzione, aggiungendo potenzialmente più dettagli.

Ulteriori casi d'uso per la modifica tramite img2img sono offerti da numerose implementazioni front-end del modello Stable Diffusion.

L'inpainting, usato per esempio estensivamente dal plug in di blender, implica la modifica selettiva di una porzione di un'immagine esistente delineata da una maschera di livello fornita dall'utente, che riempie lo spazio mascherato con il contenuto appena generato in base al prompt fornito. Al contrario, la pittura esterna estende un'immagine oltre le sue dimensioni originali, riempiendo lo spazio precedentemente vuoto con il contenuto generato in base al prompt fornito.

Utilizzo e copyright

Stable Diffusion concede liberamente agli utenti i diritti di utilizzo di qualsiasi immagine generata dal modello, a condizione che il contenuto dell'immagine non sia illegale o dannoso per le persone.

Poiché gli stili visivi e le composizioni non sono soggetti a copyright, si interpreta spesso che gli utenti di Stable Diffusion che generano immagini di

opere d'arte non dovrebbero essere considerati violazione del copyright di opere visivamente simili¹⁰¹.

Tuttavia, gli individui raffigurati nelle immagini generate possono essere protetti dai diritti della personalità se viene utilizzata la loro somiglianza.

AI Testuali

In questo caso, sebbene gli attori in campo siano molti, è difficile avere un reale competitor di Chat GPT-4 di Open AI. Spesso e volentieri le altre intelligenze di questo tipo (vedasi Bing Chat¹⁰²) presentano al loro interno le API di OPEN AI che quindi mi rimandano a questa doverosa trattazione.

Sebbene chat GPT nella fase terminale della mia tesi sia stata bloccata per ragioni di trattamento dei dati in Italia, ho basato il mio flusso di lavoro su di essa quindi auspico un accordo internazionale per una definizione dell'abuso o delle policy relative alle intelligenze artificiali.

ChatGPT

ChatGPT è un modello linguistico di grandi dimensioni messo a punto con tecniche di apprendimento automatico (di tipo non supervisionato), e ottimizzato con tecniche di apprendimento supervisionato e di deep learning¹⁰³¹⁰⁴, che è stato sviluppato per essere utilizzato come base per la creazione di altri modelli di machine learning. ChatGPT è stato addestrato a partire dai modelli Instruct GPT, (o GPT-3.5) di OpenAI, che sono l'evoluzione dei modelli di GPT-3. Gli Instruct GPT sono modelli in cui il pre-addestramento è stato ottimizzato manualmente da addestratori umani. Nello specifico ChatGPT è stato sviluppato da un GPT-3.5 utilizzando l'apprendimento supervisionato e il DRL come tecniche di ottimizzazione del modello¹⁰⁵. Il 14 marzo 2023 è stata annunciata l'introduzione di GPT-4, ovvero un modello multimodale su larga scala che può accettare input di immagini, video, audio e testo e produrre output di testo¹⁰⁶¹⁰⁷.

ChatGPT è stato lanciato il 3 novembre 2022¹⁰⁸[11] e ha attirato l'attenzione per le sue risposte dettagliate e articolate, sebbene la sua accuratezza sia stata criticata. Sia l'apprendimento supervisionato che l'apprendimento per rinforzo hanno utilizzato istruttori umani per migliorare le prestazioni del modello. Nel

primo caso il modello è stato alimentato con conversazioni nelle quali gli istruttori interpretavano entrambe le parti: l'utente e l'assistente basato su intelligenza artificiale. Nella fase di rinforzo, gli istruttori umani hanno prima valutato le risposte che il modello aveva creato nella conversazione precedente. Queste valutazioni sono state utilizzate per creare "modelli di ricompensa" sui quali il modello è stato ulteriormente perfezionato, utilizzando diverse iterazioni di Proximal Policy Optimization (PPO)¹⁰⁹¹¹⁰. Gli algoritmi di Proximal Policy Optimization presentano un vantaggio su algoritmi di Trust Region Policy Optimization annullando molte delle operazioni computazionalmente costose con prestazioni migliori¹¹¹. I modelli sono stati addestrati in collaborazione con Microsoft sulla loro infrastruttura cloud Azure¹¹².

Miglioramenti

Rispetto a InstructGPT, ChatGPT risulta significativamente migliorato:

è significativamente più verboso di altri modelli.

è significativamente più bravo a comprendere le intenzioni dell'utente

è stato ottimizzato per ridurre le "allucinazioni" ossia la creazione di contenuti inventati

Ad esempio, mentre un InstructGPT accetta la richiesta "Parlami di quando Isaac Newtown ha inventato la gravità nel 2015" come veritiera, ChatGPT utilizza le informazioni sulle pubblicazioni di Newton e le informazioni sul mondo moderno per costruire una risposta che immagina cosa accadrebbe se Newton avesse inventato, senza che vi fosse prima la gravità nel 2015. I dati di addestramento di ChatGPT includono pagine di manuali e informazioni su Internet e sui linguaggi di programmazione, come i bulletin board system e il linguaggio di programmazione C# o Python rispettivamente utili per Unity e Blender.

Limitazioni

ChatGPT soffre di molteplici limitazioni. Il modello di ricompensa di ChatGPT, progettato attorno alla supervisione umana, può essere eccessivamente ottimizzato e quindi ostacolare le prestazioni, fenomeno noto come legge di Goodhart.

Inoltre, ChatGPT ha una conoscenza limitata degli eventi accaduti dopo il 2021 e non è in grado di fornire informazioni su alcune celebrità. Nella fase di

apprendimento, gli istruttori hanno preferito risposte più lunghe, indipendentemente dalla comprensione effettiva o dal contenuto. I dati di addestramento possono anche risentire di pregiudizi algoritmici; suggerimenti che includono indicazioni vaghe di persone, come un CEO, potrebbero generare una risposta che presuppone che una persona del genere, ad esempio, sia un maschio bianco¹¹³.

ChatGPT attualmente utilizza la seconda generazione di reti neurali, la più diffusa ma la più costosa in termini di prestazioni e consumi, infatti il consumo stimato della versione ChatGPT 3 durante la fase di apprendimento è stato di circa 1287 MWh¹¹⁴ grazie all'ausilio di molteplici GPU. Un porting per la terza generazione di reti neurali (Rete neurale spiking) permette un notevole passo avanti in termini sia di prestazioni, che di riduzione dei consumi.

Progresso perché il porting dell'architettura permette un numero "illimitato" di generazioni dei dati (risposte testuali e grafiche), proprio perché le reti neurale spiking e NeuralLead distribuiscono i dati nel tempo, invece che avere un set di dati predefinito che può essere eseguito una volta sola.

Avanzamenti recenti

Nel marzo 2023 è stato lanciato un'assistente virtuale umanoide, che permette anche il riconoscimento della voce e quindi l'effettuazione delle ricerche mediante il microfono, senza la necessità di digitare il testo.¹¹⁵ L'assistente ha il supporto multilingue.

Il 9 marzo 2023 Andreas Braun, CTO di Microsoft Germany, ha annunciato durante un suo intervento all'evento 'AI in Focus - Digital Kickoff' che "Introdurremo Gpt-4 la prossima settimana. E in questo modo avremo modelli multimodali che offriranno possibilità completamente diverse, ad esempio la comprensione dei video"¹¹⁶. Se Gpt-3 operava con 175 miliardi di parametri di apprendimento automatico, mentre Gpt-4 ne ha 100.000 miliardi.¹¹⁷

Esistono estensioni gratuite dei browser che rendono ChatGPT più facilmente accessibile da altri siti web, permettono di esportare la chat o suggeriscono il testo delle query da sottoporre al sistema.¹¹⁸

Alcuni degli esempi più semplici di API dentro software cots sono l'integrazione con Microsoft Excel e con Google Sheets per la generazione di formule a partire dal linguaggio umano¹¹⁹.

Metodologia

Descrizione del progetto "Locked Up"

Locked Up è un progetto in Realtà Virtuale stand alone, per il device Oculus Quest 2 nato dalla collaborazione del Politecnico di Torino (con le figure, come registi e ideatori, di Alessandro VISCONTI e Antonio MESSINA) e Robin Studio con la Biennale di Venezia.

L'idea alla base di questo progetto è quella di trattare il tema dell'auto-isolamento. Il protagonista difatti è Robin, un adolescente che soffre di ansia sociale e si rifugia nel suo mondo sicuro: la sua camera da letto che diventa la sua realtà "fuori dal tempo". La storia si sviluppa attraverso una serie di scene e atti che alternano ciò che accade nel mondo al di fuori e quello che avviene nella stanza di Robin, permettendo all'utente di immergersi nei suoi panni e di esplorare il suo universo interiore ed esteriore.

Nel mondo reale, prima di rinchiudersi nella sua stanza Robin affronta situazioni comuni: andare a scuola, interagire con i compagni di classe, risolvere un problema alla lavagna. Queste scene mettono in luce le difficoltà che il protagonista incontra nel gestire le sue ansie e nella relazione con gli altri, le delusioni di un irraggiungibile perfezione, motivando la sua scelta di rinchiudersi nella sua stanza.

Nella camera, un ambiente interattivo, ci troviamo in una sorta di locus amoenus in cui gli oggetti presenti rivelano aspetti della personalità di Robin, dei suoi sogni e delle sue paure. Gli utenti possono interagire con quest'ultimi per sbloccare nuove scene e avanzare nella trama. Ad esempio, un semplice pupazzo potrebbe contenere importanti informazioni, mentre una foto potrebbe mettere in luce le opportunità perdute.

Man mano che la storia si sviluppa, l'utente è incoraggiato a prendere decisioni che influenzano il percorso di Robin. Queste scelte si riflettono nei due possibili finali: uno in cui riesce a superare le sue ansie e a stabilire relazioni significative con gli altri, e un altro in cui rimane intrappolato nel suo isolamento.

Il progetto è sviluppato utilizzando il motore di gioco Unity, specificatamente per il dispositivo Oculus Quest 2. L'interazione utente è al centro del design, con un'attenzione particolare all'ambientazione e all'atmosfera. La colonna sonora originale e gli effetti sonori immersivi danno vita all'esperienza, contribuendo a renderla coinvolgente ed emozionante per gli utenti.

Questa tesi dal mio punto di vista si è articolata in due fasi: una prima fase, in cui mi sono occupato esclusivamente del texturing e dell'ottimizzazione dei modelli.

Una seconda, in cui, avendo osservato una prima integrazione dell'intelligenza artificiale per l'animazione, mi è sembrato opportuno approfondirne l'uso per tutto il processo creativo (essendo questo aspetto, quello del rigging AI già trattato per questo progetto rimanderò alla trattazione del collega Antonio Messina e non mi soffermerò su di esso)

Entrando nel merito, dopo un lavoro individuale di tipo tradizionale con tool convenzionali quali photoshop e substance painter e dopo una fase generativa fatta con textures procedurali ad imitare la pennellata su blender ho deciso di ricercare e usare quando possibile l'intelligenza artificiale in ottica di snellire ed automatizzare, con tool quanto più gratuiti, il processo generativo mantenendo lo stile che si voleva raggiungere.

Come già accennato nell'introduzione questa tesi ha come scopo quello di mostrare i risultati ottenuti con l'ausilio dell'intelligenza artificiale lungo tutto il processo creativo dell'esperienza VR, con un focus particolare sulle texture.

Lo scopo del processo non è tanto arrivare con gli strumenti e i plug-in di questo inizio di 2023 a un prodotto finito ma quanto più di dimostrare un flusso di lavoro agile che porti ad un prototipo usabile con rapidità e che consenta quindi di valutare in maniera più efficace le modifiche da apportare al progetto.

Adesso andrò, sezione per sezione, a rivedere con l'ausilio dell'intelligenza artificiale i punti salienti dell'esperienza cercando di dare il focus lungo tutto percorso creativo, ad una specifica parte del progetto.

Gli esempi portati andranno ad estendere inoltre quello che è stato già fatto in ottica di integrazione con l'esperienza allo stato attuale. Sebbene alcune parti siano già in uno stadio avanzato, ad esempio quella dei dialoghi o dello

storytelling. Per completezza di trattazione ho deciso di includere anche quest'ultime che mi sono state utili per generare contenuto per descrizioni più approfondite da fornire per la concettualizzazione artistica o per generare codice per la modellazione.

Presentazione del lavoro svolto in precedenza

*metodologia procedurale a nodi e hand painting con Suite Adobe
Substance e Photoshop.*

La fase di texturing, come anticipato pocanzi, è stata oggetto del mio studio ancora prima che la generazione attraverso l'intelligenza artificiale fosse il focus di questa tesi.

I vincoli del progetto sono sempre stati piuttosto chiari: Modelli con pochi poligoni per rendere meno pesante possibile la fruizione con Oculus Quest 2, essendo un'applicazione nativa, per i quali pertanto è stata svolta un'operazione di ottimizzazione.

Texture senza normali o metallici con le luci dipinte per andare incontro ad un look indie e migliorare la resa di un progetto low-poli, perciò sono state disegnate in modo ricco e pennellato per dare questa estetica di digital art.

Prima di sposare la generazione con l'AI le mie attenzioni si sono rivolte a ottimizzare il processo quanto più possibile in termini di tempo, sia per questioni di scadenze di progetto che per resa estetica e facilità di revisione e modifica.

La tecnica di project from view è stato il mio primo obiettivo, con questa metodologia il progetto è più veloce da visualizzare come bozza 3D e i risultati sono più vicini agli artwork disegnati da Edoardo Audino e dalle immagini nel Mood-Board.



Jinx da Arcane – Netflix 2021

Di grande spunto creativo sono state le grafiche e le ambientazioni della serie netflix Arcane

Arcane ha una particolarità piuttosto peculiare: le luci, le ombre e in generale tutte le inquadrature sono state disegnate a partire dalla prospettiva della camera stessa.

Da questo ragionamento, in principio, ho sviluppato un flusso di lavoro incentrato su tale metodologia:

in primis inquadravo il modello da un'angolazione con la camera, qui ad esempio

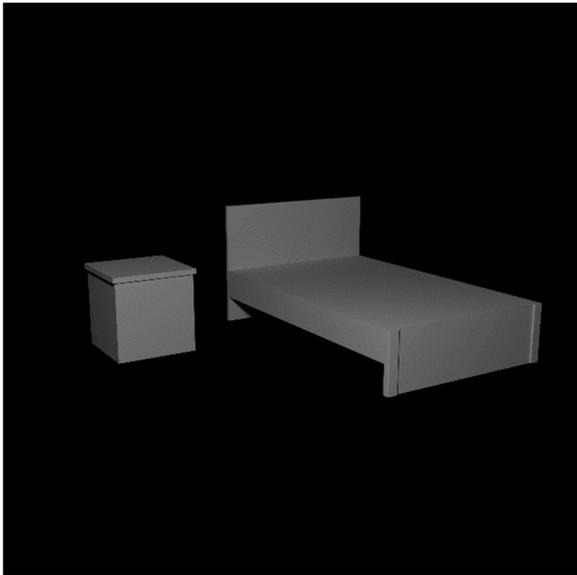


Figura 2



Figura 1

nella stanza da letto di robin ho creato una vista del letto e del comodino.

A questo punto andavo con Photoshop come si vede dalla seconda immagine a generare delle textures disegnate dall'angolazione scelta andando a ricercare uno stile che fosse quanto più simile al cubo disegnato qui sotto.

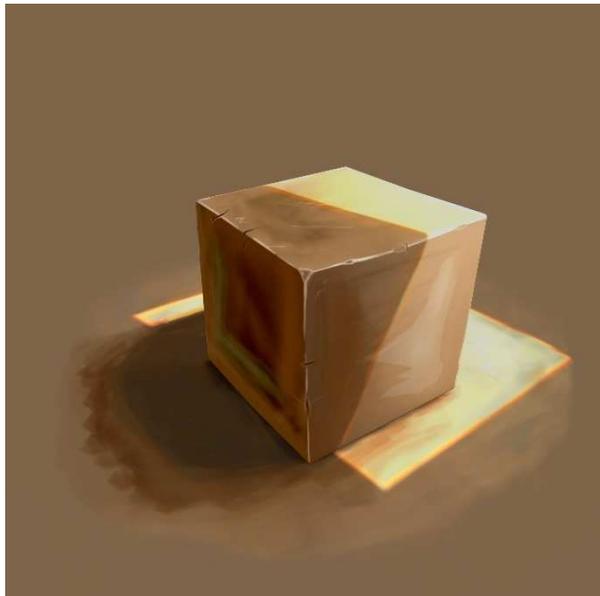


Figura 3

Inutile sottolineare quanto fosse lenta la procedura e quanto fosse necessario avere più viste per poter integrare efficacemente un modello, (il letto ed il comodino non presentano grandi problematiche avendo praticamente solo due lati visibili dalla prospettiva) ma un'esperienza VR pone l'utente di fronte agli oggetti e a potervi girare intorno. Avendo essi pochi poligoni le texture devono essere quanto più efficaci possibili, ed essendo lo stile pennellato non avevo molto margine di manovra con le mappe delle normali che avrebbero rischiato di rovinare questo look, essendo la granularità della pennellata a definire luci ed ombre non l'ambiente stesso.

Dopodiché, conoscendo abbastanza bene la dinamica dei nodi di blender, ho deciso di applicare un nuovo flusso di lavoro incentrato su un texturing tradizionale fatto per agilità con i fast materials su substance painter passandolo attraverso un filtro nodale direttamente in blender che mi consentisse di generare la texture pennellata del modello. Per ottenerla, però, sono riuscito solo in Eevee poiché come motore di rendering possiede un'illuminazione non ray-traced cosa che avviene invece in Cycles.

La combinazione di questi fattori mi rende impossibile applicare le textures direttamente per fare un Bake ed esportare il risultato ed il modello su unity. Perciò, sempre con il project from view, devo ri-applicare i render delle viste e fare il bake finale delle textures del modello per ottenere le mie esportazioni. Di seguito mostro l'architettura nodale, non entrando nel merito di ogni singolo nodo poiché non è lo scopo di questa tesi però vorrei presentare quella che è la logica che sottende questa architettura nodale

Queste due strutture di nodi sono, in buona sostanza, uno shader “Before” che

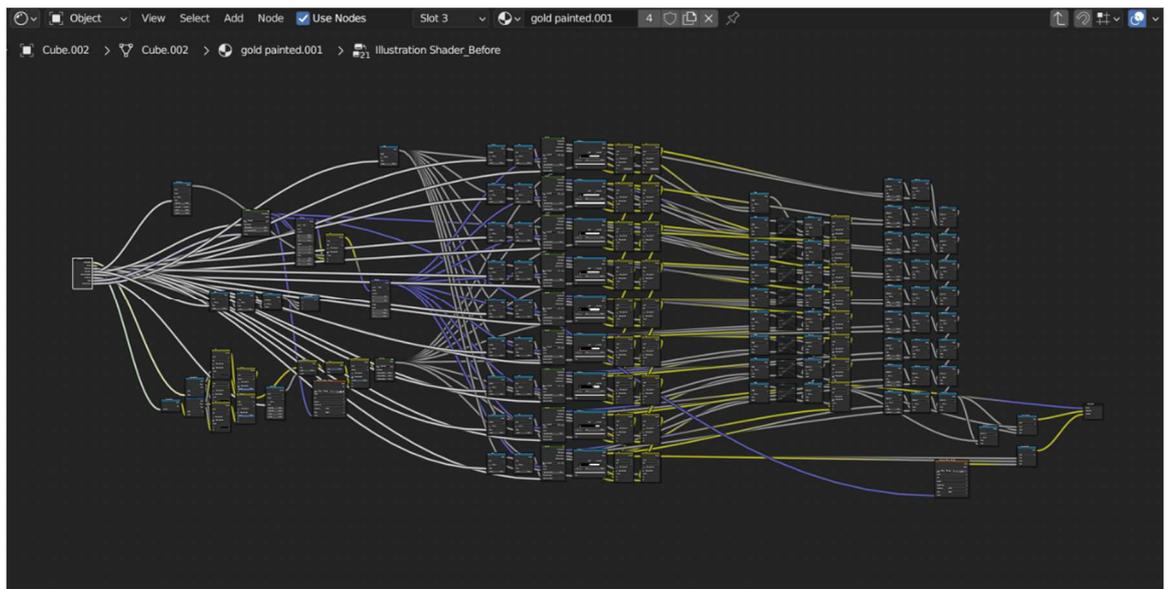


Figura 5

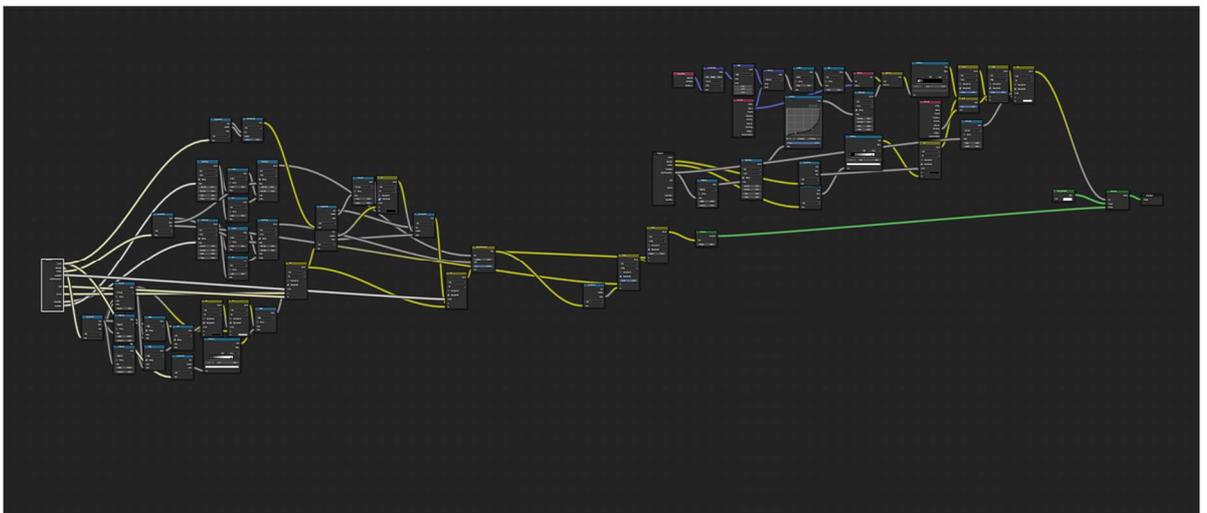


Figura 4

va a generare delle aree di colore sul modello a partire dalla texture di input, e uno shader “After” che genera, a partire da una serie di maschere in bianco e nero di pennellate, delle pennellate del colore di sottofondo unito al contributo della luce.

Processo di integrazione dell'IA nel workflow

Design dell'esperienza

"In principio, il team di sviluppo creò l'idea di un gioco e il design del suo mondo."

Per la fase di concettualizzazione e scripting mi sono avvalso di Chat GPT-4, a differenza del più anziano 3.5 è meno rapido, ma passa da un prompt con capienza massima di 3000 parole a 25000, un netto avanzamento.

A livello di “memoria” dà solidità mostrando di essere molto più stabile: Ricorda quello che gli viene detto anche a distanza di numerose iterazioni e riesce a collegare anche due concetti diversi riuscendo ad integrare in maniera efficace i vari input forniti, infine, questa versione più recente, dà accesso all’input visivo che consente di inserire immagini da far decifrare

Ho sperimentato un po’ in questa fase ipotizzando diverse situazioni operative, e diversi stadi di avanzamento del progetto così da capire dove potersi inserire al meglio.

Per prima cosa ho valutato il risultato della generazione a partire da un concetto molto basilare “creare un’esperienza VR che possa far immedesimare l’utente per sensibilizzare sul tema degli Hikikomori”

Siccome chat GPT-4 ha un database che dopo il 2019 è piuttosto lacunoso sono andato alla ricerca, via browser, di alcuni articoli di trattazione scientifica in merito ai cardini di un’esperienza VR e al tema dell’Hikikomori, e di come, chi abbia sofferto di questa patologia, veda il mondo.

Ho chiesto poi un riassunto di ciò che ho inserito e, senza alcuna difficoltà, ho constatato che il training è andato a buon fine.

A questo punto, ho chiesto di generare prima delle guideline per questa esperienza di cui riporto di seguito il risultato:

<p><i>“Title: "Robin's World: A Hikikomori Experience”</i></p>	<p><i>“Titolo: "Il mondo di Robin: Un'esperienza Hikikomori”</i></p>
<p><i>Introduction</i></p>	<p><i>Introduzione</i></p>
<p><i>Briefly explain the concept of Hikikomori and the purpose of the VR experience.</i></p>	<p><i>Spiegare brevemente il concetto di Hikikomori e lo scopo dell'esperienza VR.</i></p>
<p><i>Introduce the protagonist, Robin, as a young individual experiencing Hikikomori.</i></p>	<p><i>Presentare il protagonista, Robin, come un giovane individuo che sperimenta l'Hikikomori.</i></p>
<p><i>Robin's Room</i></p>	<p><i>La stanza di Robin</i></p>
<p><i>The experience begins in Robin's room, which is dimly lit and filled with various items related to their hobbies and interests (e.g., books, video games, art supplies).</i></p>	<p><i>L'esperienza inizia nella stanza di Robin, poco illuminata e piena di oggetti legati ai suoi hobby e interessi (ad esempio, libri, videogiochi, materiale artistico).</i></p>
<p><i>The room is cluttered, with a bed, desk, and computer in the center. Curtains are drawn, blocking out sunlight and any view of the outside world.</i></p>	<p><i>La stanza è disordinata, con un letto, una scrivania e un computer al centro. Le tende sono tirate e bloccano la luce del sole e la vista del mondo esterno.</i></p>
<p><i>The user can interact with various items in the room to learn more about Robin's interests and feelings of isolation.</i></p>	<p><i>L'utente può interagire con i vari oggetti presenti nella stanza per conoscere meglio gli interessi di Robin e i suoi sentimenti di isolamento.</i></p>
<p><i>Daily Routine</i></p>	<p><i>La routine quotidiana</i></p>
<p><i>The user follows Robin through their daily routine, which mostly consists of staying in their room, avoiding social</i></p>	<p><i>L'utente segue Robin nella sua routine quotidiana, che consiste per lo più nel rimanere nella sua stanza, evitando</i></p>

<p><i>situations, and engaging in solitary activities.</i></p> <p><i>Periodically, Robin receives messages or phone calls from concerned friends or family members, which they can choose to ignore or respond to.</i></p> <p><i>Struggles and Challenges</i></p>	<p><i>situazioni sociali e dedicandosi ad attività solitarie.</i></p> <p><i>Periodicamente, Robin riceve messaggi o telefonate da amici o familiari preoccupati, che può scegliere di ignorare o rispondere.</i></p> <p><i>Lotte e sfide</i></p>
<p><i>The user witnesses Robin's internal struggles, such as feelings of loneliness, anxiety, and depression.</i></p> <p><i>Robin faces various challenges, such as the fear of leaving their room, interacting with others, and facing the outside world.</i></p> <p><i>Guided Self-Reflection</i></p>	<p><i>L'utente assiste alle lotte interne di Robin, come i sentimenti di solitudine, ansia e depressione.</i></p> <p><i>Robin affronta varie sfide, come la paura di uscire dalla propria stanza, di interagire con gli altri e di affrontare il mondo esterno.</i></p> <p><i>Auto-riflessione guidata</i></p>
<p><i>Throughout the experience, Robin's thoughts and feelings are displayed as text or voiced by a narrator, allowing the user to better understand their perspective and emotions.</i></p> <p><i>The user is prompted to reflect on their own experiences and emotions, drawing parallels between their life and Robin's.</i></p> <p><i>Possible Outcomes</i></p>	<p><i>Nel corso dell'esperienza, i pensieri e i sentimenti di Robin vengono visualizzati sotto forma di testo o di voce narrante, consentendo all'utente di comprendere meglio la sua prospettiva e le sue emozioni.</i></p> <p><i>L'utente è invitato a riflettere sulle proprie esperienze ed emozioni, tracciando paralleli tra la sua vita e quella di Robin.</i></p> <p><i>Possibili esiti</i></p>
<p><i>The VR experience could offer multiple endings, depending on the choices made by the user throughout the story.</i></p> <p><i>Some endings might show Robin seeking help, opening up to others, and gradually overcoming their Hikikomori, while others might show</i></p>	<p><i>L'esperienza VR potrebbe offrire diversi finali, a seconda delle scelte fatte dall'utente nel corso della storia.</i></p> <p><i>Alcuni finali potrebbero mostrare Robin che cerca aiuto, si apre agli altri e supera gradualmente il suo</i></p>

Robin continuing to isolate themselves and struggle with their mental health. This VR experience aims to create empathy and understanding for those experiencing Hikikomori by allowing users to step into the shoes of someone living with the condition."

Hikikomori, mentre altri potrebbero mostrare Robin che continua a isolarsi e a lottare con la sua salute mentale. Questa esperienza VR mira a creare empatia e comprensione per coloro che sperimentano l'Hikikomori, consentendo agli utenti di calarsi nei panni di chi vive questa condizione".

In questa fase ho chiesto di generare un abstract che integrasse tutto ciò che è stato suggerito ottenendo questo output:

"Title: Robin's Sanctuary

Description:

As you enter Robin's room, you are immediately enveloped by the dim glow of a single desk lamp, casting soft shadows on the cluttered surroundings. The walls, painted in a muted shade of gray-blue, are adorned with an eclectic mix of posters, drawings, and shelves filled with books, DVDs, and memorabilia from various movies and video games.

"Titolo: Santuario di Robin

Descrizione:

Entrando nella stanza di Robin, si viene immediatamente avvolti dal tenue bagliore di un'unica lampada da tavolo, che proietta ombre morbide sul disordine dell'ambiente circostante. Le pareti, dipinte in una tenue tonalità di grigio-blu, sono adornate da un eclettico mix di poster, disegni e scaffali pieni di libri, DVD e cimeli di vari film e videogiochi.

The room's centerpiece is a well-worn bed, covered in a disheveled comforter with a pattern of stars and galaxies. On the bed, you find a multitude of pillows and stuffed animals, offering a sense of comfort and familiarity in this enclosed space.

To the left of the bed, there is a large desk overflowing with papers, art supplies, and an open laptop, its screen displaying a paused video game. A pair of headphones rests beside the laptop, tangled in a mess of cords. A half-empty cup of cold coffee sits next to a stack of books, their spines creased from frequent use.

In the far corner, a small television sits on a cluttered shelf, surrounded by various gaming consoles and controllers. A worn-out bean bag chair lies in front of the TV, its fabric faded from countless hours spent in its embrace.

Opposite the desk, a tall bookshelf filled to capacity with books, figurines, and other collectibles leans against the wall, a testament to Robin's eclectic interests. The floor is covered in a soft, shaggy carpet, which adds a cozy feel to the room despite the clutter.

Al centro della stanza c'è un letto ben usurato, coperto da un piumino spettinato con un motivo di stelle e galassie. Sul letto si trova una moltitudine di cuscini e animali di peluche, che offrono un senso di comfort e familiarità in questo spazio chiuso.

A sinistra del letto c'è una grande scrivania traboccante di fogli, materiale artistico e un computer portatile aperto, il cui schermo mostra un videogioco in pausa. Un paio di cuffie è appoggiato accanto al portatile, aggrovigliato in un groviglio di cavi. Una tazza di caffè freddo mezza vuota si trova accanto a una pila di libri, i cui dorsi sono sgualciti dall'uso frequente.

Nell'angolo in fondo, un piccolo televisore è appoggiato su uno scaffale disordinato, circondato da varie console di gioco e controller. Davanti al televisore si trova una poltrona a sacco logora, il cui tessuto è sbiadito dalle innumerevoli ore trascorse nel suo abbraccio.

Di fronte alla scrivania, un'alta libreria piena di libri, statuette e altri oggetti da collezione è appoggiata alla parete, a testimonianza degli interessi eclettici di Robin. Il pavimento è ricoperto da un morbido tappeto, che conferisce un'atmosfera accogliente alla stanza nonostante il disordine.

The window, covered by heavy curtains, allows only a sliver of sunlight to enter the room, leaving the outside world obscured from view. It's clear that this room is Robin's sanctuary, a place where they can retreat from the world and immerse themselves in their interests, isolated from the pressures of social interaction."

La finestra, coperta da pesanti tende, lascia entrare solo una piccola parte della luce del sole, lasciando il mondo esterno oscurato alla vista. È chiaro che questa stanza è il santuario di Robin, un luogo dove ritirarsi dal mondo e immergersi nei propri interessi, isolati dalle pressioni dell'interazione sociale".

L'abstract prodotto mostra come l'output sia coerente e un'ottima base di partenza per futuri sviluppi e modifiche anche con l'IA stessa, non si discosta molto da quella che è l'idea già definita dell'esperienza pertanto non è mia cura andare nel dettaglio di una riscrittura ma ha buoni spunti per immedesimarsi nel personaggio e delinea dei buoni punti per un'esperienza VR immersiva come fornito dagli articoli in precedenza.

Come secondo uso, in fase di concept, mi sono focalizzato sulla parte in cui il protagonista dovrà sperimentare un attacco di panico. Essendo questa esperienza poco efficace, soprattutto come artwork e estetica, ho chiesto spunti per migliorare l'esperienza.

Non tutti i suggerimenti presentati sono applicabili su Oculus Quest 2 ma sono sicuramente spunti molto efficaci.

A partire da essi ho generato una descrizione della scena e dell'ambiente che ho poi usato nella fase successiva.

Non mi soffermerò oltre su questa sezione, nonostante abbia condotto altri esperimenti piuttosto efficaci, come il cambio di tono o di registro del protagonista o dei vari NPC possibili, la comprensione del testo o il riassunto dello script per vedere l'aderenza con l'abstract. Non essendo questo lo scopo dell'esperienza, però lo lascio come spunto di revisione o come base per la creazione di nuovo contenuto.

“To create an immersive VR experience that effectively conveys the emotional intensity of the scene, we can make use of the following suggestions and descriptions:

Spatial audio: Utilize 3D spatial audio to make the sounds in the environment feel more realistic. The cicadas outside, the teacher's voice, and the classmates' murmuring and laughter should all be positioned correctly in the virtual space to create a sense of presence.

Per creare un'esperienza VR immersiva che trasmetta efficacemente l'intensità emotiva della scena, possiamo avvalerci dei seguenti suggerimenti e descrizioni:

Audio spaziale: Utilizzare l'audio spaziale 3D per rendere più realistici i suoni dell'ambiente. Le cicale all'esterno, la voce dell'insegnante, i mormorii e le risate dei compagni di classe devono essere posizionati correttamente nello spazio virtuale per creare un senso di presenza.

Dynamic lighting and visual effects: Gradually increase the intensity of the lighting in the classroom as Robin's anxiety builds up. This can be accompanied by subtle visual distortions like blurring, tunnel vision, or a reddening of the edges of the field of view to convey the physical sensations of a panic attack.

Interactivity: Encourage the user to engage with the environment by providing interactive elements, like grabbing the chalk, writing on the blackboard, or looking around the classroom. These interactions can help create a more vivid sense of presence in the virtual world.

Real-time heartbeat tracking: If the VR headset is equipped with biometric sensors, the user's real-time heartbeat can be integrated into the experience. As Robin's anxiety increases, the user's heartbeat can be amplified and made audible, creating an even more immersive and personal experience.

Reactive character animations and dialogues: The teacher and classmates should react to Robin's actions (or lack thereof) with appropriate animations and dialogue. This

illuminazione dinamica ed effetti visivi: Aumentare gradualmente l'intensità dell'illuminazione dell'aula man mano che l'ansia di Robin aumenta. Questo può essere accompagnato da sottili distorsioni visive come sfocatura, visione a tunnel o arrossamento dei bordi del campo visivo per trasmettere le sensazioni fisiche di un attacco di panico.

Interattività: Incoraggiare l'utente a interagire con l'ambiente fornendo elementi interattivi, come afferrare il gesso, scrivere sulla lavagna o guardarsi intorno in classe. Queste interazioni possono contribuire a creare un senso di presenza più vivido nel mondo virtuale.

Tracciamento del battito cardiaco in tempo reale: Se le cuffie VR sono dotate di sensori biometrici, il battito cardiaco in tempo reale dell'utente può essere integrato nell'esperienza. Quando l'ansia di Robin aumenta, il battito cardiaco dell'utente può essere amplificato e reso udibile, creando un'esperienza ancora più coinvolgente e personale.

Animazioni e dialoghi reattivi dei personaggi: L'insegnante e i compagni di classe devono reagire alle azioni (o alla mancanza di azioni) di Robin con animazioni e dialoghi appropriati. Questo può includere un comportamento più

can include more aggressive or impatient behavior if Robin fails to follow instructions, or sympathetic looks and comments if Robin is struggling.

Difficulty progression: As the math problem becomes more complex, the environment and characters should react accordingly. The teacher's impatience, classmates' laughter, and the overall intensity of the situation should increase as the equation becomes more difficult, heightening the sense of pressure on Robin.

Breathing exercises and calming techniques: To help users manage their anxiety during the experience, consider integrating calming techniques or guided breathing exercises. This can be done through subtle visual or audio cues, like a calming voice guiding the user to take deep breaths, or visual indicators that show the user when to inhale and exhale.

Debriefing and reflection: After the experience, provide users with a debriefing area where they can reflect on their emotions and thoughts during the simulation. This can be a virtual environment, like a calming garden or a quiet room, where users can take a moment to process the experience and learn from it.

aggressivo o impaziente se Robin non segue le istruzioni, oppure sguardi e commenti compassionevoli se Robin è in difficoltà.

Progressione della difficoltà: Man mano che il problema matematico diventa più complesso, l'ambiente e i personaggi devono reagire di conseguenza. L'impazienza dell'insegnante, le risate dei compagni e l'intensità generale della situazione devono aumentare man mano che l'equazione diventa più difficile, aumentando il senso di pressione su Robin.

Esercizi di respirazione e tecniche di rilassamento: Per aiutare gli utenti a gestire l'ansia durante l'esperienza, si può considerare l'integrazione di tecniche di rilassamento o di esercizi di respirazione guidati. Ciò può essere fatto attraverso sottili indicazioni visive o audio, come una voce calmante che guida l'utente a fare respiri profondi, o indicatori visivi che mostrano all'utente quando inspirare ed espirare.

Debriefing e riflessione: Dopo l'esperienza, fornire agli utenti un'area di debriefing in cui possano riflettere sulle emozioni e i pensieri provati durante la simulazione. Può trattarsi di un ambiente virtuale, come un giardino rilassante o una stanza tranquilla, dove gli utenti possono prendersi un momento per elaborare l'esperienza e trarne insegnamento.

By incorporating these suggestions and descriptions into the VR experience, users can fully immerse themselves in Robin's emotional journey and gain a better understanding of the challenges faced by those who suffer from panic attacks.”

Incorporando questi suggerimenti e descrizioni nell'esperienza VR, gli utenti possono immergersi completamente nel viaggio emotivo di Robin e comprendere meglio le sfide affrontate da chi soffre di attacchi di panico

Concettualizzazione Artistica

"Poi il team di sviluppo disse: 'Generiamo un'estetica che separi lo stile dell'esperienza dalle altre.'"

In questa fase, insieme a quella di texturing, essendo esse stesse molto legate tra di loro, ho speso la maggior parte del tempo al fine di arrivare quanto più vicino possibile ai desiderata stilistici dati come compito.

In maniera empirica, allenando un prompt di Chat GPT-4 da comandi che hanno prodotto buoni risultati dentro Midjourney e Dall-E (simili a quelli ottenuti da immagini vicine al risultato atteso) ho generato, a partire da semplici frasi, degli spunti creativi per l'esperienza. Avvalendomi dell'altro prompt creato in precedenza mi sono fatto generare delle descrizioni dell'ambiente, in questo caso della stanza e dell'attacco di panico.

Per questione di avanzamento tecnologico e stilistica come menzionato nella bibliografia al momento Midjourney ottiene migliori risultati in termini di estetica "disegnata"; pertanto, i risultati migliori (quelli che andrò a presentare in questa sezione) vengono da questa intelligenza artificiale.

Di seguito, nelle prossime pagine, presenterò alcuni esempi di prompt generati su Chat GPT usati su Midjourney con l'immagine di risultato ottenuta a partire dalla selezione delle quattro opzioni proposte. L'ultima casistica mostra come lo stesso prompt possa generare effetti molto diversi o lontani tra loro ma anche come sia importante la scelta ed il gusto del creativo per delineare al meglio quello che si ritiene più adatto come chiave stilistica dell'esperienza.



“Digital illustration that captures the essence of a panic attack within a classroom setting, using a stylistic approach inspired by an animated series known for its dark, moody atmosphere, rich color palette, painterly textures, and atmospheric lighting that skillfully blends Victorian-era settings with futuristic elements. In this intense and unsettling scene, the classroom walls appear to close in on the protagonist, while desks and chairs distort and shift, trapping and cornering them. The blackboard, once a source of clarity, now blurs and recedes, with math problems transforming into complex, nonsensical equations. Educational posters and diagrams on the walls warp into intimidating, taunting images that come to life and further disorient the protagonist. The windows, once bright, become foggy and darkened, casting eerie shadows throughout the room, amplifying the sense of dread. The teacher's desk looms larger and more imposing, with its items devolving into chaotic disarray. The bookshelves tower over the protagonist, books spilling out and shifting. The wooden floor creaks and warps, giving a sensation of instability and sinking. Emphasize the protagonist's feelings of fear, disorientation, and helplessness with a dark, brooding color palette, dramatic contrasts, and the rich, painterly textures characteristic of the animated series. The illustration should evoke a deeply scary and intense experience, capturing the essence of a panic attack through the unique blend of Victorian-era aesthetics and futuristic design elements, with a strong focus on distorted elements and a dark, oppressive atmosphere --ar 16:9 --s 5 --q 5 --v 5”



Digital illustration that captures the essence of a panic attack within a classroom setting, using a stylistic approach inspired by an animated series known for its intensely dark, brooding atmosphere, heavily distorted visuals, saturated color palette, painterly textures, and atmospheric lighting reminiscent of Victorian-era settings combined with futuristic elements. In this menacing scene, the classroom walls appear to close in on the protagonist, while desks and chairs shift, trapping and cornering them. The other students lean in with faces grotesquely distorted in cruel laughter. The blackboard, once a source of clarity, now blurs and recedes, with math problems transforming into complex, nonsensical equations. Educational posters and diagrams on the walls warp into sinister, taunting images that pulsate with a malevolent energy. The windows, once bright, become foggy and darkened, casting eerie shadows throughout the room. The teacher's desk looms larger and more imposing, with its items devolving into disarray. The bookshelves tower over the protagonist, books spilling out and shifting, reflecting overwhelming expectations. The wooden floor creaks and warps, giving a sensation of instability and sinking. Emphasize the protagonist's feelings of fear, disorientation, and helplessness with an oppressively dark atmosphere and extremely distorted visuals, characteristic of the animated series. Enhance the scene with saturated colors, strong contrasts, and the rich, painterly textures, creating a vivid and intense experience inspired by the unique blend of Victorian-era aesthetics and futuristic design elements, all steeped in a hauntingly dark and gloomy tone --ar 16:9 --s 5 --q 5 --v 5



Digital illustration that captures the menacing and distorted appearance of the other students in a classroom during the protagonist's panic attack, with a style that emphasizes dark, vibrating tones and atmospheric lighting. In this haunting scene, portray the students with grotesquely distorted faces, twisted grins, and malicious laughter echoing through the room. Their eyes should gleam with mockery as they lean in, surrounding the protagonist, and contributing to the overwhelming sense of entrapment. Make their uniforms and clothing appear to shift and blend with the dark, pulsating atmosphere, enhancing the unsettling vibe of the scene. Illustrate the students' postures and gestures as exaggerated and unnerving, with some pointing accusatively at the protagonist or whispering sinisterly to one another. Use exaggerated shadows and eerie, unnatural lighting to further emphasize the malicious intent of the students, making them appear as twisted, distorted figures that feed off the protagonist's panic and vulnerability. Focus on capturing the sense of dread and isolation the protagonist experiences, using bold colors, strong contrasts, and rich textures to create an intense and immersive visual experience --s 5 --q 5 --v 5



digital illustration that captures the menacing and distorted appearance of the other students in a classroom during the protagonist's panic attack, with a style inspired by a popular animated series known for its bold colors, painterly textures, and atmospheric lighting. Emphasize the dark, vibrating tones characteristic of the series to create a sinister atmosphere. In this haunting scene, portray the students with grotesquely distorted faces, twisted grins, and malicious laughter echoing through the room. Their eyes should gleam with mockery as they lean in, surrounding the protagonist, and contributing to the overwhelming sense of entrapment. Make their uniforms and clothing appear to shift and blend with the dark, pulsating atmosphere, enhancing the unsettling vibe of the scene. Illustrate the students' postures and gestures as exaggerated and unnerving, with some pointing accusatively at the protagonist or whispering sinisterly to one another. Use exaggerated shadows and eerie, unnatural lighting to further emphasize the malicious intent of the students, making them appear as twisted, distorted figures that feed off the protagonist's panic and vulnerability. All the students are staring directly at the camera, creating an uncomfortable and unnerving connection with the viewer. Focus on capturing the sense of dread and isolation the protagonist experiences, using bold colors, strong contrasts, and rich textures reminiscent of the animated series to create an intense and immersive visual experience. --ar 3:4 --s 5 --q 5 --v 5

Inoltre ho sperimentato l'applicazione di stile avvalendomi del prompt di Midjourney. A partire dai contenuti generati dall'artista, le concept art e il moodboard, sono state generate immagini in linea con il progetto che ritengo di grande aiuto per la realizzazione di contenuti.

Come spiegato in precedenza ho quindi linkato il contenuto necessario per il trasferimento di stile, per una maggiore efficacia in questo ho anche inserito il link dell'immagine a Chat GPT per evidenziare i caratteri di stile necessari a ricreare il medesimo risultato e ho combinato il trasferimento di stile con il prompt.

Passando ad una ricerca sui contenuti simili ho valutato i prompt generati dagli utenti e stilato una piccola lista che ho passato su Chat GPT per ottenere quindi il nuovo prompt da passare a Midjourney.

Di questa fase ho messo i contributi nella sezione galleria della trattazione con i prompt ad esso allegati per evitare di appesantire troppo questa sezione anche perché, alcuni di essi, sono la base di partenza per processi che vedremo in seguito.

Modelling e Texturing

"Poi il team di sviluppo disse: 'Applichiamo l'estetica per generare il mondo di gioco con personaggi, ambienti, oggetti e animali.'"

In questa fase mi sono avvalso di due tecniche fondamentali: la proiezione della camera per la realizzazione dei modelli e la generazione di oggetti con alcuni tools AI sempre a partire da contributi di immagine.

Sebbene la prima citata non sia propriamente una tecnica AI ho sfruttato direttamente le immagini generate su midjourney e ho applicato un project from view a partire dalla camera su Blender.

Questo metodo generativo velocizza di molto la fase di modellazione, si può sempre tenere sott'occhio il numero di poligoni e si ottiene, alla fine del processo creativo, un modello già parzialmente texturizzato. Per modelli molto regolari, come pareti, banchi, libri etc potrebbe già essere una soluzione definitiva.

Va però notato che, spesso e volentieri, i punti di fuga di un'immagine AI non sono propriamente definiti e il senso della proiezione richiede del lavoro ulteriore per sistemare poi le texture dell'oggetto in questione.

Un esempio di questo processo può essere la generazione del modello della lavagna



O in maniera ancora più evidente un banco a partire da un'immagine generata da un prompt di stile



Un progetto per passare modelli invece “meno geometrici” a partire da immagini è PIFuHD, un software AI ancora in fase di sviluppo che produce,



almeno per la parte della figura umana buoni risultati, in questo caso la procedura al momento passa da un drive al link: <https://shunsukesaito.github.io/PIFuHD/> e qui, con una serie di passaggi ed installazioni, possiamo integrare figure umane a dimensione naturale generate a partire da immagini di prompt

Un'altra applicazione che ho usato, sempre free access è Avatar SDK <https://avatarsdk.com/> per generare immagini di volti e busti.

Con questo tool è bastato importare il viso realizzato su Midjourney a partire da un prompt di un personaggio per vederlo trasformato in una modello 3d.

La topologia del modello è esportata in Obj e il modello presenta numerosi poligoni ma il problema è risolvibile con un po' di retopology di Blender.

Ciò che si nota di questo approccio è che elementi come gli occhiali risultino “attaccati” al viso e le proporzioni facciali sono diverse; tuttavia, dato il tempo di generazione il modello acquista decisamente più valore.



Per quanto riguarda la generazione di texture invece mi soffermerò principalmente sul plug in di Stable Diffusion, spiegandone il funzionamento in breve e mostrando le modalità operative e i risultati.

Il concetto alla base di questo plug in è il project from view delle texture e la generazione di esse da un'immagine di contributo.

Riassumendo brevemente si prende un prompt testuale molto stringato, gli si assegna, come la stessa Stable Diffusion fa, uno stile (in questo caso ho trovato molto calzante lo stile illustrazione fantasy che rende un effetto pennellato) e lo si unisce ad un'immagine generata a partire dalla visuale della finestra di Blender. Questa procedura quindi genera una nuova UV Map proiettata dalla visuale stessa e un nuovo materiale con l'immagine così prodotta.

Questo Plug in è estremamente efficace per questo progetto: per la semplicità d'uso e per la possibilità di blender di avvalersi del texture paint nativo.

L'altro aspetto positivo di questo approccio è che tutti i materiali vengono generati a partire da un nodo "UV map" che li lega appunto alla loro "UV map generatrice", pertanto i modelli da esportare avranno null'altro da fare che un

bake su un nuovo materiale di tutte le textures e senza preoccuparsi dell'unpackage dell'UV mapping.

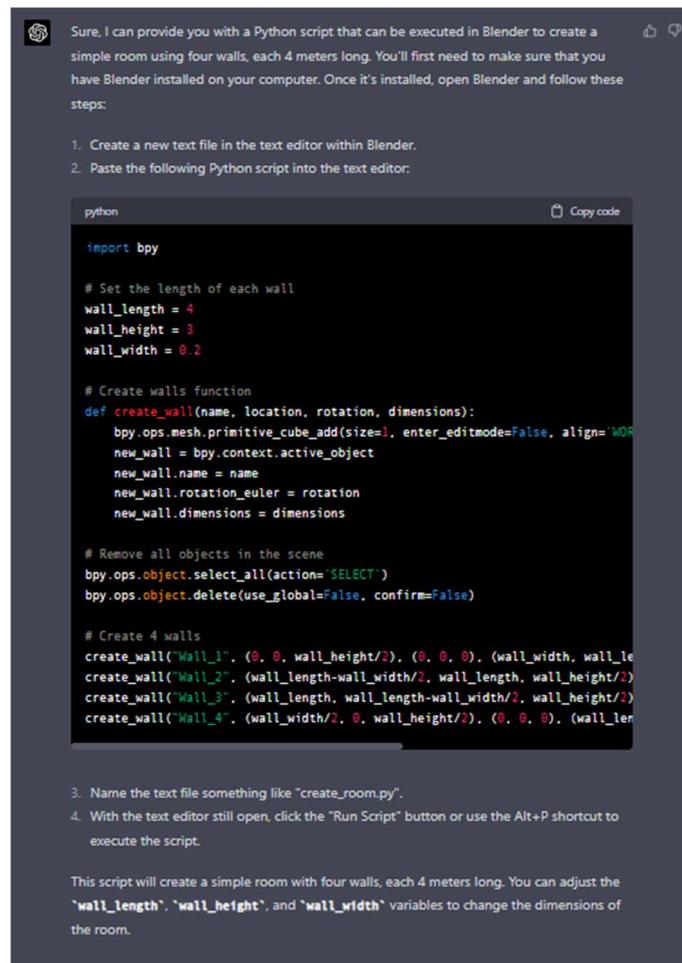
La potenza di questo plug-in non termina qui, in quanto consente anche di ottenere risultati fotografici (nonostante questa sia una casistica non presentata per questo progetto che, come dicevo, ha come obiettivo uno stile pennellato) e può rendere davvero bene questo look fotorealistico, sebbene nel disegno, soprattutto creativo sia molto più performante.



Fase di Generazione e Prototipazione del Gioco

"Poi il team di sviluppo disse portiamo tutto su Unity, e valutò Unreal engine."

in questa fase, si valuta il plug-Open AI di Chat GPT in di Unity per la realizzazione di azioni tramite comandi prompt, si utilizza l'intelligenza artificiale per creare codice per Unity o per Blender da importare in unity. E si esaminano plug-in per la realizzazione di contenuti da prompt per Unreal Engine.



Sure, I can provide you with a Python script that can be executed in Blender to create a simple room using four walls, each 4 meters long. You'll first need to make sure that you have Blender installed on your computer. Once it's installed, open Blender and follow these steps:

1. Create a new text file in the text editor within Blender.
2. Paste the following Python script into the text editor:

```
python
import bpy

# Set the length of each wall
wall_length = 4
wall_height = 3
wall_width = 0.2

# Create walls function
def create_wall(name, location, rotation, dimensions):
    bpy.ops.mesh.primitive_cube_add(size=1, enter_editmode=False, align='WORLD', location=location, rotation=rotation)
    new_wall = bpy.context.active_object
    new_wall.name = name
    new_wall.rotation_euler = rotation
    new_wall.dimensions = dimensions

# Remove all objects in the scene
bpy.ops.object.select_all(action='SELECT')
bpy.ops.object.delete(use_global=False, confirm=False)

# Create 4 walls
create_wall("Wall_1", (0, 0, wall_height/2), (0, 0, 0), (wall_width, wall_length, wall_height/2))
create_wall("Wall_2", (wall_length-wall_width/2, wall_length, wall_height/2), (0, 0, 0), (wall_width, wall_length, wall_height/2))
create_wall("Wall_3", (wall_length, wall_length-wall_width/2, wall_height/2), (0, 0, 0), (wall_width, wall_length, wall_height/2))
create_wall("Wall_4", (wall_width/2, 0, wall_height/2), (0, 0, 0), (wall_length, wall_width, wall_height/2))
```

3. Name the text file something like "create_room.py".
4. With the text editor still open, click the "Run Script" button or use the Alt+P shortcut to execute the script.

This script will create a simple room with four walls, each 4 meters long. You can adjust the 'wall_length', 'wall_height', and 'wall_width' variables to change the dimensions of the room.

"Poi il team di sviluppo disse: il prototipo è pronto e vide che era cosa buona e giusta, (ma andava modificato)"

Risultati

Come si evince dalla precedente sezione è stato possibile ottenere supporto con l'intelligenza artificiale in tutte le fasi sopra elencate.

Andando con ordine:

partendo dalla fase di prototipazione concettuale è piuttosto semplice mostrare come l'output verboso di Chat GPT anche senza addestramento possa essere decisamente di buona qualità e utile per creare contenuto rapidamente.

Possiamo dire che nel caso peggiore possiamo ottenere un Lorem Ipsum di contesto che comunque può essere indirizzato per ottenere uno spunto più specifico ed essere aggiustato manualmente su un programma di testo.

Chat GPT inoltre è piuttosto efficiente a fare cambi di registro o ad esempio passare da una prima ad una terza persona o un tempo verbale rendendo velocemente e in maniera automatica una procedura tediosa e ripetitiva.

Ad esempio, il solo scrivere manualmente i dialoghi per uno scrittore che usa una tastiera, ha un flusso costante di idee e non fa ricerche approfondite per un totale di 1000 parole si potrà ottenere in circa 25 minuti; uno scrittore che scrive a mano dovrebbe impiegare circa 50 minuti¹²⁰.

Sebbene spesso sia necessario dare in pasto informazioni piuttosto specifiche su di un dato argomento per addestrare il prompt in realtà possiamo assumere che questo tempo di ricerca possa essere integrato in quello che è il normale flusso di lavoro. È piuttosto ovvio immaginare come non sia possibile stilare un articolo o redigere uno script se non si hanno le competenze specifiche in merito. La lettura dell'articolo o delle fonti, in ogni caso, sebbene sia sempre consigliata si può anche demandare interamente all'AI o eventualmente semplificata, sintetizzata con un riassunto o una riscrittura nel caso non si comprendessero concetti o spiegazioni. Il paradosso applicativo è che il creativo può effettivamente andare anche nello specifico di concetti che lui stesso non comprende o su cui non ha svolto studi a patto che la trattazione per l'apprendimento usata sia affidabile e validata in partenza.

Tralasciando la questione della validazione e dell'affidabilità dell'articolo (il che aprirebbe ulteriori altre questioni) entrando nel merito creativo l'artista digitale ha a disposizione un suggeritore che potenzialmente ha un bagaglio di conoscenze maggiore nell'ambito trattato dello stesso autore. Questa questione porta un'implicazione immediata: il collasso della figura del "artigiano digitale".

Cercando di esprimere meglio questo concetto definisco “artigiano digitale” quella figura che a partire da un input di un committente esegue quelle che sono le specifiche del progetto con il maggiore virtuosismo tecnico possibile.

Uno sceneggiatore, nel merito di questa sezione, potrebbe perdere la necessità di avere qualcuno che espanda le sue idee e realizzi la trattazione dettagliata, non avrà quindi bisogno di riscrivere per modificare e ogni modifica perderà quindi il peso della digitazione che citavo in precedenza.

È quindi difficile immaginare un futuro in cui queste figure professionali non si modifichino per sfruttare al massimo la creatività e il gusto compositivo a scapito della scrittura generata integralmente dalle mani del creativo.

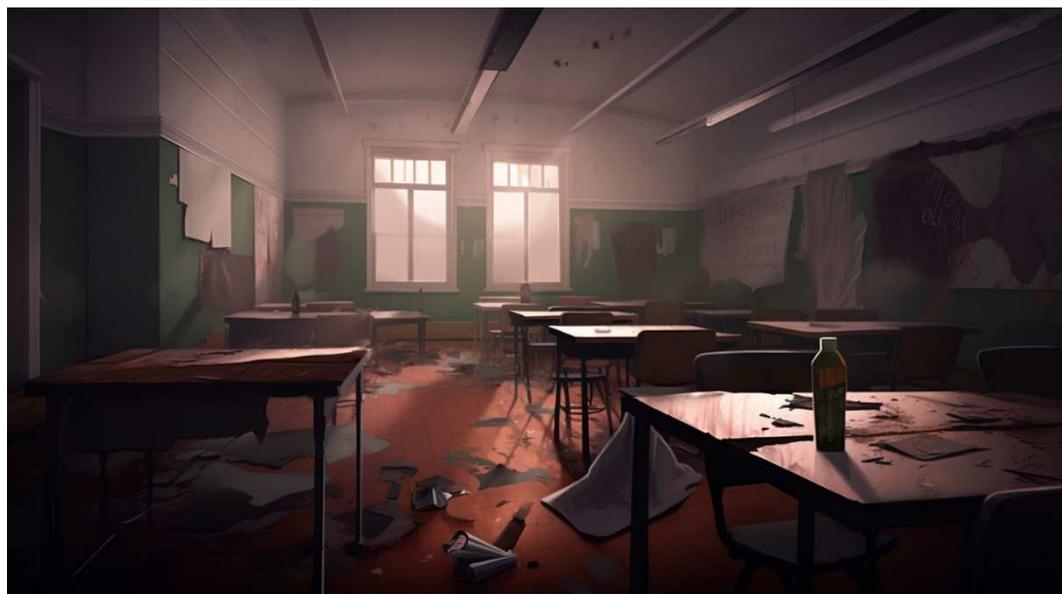
Al momento il problema principale in merito risiede in una poco chiara gestione dei diritti in merito a ciò che verrà generato dall’AI. La questione su questo copyright, come avevo già anticipato, in America è già definita e non è possibile quindi esercitare alcun diritto di autore su contenuti generati da AI.

Il che apre diverse questioni etiche e professionali sul futuro, se sia o meno corretto mettere diritto d’autore su un libro generato con l’intelligenza artificiale.

Il GDPR attualmente garantisce ad esempio il diritto d’autore sulle composizioni, pertanto, almeno in Europa, la trattazione è ancora aperta.

Nella fase di generazione di immagini quello che è stato esplicitato sopra è diventato ancora più evidente come si evince dalle due immagini una generata dalle mani talentuose di Edoardo Audino l’altra dall’intelligenza artificiale integralmente. Sebbene chi sta scrivendo questa tesi abbia competenze di disegno e composizione creativa spesso e volentieri non ha la capacità di disegnare determinati artwork, come potrebbe non averne chi ha intenzione di farsi ispirare per realizzare un contenuto per un’esperienza in realtà virtuale.

Lungi da me criticare il lavoro creativo svolto ma appare evidente come questo strumento possa sconvolgere il mercato e la creazione di contenuto impostando nuovi standard.



Le luci sono coerenti, le proporzioni pure, i colori ben armonizzati e si percepisce ciò che si voleva comunicare anche nell'immagine generata dall'intelligenza artificiale. Assumendo inoltre che è stata generata interamente da prompt addestrato a partire da semplici parole quali “rappresentami la classe di Robin come potrebbe vederla durante l'attacco di panico” possiamo notare come anche avessimo generato 20 prompt per un totale di 40 minuti di tempo generativo, con lo stesso tempo non riusciremmo ad avvicinarci neanche

lontanamente ad un tale risultato, certo riusciremmo ad abbozzare una scena ma non avremmo esplorato 80 possibilità ad alta definizione (20x4 a prompt).

Un'ipotesi non valutata, al momento disponibile solo in America in beta, è Leonardo.AI che genera già i layer dell'immagine e permetterà quindi un passaggio diretto su photoshop cosa che consentirà il massimo della scalabilità dell'output.

In quest'ottica è importante però integrare anche le capacità di selezione dei contributi per unire tutti gli output e ritoccare le immagini per renderle aderenti alla resa prevista.



Come si evince da quest'immagine sebbene sia un ottimo risultato dal punto di vista visivo non è in linea con l'estetica degli artwork di Audino.

Al contrario alcune immagini sono in linea con l'estetica del disegno ma sono carenti in termini di colori e alcuni oggetti generati sono fuori tema.

QAnche qui, sebbene in maniera meno evidente dal punto di vista del testo scritto, l'artigiano tecnicamente preparato ma con poco senso artistico rischia molto dal punto di vista lavorativo.

Vero è che, come dicevo il processo di integrazione del contenuto generato, potrebbe generare una nuova professione che richiede comunque l'occhio esperto di un addetto ai lavori per riuscire a uniformare il tutto in un'immagine unica coerente.

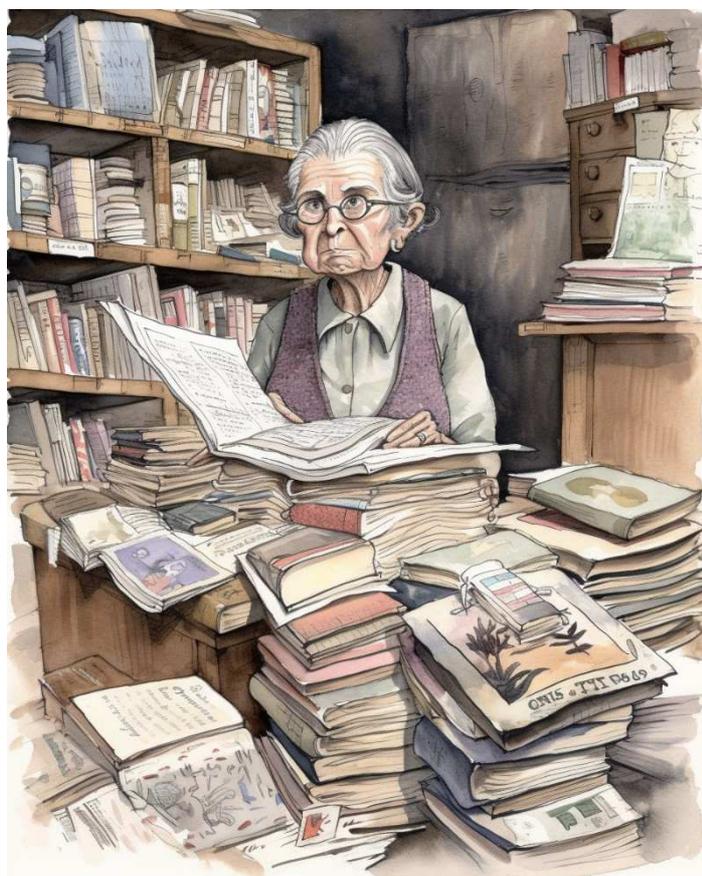


Al momento va notato che gli errori generativi sulle figure umane sono diminuiti. Midjourney v5 ha risolto molto in questo senso, le dita generate o le figure umane in generale può capitare abbiano alcune deformità ma il dato è stato molto ridimensionato, basti pensare ai deep-fake di Papa Francesco con un piumino da Trapper che hanno ingannato i media¹²¹.

Non è distante il giorno in cui non si vedrà differenza tra una foto o un dipinto rispetto ad un'opera artigianale, ma senz'altro ad oggi gli errori esistono e tanti. Per ottenere l'immagine della classe menzionata sopra mi ci sono voluti diversi prompt e iterazioni e anche il risultato finale sebbene soddisfacente non era esattamente quello che avevo in mente.

Certo anche questa ricerca senz'altro potrebbe sostituire il lavoro del bozzetto o influenzarlo molto ma la chiara visione artistica se la si ha in mente spesso e volentieri non trova esatta realizzazione nell'output del prompt.

Se andiamo difatti ad analizzare ad esempio:



*Prompt: A striking and evocative watercolor cartoon-style illustration of an intimidating and steadfast elderly female teacher, surrounded by books, artifacts, and educational materials in a **well-ordered, structured classroom**. The subject is depicted with a stern, unwavering gaze, her eyes filled with years of knowledge and experience, while her attire reflects the authoritative presence of a seasoned educator. The art director meticulously oversees the creation of this illustration, ensuring that the color palette is bold and commanding, with deeper, more intense shades evoking a sense of discipline and respect. The composition is carefully balanced, with the teacher taking center stage, while the **surrounding elements create a sense***

***of depth and structure.** The illustrative style merges the charm of watercolor textures with the dynamic nature of cartoon aesthetics, resulting in a powerful and engaging artwork that captures the spirit and essence of a formidable educator. The scene is brought to life using a high-resolution digital canvas, with a Wacom Cintiq Pro tablet and stylus for precise control and intuitive creation. The final image is rendered in a versatile aspect ratio of 4:5, ensuring adaptability for various formats and applications. --ar 4:5 --q 2 --v 5 --*

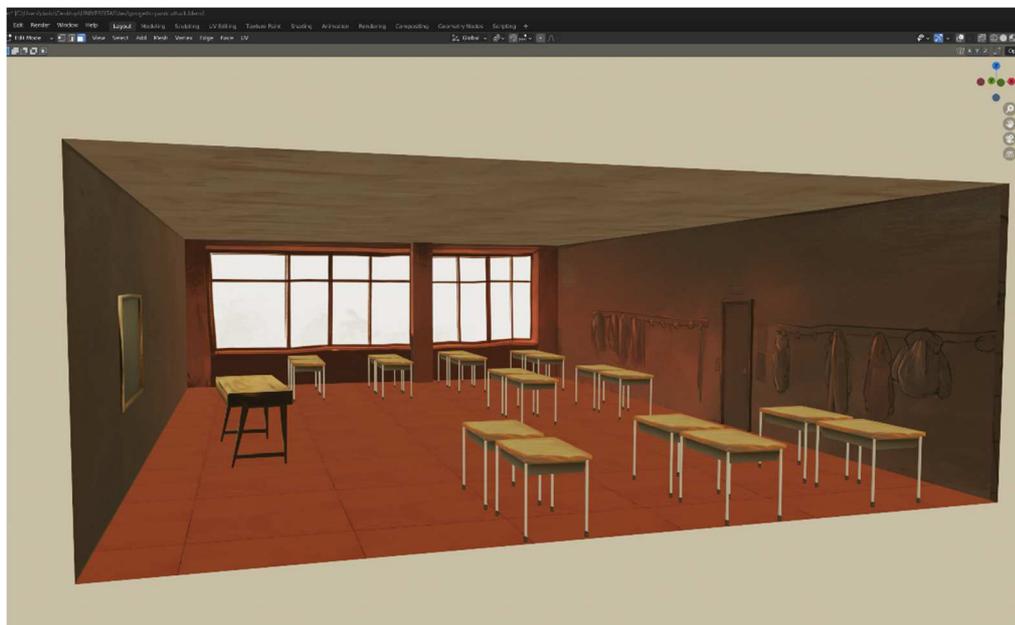
Come si nota dalle parti in grassetto queste richieste ad esempio sono state ignorate, sebbene l'esito però dal prompt sia migliorato visto che si è generato con chat gpt dalla semplice richiesta "Genera un prompt di testo per rappresentare la professoressa di Robin che mi hai appena descritto nello stile che ti ho indicato in precedenza"

In questo caso anche la richiesta alla base è stata vaga, e il risultato tuttavia ha centrato il compito di rappresentare l'insegnante a partire da una descrizione fornita. In questo caso la descrizione arrivava direttamente dalla generazione di una descrizione del personaggio interna a Chat GPT (si potrebbe passare anche una descrizione eventualmente anche abbozzata delle caratteristiche del personaggio).

Quello che si nota però è che, se da solo vado a inserire il testo nel prompt, spesso, non ottengo il risultato vicino ai miei desiderata.

Per quanto riguarda la fase di modellazione e di Texturing basta osservare ciò che è stato possibile ottenere in poco tempo:

Come si evince dall'immagine, il risultato nell'engine 3d è molto vicino all'artwork di Audino e gli oggetti sono tutti con pochi poligoni. Come da specifica



Questo risultato una volta compreso il flusso operativo, è frutto di poco tempo. Per generare un modello low poli texturizzato a partire da un project from view stiamo parlando dell'ordine di un ora a seconda della complessità, mentre per la fase di texture è probabile che la generazione con stable diffusion per il rapid texturing possa produrre in una decina di minuti un'alternativa efficace ed una risposta alla generazione procedurale di textures.

Un aspetto poco trattato ma che ugualmente può fornire utili spunti è il fatto che blender possa includere anche al suo interno direttamente un prompt per la

generazione di oggetti con linguaggio naturale, che converte e trascrive del testo su chat gpt per generare in codice python. In questo modo possiamo convertire quello che è un'interfaccia grafica con una curva di apprendimento spesso molto ripida in un tool fruibile da testo. E far eseguire comandi anche e spesso molto complessi composti cioè da molte fasi o dalla generazione che “fatta a mano” richiederebbe molto più tempo. Un esempio banale è una torre di cubi animata con il rigid body sopra un piano che faccia da collider ma si possono generare anche output più complessi come, ad esempio, questa navicella spaziale¹²².

La generazione con il project from view garantisce inoltre il fatto che le proporzioni degli oggetti sono mantenute, e l'estetica può essere controllata dall'inizio della generazione, lavorando quindi praticamente a ritroso.

Dal punto di vista della procedura il tutto aiuta molto, snellisce il flusso operativo e lascia l'attenzione all'ottimizzazione del modello.

In ultimo il codice su Unity è piuttosto standard, spesso e volentieri non si ottiene un “play” che dia il risultato che si voleva comunicare specie per comandi più complessi ma snellisce quelli che sono i task base o i copia incolla standard, ad esempio, il setting di una camera in first person in VR, generare piccole animazioni, posizionare a random oggetti e luci.

Scomponendo in sottotask e facendo del lavoro di taglia e cuci però è possibile generare anche codici piuttosto articolati. Tuttavia, è ancora lontano il momento in cui un non programmatore possa digitando frasi di testo ottenere un'esperienza di gioco completa.

Unreal Engine ha un legame ancora più stretto con Chat GPT e dalla sua un ulteriore vantaggio una grande libreria di oggetti ad alta risoluzione adatta alla generazione di esperienze di questo tipo, la trattazione tuttavia non può ottenere i risultati richiesti poiché Nahmic non riesce a girare in maniera autonoma su Oculus Quest 2 nativamente. Però quello che è importante specificare è che l'esperienza VR con visori più performanti e macchine più performanti grazie a questa Game Engine potrebbe assumere i connotati di un'opera prima.

In questo caso, con unreal potremmo non preoccuparci della dimensione dei poligoni degli oggetti, potremmo quindi generare texture più complesse e ricche e far partire il tutto real time immergendoci nelle “parole” del creatore.

Conclusioni

Ad oggi nel 2023 è possibile integrare a basso costo e con un buon risparmio di tempo nel flusso di lavoro generativo di un'esperienza VR, l'intelligenza artificiale.

Da essa ci si può far aiutare dall'inizio, quando ancora il progetto è solo un'idea creativa nella mente dello sviluppatore, per farla nascere. Ma si può giungere ben oltre, e a 360° si può, usando quanto più possibile il linguaggio naturale, abbattendo quindi i paradigmi di comunicazione con il calcolatore, comunicare con i programmi e ottenere i risultati attesi.

Non vanno dimenticati però i problemi etici e di diritto d'autore che stanno nascendo quando ci si trova di fronte a questi nuovi strumenti.

Un risultato di un flusso di lavoro del genere già adesso o con gli sviluppi futuri magari tra un anno, o magari dieci, potrà essere stravolgente dal punto di vista operativo andando a schiacciare la classe dell'"artigiano" digitale.

Perciò è quanto più necessario affidarsi a creativi digitali, esperti del settore, che possano supervisionare e sfruttare le possibilità offerte da questi strumenti.

L'IA ci pone di fronte ad un quesito importante: può un'intelligenza artificiale essere un artista, un designer, un compositore? O sarà sempre l'uomo a tirare queste redini creative?

A questo quesito mi sento di rispondere dopo questo percorso universitario in Ingegneria del cinema e dei mezzi di comunicazione, che chi crea, da quando l'uomo primitivo incidava le rocce o scriveva sulle pareti delle grotte, aveva una necessità, un bisogno intrinseco e viscerale: comunicare.

L'arte, il cinema, i videogiochi, le esperienze in VR sono solo uno dei tanti metodi in cui si cerca di passare un messaggio anche un virtuosismo o un decostruttivismo.

Con questo piccolo contributo spero di aver dimostrato che ancora l'uomo, quello creativo, quasi demiurgo, è necessario in quanto un'esperienza, un'idea esiste per gli altri quando vi è il bisogno umano di comunicare. E quindi l'IA è lo scalpello forse (per ora) migliore della storia.

-
- ¹ McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*.
- ² Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460.
- ³ Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
- ⁴ Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210-229.
- ⁵ Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- ⁶ LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- ⁷ Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.
- ⁸ Jackson, P. (1986). *Introduction to expert systems*. Addison-Wesley.
- ⁹ Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- ¹⁰ Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Viking.
- ¹¹ Bostrom, N. (2014). *Super intelligence: Paths, dangers, strategies*. Oxford University Press.
- ¹² Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. Knopf.
- ¹³ LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444
- ¹⁴ Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- ¹⁵ Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- ¹⁶<https://designguidelines.withgoogle.com/cardboard/designing-for-google-cardboard/introduction.html>
- ¹⁷ <https://developer.oculus.com/>
- ¹⁸<https://unity.com/solutions/virtual-reality>
- ¹⁹ <https://www.unrealengine.com/en-US/vr>
- ²⁰ <https://docs.unity3d.com/Manual/VROverview.html>
- ²¹ <https://docs.unrealengine.com/en-US/Platforms/VR/index.html>
- ²² <https://www.blender.org/>
- ²³ <https://www.autodesk.com/products/maya/overview>
- ²⁴ <https://www.autodesk.com/products/3ds-max/overview>
- ²⁵ <https://www.adobe.com/products/photoshop.html>
- ²⁶ <https://www.substance3d.com/products/substance-painter/>
- ²⁷ <https://quixel.com/mixer>
- ²⁸ <https://developer.oculus.com/>
- ²⁹ <https://developer.vive.com/resources/>
- ³⁰ <https://partner.steam>
- ³¹ <https://www.oculus.com/experiences/quest/>
- ³² <https://store.steampowered.com/app/250820/SteamVR/>
- ³³<https://developer.oculus.com/documentation/unity/unity-intro-vr/>
- ³⁴<https://developer.vive.com/resources/knowledgebase/vive-hand-tracking-sdk/>
- ³⁵ <https://partner.steam>
- ³⁶ <https://developer.nvidia.com/optix-denoiser>
- ³⁷ https://www.researchgate.net/publication/221605467_3D_mesh_segmentation_a_survey
- ³⁸ <https://www.sciencedirect.com/science/article/pii/S0097849312001707>
- ³⁹ <https://www.sciencedirect.com/science/article/pii/S0097849312001707>
- ⁴⁰ <https://dl.acm.org/doi/10.1145/3272127.3275090>
- ⁴¹ <https://dl.acm.org/doi/10.1145/3272127.3275090>
- ⁴²<https://arxiv.org/abs/1710.10196>
- ⁴³ <https://ieeexplore.ieee.org/abstract/document/7410402>

-
- ⁴⁴ <https://arxiv.org/abs/1508.06576> [4] <https://dl.acm.org/doi/10>
- ⁴⁵ <https://dl.acm.org/doi/10>
- ⁴⁶ https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Gatys_Image_Style_Transfer_CVPR_2016_paper.html
- ⁴⁷ <https://arxiv.org/abs/1603.08155>
- ⁴⁸ <https://arxiv.org/abs/1705.06830>
- ⁴⁹ <https://arxiv.org/abs/1604.08610>
- ⁵⁰ <https://arxiv.org/abs/1508.06576>
- ⁵¹ <https://dl.acm.org/doi/10.1145/3130800.3130891>
- ⁵² <https://dl.acm.org/doi/10.1145/3173574.3173616>
- ⁵³ <https://dl.acm.org/doi/10.1145/566570.566576>
- ⁵⁴ <https://dl.acm.org/doi/10.1145/383259.383296>
- ⁵⁵ https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Gatys_Image_Style_Transfer_CVPR_2016_paper.html
- ⁵⁶ <https://arxiv.org/abs/1603.08155>
- ⁵⁷ <https://arxiv.org/abs/1705.06830>
- ⁵⁸ <https://arxiv.org/abs/1604.08610>
- ⁵⁹ <https://arxiv.org/abs/1604.08610>
- ⁶⁰ <https://arxiv.org/abs/1703.07511>
- ⁶¹ <https://arxiv.org/abs/1804.04732>
- ⁶² <https://arxiv.org/abs/1703.06868>
- ⁶³ <https://unity.com/>
- ⁶⁴ <https://unity.com/solutions/multiplatform>
- ⁶⁵ <https://learn.unity.com>
- ⁶⁶ <https://assetstore.unity.com/>
- ⁶⁷ <https://unity.com/solutions/visual-scripting>
- ⁶⁸ <https://unity.com/solutions/immersive>
- ⁶⁹ <https://docs.unity3d.com/Manual/SinglePassStereoRendering.html>
- ⁷⁰ <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@1.0/manual/index.html>
- ⁷¹ <https://unity.com/products/unity.artengine>
- ⁷² <https://docs.unity3d.com/Manual/AudioSpatializerSDK.html>
- ⁷³ <https://unity.com/products/unity-teams>
- ⁷⁴ <https://unity.com/machine-learning>
- ⁷⁵ Fonti: [1] <https://www.unrealengine.com/>
- ⁷⁶ <https://www.unrealengine.com/en-US/raytracing>
- ⁷⁷ <https://www.unrealengine.com/en-US/platforms>
- ⁷⁸ <https://docs.unrealengine.com/en-US/ProgrammingAndScripting/Blueprints/index.html>
- ⁷⁹ <https://www.unrealengine.com/marketplace/en-US/store>
- ⁸⁰ <https://docs.unrealengine.com/en-US/SharingAndReleasing/XRDevelopment/index.html>
- ⁸¹ <https://docs.unrealengine.com/en-US/RenderingAndGraphics/VR/index.html>
- ⁸² <https://docs.unrealengine.com/en-US/InteractiveExperiences/Physics/index.html>
- ⁸³ <https://www.unrealengine.com/en-US/industry/architecture>
- ⁸⁴ <https://docs.unrealengine.com/en-US/Audio/index.html>
- ⁸⁵ <https://www.unrealengine.com/en-US/blog/collaborating-in-unreal-engine-4>
- ⁸⁶ <https://developer.nvidia.com/deep-learning-sdk>
- ⁸⁷ <https://huggingface.co/spaces/dalle-mini/dalle-mini>
- ⁸⁸ https://rudalle.ru/en/demo_emoji
- ⁸⁹ <https://labs.openai.com/>
- ⁹⁰ <https://www.midjourney.com/home/?callbackUrl=%2Fapp%2F>
- ⁹¹ <https://www.lmu.de/en/newsroom/news-overview/news/revolutionizing-image-generation-by-ai-turning-text-into-images.html>
- ⁹² <https://github.com/CompVis/stable-diffusion>

-
- ⁹³ <https://stability.ai/blog/stable-diffusion-announcement>
- ⁹⁴ Rombach; Blattmann; Lorenz; Esser; Ommer (June 2022). High-Resolution Image Synthesis with Latent Diffusion Models (PDF). International Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA.
- ⁹⁵ <https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/>
- ⁹⁶ <https://github.com/AUTOMATIC1111/stable-diffusion-webui-feature-showcase>
- ⁹⁷ <https://huggingface.co/CompVis/stable-diffusion-v1-4>
- ⁹⁸ <https://huggingface.co/CompVis/stable-diffusion-v1-4>
- ⁹⁹ <https://stability.ai/blog/stable-diffusion-public-release>
- ¹⁰⁰ <https://github.com/ShieldMnt/invisible-watermark/blob/9802ce3e0c3a5ec43b41d503f156717f0c739584/README.md>
- ¹⁰¹ <https://huggingface.co/spaces/CompVis/stable-diffusion-license>
- ¹⁰² <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>
- ¹⁰³ Stiennon, Nisan, Ouyang, Long e Wu, Jeffrey, Learning to summarize with human feedback, in Advances in Neural Information Processing Systems
- ¹⁰⁴ Leo Gao, John Schulman e Jacob Hilton, Scaling Laws for Reward Model Overoptimization, in arXiv:2210.10760 [cs, stat]
- ¹⁰⁵ W. Bradley Knox e Peter Stone, Augmenting Reinforcement Learning with Human Feedback (PDF), Proceedings of the ICML Workshop on New Developments in Imitation Learning.
- ¹⁰⁶ GPT-4 Technical Report [Report tecnico di GPT-4] (PDF), su cdn.openai.com, pp. 98.
- ¹⁰⁷ ChatGPT: Optimizing Language Models for Dialogue, su OpenAI, 30 novembre 2022
- ¹⁰⁸ OpenAI, ChatGPT: Optimizing Language Models for Dialogue
- ¹⁰⁹ ChatGPT proves AI is finally mainstream and things are only going to get weirder, su The Verge.
- ¹¹⁰ Proximal Policy Optimization (PPO) Explained, su towardsdatascience.com.
- ¹¹¹ ChrisHMSFT, Azure OpenAI Service models - Azure OpenAI, su learn.microsoft.com.
- ¹¹² ChatGPT: The Most Advanced AI Chatbot in 2022, su ChatGPT Pro
- ¹¹³ Kasper Groes Albin Ludvigsen, The Carbon Footprint of ChatGPT, su Medium, 22 dicembre 2022.
- ¹¹⁴ ChatGPT ora ha un volto e si chiama Alice. Ora potrai conversare con l'AI parlando al microfono, su redhotcyber.com
- ¹¹⁵ Microsoft pronta a integrare la nuova versione di ChatGpt, su ansa.it, 13 marzo 2023.
- ¹¹⁶ Microsoft sta espandendo la sua IA su supercomputer da centinaia di milioni di dollari. Quasi pronto GPT-4 per i video e gli audio, su dday.it.
- ¹¹⁷ Le migliori estensioni browser per sfruttare al massimo ChatGpt, su wired.it, 14 marzo 2023.
- ¹¹⁸ Come Integrare ChatGPT in Excel e Google Sheets, su punto-informatico.it.
- ¹¹⁹ ChatGPT Plus: ecco come utilizzarlo gratis, su html.it.
- ¹²⁰ <https://www.schemaninja.com/it/how-long-does-it-take-to-write-1000-words/>
- ¹²¹ <https://www.rainews.it/articoli/2023/03/papa-francesco-col-piumino-bianco-da-trapper-fa-milioni-di-clic-la-foto-fake--d054802c-30ae-4632-9d5a-193567a29726.html>
- ¹²² <https://www.youtube.com/watch?v=NmPFKVRrh7c>