POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Biomedica



Tesi di Laurea Magistrale

Correzione di artefatti nelle immagini di microscopia confocale a fluorescenza: dall'euristica all'Intelligenza Artificiale

Relatori Prof. Massimo SALVI Prof.ssa Kristen Mariko MEIBURGER Ing. Francesco BRANCIFORTI Candidata

Maura MAGGIORE

Luglio 2023

Indice

Elenco delle figure					
1	Introduzione				
	1.1	Imagin	ng ottico medico	9	
	1.2	Micros	scopia Confocale a Fluorescenza	10	
	1.3	Strum	entazione e schema di funzionamento	11	
	1.4	Immag	gini di microscopia confocale a fluorescenza	12	
		1.4.1	Artefatti nelle immagini di microscopia confocale a fluorescenza	13	
2	Mat	teriali	e Metodi	15	
	2.1	Datas	et di immagini	15	
		2.1.1	Dataset 1: Cardiosfere	15	
		2.1.2	Dataset 2: Nuclei U2OS	19	
		2.1.3	Dataset 3: Actina e Mitocondri	20	
	2.2	Analis	si degli artefatti	21	
	2.3	Algori	tmo	25	
		2.3.1	Pipeline algoritmo generico	26	
		2.3.2	Funzione di incremento della luminosità degli oggetti $\ .\ .\ .$	27	
		2.3.3	Denoising	31	
		2.3.4	Miglioramento della nitidezza	35	
	2.4	Algori	tmo specializzato nella correzione dell'autofluorescenza	37	
		2.4.1	Calibrazione dell'autofluorescenza	39	
	2.5	Gener	azione di immagini prive di artefatti tramite GAN \ldots .	44	
		2.5.1	GAN: Reti Neurali Generative Avversarie	44	
		2.5.2	Architettura Pix2Pix	47	
		2.5.3	Settaggio iperparametri del modello Pix2Pix utilizzato per		
			la generazione di immagini corrette	49	
	2.6	Metric	che di validazione	51	
		2.6.1	CII	51	
		2.6.2	MS-SSIM	52	
		2.6.3	MANIQA	53	

3	Risultati		55
	3.1	Risultati algoritmo generico	55
	3.2	Risultati algoritmo specializzato nella correzione dell'autofluorescenza	59
	3.3	Risultati modello generativo addestrato	60
	3.4	Risultati metriche di validazione	63
4	Conclusioni		67
Bi	Bibliografia		

Elenco delle figure

$1.1 \\ 1.2$	Schema di funzionamento della microscopia confocale a fluorescenza[2]. Esempi di immagini acquisite tramite tecnica di microscopia confo-	12
	cale a fluorescenza $[3, 4, 5]$.	13
2.1	Struttura delle cardiosfere analizzata tramite tecnica di microscopia confocale. Le tre sezioni raffigurano la pila osservata lungo la linea mediana (in alto a sinistra), assi X e Y (a destra). Attraverso l'applicazione della colorazione con F-actina marcata con Falloidina in bianco si apprezza il citoscheletro, mentre in blu si osservano in nuclei, evidenziati con il DAPI[3].	17
2.2	Esempio di slice estratte a campione da diversi volumi contenuti nel	
	"CARE algorithm dataset"[3]	18
2.3	Esempio di immagini nel canale blu estratte da slice contenute nelle	
	acquisizioni 3D	19
2.4	Esempio di immagini di nuclei $U2OS[4]$	20
2.5	Esempio di immagini di actina e mitocondri estratte dal GigaScience Database. Le due sotto-figure in alto rappresentano la struttura dell'actina acquisita a $60 \times$ NA mentre quelle in basso raffigurano le strutture dei mitocondri acquisite a $20 \times$ NA (basso SNR a sinistra	
2.6	e alto SNR a destra)[5] Slice estratte con ordine di profondità crescente da una delle acqui-	21
2.7	sizioni 3D di cardiosfere	23
	visibilità e rumore di background.	24
2.8	Esempi di acquisizioni in cui è visibile il fenomeno dell'autofluorescenza.	25
2.9	Diagramma di flusso che descrive gli step sui quali si basa l'algoritmo	
	generico per il processing delle immagini dei diversi dataset.	26
2.10	Esempio di funzione di incremento dei valori di intensità luminosa	00
0.11	dei pixel con il metodo del contrast stretching	28
2.11	nell'algoritmo	31

2.12	Schema di funzionamento della trasformata wavelet sulle immagini[13].	33
2.13	Esempio di effetto dell'incremento della luminosità degli oggetti con	
	e senza successivo filtraggio.	34
2.14	Esempio di trasformazione effettuata attraverso l'applicazione del- l'algoritmo su un'immagine di nuclei di cardiosfera.	36
2.15	Esempio di trasformazione effettuata attraverso l'applicazione del-	
	l'algoritmo su un'immagine di nuclei U2OS.	36
2.16	Esempio di trasformazione effettuata attraverso l'applicazione del- l'algoritmo su immagini di actina e mitocondri.	37
2.17	Diagramma di flusso relativo all'algoritmo specializzato nella corre- zione dell'autofluorescenza.	38
2.18	Esempi di maschere di segmentazione dei nuclei ottenute con il model- lo ConvUNeXt sovrapposte alle immagini processate con l'algoritmo generale dalle quali sono state ottenute	39
2.19	Esempio di estrazione dell'autofluorescenza per l'applicazione della calibrazione delle luminosità.	40
2.20	Esempio di leggi di modifica dei valori dei pixel dell'area dell'auto-	
	fluorescenza.	41
2.21	Diagramma di flusso che illustra gli step seguiti per il controllo dell'autofluorescenza.	42
2.22	Esempio di effetto risultante dalla calibrazione delle luminosità	
	dell'autofluorescenza.	43
2.23	Esempio di correzioni di intensità dell'autofluorescenza effettuate con tre diverse percentuali di riduzione	43
2.24	Schema esemplificativo della struttura di una GAN per la generazione di immagini[16]	45
2.25	Confronto tra la struttura generica encoder-decoder e la struttura del generatore di una rete Pix2Pix caratterizzata dalla presenza delle	10
0.00	skip connections. $[17]$.	48
2.26	Esempio di immagini di nuclei di cardiostera migliorate generate dal modello addestrato.	50
3.1	Risultato dell'applicazione dell'algoritmo generico alle acquisizioni di nuclei di cardiosfera.	56
3.2	Risultato dell'applicazione dell'algoritmo generico alle immagini di nuclei U2OS.	56
3.3	Risultato dell'applicazione dell'algoritmo generico alle immagini di	
	mitocondri e actina	57
3.4	Confronto tra le immagini di mitocondri e actina migliorate con	
	l'acquisizione e con l'algoritmo.	58

3.5	Risultato dell'applicazione dell'algoritmo specializzato nella corre-	
	zione dell'autofluorescenza nelle immagini dei nuclei di cardiosfera	60
3.6	Esempio di immagini sintetiche migliorate generate dal modello	
	addestrato a partire dalla loro versione originale	61
3.7	Confronto tra il risultato prodotto dall'algoritmo e il risultato	
	prodotto dalla GAN su una stessa immagine	62
3.8	Valori di CII calcolati sulle immagini del test set generate dalla GAN	
	e sulle immagini del test set processate dall'algoritmo	64
3.9	Valori di MS-SSIM calcolati sulle immagini del test set generate	
	dalla GAN e sulle immagini del test set processate dall'algoritmo	65
3.10	Valori di MANIQA calcolati sulle immagini del test set originali,	
	sulla loro versione rielaborata generata dalla GAN e sulle immagini	
	del test set processate dall'algoritmo	66

Sommario

La microscopia confocale a fluorescenza è una delle metodologie di imaging ottico medico più utilizzate in ambito diagnostico ed in particolare nel campo dell'anatomia patologica. L'analisi di strutture cellulari, lo studio del funzionamento dei vasi sanguigni e l'indagine delle cellule tumorali sono solo alcune delle applicazioni della microscopia confocale a fluorescenza, che si rivela un metodo molto valido anche per i vantaggi che presenta dal punto di vista della semplicità e non invasività del suo funzionamento. La possibilità di ottenere immagini dettagliate di oggetti biologici a livello cellulare o subcellulare è la potenzialità che contraddistingue questa tecnica. Tuttavia, la presenza di artefatti può compromettere la buona visibilità degli oggetti all'interno delle immagini acquisite, pregiudicando una successiva corretta diagnosi. Il presente lavoro di tesi si incentra sull'implementazione di un metodo automatico per la rimozione di artefatti presenti in immagini di microscopia confocale a fluorescenza.

L'obiettivo di miglioramento delle immagini sulle quali si è svolto il progetto di tesi è stato raggiunto strutturando un metodo che si concentra sull'enfatizzare la visibilità di tutti gli oggetti di interesse che risultano poco evidenti, diminuendo inoltre la presenza di rumore ed interferenze luminose, preservando contemporaneamente l'informazione rilevante dei dati di input. Le problematiche individuate sono dovute sia al processo di acquisizione che a fenomeni derivanti dalle proprietà luminose degli oggetti biologici. L'utilizzo di diversi dataset sui quali è stato sviluppato e testato il metodo ha permesso di ottenere un algoritmo generale e robusto che, infine, è stato specializzato nella correzione di acquisizioni 3D di campioni biologici affetti dal fenomeno dell'autofluorescenza.

In seguito alla verifica della validità dei risultati ottenuti dall'algoritmo si è proceduto con l'addestramento di una Generative Adversarial Network (GAN) con architettura pix2pix, col fine di produrre un modello addestrato in grado di generare immagini quanto più possibile simili a quelle risultanti dal processing effettuato tramite l'algoritmo.

La finalità dell'utilizzo del modello generativo addestrato è quella di superare i limiti dell'algoritmo nel tentativo di ottenere risultati caratterizzati da una maggiore qualità, sfruttando la capacità di generalizzazione della rete che, essendo un modello non euristico, è in grado di stabilire un mapping più generico tra il dominio delle immagini originali e il dominio delle immagini migliorate, rispetto a quello definito dall'algoritmo euristico.

I due metodi proposti hanno dunque condotto alla risoluzione del task, producendo dei risultati validi in maniera complementare.

Il presente lavoro di tesi costituisce un possibile punto di partenza per lo sviluppo di protocolli di image processing utili ad agevolare l'analisi di immagini di strutture cellulari a scopi diagnostici.

Capitolo 1 Introduzione

Nel primo capitolo viene presentata una panoramica della tipologia di immagini sulle quali si è sviluppato il lavoro di tesi, proponendo una generale descrizione delle tecniche di imaging con le quali vengono acquisite e analizzando le loro applicazioni in ambito medico e diagnostico. L'obiettivo principale del progetto di tesi è la costruzione di un algoritmo automatico per la correzione di artefatti in immagini di microscopia confocale, caratterizzato dalla conservazione del contenuto semantico dei dati sul quale esso agisce e la successiva generazione di immagini equivalenti a quelle proposte come output dall'algoritmo, tramite l'utilizzo di reti neurali generative. Per contestualizzare al meglio il lavoro, nel capitolo corrente viene descritta con maggiore attenzione la tecnica di acquisizione delle immagini costituenti i dataset di partenza, con un focus sulla strumentazione necessaria per la sua applicazione, sui fondamenti fisici su cui la tecnica si basa e sui suoi principi di funzionamento.

Infine vengono discusse le caratteristiche fondamentali dei dati registrati con la modalità analizzata.

1.1 Imaging ottico medico

L'imaging ottico in ambito medico include tutte le tecniche di acquisizione e visualizzazione di immagini di componenti del corpo umano come tessuti, organi, cellule o strutture subcellulari a differenti scale di grandezza e dettaglio. Le immagini acquisite tramite queste tecnologie sono prodotte dall'elaborazione di dati forniti da fenomeni fisici quali la diffusione, l'assorbimento, l'emissione e la riflessione di fotoni, sfruttando le proprietà ottiche da cui sono caratterizzati i tessuti biologici, in funzione degli oggetti che si vogliono analizzare.

Le tecniche di imaging ottico hanno come obiettivo quello di raccogliere informazioni di tipo funzionale, anatomico e metabolico, al fine di contribuire allo studio in vivo di tessuti patologici o sani, a scopo diagnostico.

L'imaging Ottico si contraddistingue da altre metodologie di acquisizione di immagini mediche per alcuni vantaggi che lo definiscono. Il mancato utilizzo di radiazioni ionizzanti è uno tra i più importanti benefici di questo tipo di esami, in quanto permette di ripetere l'analisi senza rischi per i pazienti sui quali vengono applicate. Le tecnologie utilizzate, inoltre, permettono di effettuare una tipologia di esame non invasivo e sfruttano strumenti con dimensioni e costi ridotti.

Alcune metodiche di imaging ottico medico sono la Tomografia a Coerenza Ottica (OCT) che sfrutta la luce infrarossa per catturare immagini tridimensionali di tessuti come la retina oftalmica oppure tessuti dermatologici. Un altro esempio è quello dell'imaging di bioluminescenza in cui viene sfruttata la luce emessa da molecole bioluminescenti o agenti chimici per registrare meccanismi biologici, riuscendo ad analizzare l'attività cellulare oppure a localizzare la presenza di cellule tumorali. L'Angiografia ad indocianina verde (ICG) e la Microscopia Confocale a Fluorescenza sono ulteriori esempi di tecnologie di imaging ottico che consentono, rispettivamente, di valutare lo stato dei vasi sanguigni e di registrare immagini di campioni cellulari e di strati sottili di tessuti vivi.

Le applicazioni diagnostiche di questa modalità di imaging sono dunque numerose e si adattano ad individuare e, di conseguenza intervenire, su patologie che spaziano da quelle in ambito cardiovascolare, a quelle oftalmiche e dermatologiche, fino a quelle di tipo neurologico e oncologico.

1.2 Microscopia Confocale a Fluorescenza

Il lavoro di tesi si è sviluppato su diversi dataset di immagini acquisite tramite tecnica di microscopia confocale a fluorescenza. La microscopia confocale a fluorescenza è una tecnica di imaging che permette di acquisire immagini di campioni biologici con risoluzione ottica tridimensionale.

La materia biologica è strutturata in quattro dimensioni di cui tre spaziali e una temporale. La microscopia confocale a fluorescenza offre numerosi vantaggi per la ricostruzione 3D delle immagini dei campioni biologici rispetto alla microscopia tradizionale. Essa, attraverso il suo funzionamento, consente di acquisire immagini ottiche di porzioni sottili dei campioni da diverse profondità, creando una pila di immagini sequenziali a diversi piani focali. Queste immagini possono poi essere utilizzate per generare una rappresentazione 3D della struttura del campione utilizzando software di analisi che si basano su tecniche che si occupano di eliminare i contributi delle aree fuori fuoco.

La differenza principale, dunque, tra gli approcci di microscopia tradizionale e quelli di microscopia confocale è l'applicazione di metodi che consentono di sopprimere le informazioni che giungono da piani fuori fuoco[1]. L'eliminazione delle informazioni fuori fuoco è una caratteristica intrinseca della microscopia confocale, che è progettata specificamente per la loro rimozione, mentre la microscopia tradizionale fornisce un'immagine che rappresenta l'intero campo visivo, comprese le informazioni fuori fuoco.

L'eliminazione delle informazioni fuori fuoco è fondamentale per ottenere immagini di migliore qualità e per agevolare la capacità di analisi e comprensione dei campioni esaminati con la microscopia confocale.

1.3 Strumentazione e schema di funzionamento

L'acquisizione di immagini tramite tecnica di microscopia confocale a fluorescenza richiede l'utilizzo di una particolare strumentazione che consenta di acquisire immagini tridimensionali ad alta risoluzione di campioni di materiale biologico fluorescenti.

Il microscopio confocale è il componente principale del sistema di acquisizione. Esso sfrutta un sistema di scansione laser per illuminare il campione in esame con un fascio di luce puntiforme ed è poi fornito di un sistema di specchi e lenti per raccogliere il segnale fluorescente generato dall'oggetto biologico.

I campioni devono essere marcati con degli specifici fluorofori nelle zone che si vogliono studiare, in modo da ottenere i segnali di risposta solo dalle regioni di interesse una volta che queste vengono colpite da radiazioni luminose caratterizzate da una precisa lunghezza d'onda in grado di eccitare i fluorofori. Per questo scopo è necessaria una sorgente di luce che normalmente è un laser allo stato solido o un laser a fibra.

La luce del laser viene fatta passare attraverso un sistema di ottica di collimazione verso uno specchio dicromatico variabile e poi è riflessa sulla lente dell'obiettivo che fa sì che il raggio focalizzi su un punto dell'oggetto biologico analizzato.

Gli specchi di scansione servono per spazzare l'eccitazione luminosa proveniente dal laser su tutti i punti del campione per poter ricostruirne l'immagine.

Il sistema di acquisizione è fondamentale per la registrazione delle immagini a fluorescenza generate dal materiale biologico. La fluorescenza prodotta passa nuovamente attraverso la lente dell'obiettivo e lo specchio dicromatico per poi giungere al rilevatore che è un fotomoltiplicatore ad elevata sensibilità, che consiste in un dispositivo elettronico utilizzato per rilevare e amplificare segnali luminosi molto deboli. La soppressione dei contributi di luce sfocati avviene grazie ad un foro di dimensioni molto piccole posto sul piano dell'immagine in corrispondenza del punto focale del campione[2].



Figura 1.1: Schema di funzionamento della microscopia confocale a fluorescenza[2].

1.4 Immagini di microscopia confocale a fluorescenza

Le immagini di microscopia confocale a fluorescenza sono immagini a colori e in particolare, i diversi colori visibili derivano dalla lunghezza d'onda della luce emessa dai fluorofori con cui i campioni vengono marcati nelle loro componenti che si vogliono studiare. Questo permette di distinguere con un alto livello di dettaglio le varie strutture della materia biologica in esame.

Normalmente le immagini acquisite con le tecniche di microscopia confocale a fluorescenza si contraddistinguono per l'elevata risoluzione che fa sì che si possano visualizzare facilmente strutture cellulari di piccole dimensioni. Questo avviene grazie all'utilizzo di obiettivi ad alta apertura numerica, dunque caratterizzati da un'alta efficienza con cui la luce proveniente dal campione viene raccolta e focalizzata sull'immagine e anche grazie all'eliminazione della luce fuori fuoco.

La possibilità di ottenere acquisizioni a più canali che permettono di osservare in maniera distinta la materia biologica a diversi livelli di profondità consente di analizzare, all'interno della stessa acquisizione, più strutture biologiche e differenti eventi dinamici che occorrono al loro interno.

La microscopia a fluorescenza è dunque una metodica di imaging molto versatile e potente.



Figura 1.2: Esempi di immagini acquisite tramite tecnica di microscopia confocale a fluorescenza[3, 4, 5].

1.4.1 Artefatti nelle immagini di microscopia confocale a fluorescenza

La qualità delle immagini ottenibili tramite metodi di imaging di microscopia a fluorescenza e la robustezza della tecnologia che ne caratterizza il sistema di acquisizione non escludono la presenza di artefatti e distorsioni all'interno dei segnali registrati.

I più comuni artefatti che si possono osservare nelle immagini a fluorescenza sono spesso dovuti a fattori di errata illuminazione dei campioni e in questi casi ad esempio si parla di fotobleaching, che è un fenomeno che determina un offuscamento degli oggetti nelle immagini a causa della perdita di fluorescenza dei fluorofori usati per marcare le strutture biologiche di interesse in seguito ad una loro irradiazione prolungata durante l'acquisizione.

In molte situazioni si riscontra la presenza di rumore di fondo che ostacola la visualizzazione delle varie strutture della materia biologica che si sta studiando, riducendo di conseguenza il rapporto segnale-rumore dell'immagine.

Talvolta anche *l'autofluorescenza* delle strutture biologiche circostanti gli oggetti sui quali si vuole concentrare l'analisi, ovvero la loro naturale emissione di luce senza l'uso di fluorofori, può costituire un ostacolo alla visualizzazione delle zone che si vogliono evidenziare attraverso l'utilizzo di specifici coloranti.

Ulteriori difficoltà possono occorrere durante il protocollo di acquisizione e in questo caso si tratta ad esempio di artefatti di movimento oppure di rumore termico che è dovuto, appunto, all'agitazione termica delle molecole presenti nel sistema e che può contribuire al deterioramento della qualità dell'immagine, determinando la presenza punti luminosi casuali o uno sfondo irregolare.

Altre criticità sono legate all'utilizzo dei fluorofori che talvolta, di seguito a una fase di ossidazione o degradazione in generale, perdono la loro capacità di emettere luce.

Infine, vi sono anche problematiche dal punto di vista della messa a fuoco dell'obiettivo che diminuiscono la risoluzione dell'immagine.

Le immagini di microscopia confocale a fluorescenza hanno un ruolo molto importante in ambito diagnostico e, in particolare, nel campo dell'anatomia patologica dove l'analisi delle strutture cellulari e dei tessuti permette di identificare molteplici patologie e di intervenire tempestivamente su esse. La correzione degli artefatti è essenziale per riuscire ad ottimizzare il percorso della diagnosi svolto dagli specialisti, riducendo la probabilità di errore sul risultato dell'esame clinico effettuato. La correzione delle criticità che si possono incontrare su questo tipo di immagini mediche avviene, sempre più diffusamente, attraverso lo studio e l'implementazione di algoritmi automatici che permettono di ottimizzare le immagini senza dover agire dal punto di vista hardware, ovvero senza la necessità di apportare modifiche sul sistema di acquisizione. Questi metodi, dunque, consentono di ottenere risultati notevoli con tempistiche molto brevi.

Il presente lavoro di tesi ha come obiettivo quello di sviluppare un metodo automatico per la correzione di artefatti presenti in immagini di microscopia confocale a fluorescenza differenti, per riuscire di migliorarne la qualità e la visibiità senza dover agire sul sistema di acquisizione.

Capitolo 2 Materiali e Metodi

Il secondo capitolo contiene un'analisi dei dataset di immagini sui quali si è lavorato per lo sviluppo e per il testing dell'algoritmo di denoising ed anche per il successivo addestramento della Rete Neurale Generativa (GAN).

La descrizione dei dataset si focalizza sulla loro composizione, esponendo il contenuto delle immagini, le loro caratteristiche, il protocollo secondo il quale sono state acquisite ed infine gli artefatti che le caratterizzano.

A seguire viene presentata una descrizione dettagliata dell'algoritmo implementato per la correzione degli artefatti presenti nelle immagini, esaminando passo per passo il modo in cui esso agisce sui dati.

Infine, viene presentata l'architettura della Rete Neurale Generativa (GAN) usata per produrre immagini migliorate con caratteristiche simili a quelle prodotte dall'algoritmo euristico, allenandola su un set di immagini estratte da uno dei dataset e successivamente processate con il metodo euristico sviluppato.

2.1 Dataset di immagini

L'algoritmo ideato per la correzione di artefatti è stato sviluppato su tre diversi dataset di immagini di microscopia confocale a fluorescenza, con lo scopo di ottenere un metodo automatico robusto funzionante su diverse tipologie di immagini. Il metodo è stato poi specializzato per la correzione di una particolare problematica presente in uno dei tre dataset.

2.1.1 Dataset 1: Cardiosfere

Le immagini contenute nel database ad accesso libero CARE algorithm dataset raffigurano oggetti biologici che sono definiti *cardiosfere*.

Le cardiosfere sono complessi tridimensionali di cellule progenitrici cardiache che

sintetizzano un ambiente in micro - scala paragonabile a quello di un sistema di cellule staminali che sono in grado di produrre una rigenerazione del miocardio nel caso in cui questo sia affetto da patologie, come ad esempio la fibrosi cardiaca [6, 7].

Il modello della cardiosfera, dunque, è un modello rappresentativo della nicchia cardiaca che si adatta molto bene a studi in ambito ingegneristico e rigenerativo del miocardio[8, 9] e permette di modellizzare i meccanismi alla base di patologie alle quali si è interessati in questo campo di ricerca. Il dataset è costituito da 90 volumi di immagini di cardiosfere di origine umana, provenienti da pazienti differenti.

In totale sono state analizzate 27 Cardiosfere e 1160 vetrini. Le pile 3D sono state acquisite attraverso l'utilizzo di due differenti laser in modo da riuscire ad evidenziare, e quindi analizzare, sia le membrane cellulari (PHAL), sia i nuclei (DAPI).

Ognuno dei voxel ha dimensione $0.345 \times 0.345 \,\mu\text{m/pixel}^3$, mentre le singole slice costituenti gli stack hanno dimensione 1024×1024 pixel, con risoluzione pari a $0.345 \,\mu\text{m/pixel}$.

I volumi 3D non sono formati tutti dallo stesso numero di slice in quanto, per riuscire ad includere l'intera acquisizione della cardiosfera nella dimensione di profondità del volume, per ogni campione è stato necessario acquisire un numero differente di fette. Il numero medio di slice che compongono i volumi 3D di immagini è pari a 112[3].

Protocollo di acquisizione Le immagini del dataset sono state ottenute attraverso un'analisi al microscopio confocale di cardiosfere umane estratte da pazienti differenti.

Le cardiosfere primarie, sottoposte poi ad analisi microscopica, sono state ricavate da 15 biopsie effettuate sull'appendice atriale destra avvenute su tre pazienti che hanno affrontato un intervento chirurgico, secondo indicazioni cliniche specifiche. La procedura è avvenuta previo consenso informato secondo un protocollo avvallato dall'Institutional Review Board. L'intervento chirurgico ha avuto luogo presso il Policlinico "Umberto I" dell'Università di Roma "La Sapienza" conformemente alla regolamentazione vigente nell'ambito di applicazione.

Le colture di prelievo sono state ricavate a seguito di una frammentazione meccanica e di una digestione enzimatica del tessuto del miocardio avvenuta con tripsina/ED-TA 0.05% per 15 minuti a temperatura ambiente su piastre di Petri rivestite di fibronectina. Trascorse quattro settimane, si è proceduto con un lavaggio EDTA e con una blanda procedura di tripsinizzazione delle cellule espiantate, che erano migrate in modo spontaneo dai frammenti di tessuto raccolti.

Una volta terminata la fase di raccolta del materiale biologico si è proceduto con lo step di analisi di microscopia. I coloranti utilizzati per mettere in evidenza la distribuzione delle cellule, i nuclei e le loro forme sono stati il DAPI e la falloidina marcata con TRITC [10].

L'acquisizione vera e propria è avvenuta tramite tecnica di microscopia confocale convenzionale, registrando le immagini in modo da ottenere le pile tridimensionali con una risoluzione sufficientemente alta.

Osservando la struttura delle cardiosfere, questa appare organizzata con un andamento ripetuto di cellule e dei loro nuclei. All'interno della struttura cellulare si notano anche delle cavità di dimensione non uniforme.



Figura 2.1: Struttura delle cardiosfere analizzata tramite tecnica di microscopia confocale. Le tre sezioni raffigurano la pila osservata lungo la linea mediana (in alto a sinistra), assi X e Y (a destra). Attraverso l'applicazione della colorazione con F-actina marcata con Falloidina in bianco si apprezza il citoscheletro, mentre in blu si osservano in nuclei, evidenziati con il DAPI[3].

I volumi sono stati salvati come file LSM con formato 8 bit.



Figura 2.2: Esempio di slice estratte a campione da diversi volumi contenuti nel "CARE algorithm dataset"[3].

L'algoritmo implementato in questo progetto agisce esclusivamente sul canale blu delle slice estratte dai volumi. Nel canale blu sono raffigurati i nuclei delle cardiosfere. L'applicazione dei metodi alle sole acquisizioni dei nuclei delle cardiosfere è stata determinata dal fatto che esse presentano delle criticità dal punto di vista della scarsa visibilità degli oggetti e della presenza di artefatti che gli altri canali non mostrano in maniera rilevante.

La figura sotto riportata mostra alcuni esempi di slice dalle quali è stato estratto il solo canale blu in cui sono rappresentati i nuclei delle cardiosfere.

Materiali e Metodi



Figura 2.3: Esempio di immagini nel canale blu estratte da slice contenute nelle acquisizioni 3D.

2.1.2 Dataset 2: Nuclei U2OS

Il secondo tra i dataset di immagini utilizzati fa parte di una raccolta di immagini provenienti da uno screening chimico "Nuclei of U2OS cells in a chemical screen" effettuato su cellule U2OS che sono cellule derivate da un osteosarcoma umano, isolate per la prima volta nel 1964.

Le cellule U2OS sono ampiamente utilizzate come modello cellulare per lo studio del metabolismo osseo, la biologia cellulare, la genetica, la biologia del cancro e molti altri ambiti di ricerca. Questo set di dati è composto da un totale di 200 acquisizioni di nuclei catturati con microscopia a fluorescenza utilizzando la colorazione Hoechst. Nello specifico, le immagini sono state salvate come file TIFF con risoluzione 520×696 pixel a 16 bit[4].



Figura 2.4: Esempio di immagini di nuclei U2OS[4].

2.1.3 Dataset 3: Actina e Mitocondri

L'ultimo dataset utilizzato per testare l'algoritmo è costituito da due collezioni di immagini raffiguranti le strutture cellulari di actina e mitocondri estratte dal GigaScience Database contenente le immagini in questione ottenute dal lavoro "Fluorescence microscopy datasets for training deep neural networks" [5]. Le acquisizioni sono avvenute utilizzando un microscopio IX83 dotato di diversi obiettivi, tra cui UplanSApo $100 \times / 1.40$ NA (Apertura Numerica) con immersione ad olio, $60 \times / 1.35$ NA con immersione ad olio e $20 \times / 0.75$ NA con immersione ad aria, forniti da Olympus. La fotocamera usata è la sCMOS Zyla 4.2-plus di Andor, mentre la sorgente luminosa sfruttata per questa applicazione è la SpectraX di Lumencor. I campioni contenuti nei due set derivano da vetrini contenenti cellule endoteliali dell'arteria polmonare bovina colorate con MitoTracker Red CMXRos (etichetta mitocondri) e AlexaFluor 488 falloidina (etichetta actina). Per ognuna delle immagini raccolte si hanno due versioni, una a basso SNR (rapporto segnale-rumore) ed una ad alto SNR. Questo è stato ottenuto settando il sistema con due tempi di acquisizione differenti. Il dataset si compone di 400 immagini suddivise nel seguente modo: per ogni tipo di Apertura Numerica utilizzata per l'acquisizione ($20 \times e$ $60\times$) si hanno a disposizione le immagini di actina e mitocondri nella verione a basso ed alto SNR. Le immagini sono state salvate come file PNG, con risoluzione 2048×2048 pixel su 16 bit[5]. Nel lavoro di tesi le immagini ad alto SNR sono

state utilizzate come riferimento per valutare l'efficacia del processing effettuato con il metodo euristico sviluppato.



Figura 2.5: Esempio di immagini di actina e mitocondri estratte dal GigaScience Database. Le due sotto-figure in alto rappresentano la struttura dell'actina acquisita a $60 \times$ NA mentre quelle in basso raffigurano le strutture dei mitocondri acquisite a $20 \times$ NA (basso SNR a sinistra e alto SNR a destra)[5].

2.2 Analisi degli artefatti

I successivi paragrafi del capitolo contengono la descrizione dettagliata dell'algoritmo implementato per migliorare le immagini dei diversi dataset, agendo principalmente sulla presenza di sbilanciamenti in termini di contrasto e luminosità, riducendo il rumore di fondo e gli effetti di interferenza dovuti a fenomeni di luminosità intrinseca degli oggetti biologici.

Per poter identificare il focus dell'algoritmo è necessario esaminare le problematiche

e gli artefatti presenti nelle immagini dei dataset.

Analizzando le immagini contenute nei dataset si nota che alcuni tra gli oggetti raffigurati appaiono poco luminosi e questo rende difficile la visualizzazione e il riconoscimento delle strutture cellulari che si vogliono esaminare. Gli artefatti o le problematiche che portano ad avere oggetti poco illuminati nelle acquisizioni di microscopia confocale a fluorescenza possono essere diverse. Il segnale di fluorescenza, ad esempio, può essere intrinsecamente debole e ciò può dipendere dalla natura del marcatore fluorescente utilizzato o dalla quantità di etichettante presente nell'oggetto biologico.

In alcuni casi gli oggetti biologici possono assorbire il fluoroforo, riducendo la quantità di segnale che raggiunge gli obiettivi del microscopio. Ciò può verificarsi se l'oggetto ha uno spessore relativamente grande o ha un'alta densità ottica. La fluorescenza può essere influenzata anche dallo scattering della luce nei casi in cui il campione ha una struttura complessa o presenta particelle o aggregati, che fanno si che la luce venga deviata dalla sua traiettoria originale, riducendo così il segnale che raggiunge il rilevatore.

Un ulteriore fattore che può determinare la scarsa luminosità degli oggetti rappresentati nelle immagini è la profondità di campo limitata che caratterizza questo metodo di acquisizione. Solo una sottile sezione dello spessore dell'oggetto può essere messa a fuoco in un dato momento e per questo gli oggetti al di fuori della messa a fuoco possono apparire poco illuminati. Il rumore termico e il rumore del rilevatore, inoltre, possono ridurre la qualità del segnale e influire sulla visibilità degli oggetti nell'immagine.

Infine, un fenomeno che occorre spesso nella microscopia confocale alfluorescenza è quello del *fotobleaching* che consiste nella perdita di fluorescenza nel tempo da parte di alcuni fluorofori utilizzati per marcare le strutture biologiche di interesse durante l'acquisizione, a causa dell'esposizione alla luce. Questo può portare a una diminuzione del segnale e dunque ad una scarsa luminosità degli oggetti rappresentati nell'immagine[11]. L'analisi delle immagini estratte dalle acquisizioni 3D del CARE algorithm dataset ha mostrato che i problemi di visualizzazione sono evidenti in maniera importante all'interno del canale blu in cui sono raffigurati esclusivamente i nuclei delle cardiosfere, mentre nei canali rosso e verde, che ritraggono le membrane cellulari, non vi sono problemi apprezzabili. Osservando le immagini è possibile affermare anche che nelle "fette" iniziali e finali dei volumi l'artefatto ha un impatto maggiore rispetto a quello percepibile nelle slice più centrali dello stack, come si evince dalla Figura 2.6 che ritrae alcune slice nel canale blu estratte in ordine crescente di profondità da uno dei volumi a disposizione.



Figura 2.6: Slice estratte con ordine di profondità crescente da una delle acquisizioni 3D di cardiosfere.

Questa problematica è riscontrabile anche nelle immagini degli ulteriori due dataset. In particolare nelle immagini raffiguranti le strutture cellulari di mitocondri e actina la visibilità è ridotta nelle acquisizioni a basso SNR e, in questo caso, l'obiettivo è ottenere una loro versione uguale o migliore a quella della loro rappresentazione ad alto SNR ottenuta prolungando i tempi di acquisizione. In aggiunta alla ridotta chiarezza degli oggetti, in molti casi, la presenza di sfondi in cui è visibile del rumore, dovuto molto spesso al sistema di acquisizione, fa si che la qualità delle immagini risulti minore rispetto a quella che si ha di fronte ad immagini con un background completamente nero e omogeneo.



Figura 2.7: Esempio di immagini di nuclei U2OS e di mitocondri con ridotta visibilità e rumore di background.

Un'analisi più approfondita effettuata sulle acquisizioni delle cardiosfere ha condotto al riconoscimento di un ulteriore artefatto prodotto da un fenomeno di fluorescenza intrinseca degli oggetti biologi: l'autofluorescenza. L'autofluorescenza è un fenomeno in cui oggetti biologici come cellule e tessuti emettono naturalmente luce fluorescente, senza essere legati a marcatori specifici. Questo evento, nel contesto della microscopia confocale a fluorescenza, può rappresentare un problema in quanto l'autofluorescenza può sovrapporsi o interferire con i segnali fluorescenti che invece si vogliono isolare ed analizzare. I fattori che possono determinare la presenza e l'entità dell'autofluorescenza sono diversi, come ad esempio la preparazione del campione da analizzare, le sue condizioni oppure il tipo di sorgente luminosa utilizzata durante l'acquisizione dell'immagine.

Le immagini dei nuclei delle cardiosfere presentano uno scenario in cui l'autofluorescenza emessa dalle strutture circostanti i nuclei interferisce con la fluorescenza irradiata da essi, compromettendone la corretta visualizzazione.

La figura successivamente riportata mostra degli esempi di immagini in cui è apprezzabile il fenomeno.

Materiali e Metodi



Figura 2.8: Esempi di acquisizioni in cui è visibile il fenomeno dell'autofluorescenza.

2.3 Algoritmo

L'analisi delle problematiche presenti nelle immagini dei dataset ha condotto all' implemenazione di un algoritmo che ha come primo obiettivo quello di incrementare quanto più possibile la luminosità degli oggetti biologici che appaiono oscurati rispetto agli altri. Un aspetto fondamentale del lavoro di tesi è quello della conservazione del contenuto semantico dei dati durante la fase di elaborazione attuata per la correzione degli artefatti. Le manipolazioni effettuate sulle immagini con l'applicazione dell'algoritmo sono state quindi pensate in maniera customizzata alle caratteristiche dei dataset, in modo da non intaccare in alcun modo la morfologia delle strutture cellulari raffigurate al loro interno e in maniera tale da non alterarne il contenuto informativo.

L'algoritmo non agisce esclusivamente sulla luminosità degli oggetti, ma anche sulla rimozione del rumore presente sul background e sull'incremento della nitidezza, in modo da ottenere degli output bilanciati in termini di qualità generale dell'immagine. Il metodo sviluppato è stato successivamente specializzato sulle immagini dei nuclei delle cardiosfere, che oltre a presentare le criticità osservate anche negli altri dataset, sono caratterizzate in maniera singolare dall'artefatto di autofluorescenza.

2.3.1 Pipeline algoritmo

Sulla base delle osservazioni esposte nel paragrafo precedente è possibile descrivere nel dettaglio l'algoritmo generico sviluppato, specificando le operazioni che esso applica sulle diverse tipologie di immagini.

Il diagramma di flusso sotto riportato illustra, macroscopicamente, gli step che definiscono il metodo euristico generale che si adatta all'elaborazione delle immagini di tutti e tre i diversi dataset.



Figura 2.9: Diagramma di flusso che descrive gli step sui quali si basa l'algoritmo generico per il processing delle immagini dei diversi dataset.

2.3.2 Funzione di incremento della luminosità degli oggetti

L'obiettivo di miglioramento delle immagini dal punto di vista dell'innalzamento dei livelli di visibilità degli oggetti sullo sfondo è stato perseguito scrivendo una routine che si occupa di manipolare i valori di intensità luminosa dei pixel ponendo attenzione alle aree chiare e scure delle figure e cercando di rispettare la dinamica del colore. In questo lavoro, con il termine *enhanced* ci si riferisce alla versione migliorata delle immagini esclusivamente basata sull'intensificazione della luminosità di tutti i gli oggetti biologici non emersi in maniera chiara nell'immagine durante l'acquisizione.

Contrast stretching La tecnica di base alla quale si ispira la funzione di incremento della luminosità degli oggetti "oscurati", necessaria per il miglioramento delle immagini di tutti e tre i differenti dataset, è quella del contrast stretching. Il Contrast stretching è un metodo attraverso il quale un'immagine viene migliorata estendendo l'intervallo di valori di intensità acquisiti dai pixel che la compongono, producendo di conseguenza, anche un miglioramento del contrasto. L'estensione della gamma di valori dei pixel avviene innalzando il valore dei singoli pixel tramite una funzione matematica che generalmente è una funzione lineare.

L'incremento, tuttavia, dipende dal contenuto semantico dell'immagine sulla quale si sta lavorando e per tale ragione è possibile agire in maniera differente in base al valore originario dei pixel. Per fare ciò si suddivide il range di valori assumibili dai pixel in regioni in cui si possono utilizzare leggi di incremento differenti sfruttando delle funzioni definite a tratti. Generalmente, la trasformazione differenziata dei pixel avviene creando una ripartizione di tutti i possibili valori di intensità in tre intervalli, così da gestire i toni scuri, intermedi e chiari dell'immagine in maniera distinta.

Di seguito è riportato un esempio di funzione esprime l'andamento di una trasformazione effettuata sui valori dei pixel di un'immagine con il metodo del contrast stretching.

$$y = \begin{cases} \alpha n & 0 \le n \le a \\ \beta(u-a) + u & a \le u \le b \\ \gamma(u-b) + v & b \le v \le L \end{cases}$$

con condizioni di continuità:

$$\alpha u = u$$
$$\beta(b-a) + u = v$$
$$27$$



Figura 2.10: Esempio di funzione di incremento dei valori di intensità luminosa dei pixel con il metodo del contrast stretching.

Il miglioramento della visibilità degli oggetti nelle immagini dei dataset su cui si è svolto il lavoro è stato ottenuto rielaborando il classico metodo del contrast stretching, come verrà illustrato in seguito. Questa scelta è stata dettata dalla necessità di ottenere un risultato soddisfacente che si adattasse il più possibile alla morfologia degli oggetti biologici raffigurati nelle immagini.

La funzione di incremento dei livelli di luminosità degli oggetti acquisisce in ingresso l'immagine originale e alcuni parametri necessari a determinare la crescita dei valori dei pixel delle immagini.

Nel caso del dataset contenente le acquisizioni 3D delle cardiosfere, dovendo agire sul solo canale blu delle slice costituenti i volumi è stato necessario effettuare la loro estrazione. In molti degli stack del CARE algorithm dataset vi sono delle slice che sono prive di informazione, ovvero che non raffigurano alcun oggetto e che contengono esclusivamente lo sfondo. In vista dell'applicazione delle routine di correzione e del futuro addestramento di una rete neurale, queste immagini sono state escluse, col fine di agire solo su quelle in cui sono visibili le cardiosfere e di conseguenza i loro nuclei. Per effettuare la rimozione si è agito imponendo una soglia sul valor medio e una sulla deviazione standard dei pixel nelle immagini, calcolando i due valori statistici sul canale che presenta luminosità media maggiore. Successivamente è avvenuta l'estrazione del canale blu contenente le rappresentazioni dei nuclei di cardiosfera. Le immagini degli altri due dataset invece, essendo già acquisizioni bidimensionali, sono state fornite in input alla funzione senza effettuare alcuna manipolazione oltre alla loro trasformazione in formato 8 bit attraverso il min-max scaling, essendo originariamente salvate con formato a 16 bit.

Come introdotto all'inizio del paragrafo, l'aumento dei valori dei pixel avviene attraverso l'uso di differenti funzioni lineari crescenti, su ispirazione del metodo del contrast stretching. I primi due, tra i parametri richiesti dalla chiamata della funzione, servono per settare gli intervalli in cui viene ripartito il range di valori acquisibili dai pixel e quindi degli interi compresi tra 0 e 255 (si è lavorato su immagini con dinamica a 8 bit).

Gli ultimi parametri che devono essere forniti in input alla funzione sono quelli necessari alla definizione delle funzioni lineari di incremento.

La routine procede secondo i seguenti due passaggi fondamentali:

1. Suddivisione dei valori dei pixel in intervalli: i valori tra 0 e 255 vengono ripartiti in tre macro-intervalli, per agire in modo differenziato sui toni scuri, intermedi e chiari delle immagini e in modo che i pixel costituenti gli oggetti che soffrono dell'artefatto, avendo valori bassi, ricadano all'interno del primo. Le suddivisioni sono state ottenute in maniera dinamica, ovvero con un settaggio delle soglie customizzato sulle singole immagini. Questo è stato fatto calcolando le soglie in funzione della luminosità media dell'immagine originale di input.

Oltre ad illuminare gli oggetti rappresentati da pixel con valori ridotti, è importante preservare il contrasto al loro interno, così da non perdere la dinamicità del colore che li contraddistingue. Per riuscire in questo obiettivo, il primo macro-intervallo è suddiviso ulteriormente in tre range, in modo da distinguere anche internamente ai nuclei le aree più chiare e quelle più scure.

2. Definizione delle funzioni lineari di incremento dei valori dei pixel: le funzioni di trasformazione utilizzate all'interno degli intervalli sono definite come segue:

$$y = \begin{cases} x & 0 \le x \le a\\ \beta(k-a) + k & a \le k \le b\\ \gamma(k-b) + h & b \le h \le c\\ \delta(u-c) + u & c \le u \le d\\ \epsilon(u-d) + v & d \le v \le L \end{cases}$$

con condizioni di continuità:

$$\beta(b-a) + k = h$$

$$\gamma(c-b) + h = u$$

$$\delta(d-c) + u = v$$

Si può osservare che la prima retta è la retta identità e quindi i valori dei pixel appartenenti al primo sotto-intervallo [0, a] non subiscono alcuna trasformazione, in modo da non aggiungere informazione dove questa non è originariamente contenuta, oltre che per rispettare le caratteristiche di contrasto degli oggetti raffigurati. I termini β , γ , δ ed ϵ rappresentano i coefficienti angolari delle rette passanti per due punti che dettano l'aumento dei valori dei pixel inclusi negli intervalli. L'entità dei coefficienti angolari delle rette dipende sia dalla larghezza del range in cui agiscono, che dal valore minimo e dal valore massimo che il pixel può raggiungere in quel range attraverso la funzione di trasformazione e derivano, dunque, dai parametri richiesti dalla chiamata della funzione.

Il grafico di esempio sotto riportato mostra un setting degli intervalli verosimile a quello effettivamente utilizzato per il processing delle immagini dei nuclei delle cardiosfere. La larghezza crescente delle ripartizioni dei valori, procedendo dalla prima all'ultima, ha lo scopo di ottenere una maggiore espansione dei valori per i pixel costituenti gli oggetti scuri ed un minore innalzamento per gli altri. Fissati i nuovi valori che raggiungeranno gli estremi del range dopo la trasformazione, il coefficiente angolare delle rette diminuisce al crescere della larghezza dell'intervallo.



Figura 2.11: Esempio di funzione di incremento di intensità dei pixel utilizzata nell'algoritmo.

2.3.3 Denoising

L'applicazione della funzione dedicata all'incremento della luminosità degli oggetti alle immagini è seguita da un passaggio di filtraggio.

In molti casi le immagini dei database presentano una situazione in cui i pixel costituenti le zone in cui è presente del rumore di fondo hanno valori di intensità paragonabili a quelli dei pixel degli oggetti che devono essere illuminati. L'applicazione della routine di incremento delle luminosità, in alcuni casi, produce, anche l'incremento della luminosità del rumore e per rimuovere quest'ultimo è necessario l'utilizzo di un filtro. I filtri tradizionali comunemente utilizzati nel campo dell'image processing non si adattano alle immagini in analisi in quanto agiscono principalmente sui valori di intensità e, essendo in questo caso rumore e informazione caratterizzati da valori di intensità molto simili, non permettono di discriminare i due per rimuovere esclusivamente il rumore.

Il tipo di filtraggio che maggiormente si adatta a questo tipo di immagini è quello

che si ottiene tramite la trasformata Wavelet, che agisce sulle frequenze dell'immagine.

La trasformata wavelet è un metodo molto efficace per l'elaborazione dei segnali, ma è particolarmente utile per applicazioni quali la compressione e il denoising di immagini e di segnali non stazionari in generale[12]. La trasformata Wavelet ha il vantaggio di non acquisire informazioni sulla frequenza globale del segnale, come avviene ad esempio nella trasformata di Fourier, ma decompone il segnale in un set di forme d'onda dette wavelets, per estrarne caratteristiche variabili nel tempo o nello spazio, che non si riuscirebbero a rilevare considerando l'intera acquisizione. Una wavelet è una forma d'onda di lunghezza limitata e definita da due proprietà di base: posizione e scala. La posizione di una wavelet indica dove essa è posizionata nel tempo o nello spazio, mentre la scala, detta anche dilatazione, definisce quanto la forma d'onda è "schiacciata" o "allungata". Esistono differenti tipi di wavelet con caratteristiche diverse ed è possibile scegliere quella che meglio si adatta al segnale di interesse. Le wavelets sono funzioni che possono essere utilizzate per l'analisi dei segnali sia nel tempo nello spazio, a seconda che il loro intervallo limitato venga concentrato in un dominio o nell'altro.

In ambito di image-processing l'analisi wavelet agisce nello spazio scomponendo l'informazione in una componente ad alta frequenza e in una componente a bassa frequenza. La frequenza delle immagini è la grandezza che valuta la velocità di variazione dei valori dei pixel nel dominio dello spazio. L'alta frequenza è associata alle rapide variazioni di intensità dei pixel che caratterizzano le discontinuità, il rumore e i bordi degli oggetti all'interno delle immagini, mentre le basse frequenze sono proprie delle aree di colore uniformi e degli sfondi, contraddistinti da modifiche graduali del colore.

Il denoising delle immagini tramite trasformata Wavelet avviene attraverso i passaggi sotto elencati:

1. Decomposizione: si seleziona il tipo di wavelet madre che più si adatta all'immagine (Haar, Morlet, Daubechies) e si scompone l'immagine originale in bande di frequenza con risoluzione progressivamente decrescente. La decomposizione è un passaggio iterativo in cui l'immagine è sottoposta ad un filtraggio passa-basso e ad un filtraggio passa-alto, separatamente in direzione verticale e orizzontale, per estrarre le componenti a bassa ed alta frequenza. I due filtraggi sono seguiti entrambi da uno step di downsampling utile per ottenere una rappresentazione più compatta delle informazioni risultanti dall'applicazione dei filtri.

La decomposizione può essere eseguita ad un numero di livelli differente, che determina un diverso livello di dettaglio delle bande di frequenza ricavate.

2. Thresholding: dopo aver estratto gli elementi a bassa ed alta frequenza a

diverse scale si applica uno step di thresholding per ridurre o eliminare il rumore e gli artefatti all'interno dell'immagine. Le techniche di thresholding disponibili sono diverse (Soft, Hard, Garrote, ecc.) ed è possibile scegliere quella che permette di ottenere il risultato che si desidera.

3. Ricostruzione: l'ultimo step del denoising tramite trasformata wavelet consiste nel ricomporre l'immagine con all'interno l'informazione in uscita dal thresholding. Questo ultimo punto del denoising si basa su un'operazione di upsampling necessaria per ottenere come output un'immagine con risoluzione pari a quella originale e, infine, sull'applicazione di filtri di ricostruzione che effettuano una combinazione delle nuove componenti a bassa ed alta frequenza che vengono infine sommate tra loro.



Figura 2.12: Schema di funzionamento della trasformata wavelet sulle immagini[13].

La trasformata wavelet è stata applicata, all'interno dell'algoritmo, utilizzando la wavelet di tipo Haar, scomponendo le immagini ad un numero di livelli pari a tre e sfruttando la tecnica di thresholding 'Garrote'. La tecnica Garrote permette di ridurre o annullare le componenti ad alta frequenza oltre la soglia, che costituiscono il rumore, mantenendo quanto più possibile invariate le basse frequenze grazie ad un controllo sulla stima dell'errore di regressione.

In particolare, la tecnica di thresholding Garrote agisce riducendo o eliminando il rumore secondo il seguente principio[14]:

$$G(x) = \begin{cases} x - \frac{\lambda^2}{x} & x > \lambda \\ 0 & x \le \lambda \end{cases}$$

dove x rappresenta il coefficiente derivante dalla trasformata wavelet dell'immagine e λ indica il valore di soglia definito secondo la regola di Visu Shrink, in cui la

soglia è l'Universal Threshold (UT) definita come:

$$\lambda = \sigma * \sqrt{2 * \log(N * M)}$$

con σ^2 varianza del rumore e N * M dimensione dell'immagine. Inoltre, il livello di rumore σ può essere stimato utilizzando la seguente formula:

$$\hat{\sigma} = \frac{\text{Median}\{|y_{i,j}|\}}{0.6745} \forall y_{i,j} \in \text{sottobanda HH}$$

in cui $y_{i,j}$ rappresentano i valori presi dalla sotto-banda ad alta frequenza HH estratta dall'immagine. Le figure sotto riportate presentano un confronto tra il risultato ottenibile senza eseguire il filtraggio e il risultato che invece si ha filtrando l'immagine in uscita dalla funzione di incremento della luminosità degli oggetti sfruttando la trasformata wavelet. Si può notare che, grazie all'eliminazione del rumore sullo sfondo ottenuto in seguito al riconoscimento dei diversi livelli di frequenza, si ottiene un'esaltazione di oggetti che precedentemente al filtraggio si uniformavano maggiormente con il background.



Figura 2.13: Esempio di effetto dell'incremento della luminosità degli oggetti con e senza successivo filtraggio.

2.3.4 Miglioramento della nitidezza

La decomposizione wavelet e la successiva ricostruzione dell'immagine tendono a causare una riduzione della nitidezza dell'immagine. Per ripristinare caratteristiche di nitidezza simili a quelle delle immagini pre-filtraggio si attua un'operazione di Sharpening sfruttando il metodo Unsharp Masking. L'Unsharp masking è una tecnica di elaborazione di immagini di tipo lineare che permette di incrementare la nitidezza attraverso le seguenti operazioni:

- 1. Creazione della versione sfocata dell'immagine tramite l'applicazione di un filtro di sfocatura (filtro Gaussiano).
- 2. Generazione dell'immagine maschera come risultato della differenza tra l'immagine originale e l'immagine sfocata, contenete i dettagli nitidi dell'immagine.
- 3. Applicazione della maschera: i dettagli vengono enfatizzati sommando all'immagine originale la maschera contenente i dettagli moltiplicata per un fattore che consente di modulare l'effetto di sharpening.

Il principio di funzionamento dell'unsharp masking è riassunto nella relazione sotto scritta, dove α è un fattore numerico compreso tra 0 e 1 che definisce l'entità del miglioramento della nitidezza.

 $\text{immagine}_{\text{sharped}} = \text{immagine}_{\text{originale}} + \alpha \cdot (\text{immagine}_{\text{originale}} - \text{immagine}_{\text{sfocata}})$

Una volta ottenuto un potenziamento della nitidezza grazie alla tecnica dell'unsharp masking è stato effettuato un filtraggio molto leggero con un filtro Gaussiano per rimuovere l'effetto "artificioso" dovuto allo sharpening.

Le figure che seguono riportano alcuni esempi che illustrano l'applicazione dell'algoritmo alle immagini dei diversi dataset.
Materiali e Metodi



Figura 2.14: Esempio di trasformazione effettuata attraverso l'applicazione dell'algoritmo su un'immagine di nuclei di cardiosfera.



Figura 2.15: Esempio di trasformazione effettuata attraverso l'applicazione dell'algoritmo su un'immagine di nuclei U2OS.



Figura 2.16: Esempio di trasformazione effettuata attraverso l'applicazione dell'algoritmo su immagini di actina e mitocondri.

2.4 Pipeline algoritmo specializzato nella correzione dell'autofluorescenza

Le immagini del CARE algorithm dataset, come accennato in precedenza, si differenziano da quelle degli altri due gruppi di immagini per la presenza di un artefatto derivante da un fenomeno conosciuto come *autofluorescenza* dovuto alle proprietà luminose delle strutture cellulari raffigurate.

Come affermato precedentemente durante l'analisi degli artefatti, in queste acquisizioni le problematiche si incontrano all'interno del canale blu delle slice che costituiscono i volumi e dunque solo dove sono osservabili le strutture dei nuclei delle cardiosfere. Per questo motivo è necessario eseguire un'estrazione del canale dalle slice contenute negli stack tridimensionali. A seguire è riportato il diagramma di flusso riformulato per l'algoritmo specializzato nella correzione dell'autofluorescenza sulle acquisizioni dei nuclei di cardiosfera.



Figura 2.17: Diagramma di flusso relativo all'algoritmo specializzato nella correzione dell'autofluorescenza.

Dopo la preparazione delle immagini dei nuclei di cardiosfera da fornire in input al processo sono state implementate la funzione di aumento della luminosità dei nuclei oscurati, il denoising per la rimozione del rumore e il miglioramento della nitidezza come avviene nell'implementazione dell'algoritmo generale precedentemente descitto e, in aggiunta, il metodo per il controllo dell'autofluorescenza.

2.4.1 Calibrazione dell'autofluorescenza

Le immagini dei nuclei delle cardiosfere presentano uno scenario in cui l'autofluorescenza emessa dalle strutture circostanti interferisce con la fluorescenza irradiata dai nuclei, compromettendone la corretta visualizzazione.

La riduzione del fenomeno è stata ottenuta attuando una calibrazione delle luminosità delle sole aree dell'immagine in cui è presente l'autofluorescenza, senza modificare le intensità dei nuclei precedentemente evidenziati. La strategia utilizzata per il controllo dell'effetto dell'autofluorescenza si basa sulle seguenti manipolazioni:

1. Generazione di maschere di segmentazione dei nuclei: attraverso l'utilizzo di una rete neurale di tipo ConvUNeXt, ovvero di una Rete Neurale Convoluzionale (CNN) che permette di ottenere, in maniera efficiente, segmentazioni di immagini mediche con un minor numero di parametri rispetto a quelli utilizzati da una UNet standard[15]. Le maschere sono state ottenute fornendo in input al modello pre-allenato le immagini prodotte in uscita dall'algoritmo di generico, ovvero le immagini processate con la funzione di incremento, con la trasformata wavelet e migliorate dal punto di vista della nitidezza.



Figura 2.18: Esempi di maschere di segmentazione dei nuclei ottenute con il modello ConvUNeXt sovrapposte alle immagini processate con l'algoritmo generale dalle quali sono state ottenute.

2. Isolamento dell'autofluorescenza: attraverso la versione complemetare della maschera di segmentazione è stata estratta dall'immagine solo l'informazione relativa al background e all'autofluorescenza, tenendo da parte i nuclei. Questo è stato fatto per riuscire a modulare l'intensità luminosa dell'autofluorescenza senza agire sui nuclei caratterizzati da una luminosità, e dunque, da una visibilità già ottimizzata con i passaggi precedenti dell'algoritmo.



Figura 2.19: Esempio di estrazione dell'autofluorescenza per l'applicazione della calibrazione delle luminosità.

3. Definizione delle soglie: le due differenti soglie sono state calcolate come combinazione lineare delle luminosità medie rispettivamente dei nuclei e dell'autofluorescenza, valutate escludendo lo sfondo. La relazione sotto riportata riassume il modo in cui è avvenuto il calcolo delle soglie.

 $soglia = peso_{nuclei} \cdot luminosità media nuclei$ $+ peso autofluorescenza \cdot luminosità media autofluorescenza$

La prima soglia costituisce il limite di intensità luminosa al di sotto della quale i valori dei pixel all'interno dell'area dell'autofluorescenza vengono ridotti secondo la legge di una retta che permette di mappare i valori ad una loro versione ridotta di una percentuale che è possibile settare in modo arbitrario in base all'effetto che si desidera ottenere. La seconda soglia, invece, definisce il l'intensità oltre la quale i pixel vengono aumentati di una percentuale a scelta, per riuscire a evidenziare alcuni dei nuclei inclusi nell'autofluorescenza non riconosciuti correttamente dalla segmentazione.

La combinazione lineare che descrive il calcolo della prima soglia è caratterizzata da un peso maggiore associato alla luminosità media dell'autofluorescenza, mentre nel caso della seconda soglia il peso maggiore è assegnato alla luminosità media dei nuclei dell'immagine.



Figura 2.20: Esempio di leggi di modifica dei valori dei pixel dell'area dell'auto-fluorescenza.

Il diagramma di flusso riportato alla pagina seguente riassume gli step sui quali si basa il metodo per il controllo dell'artefatto di autofluorescenza appena descritto.



Figura 2.21: Diagramma di flusso che illustra gli step seguiti per il controllo dell'autofluorescenza.



Figura 2.22: Esempio di effetto risultante dalla calibrazione delle luminosità dell'autofluorescenza.

L'autofluorescenza, come è possibile vedere nella figura in alto, non è stata eliminata completamente, ma la sua intensità è stata ridotta del 70%, mentre i valori oltre la seconda soglia hanno subito un aumento del 10%. La riduzione parziale è stata effettuata secondo una decisione conservativa dettata dal fatto che, in alcuni casi, in sede di diagnosi, potrebbe essere favorevole riuscire a visualizzare in sottofondo la struttura all'interno della quale i nuclei sono contenuti. Tuttavia, è possibile scegliere l'entità della riduzione a piacere in base all'effetto desiderato, come è mostrato nella figura seguente, dove è visibile il risultato ottenibile con tre differenti percentuali di riduzione dell'intensità dell'autofluorescenza.



Figura 2.23: Esempio di correzioni di intensità dell'autofluorescenza effettuate con tre diverse percentuali di riduzione.

2.5 Generazione di immagini prive di artefatti tramite GAN

Lo sviluppo dell'algoritmo per la manipolazione delle immagini di microscopia confocale a fluorescenza è stato affiancato dall'addestramento di una Rete Generativa Avversaria (GAN) con lo scopo di ottenere un modello addestrato alla generazione di immagini semanticamente simili a quelle prodotte dall'applicazione dell'algoritmo. La scelta di utilizzare un modello non euristico come quello della rete è stata dettata dal fatto che esso è in grado di stabilire un mapping più generico tra il dominio delle immagini originali e il dominio delle immagini migliorate e questo può condurre, in alcuni casi, al superamento dei limiti dell'algoritmo euristico descritto nei paragrafi precedenti. L'architettura della GAN scelta per questo compito è la pix2pix. In particolare, il task richiesto alla rete in questo caso è di tipo *image-to-image translation* in cui la GAN viene addestrata a migliorare la qualità complessiva dell'immagine, illuminando gli oggetti biologici che risultano poco chiari e agendo sulla presenza di artefatti.

2.5.1 GAN: Reti Neurali Generative Avversarie

La tecnologia basata su Intelligenza Artificiale Generativa maggiormente diffusa è quella delle Reti Neurali Generative Avversarie (GAN). Le GAN sono reti neurali capaci di generare dati quali immagini, audio e video, aventi caratteristiche molto simili a quelle dei dati reali. La struttura di una rete GAN comprende due componenti principali: il generatore e il discriminatore. Il generatore e il discriminatore lavorano insieme in una competizione durante la fase di addestramento della GAN. Il generatore è responsabile della generazione di dati artificiali (o sintetici) simili ai dati reali presenti nel set di addestramento. Normalmente il generatore è costituito da una rete neurale artificiale convoluzionale (CNN) o da una rete neurale ricorrente (RNN). Il generatore acquisisce come input del rumore casuale definito spazio latente e lo trasforma in un'immagine o un dato sintetico quanto più possibile simile a quello di reale di riferimento.

Il discriminatore ha il compito di distinguere i dati reali dai dati prodotti dal generatore. Anche il discriminatore è spesso implementato utilizzando una rete neurale, come una CNN. Il discriminatore richiede in ingresso un dato o un'immagine e restituisce una previsione che indica se l'input è reale o generato.





Figura 2.24: Schema esemplificativo della struttura di una GAN per la generazione di immagini[16].

L'addestramento di una GAN si basa sul *min-max game* che è un concetto che riassume il principio della competizione tra il generatore e il discriminatore durante il l'allenamento.

Il termine "min-max" si riferisce al fatto che, durante l'addestramento, il generatore cerca di minimizzare una funzione di loss, mentre il discriminatore cerca di massimizzarla. L'obiettivo del generatore, in questa fase, è quelli di ingannare il discriminatore, generando dati che il discriminatore non può distinguere da quelli reali. D'altra parte, il discriminatore cerca incrementare la sua capacità di discernere tra immagini o dati reali e sintetici. L'addestramento delle due reti avviene simultaneamente attraverso un processo iterativo competitivo il cui fine è raggiungere un punto di equilibrio in cui il generatore sia in grado di produrre dati che sembrano reali e il discriminatore sia difficile da "ingannare". In questo stato di equilibrio, il generatore ha appreso la distribuzione dei dati di addestramento e può generare dati nuovi e coerenti con quelli del set di dati originale.

Schematicamente, i passaggi che definiscono la fase di addestramento di una GAN sono:

- 1. Inizializzazione: i pesi del generatore e del discriminatore vengono inizializzati casualmente o utilizzando una strategia di pre-addestramento. Viene poi definita una funzione di loss necessaria a misurare la discrepanza tra le previsioni del discriminatore e le etichette corrette.
- 2. Fase di addestramento del discriminatore: viene selezionato un batch (gruppo) di immagini reali dal set di dati di addestramento. A seguire viene poi generato un batch di immagini artificiali utilizzando il generatore a cui viene fornito in input un rumore casuale dallo spazio latente.

Il discriminatore viene addestrato sulla combinazione di immagini reali e

generate per distinguere correttamente tra le due categorie. La funzione di loss viene poi calcolata per misurare l'abilità del discriminatore di distinguere le immagini reali da quelle sintetiche e in funzione del valore di loss ottenuto i pesi del discriminatore vengono aggiornati utilizzando la retropropagazione dell'errore per migliorare la capacità discriminativa.

- 3. Fase di addestramento del generatore: un nuovo batch di immagini artificiali viene generato utilizzando il generatore che acquisisce come input il rumore casuale dallo spazio latente. Il generatore cerca di ingannare il discriminatore, facendo sì che le immagini generate siano classificate come reali aggiornando i propri pesi sulla base della minimizzazione della loss del discriminatore. La funzione di loss del generatore viene poi calcolata sulla base della capacità del discriminatore di classificare erroneamente le immagini generate. Anche i pesi del generatore vengono aggiornati utilizzando la retropropagazione dell'errore. In questo step di addestramento del generatore, il discriminatore viene "bloccato" e i suoi pesi non vengono modificati.
- 4. Iterazione: gli step di addestramento delle due reti vengono ripetuti per un certo numero di iterazioni o epoche fino a raggiungere un equilibrio, definito *equilibrio di Nash*, in cui il generatore è in grado di generare immagini sempre più realistiche e il discriminatore è sempre più difficile da ingannare.

In termini matematici, il modello generativo acquisisce in input un vettore di rumore z che rappresenta un punto nello spazio rumore, derivante dalla distribuzione casuale p_z . Il generatore trasforma z attraverso la funzione di generazione G, per produrre un campione x che sia associato ad una distribuzione stimata p_g simile a quella dei dati reali e che rappresenta un'approssimazione della distribuzione reale p_d .

La funzione di loss totale, definita anche *Total Loss Function*, che guida la progressione dell'addestramento è la seguente:

$$\min_{G} \max_{D} V(G, D) = \mathbb{E}_{x \sim p_d}(x) \left[\log D(x) \right] + \mathbb{E}_{z \sim p_z}(z) \left[\log(1 - D(G(z))) \right]$$

Dove:

- G(z) sono gli esempi prodotti dal generatore a partire dal vettore di elementi casuali z chiamato spazio latente.
- D è l'output del discriminatore e può assumere valore pari a 0 o a 1, dove 0 indica che il discriminatore riconosce come falso il dato prodotto dal generatore, mentre 1 indica che il dato sintetico è stato riconosciuto come reale.
 Durante l'addestramento il discriminatore riceve in input sia esempi reali

provenienti da training set, sia esempi sintetici ottenuti dal generatore. Nel primo caso l'output del discriminatore sarà D(x), mentre nel secondo caso sarà D(G(z)).

La Total Loss Function può essere divisa in modo da definire separatamente le funzioni di perdita per il discriminatore e per il generatore.

• Loss del discriminatore:

$$\log(D(x)) + \log(1 - D(G(z)))$$

• Loss del generatore:

$$\log(1 - D(G(z)))$$

Il discriminatore cerca di massimizzare la sua funzione di perdita per migliorare la sua capacità di distinguere tra dato reale e dato sintetico, mentre il generatore cerca di minimizzare la sua funzione di loss per generare campioni sempre più simili a quelli reali.

2.5.2 Architettura Pix2Pix

L'architettura scelta per ottenere un modello addestrato al miglioramento di immagini di microscopia confocale a fluorescenza e in maniera specifica anche per la correzione della problematica dell'autofluorescenza è quella della *Pix2Pix*. Il modello della Pix2Pix è strutturato nel seguente modo:

• Generatore: l'architettura del generatore di una GAN Pix2Pix è basata su una rete neurale convoluzionale con una struttura encoder-decoder caratterizzata dalla presenza di connessioni di salto, comunemente chiamate skip connections. Nello specifico, l'encoder riceve in ingresso l'immagine originale che attraversa una serie di layer che effettuano un downsampling tramite operazioni di convoluzione. L'operazione di downsampling serve a estrarre le informazioni rilevanti dell'immagine di input, denominate *feature map*, che vengono poi utilizzate nel processo di decoding per generare l'immagine di output desiderata. Il decoder riceve le feature map dall'encoder ed effettua un upsampling facendo scorrere le mappe semantiche attraverso degli strati deconvoluzionali per ricostruire poi gradualmente l'immagine di immagine di output con risoluzione pari a quella dell'immagine di input. L'inserimento delle skip connections tra i layer dell'encoder e del decoder con stessa risoluzione permettono di stabilire collegamenti diretti tra i gli strati dell'encoder e del decoder corrispondenti, consentendo un flusso di informazioni più diretto durante la generazione dell'output. Esse hanno lo scopo di preservare informazioni di basso livello

importanti durante il processo di generazione dell'output, superando così le limitazioni dell'upsampling tradizionale e consentendo al decoder di accedere a informazioni dettagliate che sarebbero altrimenti perse durante l'operazione di downsampling.



Figura 2.25: Confronto tra la struttura generica encoder-decoder e la struttura del generatore di una rete Pix2Pix caratterizzata dalla presenza delle skip connections.[17].

• Discriminatore: il discriminatore è costituito da un classificatore PatchGAN convoluzionale, che ha il vantaggio di preservare i dettagli ad alta frequenza estratti al livello di patch dell'immagine. La sua funzione principale è quella di distinguere tra le immagini generate e le immagini reali, fornendo un feedback al generatore per guidarne l'apprendimento. Il discriminatore, per ogni patch, produce una valutazione che rappresenta la probabilità che quella specifica patch sia reale o generata. Successivamente viene applicato un filtro convoluzionale a ogni patch per produrre una singola uscita, che viene interpretata come la probabilità di autenticità della patch. Le valutazioni delle diverse patch vengono poi combinate tra loro per ottenere un risultato complessivo per l'intera immagine. Ciò può essere fatto mediante una media o un'operazione di pooling.

L'obiettivo di una Pix2Pix è generare un'immagine di output che sia strettamente correlata all'immagine di input che costituisce la "condizione" del problema e per questo motivo si parla di GAN condizionale.

L'obiettivo di una GAN condizionale può essere espresso con la seguente espressione matematica che esprime la funzione di perdita:

$$\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))]$$
48

dove il generatore (G) cerca di minimizzare questo obiettivo mentre il discriminatore (D) avversario cerca di massimizzarlo. La funzione di perdita della GAN condizionale viene inoltre combinata con la funzione di regolrizzazione L1, in modo che il generatore sia in grado di produrre immagini sintetiche (y) il più possibile simili al target (\hat{y}) , oltre ad "ingannare" il discriminatore.

$$\mathcal{L}_{L1}(y,\hat{y}) = \mathbb{E}_{x,y,z}[|y - G(x,z)|]$$

La funzione obiettivo finale e cioè la funzione obiettivo ottenuta dalla combinazione delle due loss risulta essere:

$$\mathcal{L}_{G} = \arg\min_{G} \max_{D} \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

dove λ è un iperparametro necessario per bilanciare le perdite e che determina l'importanza relativa tra le due componenti nella funzione obiettivo finale.

2.5.3 Settaggio iperparametri del modello Pix2Pix utilizzato per la generazione di immagini corrette

La rete Pix2Pix addestrata per la generazione di immagini di microscopia confocale a fluorescenza migliorate è stata addestrata su un training set costituito da 6804 immagini di nuclei di cardiosfera estratte dalle acquisizioni 3D di provenienza e dalle corrispettive versioni migliorate attraverso l'algoritmo specializzato su esse che hanno costituito il "target" della rete. Il testing della rete è avvenuto poi su un test set composto da 539 immagini. La scelta di addestrare la rete sul processing delle immagini del CARE algorithm dataset è avvenuto principalmente per verificare l'efficacia del metodo non euristico su un task caratterizzato dalla difficoltà supplementare della correzione del fenomeno dell'autofluorescenza.

Il setting della Pix2Pix è stato definito nel seguente modo:

- architettura generatore: U-Net
- architettura del discriminatore: PatchGAN
- dimensione immagini di input: 1024x1024
- numero di canali delle immagini di input: 1
- numero di epoche: 100
- batch size: 8
- learning rate: 0.0002
- funzione di loss: L_1

- ottimizzatore: Adam
- numero di filtri utilizzati nell'ultimo livello di convoluzione del generatore: 64
- numero di filtri utilizzati nel primo livello di convoluzione del discriminatore: 64





Figura 2.26: Esempio di immagini di nuclei di cardiosfera migliorate generate dal modello addestrato.

2.6 Metriche di validazione

I risultati proposti dall'applicazione dell'algoritmo alle immagini sono stati valutati, in una prima fase, in modo qualitativo, basandosi su un'attenta osservazione delle immagini risultanti dal processing e verificando che gli artefatti fossero effettivamente ridotti o eliminati, preservando al massimo l'informazione di interesse. Una volta raggiunti i risultati desiderati, si è proceduto con una validazione più oggettiva delle prestazioni dell'algoritmo specializzato anche nella correzione dell'artefatto di autofluorescenza, in modo da testare le prestazioni su un compito maggiormemte complesso. La validazione è stata effettuata utilizzando tre diverse metriche per valutare tre differenti parametri delle immagini proposte in output dall'algoritmo per comprendere le variazioni apportate dalla routine alle caratteristiche dell'immagine. Le stesse misure di valutazione sono state utilizzate per verificare la qualità delle immagine sintetiche prodotte dal modello generativo addestrato. Le metriche sono state valutate su un set di immagini uguale a quello che ha costituito il test set utilizzato per il testing della GAN. Le metriche sono state dunque utilizzate anche per verificare la qualità delle immagini prodotte con il metodo basato su intelligenza artificiale generativa, per proporre poi un confronto tra le due diverse strategie.

2.6.1 CII

Le manipolazioni effettuate sulle immagini tendono a generare un aumento del loro contrasto. Questo avviene perchè, innalzando i livelli di intensità luminosa degli oggetti e riducendo contemporaneamente la luminosità dell'autofluorescenza e il rumore presente nel background, si intensificano localmente le differenze di intensità luminosa tra le aree chiare e aree scure. Il Contrast Improvement Index (CII) è una metrica utilizzata per valutare il miglioramento del contrasto ottenuto dall'applicazione di un metodo per l'elaborazione dell'immagine e, in questo caso, permette di capire se vi è stato un miglioramento da questo punto di vista.

Il CII viene calcolato confrontando i valori massimi e minimi delle intensità dell'immagine originale con i valori massimi e minimi dell'immagine ottenuta dopo l'applicazione del metodo di elaborazione secondo la relazione[18]:

$$\text{CII} = \frac{C_{\text{enhanced}}}{C_{\text{original}}}$$

dove C_{enhanced} e C_{original} sono i valori medi del contrasto locale valutati sfruttando dei kernel di dimensione 9x9 in questo modo: si scorre tutta l'immagine con il kernel scelto e, di volta in volta, si identificano al suo interno il valore massimo e il valore minimo di intensità. Si cacalcolano poi la somma dei tutti i valori massimi e di tutti i valori minimi individuati nell'intera immagine. Si considera poi il rapporto: $\frac{somma_{massimi} - somma_{minimi}}{somma_{massimi} + somma_{minimi}}$

Il risultato del rapporto viene poi diviso per il numero di finestre (kernel) in cui è stata suddivisa l'immagine. Si calcola poi il rapporto tra i due valori medi di contrasto e si ottiene il CII. Misure di CII maggiori di 1 indicano che con l'applicazione dell'algoritmo vi è stato un incremento del contrasto, mentre valori inferiori a 1 sono relativi ad una sua diminuzione.

2.6.2 MS-SSIM

L'elaborazione delle immagini è avvenuta ponendo particolare attenzione alla conservazione dell'informazione rilevante contenuta al loro interno. Nello specifico, applicando l'algoritmo ad immagini raffiguranti nuclei e strutture cellulari in genere, le manipolazioni sono state studiate in maniera da ridurre al massimo il rischio della cancellazione di alcune di esse.

La Multi-Scale Structural Similarity Index (MS-SSIM) è una metrica per la valutazione della similitudine tra due immagini, genericamente una la versione alterata dell'altra. Il modo in cui vengono percepiti i dettagli di un'immagine dipende da differenti fattori, come la densità di campionamento del segnale dell'immagine, la capacità percettiva del sistema visivo dell'osservatore ed anche la distanza dal piano dell'immagine all'osservatore. Essendo questi fattori variabili, per riuscire a osservare il contenuto delle immagini, non è sufficiente analizzarle nella loro risoluzione originale, ma è utile considerarle a diverse scale, ovvero a diverse risoluzioni, catturando i dettagli a diverse frequenze e spazialità. I fattori presi in considerazione dalla metrica, alle differenti scale, sono la luminosità, il contrasto e la struttura dell'immagine, dove la struttura è relativa alla disposizione degli elementi visivi presenti nell'immagine e alle loro relazioni[19].

Il valore di MS-SSIM, per confrontare il contenuto di due immagini che sono una la rielaborazione dell'altra, si ottiene sottoponendole ai seguenti passaggi:

• Definizione del numero di scale (risoluzioni) M per la valutazione multi-scala.

Per ogni scala j da 1 a M:

- Applicazione di un filtro passa-basso e sottocampionamento di un fattore 2.
- Calcolo dell confronto del contrasto cj(x, y) tra l'immagine di riferimento e l'immagine rielaborata a alla scala j-esima.
- Calcolo del confronto della struttura sj(x, y) tra le due immagini alla scala j-esima.

• Calcolo del confronto della luminanza lM(x, y) tra l'immagine di riferimento e l'immagine manipolata, alla scala M.

Utilizzando i valori dei confronti di contrasto (c_j) , struttura (s_j) e luminanza (l_M) ottenuti a diverse scale, si valuta poi il valore finale dell'indice MS-SSIM con la relazione sotto riportata:

MS-SSIM
$$(x, y) = [lM(x, y)]^{\alpha_M} \prod_{j=1}^{M} [c_j(x, y)]^{\beta_j} [s_j(x, y)]^{\gamma_j}$$

dove:

 α_M è l'esponente per il confronto della luminanza alla scala M.

 β_i è l'esponente per il confronto del contrasto alla scala j.

 γ_j è l'esponente per il confronto della struttura alla scala j.

 \prod indica la moltiplicazione su tutte le scale j da 1 a M.

I Valori di MS-SSIM possono essere compresi tra 0 e 1, dove 0 indica la totale diversità tra le due immagini e 1 si riferisce alla loro totale uguaglianza.

2.6.3 MANIQA

Un ultimo aspetto che si è ritenuto opportuno valutare è quello della qualità delle immagini proposte in output dall'algoritmo in modo "No-Reference", ovvero senza la necessità di confronto con la loro versione originale.

La metrica Multi-dimension Attention Network for no-reference Image Quality Assessment (MANIQA) consente ldi effettuare una valutazione no-reference della qualità percettiva delle immagini. Questa metrica si basa su un modello di intelligenza artificiale a due rami che opera dividendo l'immagine in patch, calcolando punteggi di qualità per ciascuna patch nel primo ramo e combinando i punteggi delle patch per ottenere uno score finale di qualità nel secondo ramo. Questo approccio permette al modello di considerare sia i dettagli specifici delle singole patch che l'immagine nel suo insieme durante la valutazione della qualità[20].

• Il primo ramo del modello è responsabile della suddivisione dell'immagine in patch e dell'assegnazione di punteggi di qualità ad ognuna di esse attraverso il modello Vision Transformer (ViT). Il modello Vision Transformer permette di elaborare e comprendere le informazioni visive di un'immagine senza l'uso di operazioni di convoluzione tipiche delle reti neurali convoluzionali (CNN). L'elaborazione per generare punteggi di qualità per ogni patch dell'immagine avviene utilizzando moduli di attenzione, come il Transposed Attention Block

(TAB) e il Scale Swin Transformer Block (SSTB) che sono componenti chiave all'interno del modello, progettati per catturare le relazioni tra le diverse parti dell'immagine di input. Nello specifico, Il Transposed Attention Block (TAB) si concentra sulle interazioni tra i diversi canali dell'immagine, mentre lo Scale Swin Transformer Block (SSTB) si focalizza sulle interazioni tra le varie regioni spaziali dell'immagine.

I moduli TAB e SSTB fanno si che la rete sia in grado di catturare le connessioni tra diverse regioni dell'immagine, considerando sia l'aspetto complessivo che i dettagli specifici.

• Il secondo ramo del modello si occupa della combinazione dei punteggi ottenuti sulle patch per raggiungere un valore finale della qualità complessiva dell'immagine. Questo ramo considera il peso di ciascun punteggio delle patch e lo utilizza per calcolare lo score finale. La combinazione dei punteggi delle patch tiene conto dell'importanza relativa di ciascuna patch nell'immagine globale, consentendo di valutare la qualità complessiva.

Un punteggio finale della metrica MANIQA più vicino a 0 indica una bassa qualità percettiva dell'immagine, mentre valori più alti indicano una qualità percettiva superiore dell'immagine.

Capitolo 3 Risultati

In questo capitolo vengono esaminate le prestazioni dei due differenti approcci implementati per la correzione delle immagini.

In particolare vengono esposti i risultati ottenuti dall'applicazione dell'algoritmo euristico generale sviluppato per la correzione degli artefatti nelle immagini dei tre differenti dataset accomunati dagli stessi artefatti, esaminando le immagini ottenute dalla sua applicazione.

Successivamente sono presentati i risultati ottenuti dall'implementazione del metodo euristico specializzato alla correzione del fenomeno dell'autofluorescenza per l'elaborazione delle immagini dei nuclei di cardiosfera.

Infine è proposta un'analisi delle prestazioni del modello generativo addestrato alla produzione di immagini aventi caratteristiche simili a quelle prodotte dall'algoritmo specializzato, a partire dallo stesso set di dati originali. Il capitolo termina con un confronto tra i risultati ottenuti con l'algoritmo specializzato e con il modello generativo, effettuato attraverso la valutazione delle metriche di validazione calcolate sugli output forniti dai due metodi.

3.1 Risultati algoritmo generico

I risultatati proposti dall'algoritmo euristico sviluppato sono stati valutati, in un primo momento, in modo qualitativo, per analizzare l'effettivo incremento della qualità delle immagini dal punto di vista delle criticità individuate. L'algoritmo generico ha mostrato di essere funzionante su differenti tipologie di immagini caratterizzate da contenuto semantico, risoluzione e formato differenti.

Di seguito sono mostrati degli esempi dell'applicazione dell'algoritmo generico adattato al perfezionamento dei tre diversi gruppi di immagini che comprende la funzione per l'illuminazione degli oggetti oscurati, il denoising per la rimozione del rumore dal background e il miglioramento della nitidezza ridotta dalle manipolazioni precedenti.



Figura 3.1: Risultato dell'applicazione dell'algoritmo generico alle acquisizioni di nuclei di cardiosfera.



Figura 3.2: Risultato dell'applicazione dell'algoritmo generico alle immagini di nuclei U2OS.

Risultati



Figura 3.3: Risultato dell'applicazione dell'algoritmo generico alle immagini di mitocondri e actina.

La figura 3.3 mostra l'effetto dell'algoritmo generico sulle immagini raffiguranti le strutture cellulari di actina e mitocondri. Come accennato nella descrizione del loro dataset di appartenenza, esse sono state registrate con due metodi di acquisizioni diversi. In particolare è stata acquisita una loro versione a basso rapporto segnale-rumore ed una ad alto rapporto segnale-rumore, incrementando i tempi du acquisizione. L'algoritmo sviluppato permette di ottenere una loro versione migliorata molto simile a quella ottenuta con un processo di acquisizione caratterizzato da un tempo di registrazione più lungo, a partire da quelle a basso SNR, senza dover agire sull'acquisizione.



Figura 3.4: Confronto tra le immagini di mitocondri e actina migliorate con l'acquisizione e con l'algoritmo.

Si può osservare che il metodo è in grado di illuminare gli oggetti, rispettando il contrasto al loro interno e preservando la loro morfologia e la dinamica del colore. L'esito ottenuto dall'implementazione di questo metodo consente di evitare di intervenire sul processo di acquisizione per l'ottenimento di immagini dalla qualità visiva migliore.

3.2 Risultati algoritmo specializzato nella correzione dell'autofluorescenza

Nel paragrafo 2.4 è stata introdotta la specializzazione dell'algoritmo euristico sulle acquisizioni tridimensionali delle cardiosfere effettuata dopo aver verificato il funzionamento e la robustezza della versione generale su diverse tipologie di acquisizioni.

Le immagini estratte dai volumi contenuti nel CARE algorithm dataset si differenziano dalle altre per la presenza di una problematica ulteriore a quelle incontrate nelle acquisizioni degli altri due set di immagini, ovvero quella dell'autofluorescenza. Il metodo è stato dunque riadattato per renderlo efficace anche sulla correzione di questo tipo di interferenza luminosa che ostacola ulteriormente la visualizzazione degli oggetti di interesse che, in questo caso, sono i nuclei delle cellule staminali cardiache. Le figure di seguito mostrate illustrano l'effetto dell'adattamento dell'algoritmo a questo tipo di artefatto.





Figura 3.5: Risultato dell'applicazione dell'algoritmo specializzato nella correzione dell'autofluorescenza nelle immagini dei nuclei di cardiosfera.

3.3 Risultati modello generativo addestrato

La GAN con architettura pix2pix è stata addestrata alla generazione di immagini con caratteristiche simili a quelle prodotte dall'algoritmo specializzato sui nuclei di cardiosfere, e quindi caratterizzate anche da una calibrazione dell'intensità dell'autofluorescenza. Seguono alcuni esempi delle immagini sintetiche migliorate generate dal modello allenato.







Figura 3.6: Esempio di immagini sintetiche migliorate generate dal modello addestrato a partire dalla loro versione originale.

I confronti mostrano che la rete ha appreso efficacemente le caratteristiche delle immagini dei nuclei di cardiosfera elaborate dall'algoritmo, riuscendo anche a correggere in modo ottimale l'effetto indesiderato dell'autofluorescenza e ad incrementare l'intensità degli oggetti. Su alcuni esempi del test set, dunque su immagini non utilizzate per l'addestramento, il modello allenato mostra una capacità di illuminare gli oggetti poco visibili con un'enfasi maggiore rispetto a quanto accade con il metodo euristico. La figura in basso illustra alcuni esempi di questo fenomeno. Questo deriva dalla capacità di generalizzazione del modello non euristico.





Figura 3.7: Confronto tra il risultato prodotto dall'algoritmo e il risultato prodotto dalla GAN su una stessa immagine.

3.4 Risultati metriche di validazione

Le metriche precedentemente descritte sono state utilizzate per valutare le prestazioni dell'algoritmo euristico specializzato e del modello generativo, in modo da confrontare le loro performance ottenute sullo stesso task. Le metriche selezionate sono state calcolate sul gruppo di immagini contenute nel test set utilizzato per la GAN e valutano tre aspetti delle immagini prodotte dai due metodi:

- aumento del contrasto
- similarità strutturale tra l'immagine originale e l'output del metodo
- qualità dell'immagine di output

La decisione di valutare l'aumento del contrasto apportato dal processing delle immagini è stata definita dal fatto che l'algoritmo, e di conseguenza il modello generativo addestrato, agendo sull'incremento dei livelli di luminosità degli oggetti e sulla riduzione o eliminazione del rumore sullo sfondo e dell'effetto di autofluorescenza, apportano inevitabilmente un aumento del contrasto. Le variazioni di contrasto sono state state valutate con il Contrast Improvement Index (CII). Il valor medio di CII ottenuto sulle immagini del test set processate dall'algoritmo è di $2.64 \pm 43.2\%$, mentre il valor medio di CII ottenuto sulla loro versione generata dalla GAN è di $3.41\pm43.1\%$ circa. Questi risultati indicano che il contrasto delle immagini di output fornito dai due metodi aumenta di più del doppio rispetto alle rispettive immagini originali, mentre l'elevata deviazione standard è dovuta all'elevata variabilità delle immagini. Questo è in linea con il principio di azione dei due metodi che aumentano la luminosità degli oggetti, mentre riducono il rumore sullo sfondo. Inoltre, il fatto che il contrasto aumenti maggiormente nel caso delle immagini generate dalla GAN rispetto a quelle prodotte dal metodo euristico rispecchia il fatto che la rete è in grado di "spingere" di più sull'innalzamento dei valori di intensità dei pixel degli oggetti oscurati, come riscontrato dall'osservazione delle immagini sintetiche.



Figura 3.8: Valori di CII calcolati sulle immagini del test set generate dalla GAN e sulle immagini del test set processate dall'algoritmo.

La similarità strutturale tra l'immagine originale e la corrispettiva processata e sintetica è stata misurata attraverso il Multi-Scale Structural Similarity Index. Il valor medio di MS-SSIM nel caso dell'algoritmo risulta essere $0.86\pm7.8\%$, nel caso della GAN invece è pari a $0.79\pm4.7\%$. Il metodo euristico sviluppato tende ad essere più conservativo dal punto di vista del mantenimento del contenuto semantico dell'immagine, rispetto al modello generativo addestrato. Questo è principalmente dovuto al fatto che la calibrazione dell'autofluorescenza effettuata con l'algoritmo mira, in maniera conservativa, a ridurne l'effetto senza eliminarlo completamente, mentre il modello generativo tende a produrre immagini in cui l'autofluorescenza identificata viene quasi completamente rimossa. Tuttavia, i valori di MS-SSIM sono relativamente elevati ed indicano un'alta fedeltà di entrambi i metodi al contenuto semantico della versione originale dell'immagine.





Figura 3.9: Valori di MS-SSIM calcolati sulle immagini del test set generate dalla GAN e sulle immagini del test set processate dall'algoritmo.

L'ultima analisi effettuata sulle immagini è quella che mira a valutare la qualità dell'immagine con una modalità no-reference per comprendere quanto la versione migliorata dall'algoritmo e la sua rielaborazione prodotta dal modello genrativo addestrato, differiscono dall'originale da un punto di vista della qualità percettiva. Il valor medio del punteggio ottenuto sulle immagini originali del test set è di $0.45\pm6.9\%$, sulla loro variante prodotta dall'algoritmo è di $0.44\pm4.8\%$, mentre sulla loro versione rielaborata generata dalla GAN è pari a $0.44\pm4.2\%$. Il confronto mostra, dunque, che entrambi i metodi non introducono una significativa variazione della qualità percettiva dell'immagine, riuscendo a non degradare la qualità complessiva dell'immagine.





Figura 3.10: Valori di MANIQA calcolati sulle immagini del test set originali, sulla loro versione rielaborata generata dalla GAN e sulle immagini del test set processate dall'algoritmo.

Capitolo 4 Conclusioni

Il lavoro di tesi presentato ha avuto come scopo l'implementazione di metodi automatici per l'elaborazione di immagini di microscopia confocale a fluorescenza. L'obiettivo di produrre versioni migliorate delle acquisizioni di microscopia è quello di facilitare il lavoro degli specialisti nel campo dell'anatomia patologica, per riuscire a rilevare ed identificare con meno difficoltà alterazioni di strutture dell'organismo a livello microscopico.

Il vantaggio di effettuare un miglioramento delle immagini mediche attraverso l'applicazione di algoritmi è quello di evitare interventi sul protocollo di acquisizione, riducendo inevitabilmente i tempi degli esami medici. Questo è stato riscontrato, ad esempio, nei risultati ottenuti sulle acquisizioni delle strutture cellulari di actina e mitocondri, dove è stato possibile ottenere delle loro versioni di qualità migliorata molto simili a quelle che erano state ricavate agendo sul metodo di registrazione, aumentando i tempi di acquisizione.

Le problematiche e gli artefatti che si possono identificare nella tipologia di immagini su cui si è lavorato nella tesi possono essere diverse e per questo è stato importante cercare di strutturare metodi di image-processing che fossero sia robusti e generici dal punto di vista della funzionalità su differenti aspetti critici, che specifici per la correzione di artefatti più singolari.

Gli algoritmi euristici sviluppati, tuttavia, possono presentare delle limitazioni in termini di performance. In questo caso è stato molto utile sfruttare metodi non euristici in grado di generalizzare le manipolazioni effettuate sulle immagini e che hanno permesso di superare le prestazioni dell'euristica grazie alla loro caratteristica di andare oltre il semplice seguire una sequenza di istruzioni ben definite e ordinate che descrivono passo dopo passo come risolvere una determinata problematica.

L'intelligenza artificiale generativa, nel caso specifico del lavoro di tesi, si è rivelata una grande risorsa che ha permesso di superare le limitazioni del metodo euristico e può offrire anche la possibilità di generare dei nuovi dataset di immagini sintetiche migliorate che possono essere utili per studi futuri. In generale, è stato

Conclusioni

possibile risolvere il task prefissato sfruttando in maniera complementare i risultati dell'algoritmo euristico e i risultati del modello generativo addestrato. Il progetto di tesi, dunque, costituisce un punto di partenza per lo sviluppo di protocolli di image processing utili ad agevolare l'analisi di immagini di strutture cellulari a scopi diagnostici, producendo delle versioni migliorate di acquisizioni di microscopia confocale senza la necessità di agire sul protocollo di acquisizione. Un possibile step aggiuntivo da sviluppare in futuro potrebbe essere quello di progettare un'interfaccia grafica (GUI) attraverso la quale settare i vari parametri che l'algoritmo acquisisce in ingresso in maniera più "user-friendly" per visualizzare l'effetto delle variazioni dei parametri sull'immagine in uscita. Inoltre, si potrebbe anche pensare di combinare i risultati del metodo euristico con i risultati della rete generativa per ottenere delle immagini di output con un grado di qualità superiore derivante dalla fusione delle prestazioni delle due metodiche.

Bibliografia

- Kherlopian, A.R., Song, T., Duan, Q. et al. «A review of imaging techniques for systems biology.» In: *BMC Systems Biology* 2 (2008). URL: https://doi. org/10.1186/1752-0509-2-74 (cit. a p. 10).
- [2] Elliott AD. «Confocal Microscopy: Principles and Modern Practices.» In: *Current protocols in cytometry*. 92 (2020). URL: https://doi.org/10.1002/ cpcy.68 (cit. alle pp. 11, 12).
- Salvi, M., Morbiducci, U., Amadeo, F. et al. «Automated Segmentation of Fluorescence Microscopy Images for 3D Cell Detection in human-derived Cardiospheres». In: Sci Rep 9 (2019). URL: https://doi.org/10.1038/ s41598-019-43137-2 (cit. alle pp. 13, 16-18).
- [4] «Nuclei of U2OS cells in a chemical screen.» In: Content Broad Bioimage Benchmark Collection. (). URL: https://bbbc.broadinstitute.org/BBBC0 39 (cit. alle pp. 13, 19, 20).
- [5] Hagen, Guy and Bendesky, Justin and Machado, Rosa and Nguyen, Tram-Anh and Kumar, Tanmay and Ventura, Jonathan. «Fluorescence microscopy datasets for training deep neural networks». In: *GigaScience* 10 (mag. 2021). URL: https://doi.org/10.1093/gigascience/giab032 (cit. alle pp. 13, 20, 21).
- [6] Pesce, M., Messina, E., Chimenti, I., & Beltrami, A. P. «Cardiac Mechanoperception: A Life-Long Story from Early Beats to Aging and Failure». In: *Stem cells and development* (gen. 2017). URL: https://doi.org/10.1089/ scd.2016.0206 (cit. a p. 16).
- [7] Chimenti, I., Massai, D., Morbiducci, U., Beltrami, A. P., Pesce, M., & Messina, E. «Stem Cell Spheroids and Ex Vivo Niche Modeling: Rationalization and Scaling-Up». In: *Journal of cardiovascular translational research* 10 (apr. 2017). URL: https://doi.org/10.1007/s12265-017-9741-5 (cit. a p. 16).

- [8] Barile, L., Chimenti, I., Gaetani, R., Forte, E., Miraldi, F., Frati, G., Messina, E., & Giacomello, A. «Cardiac stem cells: isolation, expansion and experimental use for myocardial regeneration». In: *Nature clinical practice. Cardiovascular medicine* 4 (2007). URL: https://doi.org/10.1038/ncpcardio0738 (cit. a p. 16).
- [9] Gaetani, R., Feyen, D. A., Doevendans, P. A., Gremmels, H., Forte, E., Fledderus, J. O., Ramjankhan, F. Z., Messina, E., Sussman, M. A., Giacomello, A., & Sluijter, J. P. «Different types of cultured human adult cardiac progenitor cells have a high degree of transcriptome similarity». In: *Journal of cellular and molecular medicine* (2014). URL: https://doi.org/10.1111/jcmm.12458 (cit. a p. 16).
- [10] E. et al. Messina. «Isolation and expansion of adult cardiac stem cells from human and murine heart.» In: *Circulation research* 95 (ott. 2004). URL: https://doi.org/10.1161/01.RES.0000147315.71699.51 (cit. a p. 17).
- [11] Bernas, T., Zarebski, M., Dobrucki, J. W., & Cook, P. R. «Minimizing photobleaching during confocal microscopy of fluorescent probes bound to chromatin: role of anoxia and photon flux.» In: *Journal of microscopy* 215 (mag. 2004). URL: https://doi.org/10.1111/j.0022-2720.2004.01377.x (cit. a p. 22).
- [12] Othman, G., & Zeebaree, D. Q. «The Applications of Discrete Wavelet Transform in Image Processing: A Review.» In: Journal of Soft Computing and Data Mining 1 (giu. 2020). URL: https://publisher.uthm.edu.my/ ojs/index.php/jscdm/article/view/7215 (cit. a p. 32).
- [13] Parida, P., Bhoi, N. «Wavelet based transition region extraction for image segmentation». In: *Future Computing and Informatics Journal* 2 (dic. 2017). URL: https://www.sciencedirect.com/science/article/pii/S2314728 817300508 (cit. a p. 33).
- [14] Dehda Bachir e Khaled Melkemi. «Image denoising using new wavelet thresholding function». In: Journal of Applied Mathematics and Computational Mechanics 16 (giu. 2017), pp. 55–65. DOI: 10.17512/jamcm.2017.2.05 (cit. a p. 33).
- [15] Han, Z., Jian, M., Wang, G. «ConvUNeXt: An Efficient Convolution Neural Network for Medical Image Segmentation». In: *Knowledge-Based Systems* 253 (ott. 2022). URL: https://doi.org/10.1016/j.knosys.2022.109512 (cit. a p. 39).
- [16] Orhan G. Yalçın. «Image Generation in 10 Minutes with Generative Adversarial Networks». In: *Towards Data Science* (set. 2020). URL: https://towardsdatascience.com/image-generation-in-10-minutes-with-generative-adversarial-networks-c2afc56bfa3b (cit. a p. 45).

- [17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou e Alexei A. Efros. «Image-to-Image Translation with Conditional Adversarial Networks». In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (lug. 2016). URL: https://doi.org/10.48550/arXiv.1611.07004 (cit. a p. 48).
- [18] P. Siva Kumar. «IMAGE CONTRAST ENHANCEMENT USING HISTO-GRAM EQUALIZATION TECHNIQUES: REVIEW». In: (2014). URL: http: //warse.org/pdfs/2014/ijacst03332014.pdf (cit. a p. 51).
- [19] Z. Wang, Eero Simoncelli e Alan Bovik. «Multiscale structural similarity for image quality assessment». In: vol. 2. Dic. 2003, 1398–1402 Vol.2. URL: https://ieeexplore.ieee.org/document/1292216 (cit. a p. 52).
- [20] Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao Yuan Gong, Mingdeng Cao, Jiahao Wang, Yujiu Yang. «MANIQA: Multi-dimension Attention Network for No-Reference Image Quality Assessment». In: arXiv - CS - Computer Vision and Pattern Recognition (ott. 2022). URL: https://doi.org/10. 48550/arXiv.2204.08958 (cit. a p. 53).
Ringraziamenti

Giunta al termine di questo percorso, credo sia necessario un momento di riflessione per dedicare un ultimo sguardo a quanto accaduto in questi anni e per ringraziare chi è stato accanto a me fino a questo momento.

Vorrei ringraziare innanzitutto i miei relatori, in particolare il professor Salvi e l'ingegner Branciforti per avermi dato l'opportunità di lavorare a questo progetto di tesi che mi ha permesso di ampliare le mie conoscenze in questo ambito per me di grande interesse e per avermi seguito impeccabilmente nella sua realizzazione.

Un grazie immenso va ai miei genitori, senza di voi tutto questo non sarebbe stato realizzabile. Grazie Papà per i tuoi consigli preziosi e per avermi trasmesso l'amore per la scienza. Grazie Mamma per avere ogni giorno un pensiero per me e per sapermi stare vicino anche da lontano. Grazie ad entrambi per avermi sempre sostenuta ed incoraggiata durante tutti i momenti di difficoltà e per aver gioito con me ad ogni piccola vittoria.

Ai miei fratelli, Marica, Matteo e Marco, ed anche a Paola e Claudio e ai miei "nipotini" Emma, Gioele e Adele, mi avete sempre dato sicurezza e grazie a voi non mi sono mai sentita e mai mi sentirò sola.

Ringrazio i miei compagni di corso "bio" e le mie coinquiline Marilyn e Mariachiara per tutti i momenti condivisi che hanno alleggerito le giornate più grigie e noiose. Grazie Roberta, la mia amica di sempre, in grado di strapparmi un sorriso anche lontano dai banchi del liceo.

Un ringraziamento davvero speciale va a Te. Grazie Antonio. Insieme abbiamo condiviso ogni singolo momento di questo percorso fatto di successi e insuccessi, cambiamenti e progressi. Sono certa che senza il tuo amore e senza la tua pazienza nel supportarmi anche durante i miei momenti bui sarebbe stato tutto ancora più difficile.

Infine un piccolo ringraziamento a me stessa per non aver mai perso la motivazione, anche quando tutto sembrava quasi impossibile.

"Se un sogno ha così tanti ostacoli, significa che è quello giusto".