



**Politecnico  
di Torino**

# **POLITECNICO DI TORINO**

**Corso di Laurea Magistrale in Ingegneria Biomedica**

**Tesi di Laurea Magistrale**

## **Sviluppo di un modello generativo per la simulazione di immagini di tomografia a coerenza ottica nella degenerazione maculare legata all'età**

**Relatore**

Prof. Filippo Molinari

**Candidata**

Monica Sequenza

**Correlatori**

Prof.ssa Kristen Mariko Meiburger

Prof. Massimo Salvi

Prof. Michele Reibaldi

Prof. Enrico Borrelli

**Anno Accademico 2024/2025**



## Sommario

La Degenerazione Maculare Legata all'Età (AMD) rappresenta una delle principali cause di perdita della vista nei Paesi industrializzati, con impatto significativo sulla qualità della vita. La malattia si manifesta con una progressiva degenerazione della macula, compromettendo la visione centrale. La diagnosi precoce e il monitoraggio dei diversi stadi della patologia risultano fondamentali per garantire interventi terapeutici tempestivi ed efficaci; in grado di rallentare la progressione della malattia e preservare la funzione visiva residua.

In questo contesto, la Tomografia a Coerenza Ottica (OCT) rappresenta lo strumento di imaging più avanzato e diffuso per lo studio della retina. Riconoscere correttamente le caratteristiche delle immagini OCT è fondamentale per consentire ai clinici di diagnosticare la malattia e identificarne lo stadio, tuttavia, la disponibilità limitata di dataset clinici ampi e annotati riduce le opportunità di addestramento e pratica. In questo scenario, l'utilizzo di Reti Generative Avversarie (GAN) per generare immagini sintetiche realistiche rappresenta uno strumento sia per ampliare i dataset disponibili sia per supportare la formazione dei clinici.

L'obiettivo di questa tesi è lo sviluppo di una GAN per generare immagini OCT relative ai diversi stadi dell'AMD. Nello studio sono stati considerati quattro stadi della malattia: soggetti sani, stadio intermedio, forma neovascolare e atrofia geografica. I dati di ciascuna categoria sono stati trattati separatamente e suddivisi in training e test set con proporzione 80%-20%. Le immagini in input alla rete sono state sottoposte a un preprocessing finalizzato alla generazione di mappe semantiche. Le immagini sintetiche ottenute sono state valutate sia da clinici esperti, tramite una valutazione qualitativa in modalità cieca, sia mediante metriche quantitative, al fine di verificarne qualità visiva, coerenza strutturale e fedeltà delle caratteristiche patologiche. I valori ottenuti tramite PSNR, SSIM, FID indicano una moderata fedeltà strutturale (PSNR  $\sim$  22-23; SSIM  $\sim$  0.43-0.46). L'analisi FID mostra risultati migliori per i sani e atrofia geografica ( $\sim$  28). Il CII, con valori compresi tra 1.6 e 1.7, mostra un miglioramento del contrasto nelle immagini sintetiche. L'analisi delle texture ha mostrato un errore medio compreso tra circa 2.6% e 3.8%, dimostrando che la GAN riproduce con fedeltà le caratteristiche locali delle immagini negli stadi della malattia.

I risultati suggeriscono che la GAN risulta in grado di generare immagini realistiche e coerenti, con potenziale applicazione nell'ampliamento dei dataset per futuri modelli di deep learning e come supporto alla formazione clinica.

## English version

Age-related macular degeneration (AMD) is one of the leading causes of vision loss in industrialised countries, with a significant impact on quality of life. The disease manifests itself as progressive degeneration of the macula, compromising central vision. Early diagnosis and monitoring of the different stages of the disease are essential to ensure timely and effective therapeutic interventions that can slow the progression of the disease and preserve residual visual function.

In this context, Optical Coherence Tomography (OCT) is the most advanced and widely used imaging tool for studying the retina. Correctly recognising the characteristics of OCT images is essential to enable clinicians to diagnose the disease and identify its stage. However, the limited availability of large, annotated clinical datasets reduces opportunities for training and practice. In this scenario, the use of Generative Adversarial Networks (GANs) to generate realistic synthetic images is a tool for both expanding the available datasets and supporting the training of clinicians.

The aim of this thesis is to develop a GAN to generate OCT images relating to the different stages of AMD. The study considered four stages of the disease: healthy subjects, intermediate stage, neovascular form and geographic atrophy. The data for each category were processed separately and divided into training and test sets with a ratio of 80% to 20%. The images input into the network underwent preprocessing aimed at generating semantic maps. The synthetic images obtained were evaluated both by expert clinicians, through a blind qualitative assessment, and by quantitative metrics, in order to verify their visual quality, structural consistency, and fidelity of pathological characteristics. The values obtained using PSNR, SSIM, and FID indicate moderate structural fidelity (PSNR  $\sim$  22–23; SSIM  $\sim$  0.43–0.46). The FID analysis shows better results for healthy subjects and geographic atrophy ( $\sim$  28). The CII, with values between 1.6 and 1.7, shows an improvement in contrast in synthetic images. Texture analysis showed an average error of between approximately 2.6% and 3.8%, demonstrating that GAN faithfully reproduces the local characteristics of images in the stages of the disease.

The results suggest that GANs are capable of generating realistic and consistent images, with potential applications in expanding datasets for future deep learning models and as a support for clinical training.



# Indice

|                                                        |     |
|--------------------------------------------------------|-----|
| <b>Elenco delle figure</b>                             | III |
| <b>Elenco delle tabelle</b>                            | V   |
| <b>1 Introduzione</b>                                  | 1   |
| 1.1 Anatomia dell'occhio e della retina . . . . .      | 1   |
| 1.2 Inquadramento clinico dell'AMD . . . . .           | 4   |
| 1.2.1 Classificazione degli stadi . . . . .            | 5   |
| 1.2.2 Diagnosi e tecniche di imaging . . . . .         | 5   |
| 1.3 Tomografia a Coerenza Ottica (OCT) . . . . .       | 7   |
| 1.4 Intelligenza Artificiale in medicina . . . . .     | 10  |
| 1.4.1 Generative Adversarial Networks (GANs) . . . . . | 11  |
| 1.4.2 Pix2Pix . . . . .                                | 13  |
| <b>2 Materiali e metodi</b>                            | 16  |
| 2.1 Dataset . . . . .                                  | 16  |
| 2.2 Preprocessing . . . . .                            | 17  |
| 2.3 Parametri della GAN . . . . .                      | 21  |
| 2.4 Metriche di valutazione . . . . .                  | 22  |
| <b>3 Risultati</b>                                     | 28  |
| 3.1 Risultati GAN . . . . .                            | 28  |
| 3.2 Analisi qualitativa . . . . .                      | 32  |
| 3.3 Analisi quantitativa . . . . .                     | 34  |
| <b>4 Discussione e Conclusione</b>                     | 39  |
| 4.1 Discussione . . . . .                              | 39  |
| 4.2 Sviluppi futuri . . . . .                          | 40  |
| <b>Bibliografia</b>                                    | 42  |

# Elenco delle figure

|      |                                                                                                                         |    |
|------|-------------------------------------------------------------------------------------------------------------------------|----|
| 1.1  | Occhio umano . . . . .                                                                                                  | 2  |
| 1.2  | Sezione OCT della retina . . . . .                                                                                      | 4  |
| 1.3  | Interferometro di Michelson-Morley . . . . .                                                                            | 8  |
| 1.4  | Rappresentazione funzionamento TD-OCT, SD-OCT e SS-OCT . .                                                              | 10 |
| 1.5  | Algoritmo GAN . . . . .                                                                                                 | 12 |
| 1.6  | Schematizzazione del funzionamento di una GAN . . . . .                                                                 | 13 |
| 1.7  | Schema Pix2Pix . . . . .                                                                                                | 15 |
| 2.1  | Immagine originale soggetto sano (HEA_ID300 slice 24) . . . . .                                                         | 17 |
| 2.2  | Immagine processata di un soggetto sano, ritagliata e con padding<br>(HEA_ID300 slice 24) . . . . .                     | 18 |
| 2.3  | Step di preprocessing applicati . . . . .                                                                               | 18 |
| 2.4  | Esempio di immagine accoppiata fornita in input alla GAN (HEA_ID288<br>slice 25) . . . . .                              | 20 |
| 2.5  | Esempio di immagine accoppiata fornita in input alla GAN durante<br>la fase di inference (HEA_ID300 slice 25) . . . . . | 21 |
| 2.6  | Interfaccia grafica valutazione clinica: valutazione . . . . .                                                          | 26 |
| 2.7  | Interfaccia grafica valutazione clinica: originale Vs immagine sintetica                                                | 27 |
| 3.1  | Risultati GAN stadio sano . . . . .                                                                                     | 29 |
| 3.2  | Risultati GAN stadio intermedio . . . . .                                                                               | 30 |
| 3.3  | Risultati GAN stadio neovascolare . . . . .                                                                             | 31 |
| 3.4  | Risultati GAN stadio atrofia geografica . . . . .                                                                       | 32 |
| 3.5  | Istogramma della validazione della prima GUI . . . . .                                                                  | 33 |
| 3.6  | Andamento delle feature di texture del primo ordine: Energia . . .                                                      | 35 |
| 3.7  | Andamento delle feature di texture del primo ordine: Kurtosis . . .                                                     | 35 |
| 3.8  | Andamento delle feature di texture del primo ordine: Mean Absolute<br>Deviation . . . . .                               | 36 |
| 3.9  | Andamento delle feature di texture del primo ordine: Uniformity . .                                                     | 36 |
| 3.10 | Andamento delle feature di texture del secondo ordine: Contrast,<br>Difference Average e Difference Entropy . . . . .   | 37 |

|      |                                                                                                                            |    |
|------|----------------------------------------------------------------------------------------------------------------------------|----|
| 3.11 | Andamento delle feature di texture del secondo ordine: Difference<br>Variance, Inverse Variance, Autocorrelation . . . . . | 37 |
| 3.12 | Andamento delle feature di texture del secondo ordine: Correlation,<br>Joint Average, Joint Energy . . . . .               | 38 |
| 3.13 | Andamento delle feature di texture del secondo ordine: Joint Energy,<br>MCC, Sum Average . . . . .                         | 38 |

# Elenco delle tabelle

|     |                                                                                                                                 |    |
|-----|---------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Strati istologici della retina . . . . .                                                                                        | 3  |
| 1.2 | Tecniche di imaging per la diagnosi e il monitoraggio dell'AMD . .                                                              | 7  |
| 2.1 | Distribuzione dei pazienti nei set di training e test per ciascuna categoria diagnostica. . . . .                               | 17 |
| 2.2 | Parametri principali della cGAN Pix2Pix. . . . .                                                                                | 22 |
| 3.1 | Distribuzione delle valutazioni da 1 a 5 per i diversi stadi . . . . .                                                          | 32 |
| 3.2 | Risultati delle immagini reali ottenuti dalla valutazione soggettiva delle immagini reali e sintetiche. . . . .                 | 33 |
| 3.3 | Risultati delle immagini sintetiche ottenuti dalla valutazione soggettiva delle immagini reali e sintetiche. . . . .            | 33 |
| 3.4 | Risultati CII, FID, SSIM, PSNR . . . . .                                                                                        | 34 |
| 3.5 | Risultati feature di texture del primo ordine . . . . .                                                                         | 34 |
| 3.6 | Valori medi $\pm$ deviazione standard delle feature di texture di secondo ordine (GLCM) per ciascuno stadio patologico. . . . . | 37 |

# Capitolo 1

## Introduzione

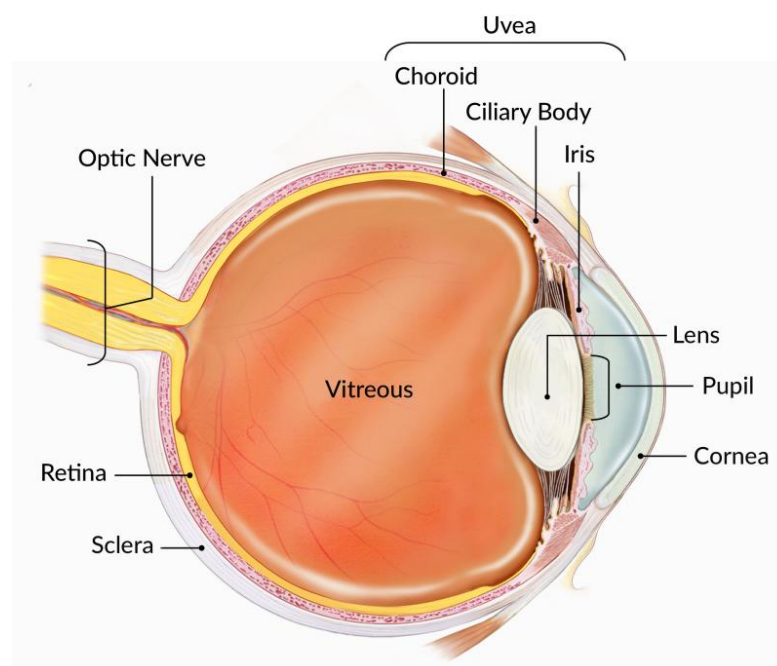
L'obiettivo di questo progetto di tesi è lo sviluppo di un'architettura Generative Adversarial Network (GAN) per la generazione di immagini di tomografia a Coerenza Ottica (OCT) relative a paziente affetti da degenerazione maculare legata all'età (AMD), nei diversi stadi di progressione della patologia. La creazione di un dataset sintetico ha lo scopo di supportare i clinici nell'apprendimento e nel processo diagnostico, e apre la strada a futuri sviluppi nell'ambito di applicazioni di deep learning e machine learning.

In questo capitolo introduttivo verranno presentate l'anatomia e le principali strutture dell'occhio e sarà descritta la patologia oggetto di studio. Verrà illustrata la tecnica di imaging oggi maggiormente utilizzata per la diagnosi, l'OCT, per poi concludere con un approfondimento sull'impiego dell'intelligenza artificiale in medicina con riferimento alle GAN e all'architettura Pix2Pix, adottata per il progetto.

### 1.1 Anatomia dell'occhio e della retina

L'occhio umano è l'organo deputato alla percezione visiva. La luce proveniente dall'ambiente esterno attraversa tutte le strutture ottiche dell'occhio, schematizzate nella Figura 1.1, dalla cornea fino a raggiungere la retina, dove l'energia luminosa viene convertita in segnali elettrici, che tramite il nervo ottico sono trasmessi al cervello [1].

La retina è costituita da dieci strati distinti, ognuno con una funzione specifica nella trasduzione e nel processamento visivo, come mostrato di seguito in Tabella 1.1 [2].



**Figura 1.1:** Schematizzazione occhio umano [1]

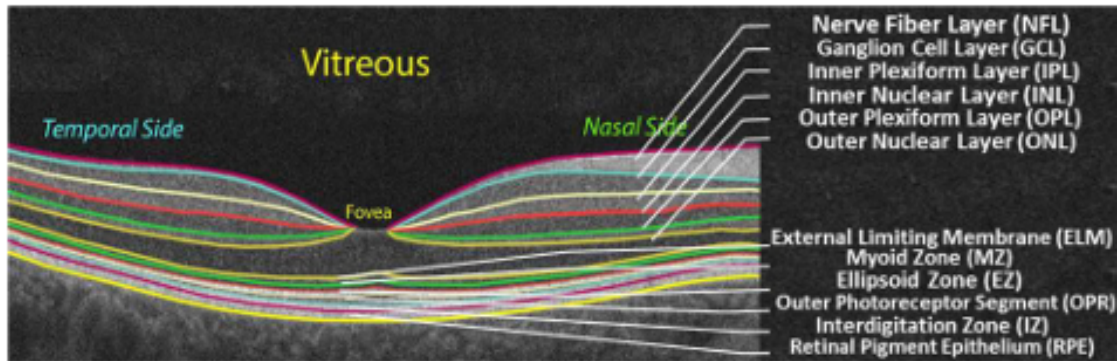
| Strato                               | Funzione principale                                                    |
|--------------------------------------|------------------------------------------------------------------------|
| Epitelio pigmentato retinico (RPE)   | Assorbimento luce, supporto metabolico dei fotorecettori               |
| Strato dei fotorecettori             | Trasduzione del segnale luminoso in impulso elettrico                  |
| Membrana limitante esterna (ELM)     | Barriera strutturale tra fotorecettori e cellule di Müller             |
| Strato nucleare esterno (ONL)        | Contiene i nuclei di coni e bastoncelli                                |
| Strato plessiforme esterno (OPL)     | Sinapsi tra fotorecettori, cellule bipolari e orizzontali              |
| Strato nucleare interno (INL)        | Contiene nuclei di cellule bipolari, amacrine, orizzontali e di Müller |
| Strato plessiforme interno (IPL)     | Sinapsi tra cellule bipolari, amacrine e gangliari                     |
| Strato delle cellule gangliari (GCL) | Contiene corpi cellulari delle cellule gangliari                       |
| Strato delle fibre nervose (NFL)     | Assoni delle cellule gangliari verso il nervo ottico                   |
| Membrana limitante interna (ILM)     | Barriera tra retina e corpo vitreo                                     |

**Tabella 1.1:** Rappresentazione dei dieci strati istologici della retina e delle rispettive funzioni principali

L'organizzazione stratificata e altamente specializzata della retina fa sì che le alterazioni anche di un singolo strato possono determinare lo sviluppo di importanti patologie oculari, tra cui la degenerazione maculare legata all'età (AMD), che interessa la regione della macula.

Nella retina troviamo due tipologie di fotorecettori: i coni, che permettono la visione a colori e l'alta definizione, e i bastoncelli, sensibili alla luce debole e al movimento, ma incapaci di distinguere il colore. La macula è la regione centrale della retina fondamentale per la visione ad alta risoluzione e la discriminazione fine dei dettagli. Al centro della macula si trova la fovea, una piccola depressione in cui sono presenti in modo concentrato i coni che stabiliscono una connessione sinaptica uno-a-uno con le cellule bipolari, per ottenere la massima acuità visiva. Spostandoci verso la zona periferica della macula, diminuisce progressivamente la densità dei coni e aumenta la presenza dei bastoncelli e in questa zona, a discapito dell'acuità visiva, migliora la percezione spaziale e la capacità di rilevare il movimento nell'ambiente circostante.

L'OCT ha rivoluzionato la diagnosi e la gestione delle patologie retiniche, permettendo di ottenere rapidamente immagini della retina visualizzando i diversi strati retinici e localizzando con precisione i processi patologici associati. In Figura 1.2 è riportata una scansione OCT in cui sono evidenziati i vari strati retinici [3].



**Figura 1.2:** Tomografia a coerenza ottica (OCT) della retina, con strati retinici evidenziati e nominati mediante etichette [4].

## 1.2 Inquadramento clinico dell'AMD

La degenerazione maculare rappresenta la principale causa di cecità nella popolazione al di sopra dei 55 anni. Si tratta di una patologia che si sviluppa con l'invecchiamento, e per questo è spesso indicata come denominata degenerazione maculare legata all'età (AMD) [5]. L'AMD, conseguenza del progressivo deterioramento o dal danneggiamento della macula, comporta una perdita della visione centrale, con conseguenti difficoltà nella visione nitida [6]. L'invecchiamento della popolazione renderà l'AMD una sfida rilevante per la sanità pubblica: si stima che, entro il 2040, le persone affette saranno circa 300 milioni. Questo avrà un impatto economico significativo, dovuto alla natura cronica della malattia e alla lunga gestione della malattia [7]. Dal punto di vista clinico si distinguono due forme principali:

- La forma secca o atrofica, è la più comune con una lenta progressione della malattia
- La forma umida o neovascolare, è meno frequente ma comporta una rapida evoluzione della malattia dovuto alla crescita di vasi sanguigni anomali sotto la retina

La causa dell'AMD è definita multifattoriale, è possibile fare una distinzione tra fattori di rischio non modificabili come l'età, il sesso femminile, il colore chiaro

dell'iride e una predisposizione genetica; e fattori modificabili come il fumo di sigaretta (il più rilevante perchè aumenta lo stress ossidativo e l'infiammazione). Uno stile di vita sano contribuisce ad essere un fattore protettivo dalla progressione della patologia.

### 1.2.1 Classificazione degli stadi

Per la classificazione dei diversi stadi dell'AMD sono stati considerati i criteri dell'Age-Related Eye Disease Study (AREDS) [8] che prendono in considerazione la dimensione dei drusen (depositi extracellulari di materiale proteico-lipido che si accumula tra RPE e la membrana di Bruch della retina) e le alterazioni RPE. La classificazione distingue i seguenti stadi:

- Nessuna AMD: assenza di drusen e alterazioni RPE
- Early AMD: sono presenti drusen di dimensioni comprese tra  $\geq 63 \mu\text{m}$  e  $< 125 \mu\text{m}$  ma non sono presenti significative alterazioni pigmentarie
- Intermediate AMD: sono presenti drusen di dimensioni maggiori di  $125 \mu\text{m}$  e/o alterazioni dell'RPE
- Late AMD: comprende la forma neovascolare (MNV) e la forma di atrofia geografica (GA); in questa fase si verifica la perdita visiva più significativa

La classificazione del paziente in base allo stadio di progressione della patologia consente di valutare il rischio di progressione dell'AMD, guidando le scelte cliniche permettendo interventi terapeutici mirati.

### 1.2.2 Diagnosi e tecniche di imaging

L'avanzamento della tecnologia per l'imaging medico ha rivoluzionato la metodologia di diagnosi.

Il primo esame utilizzato, considerato come gold standard, è la Color Fundus Photography (CFP). Questo metodo, ampiamente impiegato per la sua non invasività, non consente però di valutare il flusso vascolare nè le strutture retiniche profonde; altra tecnica utilizzata è la near-infrared reflectance imaging (NIR) che fornisce un'analisi dettagliata della retina in modo non invasivo. Tuttavia, l'interpretazione delle immagini richiede una maggiore esperienza da parte del clinico, poichè alcune alterazioni retiniche possono risultare poco evidenti. Tecniche come la fundus autofluorescence (FAF) sono importanti per la diagnosi di stadi di avanzamento come quello dell'atrofia geografica (GA), poichè fornisce immagini ad alto contrasto che evidenziano alterazioni dell'RPE senza l'utilizzo di mezzi di contrasto, ma deve

essere integrata con altre tecniche poichè non permette di visualizzare neovascolarizzazioni.

Le tecniche angiografiche con mezzo di contrasto impiegate per analizzare le neovascolarizzazioni sono :

- Fluorescein angiography (FA) considerata il gold standard che permette di identificare le neovascolarizzazioni, i drusen e le aree di atrofia. Ma è una tecnica invasiva con un tempo di acquisizione delle immagini molto lungo.
- Indocyanine green angiography (ICGA) è una tecnica invasiva poichè richiede un'iniezione endovenosa ma comporta un rischio minore di reazione allergica, consente di visualizzare la circolazione coroidale e le lesioni vascolari profonde.

La tecnica di imaging che ha rivoluzionato la diagnosi è la Tomografia a Coerenza Ottica (OCT) che risulta essere una tecnica non invasiva che fornisce immagini di alta qualità della retina e dell'epitelio pigmentato retinico (RPE), permette di rilevare drusen, liquidi intra-retinici e sotto-retinici, alterazioni strutturali della retina ma non fornisce informazioni sul flusso vascolare; quindi si introduce la OCT angiography (OCTA) per una diagnosi completa che permette di analizzare la vascolarizzazione retinica senza l'utilizzo di un mezzo di contrasto. Di seguito, nella Tabella 1.2 sono riassunte le principali tecniche descritte.

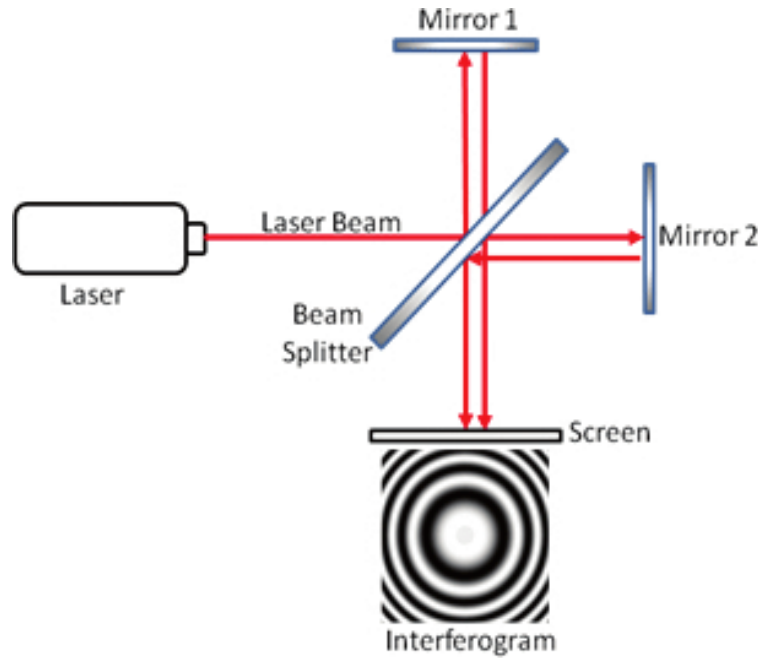
| <b>Tecnica</b>                              | <b>Vantaggi</b>                                                                      | <b>Limiti</b>                                                                |
|---------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| <b>Color Fundus Photography (CFP)</b>       | Non invasiva, veloce, utile per documentare drusen e lesioni superficiali            | Limitata alle strutture superficiali, non valuta il flusso vascolare         |
| <b>Near-Infrared Reflectance (NIR)</b>      | Esamina retina profonda, RPE e coroide; non invasiva                                 | Interpretazione meno intuitiva, ridotta sensibilità per lesioni superficiali |
| <b>Fundus Autofluorescence (FAF)</b>        | Ottima per valutare atrofia geografica e metabolismo RPE; non invasiva               | Non visualizza direttamente la neovascolarizzazione                          |
| <b>Fluorescein Angiography (FA)</b>         | Gold standard per neovascolarizzazioni; alta risoluzione del circolo retinico        | Invasiva, rischio di reazioni allergiche, acquisizione lunga                 |
| <b>Indocyanine Green Angiography (ICGA)</b> | Ottimale per lo studio della coroide e CNV occulte                                   | Invasiva, possibili reazioni allergiche                                      |
| <b>Optical Coherence Tomography (OCT)</b>   | Non invasiva, alta risoluzione strutturale, identifica drusen, fluidi, distacchi RPE | Non fornisce informazioni dirette sul flusso vascolare                       |
| <b>OCT Angiography (OCTA)</b>               | Non invasiva, valuta flusso vascolare retinico e coroideo, utile per CNV             | Non mostra leakage, artefatti di movimento, necessita OCT complementare      |

**Tabella 1.2:** Tecniche di imaging per la diagnosi e il monitoraggio dell'AMD

### 1.3 Tomografia a Coerenza Ottica (OCT)

La Tomografia a Coerenza Ottica (OCT) è una tecnica di imaging ottico non invasiva. Negli ultimi anni, la sua introduzione in abito clinico ha rivoluzionato i trattamenti oftalmologici, poichè le immagini ad alta risoluzione della retina hanno consentito analisi qualitative e quantitative degli strati retinici, della vascolarizzazione e delle lesioni [9]. Il principio di funzionamento alla base dell'OCT è simile a quello dell'ecografia medica: entrambe le tecniche si basano sull'emissione di un'onda verso il tessuto bersaglio, e sull'analisi delle onde riflesse, la cui differenza di tempo di ritorno consente di determinare la profondità delle strutture esaminate. Le due tecniche utilizzano però onde differenti: l'ecografia utilizza onde ultrasonore, mentre l'OCT utilizza onde elettromagnetiche, vicine all'infrarosso (con lunghezze d'onda

tra 800-1300 nm), con velocità molto elevate che rende difficile la misura diretta del ritardo temporale. Per risolvere questo problema viene utilizzato l'interferometro di Michelson [10]. L'interferometro di Michelson è uno strumento ottico per misurare differenze di percorso ottico. Come mostrato in Figura 1.3, un fascio di luce incidente viene diviso da uno specchio semiriflettente in due fasci che percorrono cammini ottici differenti; riflettendosi su due specchi distinti e ricombinandosi successivamente sul rivelatore. L'interferenza generata sul rivelatore dipende dalla differenza di cammino ottico tra i due fasci. Si ha interferenza costruttiva tale differenza è un multiplo intero della lunghezza d'onda della luce, producendo un segnale di massima intensità; si parla invece di interferenza distruttiva quando la differenza di cammino ottico corrisponde a un numero dispari di mezze lunghezze d'onda, con conseguente annullamento reciproco delle onde e generazione di un segnale prossimo allo zero. L'alternanza tra interferenza costruttiva e distruttiva consente di misurare con elevata precisione le variazioni di cammino ottico [11].



**Figura 1.3:** Interferometro di Michelson-Morley [11]

Nella progettazione di un OCT è importante la scelta della sorgente luminosa che determina la risoluzione assiale del dispositivo. Quest'ultima è correlata alla lunghezza di coerenza della sorgente,  $l_c$ , definita come la distanza lungo l'asse ottico nella quale il fascio mantiene una relazione di fase coerente per un'interferenza costruttiva. La lunghezza di coerenza, considerando uno spettro gaussiano è definita

come nell'Eq 1.1 :

$$l_c = \frac{2 \ln 2}{\pi} \cdot \frac{\lambda_0^2}{\Delta \lambda} \quad (1.1)$$

dove  $\lambda_0$  rappresenta la lunghezza d'onda centrale della sorgente e  $\Delta \lambda$  la lunghezza spettrale a metà altezza.

La risoluzione spaziale,  $\delta z$  è limitata dalla coerenza temporale della sorgente ed è definita come l'Eq 1.2:

$$\delta z \approx \frac{2 \ln 2}{\pi} \cdot \frac{\lambda_0^2}{n \Delta \lambda} \quad (1.2)$$

dove  $n$  è l'indice di rifrazione del mezzo di propagazione. Nei dispositivi moderni varia tra  $3 \mu\text{m}$  e  $7 \mu\text{m}$  a seconda della sorgente e dell'applicazione.

La risoluzione laterale  $\delta x$ , come mostrato dall'eq 1.3 dipende dalle proprietà ottiche del sistema di scansione e in particolare dalla dimensione del fascio e dall'apertura della lente della messa a fuoco NA. Solitamente  $\delta x$  varia tra  $10 \mu\text{m}$  e  $20 \mu\text{m}$

$$\delta x \approx \frac{2}{NA} \cdot 1.22 \lambda_0 \quad (1.3)$$

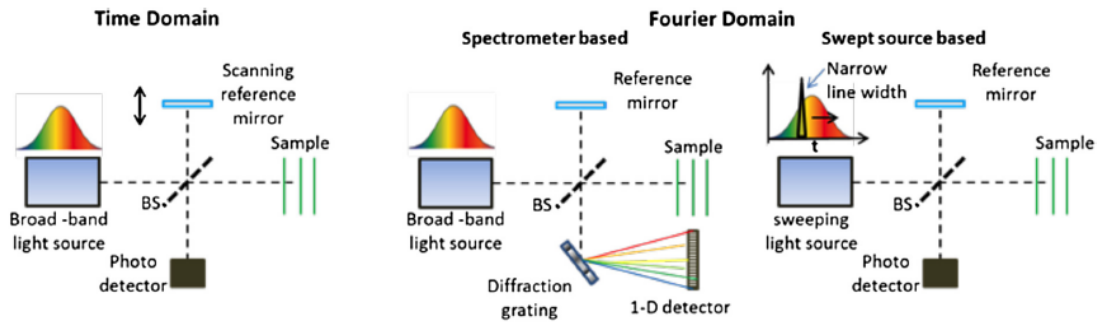
I sistemi OCT sono suddivisi in due categorie in base alla modalità di acquisizione del segnale: Time Domain OCT (TD-OCT) e Fourier Domain OCT (FD-OCT). La prima implementazione, la TD-OCT, la cui schematizzazione è mostrata in Figura ??, viene settata manualmente la lunghezza del percorso ottico del braccio di riferimento dell'interferometro. Questa scansione longitudinale permette di campionare l'intensità del segnale interferometrico e di individuare, per ciascuna posizione assiale, la condizione di interferenza costruttiva. In questo modo è possibile determinare la profondità del campione. Il segnale registrato per una singola posizione trasversale, noto come A-scan rappresenta il profilo di riflettività lungo l'asse ottico. La ripetizione di più A-scan, con uno spostamento laterale del fascio incidente, permette di ottenere una sezione bidimensionale del campione chiamata B-scan [10].

La FD-OCT rappresenta la seconda generazione dei sistemi OCT, in cui le informazioni spettrali del segnale interferometrico direttamente e l'A-scan viene ricostruito tramite la trasformata di Fourier; non è quindi necessaria la scansione manuale della lunghezza del percorso ottico. L'acquisizione può avvenire secondo due approcci la cui schematizzazione è mostrata in Figura 1.4:

- Spectral-Domain OCT (SD-OCT) è presente uno spettrometro (sostituisce il rilevatore puntale), che separa spazialmente le diverse componenti del fascio in base alla lunghezza d'onda tramite un elemento diffrattivo. Ogni acquisizione fornisce un interferogramma spettrale. La sorgente luminosa utilizzata è ad ampia banda spettrale per garantire una buona risoluzione [10].

- Swept Source OCT (SS-OCT) in questa configurazione è utilizzata una sorgente a banda stretta con una lunghezza d'onda che varia su un ampio intervallo spettrale. Ogni componente spettrale viene rilevata sequenzialmente da un sensore ottico, un fotorilevatore ad alta velocità [10].

In entrambe le varianti di FD-OCT l'interferometro riceve simultaneamente le informazioni relative a tutte le profondità del campione; la loro elaborazione tramite trasformata di Fourier permette poi di ottenere l'A-scan. Rispetto alla TD-OCT, in queste tecniche c'è un incremento del rapporto segnale-rumore e una maggiore velocità di acquisizione, quest'ultimo fattore è molto importante in ambito oftomologico dove il campione (l'occhio) è soggetto a continui movimenti.[10].



**Figura 1.4:** Rappresentazione funzionamento TD-OCT, SD-OCT e SS-OCT [12]

## 1.4 Intelligenza Artificiale in medicina

L'intelligenza artificiale (AI) fa riferimento a un insieme di tecnologie che permettono di simulare comportamenti umani, come i processi decisionali, l'apprendimento e la risoluzione di problemi o compiti complessi [13]. Un sottoinsieme dell'AI è il machine learning (ML), che comprende una serie di algoritmi basati su metodi statistici in grado di apprendere e prendere decisioni [14]. Il deep learning (DL), rappresenta un'evoluzione del ML, si basa su reti neurali artificiali. Rispetto ai modelli di ML, il DL permette di lavorare su grandi volumi di dati non strutturati ed estrarre automaticamente caratteristiche complesse dai dati [13]. I recenti progressi tecnologici hanno reso possibile l'integrazione dell'AI in diversi ambiti della medicina, tra cui la diagnosi precoce delle patologie, la pianificazione di trattamenti personalizzati, l'imaging biomedico e il supporto alla ricerca clinica tradizionale e traslazionale. Tali innovazioni hanno portato benefici significativi sia in termini economici sia nella qualità dell'assistenza sanitaria e nella relazione medico-paziente[15]. Tra le architetture del DL utilizzate in medicina troviamo le

Generative Adversarial Networks (GANs), utilizzate per la generazione di immagini mediche sintetiche, trasformazione delle immagini da una modalità diagnostica all'altra, per la rimozione di rumore dalle immagini ed altre applicazioni. La sintesi di nuove immagini ha permesso di superare i limiti delle tecniche tradizionali di data augmentation (come rotazioni o traslazioni), consentendo la creazione di dataset finale ampio, bilanciato e variegato, fondamentali per il training dei modelli di DL. In questa tesi, è stata utilizzata una GAN per la generazione di immagini OCT per i diversi stadi AMD [16].

### 1.4.1 Generative Adversarial Networks (GANs)

Le GAN sono un modello di DL introdotto nel 2014 da Ian Goodfellow e colleghi. Si basano su un approccio ispirato alla teoria dei giochi a somma zero, in cui due reti neurali artificiali, denominate Generatore ( $G$ ) e Discriminatore ( $D$ ), lavorano in competizione e vengono addestrate congiuntamente. Il Generatore riceve in input un vettore di rumore casuale  $z$ , con una distribuzione nota  $p(z)$ , tipicamente una distribuzione gaussiana standard, e produce un dato sintetico  $G(z)$  che cerca di imitare la distribuzione dei dati reali  $p_{\text{data}}(x)$  [17]. Il Discriminatore riceve in input sia i dati reali  $x \sim p_{\text{data}}(x)$  che dati sintetici  $G(z) \sim p_g(x)$  e stima la probabilità che il dato provenga dal dataset reale [17]. L'addestramento avviene in modo iterativo attraverso l'ottimizzazione min-max della funzione obiettivo (Eq. 1.4):

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1.4)$$

L'obiettivo del Discriminatore è massimizzare la probabilità di classificare correttamente i dati reali e sintetici, mentre il Generatore è addestrato per minimizzare la probabilità che i dati generati siano riconosciuti come falsi. Nelle prime fasi dell'addestramento, i dati del Generatore sono di bassa qualità, e il Discriminatore riesce a classificare correttamente tutti i dati, questo porta il termine  $\log(1 - D(G(z)))$  a saturazione ad un gradiente troppo debole per permettere un apprendimento efficace. Per risolvere il problema viene utilizzata per il Generatore la seguente espressione  $\log(D(G(z)))$ , questo permette di avere gradienti più significativi nelle prime fasi del training e migliorando la stabilità della convergenza [17]

---

**Algorithm 1** Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator,  $k$ , is a hyperparameter. We used  $k = 1$ , the least expensive option, in our experiments.

---

**for** number of training iterations **do**

**for**  $k$  steps **do**

- Sample minibatch of  $m$  noise samples  $\{z^{(1)}, \dots, z^{(m)}\}$  from noise prior  $p_g(z)$ .
- Sample minibatch of  $m$  examples  $\{x^{(1)}, \dots, x^{(m)}\}$  from data generating distribution  $p_{\text{data}}(x)$ .
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[ \log D(x^{(i)}) + \log (1 - D(G(z^{(i)}))) \right].$$

**end for**

- Sample minibatch of  $m$  noise samples  $\{z^{(1)}, \dots, z^{(m)}\}$  from noise prior  $p_g(z)$ .
- Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (1 - D(G(z^{(i)}))).$$

**end for**

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

---

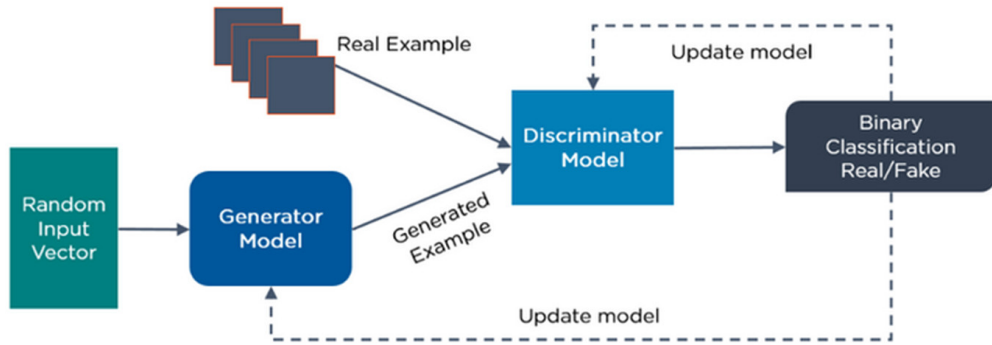
**Figura 1.5:** Algoritmo GAN [17]

L'algoritmo di addestramento proposto da Goodfellow e colleghi, come mostrato in Figura 1.5, utilizza la discesa del gradiente stocastica su mini-batch. Vengono eseguite più iterazioni di aggiornamento di  $D$  rispetto a  $G$  per mantenere  $D$  vicino al suo ottimo, dato un  $G$  fissato. Dal punto di vista teorico, fissato un Generatore esiste un Discriminatore ottimale  $D^*(x)$ , definito come (Eq. 1.5):

$$D^*(x) = \frac{p_{\text{data}}(x)}{p_{\text{data}}(x) + p_g(x)} \quad (1.5)$$

Sostituendo il  $D^*$  nella funzione obiettivo (Eq. 1.4), si ottiene la formulazione che misura la distanza tra le distribuzioni dei dati reali e quella dei dati generati, attraverso la Jensen-Shannon divergence (JSD). Questa divergenza è nulla solo quando  $p_{\text{data}} = p_g$ , quando il Generatore ha appreso la distribuzione dei dati reali. In questo caso la funzione obiettivo raggiunge il suo valore minimo pari a  $-\log(4)$ , e il Discriminatore restituisce una probabilità pari a 0.5, rappresenta l'incapacità di distinguere tra dati reali e dati generati. Tuttavia, nella pratica non è sempre possibile riuscire a raggiungere la convergenza perfetta del modello. Una delle criticità delle GAN è la necessità di avere una buona sincronizzazione tra l'aggiornamento di  $G$  e  $D$ . Un eccessivo aggiornamento di  $G$  può portare alla produzione di output quasi identici, cosiddette mode collapse (noto anche come "Helvetica

scenario”). Nonostante queste limitazioni, il framework delle GAN presenta dei vantaggi computazionali: l’addestramento che si basa solo sulla backpropagation, eliminando la necessità dei precedenti modelli generativi di tecniche di inferenza esplicita. Inoltre, il Generatore riceve solamente i gradienti dal Discriminatore, non i dati reali, riducendo così il rischio di overfitting e migliorando la generalizzazione. Il framework delle GAN può essere esteso introducendo variabili condizionali in ingresso a entrambe le reti, generando così le Conditional GAN (cGAN). Questo approccio permette di controllare le caratteristiche dei dati generati, rendendo le GAN strumenti versatili e applicabili in molteplici contesti, in particolare in quello biomedico. In Figura 1.6 è riassunto lo schema di funzionamento di una GAN



**Figura 1.6:** Schematizzazione del funzionamento di una GAN [18]

### 1.4.2 Pix2Pix

La variante Pix2Pix è introdotta da Isola e colleghi nel 2017, appartiene alla famiglia delle Conditional GAN ed è utilizzata per il task di image-to-image translation, ovvero la mappatura supervisionata tra un’immagine in input e la corrispondente immagine in output. In questa architettura il Generatore riceve in input una funzione condizionata  $G : x \rightarrow y$ , dove  $x$  è l’immagine in input e  $y$  l’immagine target.

L’addestramento è supervisionato, viene utilizzato un dataset allineato di coppie di immagini  $(x, y)$ , consente al modello di apprendere le relazioni strutturate tra domini visivi. L’architettura Pix2Pix è composta da un Generatore di tipo U-Net, una rete di tipo encoder-decoder, con la presenza di skip connections tra i layer simmetrici di encoding e decoding. Queste connessioni permettono il trasferimento di informazioni a bassa frequenza e non consentono la perdita di informazioni durante le operazioni di downsampling e rendendo più facile la ricostruzione dell’immagine in output. Il Discriminatore è una rete PatchGAN, va ad analizzare la verosimiglianza su patch di dimensioni inferiori rispetto a quella dell’immagine intera. Questo permette di analizzare la coerenza delle texture locali nelle immagini e la presenza

dei dettagli strutturali. La funzione obiettivo della Pix2Pix estende quella della GAN tradizionale, perché combina la loss avversaria condizionata con una loss  $L_1$  tra immagine generata e immagine target, come mostrato in Eq. (1.6), misura la distanza media assoluta pixel per pixel tra l'immagine generata e quella reale. Il termine  $L_1$  è pesato da un iperparametro  $\lambda$ , che regola il bilanciamento tra realismo visivo e accuratezza strutturale. La funzione complessiva è così definita:

$$G^* = \arg \min_G \max_D \mathbb{E}_y[\log D(y)] + \mathbb{E}_{x,z}[\log(1 - D(G(x, z)))] + \lambda \mathbb{E}_{x,y,z} [\|y - G(x, z)\|_1] \quad (1.6)$$

Per l'ottimizzazione dei pesi viene utilizzato l'algoritmo Adaptive Moment Estimation (ADAM), che calcola una media mobile esponenziale dei gradienti e dei gradienti al quadrato, adattando i tassi di apprendimento per ciascun peso della rete [19]. Gli aggiornamenti dei pesi seguono le seguenti Eq (1.7).

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\ \hat{m}_t &= \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \\ \theta_{t+1} &= \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \end{aligned} \quad (1.7)$$

Dove  $\theta_t$  rappresenta i pesi della rete al passo  $t$ ,  $g_t$  il gradiente,  $m_t$  e  $v_t$  le medie mobili del primo e secondo momento,  $\hat{m}_t$  e  $\hat{v}_t$  le correzioni del bias,  $\eta$  il learning rate e  $\epsilon$  un piccolo valore positivo per garantire stabilità numerica.

I principali parametri di Adam sono:  $\eta$  (tipicamente 0.001),  $\beta_1$  (0.9),  $\beta_2$  (0.999) e  $\epsilon$  ( $10^{-8}$ ) [20]. Questo algoritmo di ottimizzazione è molto utilizzato per la sua efficienza computazionale, capacità di convergere rapidamente a soluzioni ottimali anche per problemi più complessi.

In Figura 1.7 è schematizzato il funzionamento e la struttura di una rete GAN Pix2Pix. In conclusione, Pix2Pix è efficace in contesti supervisionati di image-to-image traslation, permette generare immagini strutturalmente coerenti con l'input. Tuttavia, la necessità di avere un dataset annotato costituisce una potenziale limitazione, soprattutto in alcuni contesti come quello medico in cui i dati sono potrebbero essere difficili da ottenere [21].

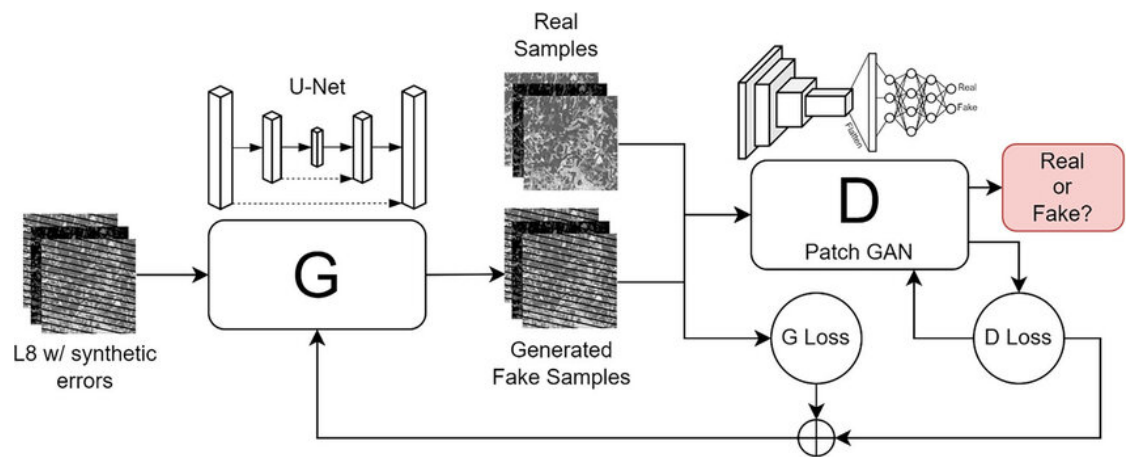


Figura 1.7: Schema pix2pix [22]

## Capitolo 2

# Materiali e metodi

In questo capitolo sarà presentato il dataset utilizzato per lo studio, descritto il preprocessing applicato alle immagini per prepararle alla fase di training, illustrata l'architettura della GAN utilizzata con i parametri di configurazione e, infine, riportate le metriche di valutazione scelte per la valutazione delle immagini generate.

### 2.1 Dataset

Il Dataset utilizzato in questo studio è stato fornito dal Dipartimento di Oftmologia, dell'ospedale Città della Salute e della Scienza di Torino, comprende esami OCT relativi a pazienti affetti da AMD in diversi stadi della patologia.

Il campione include soggetti con un'età compresa tra i 70 e 85 anni, le acquisizioni sono state eseguite nel periodo 2023 e 2024, per un totale di 739 esami.

La distribuzione dei pazienti per ogni classe in esame è la seguente:

- 163 Pazienti sani
- 272 Pazienti stadio intermedio
- 154 Pazienti stadio neovascolare
- 150 Pazienti stadio atrofia geografica

Le immagini sono state classificate dai clinici che hanno eseguito una precedente valutazione clinica del paziente.

Le acquisizioni sono state effettuate utilizzando il dispositivo Heidelberg Spectralis HRA+OCT device (Heidelberg Engineering, Heidelberg, Germany).

Ogni esame è composto da 49 immagini per paziente, corrispondenti a diverse sezioni tomografiche trasversali della retina per un totale complessivo di 36.211 immagini. Le immagini sono fornite in formato .png, con una risoluzione di 1008 x

496 pixel e codifica a colori RGB.

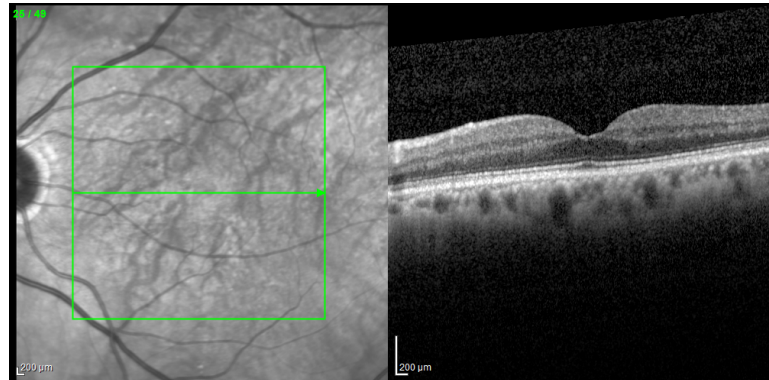
Il dataset è stato suddiviso in training set e test set con una proporzione 80-20; la suddivisione è stata effettuata prendendo in considerazione il paziente e non le singole immagini in modo tale che tutte le immagini relative a un paziente fossero assegnate esclusivamente a uno dei due insiemi. La divisione finale risultante è mostrata in Tabella 2.1:

| Categoria          | Training set | Test set |
|--------------------|--------------|----------|
| Intermedio         | 218          | 54       |
| Neovascolare       | 123          | 31       |
| Geographic Atrophy | 120          | 30       |
| Sani               | 130          | 33       |

**Tabella 2.1:** Distribuzione dei pazienti nei set di training e test per ciascuna categoria diagnostica.

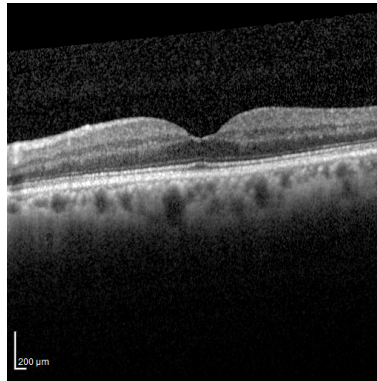
## 2.2 Preprocessing

Le immagini originali fornite sono come rappresentate in Figura 2.1.



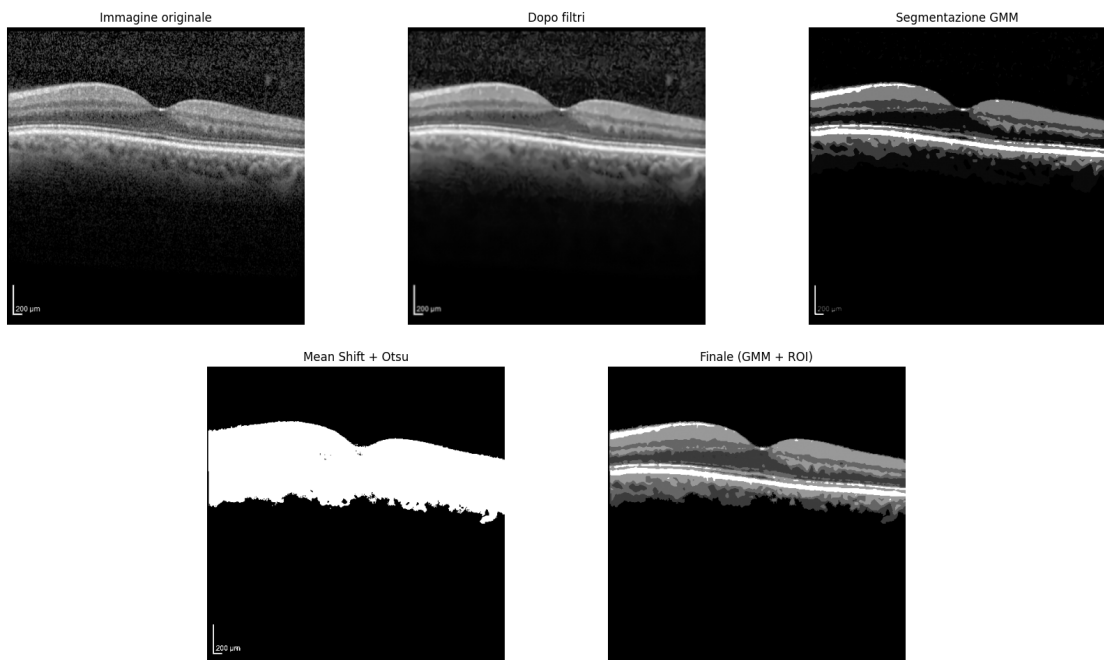
**Figura 2.1:** Immagine originale soggetto sano (HEA\_ID300 slice 24)

La regione di interesse per lo studio corrisponde alla parte destra delle immagini; quindi l'immagine originale è stata ritagliata a dimensioni di 512x496 pixel. Successivamente, è stato applicato un padding per ottenere un'immagine finale di dimensioni 512x512 pixel, come illustrato in Figura 2.2.



**Figura 2.2:** Immagine processata di un soggetto sano, ritagliata e con padding (HEA\_ID300 slice 24)

A partire dalle immagini così ottenute, l'obiettivo è l'elaborazione delle immagini per ottenere una mappa semantica necessaria per il training della GAN nella fase successiva. A tal fine, sono stati implementati i seguenti step di preprocessing come mostrato in Figura 2.3.



**Figura 2.3:** Step di preprocessing applicati

Alle immagini è stato applicata una sequenza di filtri per ridurre il rumore presente e preservare le strutture anatomiche di interesse:

- Filtro gaussiano con un  $\sigma = 0.5$ , per attenuare il rumore ad alta frequenza e le piccole fluttuazioni di intensità, migliorando l'omogeneità delle aree uniformi.
- Algoritmo di Fast Non-Local Means Denoising (Fast NLM), il quale riduce la presenza di rumore preservando i dettagli delle strutture presenti. L'algoritmo sostituisce il valore di ciascun pixel con la media pesata dei valori di pixel di patch simili, individuati nell'area di ricerca, questo permette di individuare la similarità strutturale e mantenere le caratteristiche salienti dell'immagine.

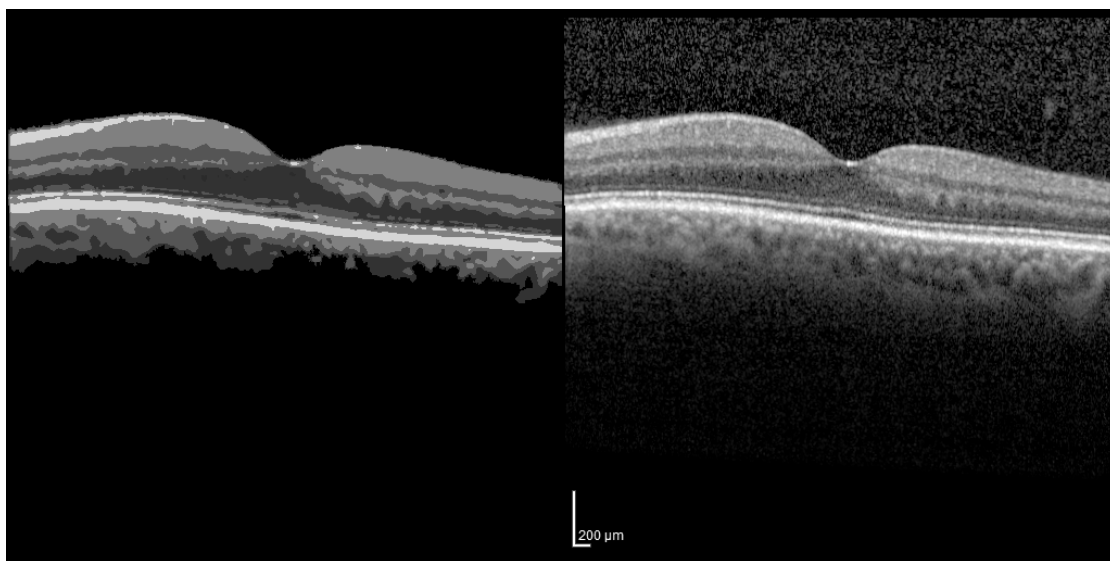
L'immagine così ottenuta è stata segmentata con due differenti procedure, eseguite parallelamente:

- Gaussian Mixture Model (GMM): questo algoritmo modella la distribuzione dei livelli di intensità dell'immagine come combinazione pesata di gaussiane multivariate. Per motivi computazionali, il modello GMM è stato allenato globalmente considerando tutte le 49 immagini relative al paziente. Per ciascuna immagine, ogni pixel viene assegnato alla componente gaussiana con la massima probabilità a posteriori. La segmentazione risultante viene convertita in una mappa a livello di grigi discreti, con valori di intensità di grigio definiti al fine di ottenere una codifica coerente tra le diverse slice del soggetto. I livelli di grigio assegnati ai cluster sono: 43, 50, 85, 128, 170, 213 e 255 corrispondenti ai sette cluster identificati.
- Mean Shift + Otsu, questa procedura permette di ottenere una segmentazione preliminare utile per stimare la Regione di Interesse (ROI), corrispondente alla retina. L'algoritmo di Mean Shift Filtering è un metodo di clustering non parametrico basato sulla densità nello spazio delle feature, che raggruppa i pixel con caratteristiche simili in regioni a intensità quasi uniforme, considerando sia la posizione spaziale che le informazioni cromatiche. L'immagine risultante viene convertita in scala di grigi per l'applicare dell'algoritmo di Otsu, che determina automaticamente la soglia ottimale per la segmentazione binaria massimizzando la varianza interclasse. Il risultato finale è una maschera binaria in cui la retina, la ROI, viene correttamente separata dallo sfondo.

Infine la segmentazione ottenuta mediante il GMM è stata combinata con quella prodotta dalla procedura Mean Shift+Otsu al fine di ottenere la segmentazione finale. Su quest'ultima è stato applicato un post processing con l'obiettivo di rimuovere elementi non appartenenti alla ROI; è stata individuata la regione con l'area massima all'interno dell'immagine e successivamente mediante un algoritmo di distance transform sono state eliminate le aree spurie, assegnandole al valore di sfondo. Il risultato della segmentazione finale è illustrato in Figura 2.3.

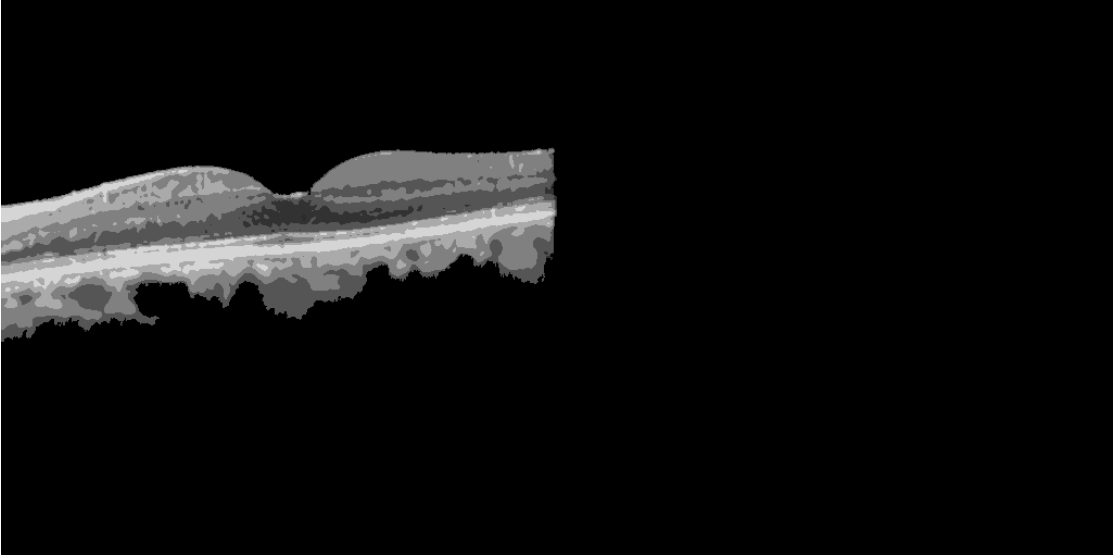
Successivamente, le immagini del dataset sono state preparate per il training della GAN, la quale richiede in input coppie di immagini in cui: nella parte sinistra

è presente la mappa semantica ottenuta con il procedimento precedentemente descritto; nella parte destra è collocata la corrispondente immagine originale. In questo modo si ottengono coppie di input strutturate come mostrato Figura 2.4.



**Figura 2.4:** Esempio di immagine accoppiata fornita in input alla GAN (HEA\_ID288 slice 25)

Per la fase di inference, le immagini appartenenti al test set sono state preparate accoppiando la mappa semantica (a sinistra) con un riquadro nero di uguali dimensioni (a destra), come mostrato in Figura 2.5.



**Figura 2.5:** Esempio di immagine accoppiata fornita in input alla GAN durante la fase di inference (HEA\_ID300 slice 25)

## 2.3 Parametri della GAN

In questo lavoro è stata utilizzata una cGAN basata su un'architettura Pix2Pix, con delle modifiche alla repository ufficiale [21]. Inoltre la rete adotta una struttura piramidale.

Il generatore è utilizza una rete U-Net a  $256 \times 256$ , con un percorso encoding-decoding con skip connection e il discriminatore adotta una struttura PatchGAN con patch di dimensioni  $70 \times 70$ . Le immagini in input e output hanno un singolo canale e la loro dimensione è  $512 \times 512$ .

La funzione di perdita combina una loss avversaria (LSGAN), una loss percettiva basata su feature VGG e loss multilivello su diverse mappe di feature ( $L_2, L_3, L_4$ ). Inoltre sono utilizzati termini di regolarizzazione derivanti da convoluzioni e operatori di Sobel. Ciascun termine è ponderato da specifici iperparametri per ottimizzare il bilanciamento tra realismo, fedeltà strutturale e qualità dei dettagli locali.

Per l' addestramento è stato impostata la normalizzazione dei layer convoluzionali di tipo batch normalization, con batch size pari a 8, i pesi della rete sono inizializzati secondo una distribuzione normale e un numero di epoche pari a 100 con learning rate costante, seguite da 100 epoche con decrescita lineare di learning rate.

Per l'ottimizzazione dei pesi è utilizzato l'algoritmo Adam, con learning rate iniziale pari a 0.0002 e  $\beta = 0.5$ .

Per la fase di test, il modello è stato utilizzato in modalità eval, caricando le immagini da un unico dominio, generando immagini in una sola direzione. Il generatore

utilizzato è lo stesso descritto nella fase di training senza ulteriori modifiche. Le immagini in input e output hanno un singolo canale e dimensione  $512 \times 512$ , il batch size posto pari a 1 e il numero di thread è impostato a 0 per garantire la riproducibilità.

Durante l'inferenza non è eseguita nessuna ottimizzazione dei pesi e non è calcolata la funzione di perdita.

La Tabella 2.2 riassume i parametri descritti in questa sezione.

| Componente            | Parametro / Configurazione                                                                                                       |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Generatore            | U-Net a $256 \times 256$                                                                                                         |
| Discriminatore        | PatchGAN, patch size $70 \times 70$                                                                                              |
| Input/Output          | 1 canale, dimensione $512 \times 512$                                                                                            |
| Batch size            | 8                                                                                                                                |
| Funzione di perdita   | LSGAN + loss percettiva VGG + loss multi-livello ( $L_2$ , $L_3$ , $L_4$ )<br>+ termini di regolarizzazione convoluzione e Sobel |
| Normalizzazione       | Batch Normalization                                                                                                              |
| Inizializzazione pesi | Distribuzione normale                                                                                                            |
| Ottimizzatore         | Adam, learning rate 0.0002, $\beta = 0.5$                                                                                        |
| Numero epoche         | 200 (100 con LR costante + 100 con LR lineare decrescente)                                                                       |

**Tabella 2.2:** Parametri principali della cGAN Pix2Pix.

## 2.4 Metriche di valutazione

La valutazione delle immagini generate dalla GAN sono state effettuate calcolando le seguenti metriche:

Peak Signal-to-Noise Ratio (PSNR), calcola la differenza di intensità tra l'immagine originale  $I$  e l'immagine generata  $\hat{I}$  ed è definita come in Eq. (2.1):

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{L^2}{\text{MSE}} \right) \quad (2.1)$$

dove  $L$  è il valore massimo dei pixel e MSE, definito come nell'Eq (2.2) è l'errore quadratico medio tra le due immagini

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I(i, j) - \hat{I}(i, j))^2 \quad (2.2)$$

Tipicamente questo indice assume valori che vanno da 20 dB (per immagini di bassa qualità) a oltre i 40 dB, più alto è il suo valore meno distorsioni sono presenti nelle immagini generate [23].

Structural Similar Index (SSIM), definita nell'Eq (2.3), permette di valutare la somiglianza strutturale tra l'immagine originale e l'immagine generata tramite la luminanza  $l$ , il contrasto  $c$  e la struttura  $s$  come definiti nell'Eq. (2.4) (2.5) (2.6)

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (2.3)$$

dove tipicamente  $\alpha = \beta = \gamma = 1$ , e

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (2.4)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (2.5)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (2.6)$$

$\mu_x, \mu_y$  sono le medie dei pixel,  $\sigma_x, \sigma_y$  le deviazioni standard locali,  $\sigma_{xy}$  la covarianza locale, e  $C_1, C_2, C_3$  costanti di stabilizzazione. Valori tipici: SSIM varia tra 0 e 1, dove 1 indica perfetta somiglianza[23].

Frechet Inception Distance (FID) misura la distanza tra le distribuzioni gaussiane delle caratteristiche estratte da un modello pre-addestrato (Inception) sulle immagini originali ( $X$ ) e generate ( $Y$ ) come mostrato nell'Eq (2.7):

$$\text{FID}(X, Y) = \|\mu_X - \mu_Y\|^2 + \text{Tr}(\Sigma_X + \Sigma_Y - 2(\Sigma_X \Sigma_Y)^{1/2}) \quad (2.7)$$

dove  $\mu$  e  $\Sigma$  rappresentano la media e la matrice di covarianza delle features [23].

Contrast Improvement Index (CII) permette di valutare il miglioramento del contrasto tra le immagini generate e le immagini originali, il calcolo viene effettuato in maniera globale utilizzando un kernel quadrato di dimensioni  $k \times k$  ed è stata considerata l'intera immagine, come mostrato nell'Eq. (2.8)[24]

$$\text{CII} = \frac{C_{\text{proposed}}}{C_{\text{original}}} \quad (2.8)$$

Feature di texture del primo ordine [25] sono operatori statistici che descrivono proprietà statistiche globali della distribuzione dei pixel non tenendo conto delle relazioni spaziali tra pixel. Definita la distribuzione normalizzata dei livelli di

intensità  $P(i)$ , come mostrato nell'Eq (2.9), determina la probabilità di occorrenza del livello di intensità  $i$  dell'immagine, mentre  $N$  rappresenta il numero di pixel totali e  $n_i$  il numero di pixel con valore  $i$

$$P(i) = \frac{n_i}{N} \quad (2.9)$$

$P(i)$  è utilizzata per il calcolo delle seguenti metriche:

- Mean:  $\sum_i i P(i)$   
fornisce informazioni sul livello medio di grigi nella regione di interesse.
- Mean Absolute Deviation (MAD):  $\sum_i |i - \mu| P(i)$   
misura la variabilità media dei pixel rispetto alla media, valori alti indicano una maggiore dispersione dei livelli di grigio e di conseguenza maggiore eterogeneità della texture.
- Root Mean Squared (RMS):  $\sqrt{\sum_i i^2 P(i)}$   
misura l'energia complessiva dell'immagine, valori elevati indicano una maggiore intensità media dei pixel.
- Kurtosis:  $\sum_i (i - \mu)^4 P(i) / \sigma^4$   
misura la distribuzione dei valori intorno alla media, valori elevati indicano una maggiore probabilità di valori estremi, mentre valori bassi indicano una distribuzione più uniforme attorno alla media.
- Uniformity:  $\sum_i P(i)^2$   
valuta quanto la distribuzione dei pixel è uniforme, valori elevati indicano immagini omogenee e regolari, valori bassi indicano texture complesse o irregolari.

dove  $\mu$  e  $\sigma$  rappresentano rispettivamente la media e la deviazione standard dei valori di intensità.

Feature di texture del secondo ordine [26] considerano le relazioni spaziali tra pixel vicini e vengono calcolate lungo determinate direzioni, si basano sulla matrice di co-occorrenza dei livelli di grigio (GLCM). Queste feature permettono di analizzare la texture locale ottenendo informazioni sulla disposizione spaziale dei pixel, variabilità delle strutture, pattern ripetitivi. Definita la matrice  $P(i, j)$ , rappresenta la probabilità congiunta di occorrenza dei livelli di intensità  $i$  e  $j$ , sono state calcolate le seguenti metriche:

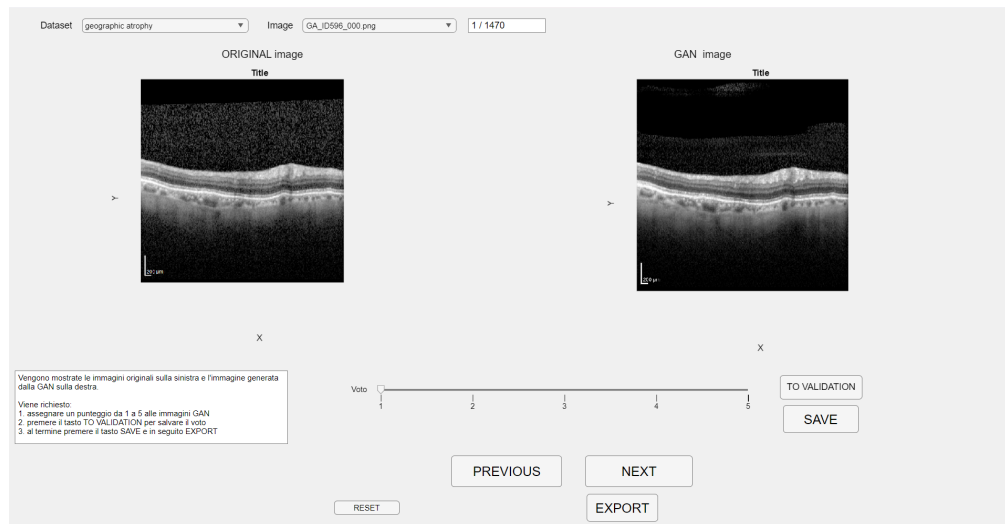
- Contrast:  $\sum_{i,j} (i - j)^2 P(i, j)$   
misura la variazione di intensità tra pixel vicini, valori alti indicano texture con forti contrasti locali, valori bassi indicano regioni più uniformi

- Correlation:  $\frac{\sum_{i,j} (i-\mu_x)(j-\mu_y)P(i,j)}{\sigma_x\sigma_y}$   
valuta la correlazione lineare tra valori di pixel vicini
- Autocorrelation:  $\sum_{i,j} i j P(i,j)$   
indica la somiglianza tra intensità di pixel vicini, valori alti indicano regioni con pattern ripetuti
- Difference Average:  $\sum_k k p_{x-y}(k)$   
media delle differenze tra coppie di pixel adiacenti.
- Difference Entropy:  $-\sum_k p_{x-y}(k) \log p_{x-y}(k)$   
misura la complessità delle differenze tra pixel vicini, valori alti indicano texture più irregolare e complessa
- Difference Variance:  $\sum_k (k-)^2 p_{x-y}(k)$   
valuta la dispersione delle differenze tra pixel vicini
- Joint Energy:  $\sum_{i,j} P(i,j)^2$   
analogia all'energia di primo ordine
- Joint Entropy:  $-\sum_{i,j} P(i,j) \log P(i,j)$   
misura la complessità dei pattern locali
- Maximum Correlation Coefficient (MCC):  $\sqrt{1 - \lambda_2}$   
misura la correlazione massima tra i pixel, legata alla struttura globale della matrice GLCM
- Sum Average:  $\sum_k k p_{x+y}(k)$   
media delle somme delle intensità di coppie di pixel, fornisce informazioni sulla distribuzione locale dei livelli di grigio
- Sum Entropy:  $-\sum_k p_{x+y}(k) \log p_{x+y}(k)$   
misura la complessità della somma dei livelli di grigio vicini, valori alti indicano texture complesse e variabili.

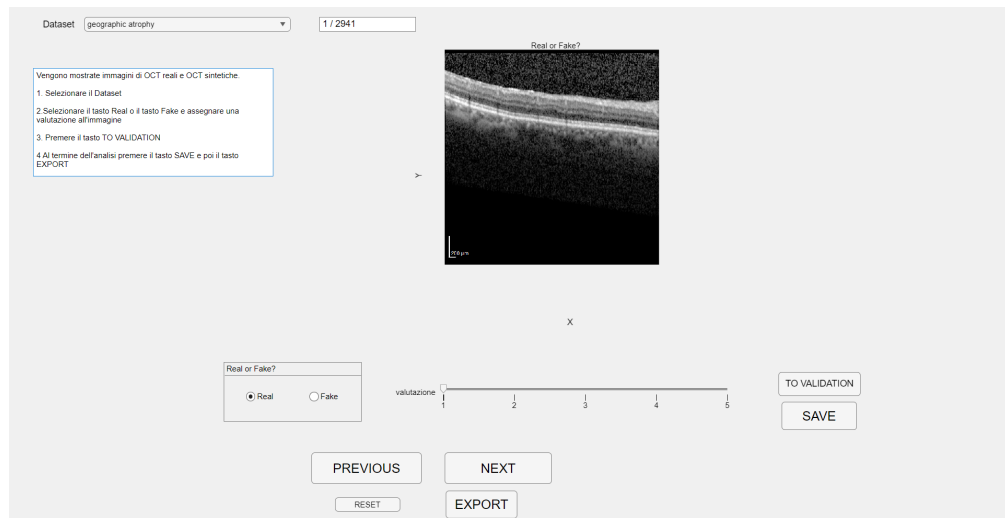
La valutazione qualitativa eseguita dai clinici è stata effettuata tramite due interfacce grafiche (GUI).

Nella prima GUI, in Figura 2.6, viene presentata all'utente, sulla sinistra, l'immagine originale e, sulla destra, la corrispondente immagine sintetica. All'utente viene richiesto l'assegnazione di un punteggio di qualità compreso tra 1 e 5, in base alla somiglianza visiva e alla percezione generale dell'immagine. Mediante i menù a tendina posti nella parte superiore dell'interfaccia, è possibile selezionare lo stadio di valutazione e scegliere e un immagine appartenente a quello stadio. Tutte le

valutazioni vengono automaticamente salvate in un file Excel. Nella seconda GUI, in Figura 2.7. All'utente viene chiesto di identificare l'immagine mostrata come reale (immagini originali) oppure fake (immagine è generata dalla GAN). In seguito, viene richiesto di assegnare un punteggio di confidenza compreso tra 1 e 5, utilizzato come misura della sicurezza soggettiva del valutatore nella risposta fornita. Tutte le risposte vengono automaticamente salvate in un file Excel. Il processo viene ripetuto per tutti gli stadi della patologia andando a selezionare, utilizzando la tendina presente in alta sinistra lo stadio soggetto di valutazione.



**Figura 2.6:** Interfaccia grafica valutazione clinica: valutazione



**Figura 2.7:** Interfaccia grafica per la valutazione clinica originale Vs immagine sintetica

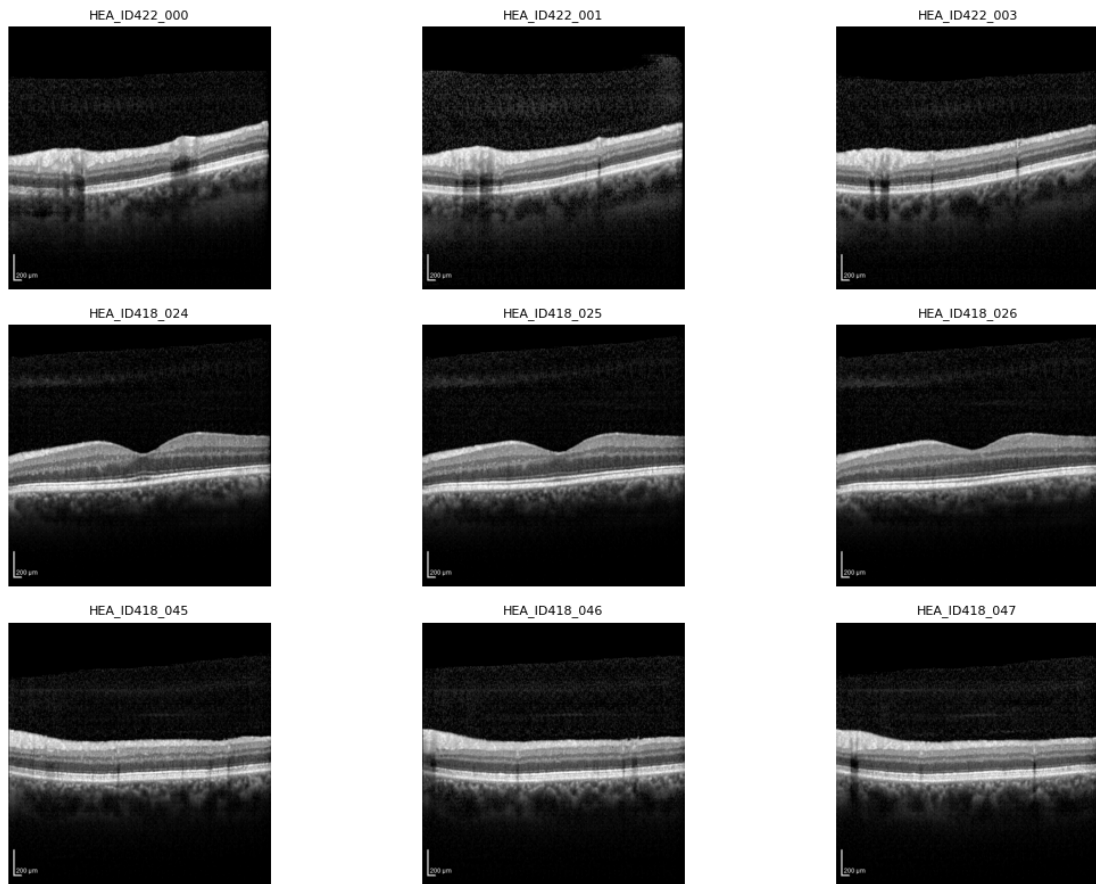
## Capitolo 3

# Risultati

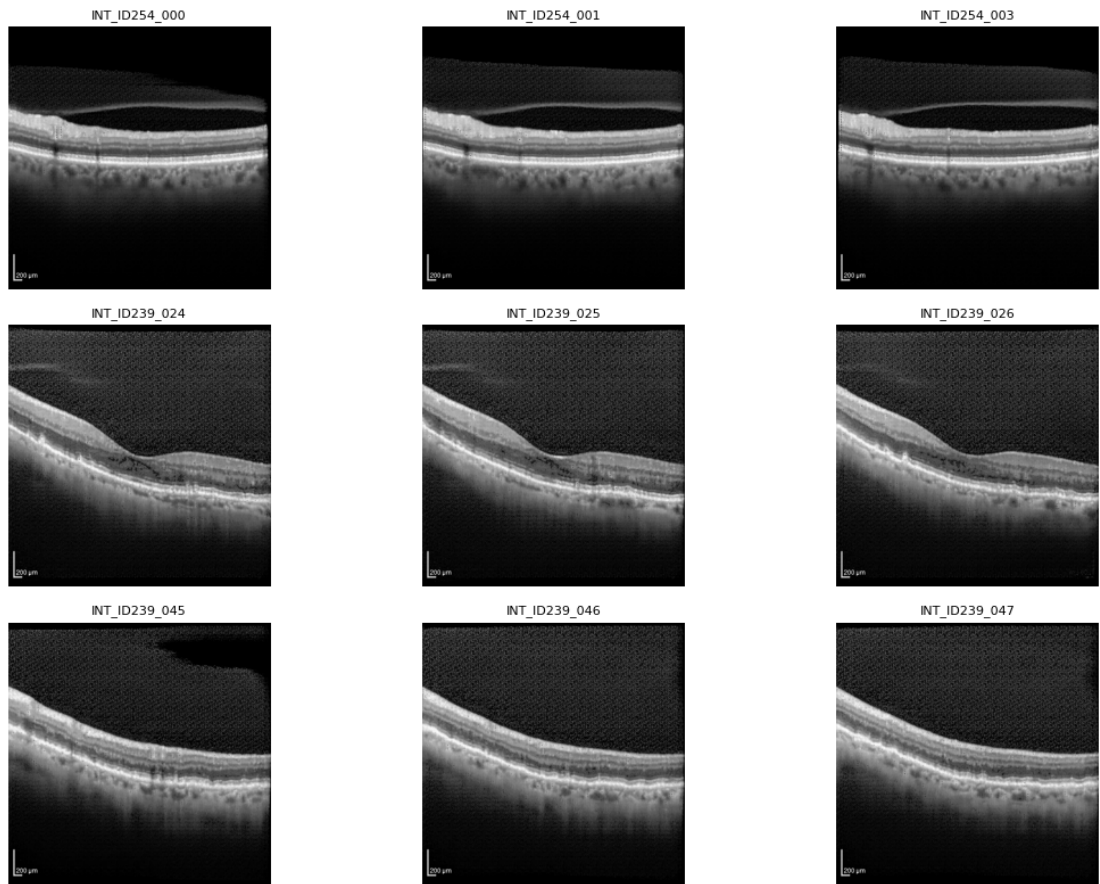
In questo capitolo verranno mostrate alcune immagini prodotte dalla GAN e i risultati delle analisi qualitative e quantitative delle metriche precedentemente illustrate. Questa presentazione permette di supportare successivamente l'interpretazione e la discussione dei risultati.

### 3.1 Risultati GAN

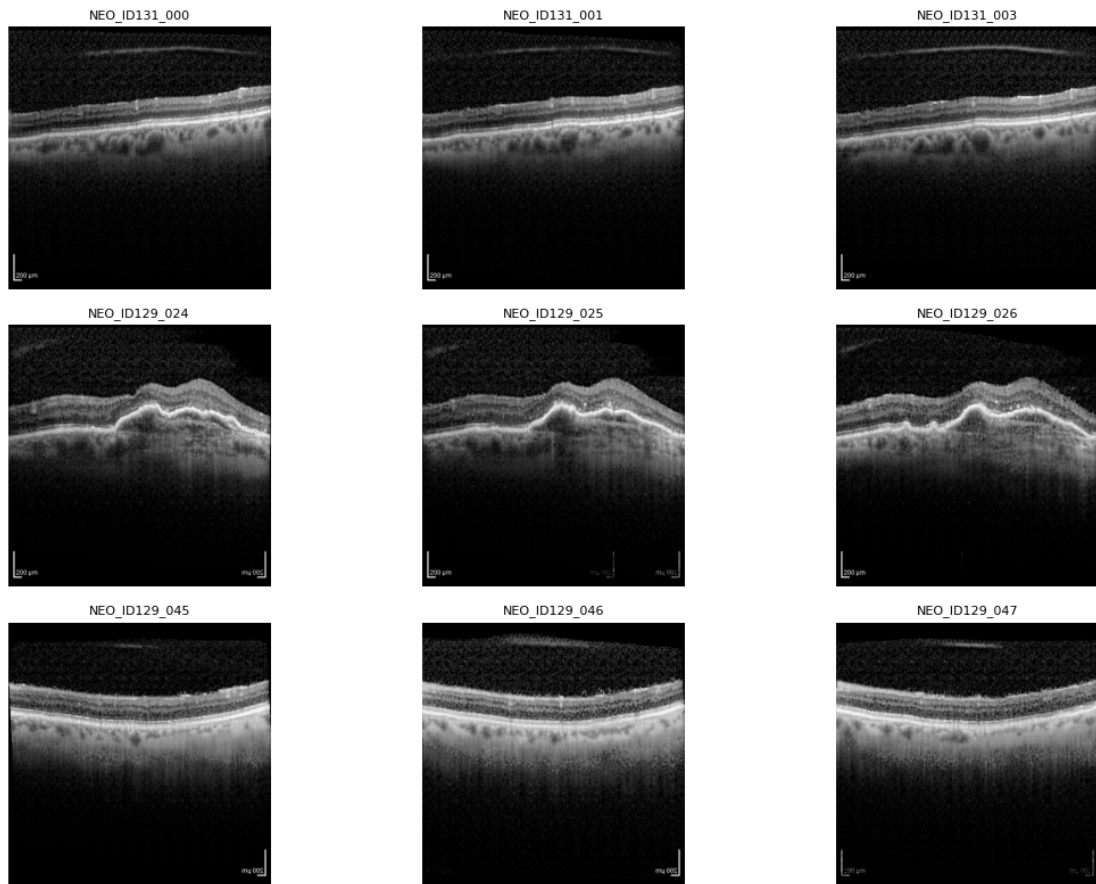
In Figura 3.1, 3.2, 3.3, 3.4 sono riportate le immagini sintetiche ottenute per i diversi stadi patologici: sano, intermedio, neovascolare e atrofia geografica. Per ciascuno stadio sono mostrate le immagini corrispondenti alle slice numero 0, 1, 3, 24, 25, 26, 45, 46, 47.



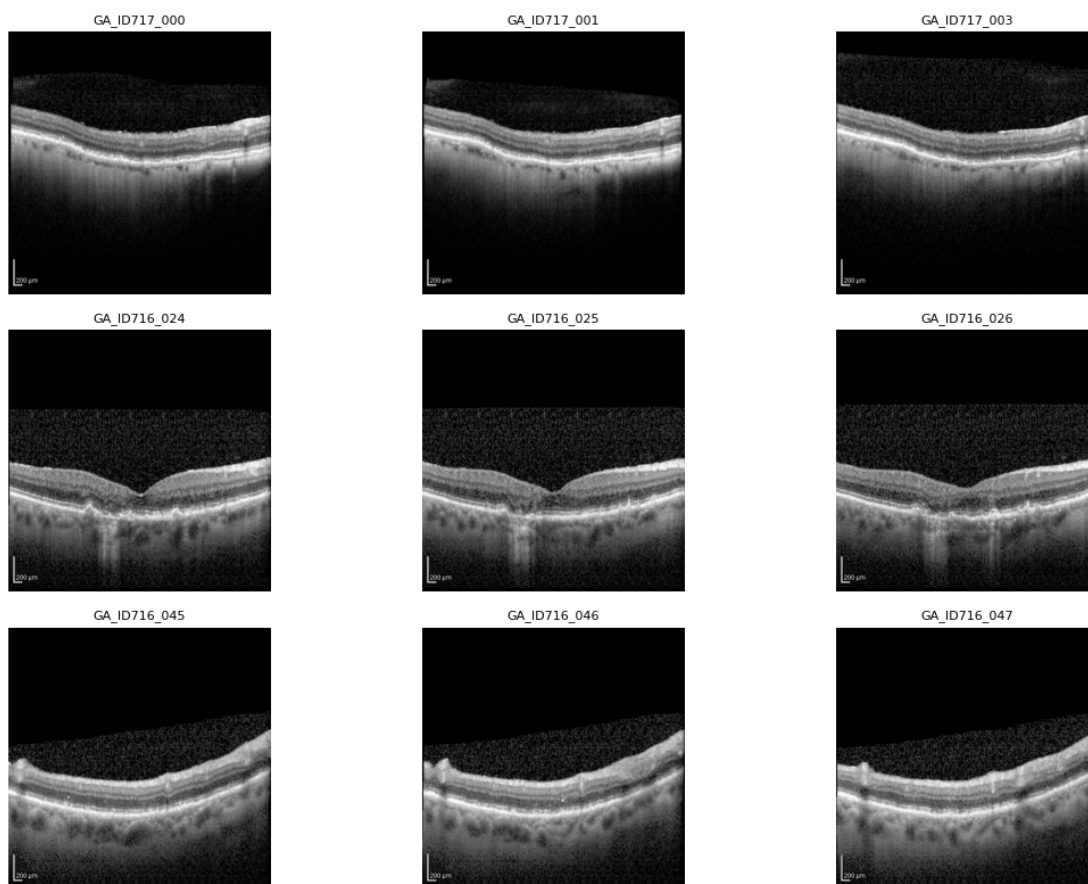
**Figura 3.1:** Risultati GAN stadio sano



**Figura 3.2:** Risultati GAN stadio intermedio



**Figura 3.3:** Risultati GAN stadio neovascolare



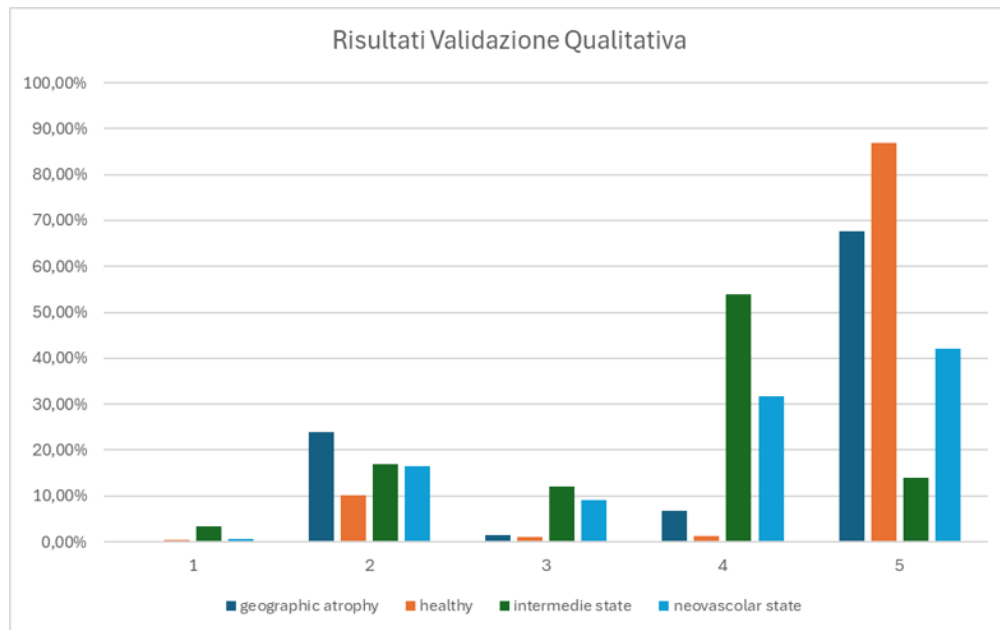
**Figura 3.4:** Risultati GAN stadio atrofia geografica

## 3.2 Analisi qualitativa

Nella Tabella 3.1 sono riportati i risultati ottenuti dalla valutazione delle immagini sintetiche tramite la prima GUI, mentre in Figura 3.5 è rappresentato l'istogramma dei risultati.

| Stadio             | 1     | 2      | 3      | 4      | 5      |
|--------------------|-------|--------|--------|--------|--------|
| Sani               | 0.07% | 10.10% | 1.14%  | 1.33%  | 86.92% |
| Intermedio         | 3.28% | 16.96% | 11.99% | 53.86% | 13.90% |
| Neovascolare       | 0.68% | 16.51% | 9.00%  | 31.79% | 42.02% |
| Atrofia geografica | 0.07% | 23.87% | 1.58%  | 6.81%  | 67.68% |

**Tabella 3.1:** Distribuzione delle valutazioni da 1 a 5 per i diversi stadi



**Figura 3.5:** Istogramma della validazione della prima GUI

Di seguito sono riportati i risultati ottenuti dalla valutazione delle immagini mediante la seconda GUI. In Tabella 3.2 sono presentati i risultati relativi alle immagini originali, mentre in Tabella 3.3 sono riportate le analisi ottenute dalle immagini sintetiche.

| Tipo immagine | Risultati [%] | Val. 1 | Val. 2 | Val. 3 | Val. 4 | Val. 5 |
|---------------|---------------|--------|--------|--------|--------|--------|
| <b>Fake</b>   | 33%           | 0%     | 0%     | 13%    | 9%     | 78%    |
| <b>Real</b>   | 67%           | 0%     | 0%     | 8%     | 2%     | 91%    |

**Tabella 3.2:** Risultati delle immagini reali ottenuti dalla valutazione soggettiva delle immagini reali e sintetiche.

| Tipo immagine | Risultati [%] | Val. 1 | Val. 2 | Val. 3 | Val. 4 | Val. 5 |
|---------------|---------------|--------|--------|--------|--------|--------|
| <b>Fake</b>   | 70%           | 0%     | 6%     | 11%    | 8%     | 75%    |
| <b>Real</b>   | 30%           | 0%     | 9%     | 9%     | 4%     | 78%    |

**Tabella 3.3:** Risultati delle immagini sintetiche ottenuti dalla valutazione soggettiva delle immagini reali e sintetiche.

### 3.3 Analisi quantitativa

La Tabella 3.4 riporta i risultati ottenuti dall'analisi dell'indice CII,SSIM,PSNR e FID.

Per ciascuno stadio patologico sono riportati il valore medio e la deviazione standard per gli indici PSNR, SSIM. In particolare, l'indice CII è calcolato considerando tutte le immagini appartenenti al medesimo gruppo, mentre gli indici SSIM e PSNR sono ottenuti confrontando ciascuna immagine generata dalla GAN con la corrispondente immagine reale. Per il FID viene fornito esclusivamente il valor medio.

| Stadio                    | CII             | FID   | SSIM            | PSNR             |
|---------------------------|-----------------|-------|-----------------|------------------|
| <b>Sani</b>               | $1.59 \pm 0.27$ | 28.59 | $0.46 \pm 0.10$ | $23.19 \pm 2.73$ |
| <b>Intermedio</b>         | $1.65 \pm 0.24$ | 55.89 | $0.44 \pm 0.09$ | $22.13 \pm 2.39$ |
| <b>Neovascolare</b>       | $1.66 \pm 0.25$ | 43.12 | $0.44 \pm 0.09$ | $22.62 \pm 2.33$ |
| <b>Atrofia Geografica</b> | $1.69 \pm 0.22$ | 29.08 | $0.43 \pm 0.08$ | $22.36 \pm 2.14$ |

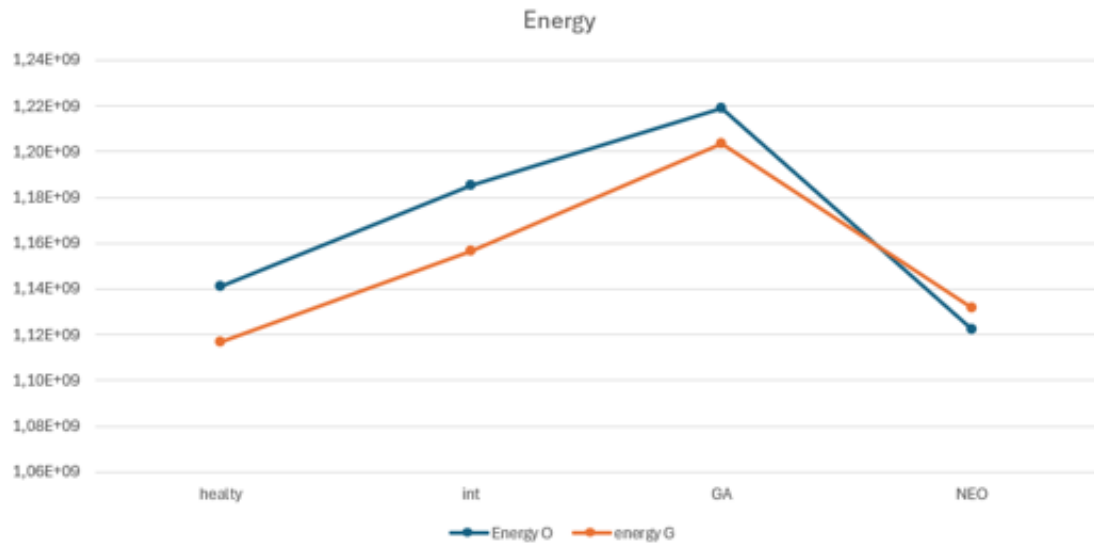
**Tabella 3.4:** Valori medi  $\pm$  deviazione standard delle metriche CII, SSIM e PSNR (solo media per FID) per ciascuno stadio patologico.

La Tabella 3.5 riporta gli errori relativi ottenuti dalle feature di texture del primo ordine, indicando per ciascuna feature il valore medio e la deviazione standard calcolati per ogni stadio patologico.

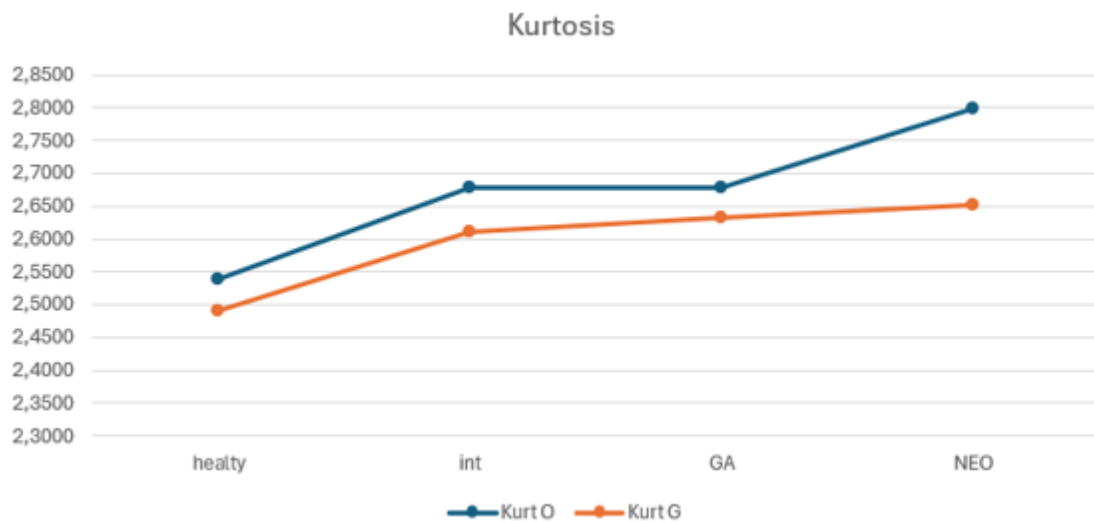
| Stadio                    | Energy            | MAD               | RMS               | Kurtosis          | Uniformity        |
|---------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| <b>Sani</b>               | $0.033 \pm 0.029$ | $0.026 \pm 0.025$ | $0.016 \pm 0.014$ | $0.040 \pm 0.037$ | $0.068 \pm 0.556$ |
| <b>Intermedio</b>         | $0.057 \pm 0.035$ | $0.031 \pm 0.025$ | $0.029 \pm 0.018$ | $0.053 \pm 0.048$ | $0.003 \pm 0.037$ |
| <b>Atrofia Geografica</b> | $0.038 \pm 0.032$ | $0.026 \pm 0.021$ | $0.019 \pm 0.016$ | $0.041 \pm 0.034$ | $0.003 \pm 0.022$ |
| <b>Neovascolare</b>       | $0.004 \pm 0.031$ | $0.033 \pm 0.024$ | $0.020 \pm 0.015$ | $0.060 \pm 0.049$ | $0.038 \pm 0.430$ |

**Tabella 3.5:** Valori medi  $\pm$  deviazione standard delle feature di texture di primo ordine per ciascuno stadio patologico.

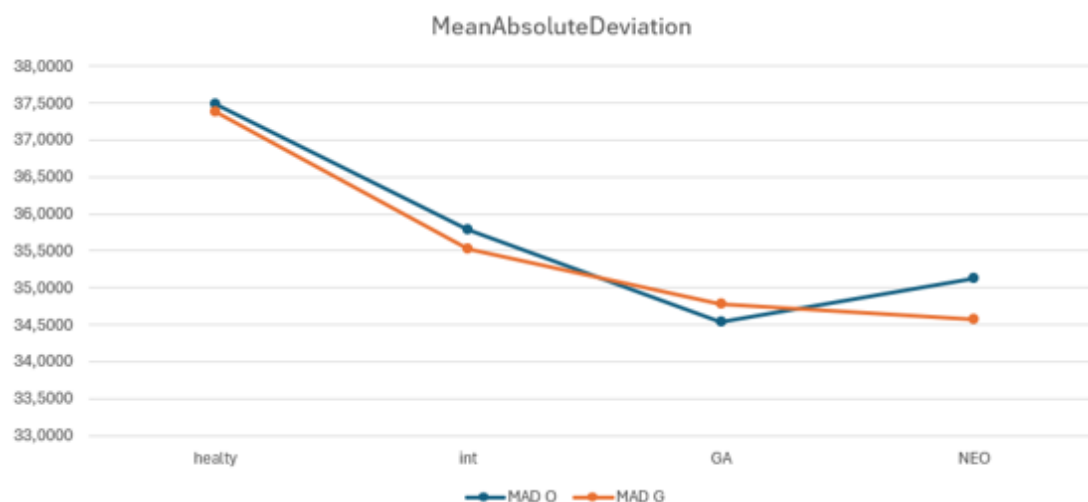
Di seguito in Figura 3.6, 3.7, 3.8, 3.9 sono riportati rispettivamente i grafici di Energia, Kurtosis, Mean Absolute Deviation, Uniformity. Tali grafici mostrano l'andamento delle feature di texture di primo ordine, messe a confronto con i valori corrispondenti calcolati sulle immagini originali



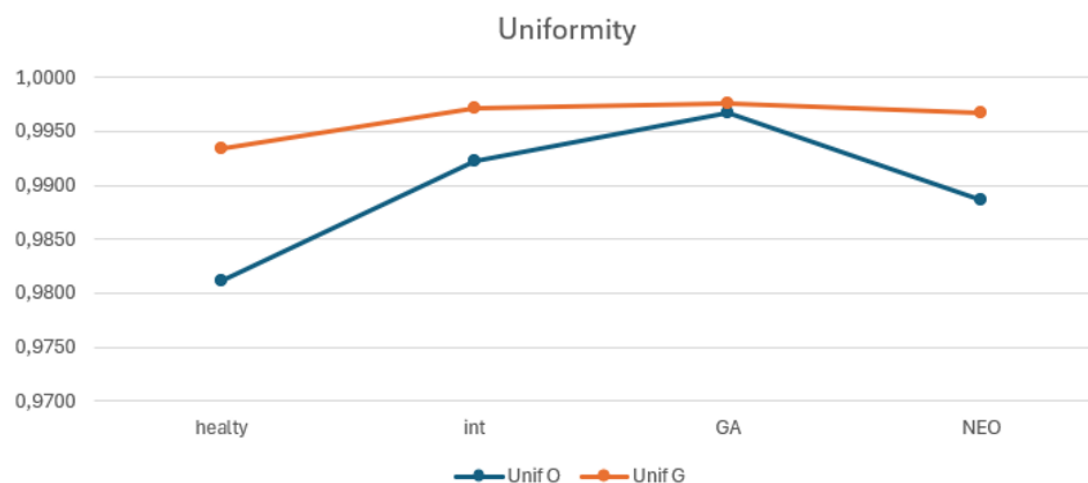
**Figura 3.6:** Andamento delle feature di texture del primo ordine: Energia



**Figura 3.7:** Andamento delle feature di texture del primo ordine: Kurtosis



**Figura 3.8:** Andamento delle feature di texture del primo ordine: Mean Absolute Deviation



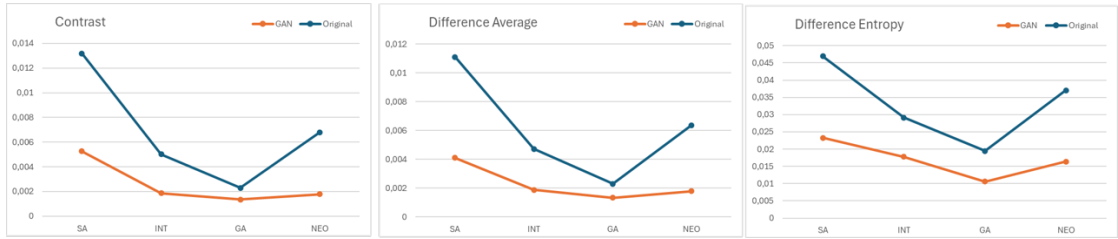
**Figura 3.9:** Andamento delle feature di texture del primo ordine: Uniformity

La Tabella 3.6 riporta gli errori relativi ottenuti dalle feature di texture del secondo ordine indicando per ciascuna feature il valore medio e la deviazione standard calcolati per ogni stadio patologico.

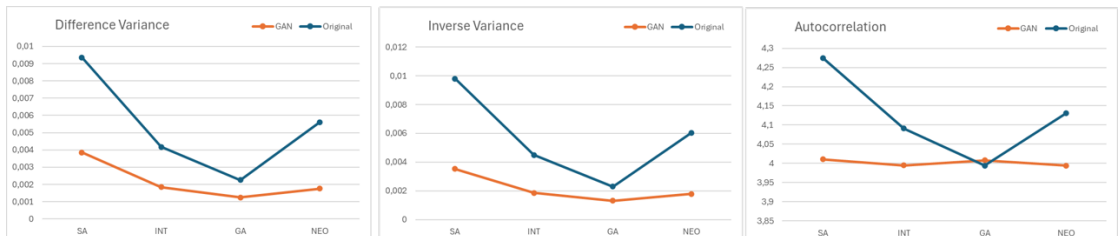
| Feature             | Sani             | Intermedio       | Atrofia Geografica | Neovascolare     |
|---------------------|------------------|------------------|--------------------|------------------|
| Contrast            | $1.31 \pm 2.99$  | $0.86 \pm 4.46$  | $0.86 \pm 1.46$    | $1.20 \pm 2.65$  |
| Difference average  | $1.31 \pm 2.99$  | $0.86 \pm 4.46$  | $0.86 \pm 1.46$    | $1.20 \pm 2.65$  |
| Difference Entropy  | $1.14 \pm 2.79$  | $0.77 \pm 3.75$  | $0.77 \pm 1.49$    | $1.02 \pm 2.14$  |
| Difference Variance | $1.308 \pm 2.99$ | $0.86 \pm 4.45$  | $0.86 \pm 1.46$    | $1.20 \pm 2.65$  |
| Inverse Variance    | $1.31 \pm 2.99$  | $0.86 \pm 4.46$  | $0.86 \pm 1.46$    | $1.20 \pm 2.65$  |
| Autocorrelation     | $0.01 \pm 0.09$  | $0.004 \pm 0.06$ | $0.003 \pm 0.11$   | $0.07 \pm 0.07$  |
| Correlation         | $0.48 \pm 0.81$  | $0.78 \pm 0.84$  | $0.79 \pm 1.98$    | $0.59 \pm 1.13$  |
| Joint Average       | $0.008 \pm 0.07$ | $0.001 \pm 0.04$ | $0.001 \pm 0.03$   | $0.004 \pm 0.05$ |
| Joint Energy        | $0.157 \pm 1.31$ | $0.004 \pm 0.75$ | $0.004 \pm 0.3$    | $0.09 \pm 1.01$  |
| Joint Entropy       | $1.208 \pm 2.9$  | $0.80 \pm 4.05$  | $0.80 \pm 1.7$     | $1.1 \pm 2.49$   |
| MCC                 | $0.491 \pm 0.83$ | $0.790 \pm 0.86$ | $0.79 \pm 1.99$    | $0.61 \pm 1.24$  |
| Sum Average         | $0.009 \pm 0.07$ | $0.001 \pm 0.04$ | $0.001 \pm 0.03$   | $0.005 \pm 0.05$ |
| Sum Entropy         | $1.23 \pm 3.12$  | $0.8 \pm 3.97$   | $0.8 \pm 1.68$     | $1.1 \pm 2.46$   |

**Tabella 3.6:** Valori medi  $\pm$  deviazione standard delle feature di texture di secondo ordine (GLCM) per ciascuno stadio patologico.

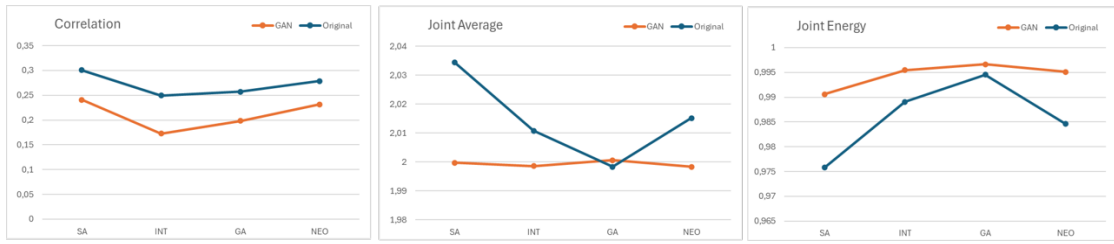
In Figura 3.10, 3.11, 3.12, 3.13 sono riportati i grafici degli andamenti delle feature di texture di secondo ordine, messe a confronto con i valori corrispondenti calcolati sulle immagini originali.



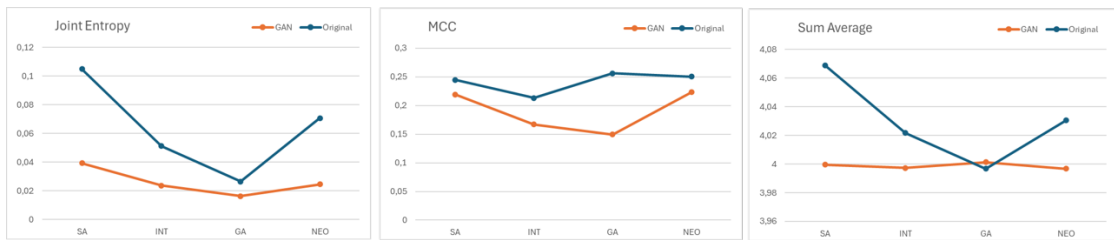
**Figura 3.10:** Andamento delle feature di texture del secondo ordine: Contrast, Difference Average e Difference Entropy



**Figura 3.11:** Andamento delle feature di texture del secondo ordine: Difference Variance, Inverse Variance, Autocorrelation



**Figura 3.12:** Andamento delle feature di texture del secondo ordine: Correlation, Joint Average, Joint Energy



**Figura 3.13:** Andamento delle feature di texture del secondo ordine: Joint Entropy, MCC, Sum Average

## Capitolo 4

# Discussione e Conclusione

In questo capitolo verranno discussi i risultati riportati nel capitolo 3 e saranno presentati possibili sviluppi futuri del lavoro.

### 4.1 Discussione

Dall'analisi qualitativa della prima GUI, lo stadio sano riscontra la percentuale più alta di immagini valutate con punteggio 5, pari a 86,92% , seguito dallo stadio atrofia geografica con una percentuale del 67,68, lo stadio neovascolare mostra una percentuale di 42,02. Infine lo stadio intermedio presenta la percentuale maggiore per le immagini valutate con punteggio 4, pari a 53,86%

L'analisi della seconda GUI mostra che il 30% delle immagini sintetiche sono state identificate come reali. Il 75% delle immagini classificate come sintetiche ha ricevuto una valutazione pari a 5 sulla scala di sicurezza; questo suggerisce un elevato livello di confidenza del valutatore. Tale risultato è stato ottenuto da un medico validatore con una conoscenza preliminare delle immagini.

In conclusione all'analisi qualitativa le immagini sintetiche risultano verosimili ad un primo impatto visivo, ma, a seguito di un'osservazione più attenta, possono essere facilmente riconosciute come artificiali.

I risultati ottenuti dall'analisi quantitativa dimostrano la capacità del modello di generare delle immagini realistiche con una buona qualità complessiva e un generale aumento del contrasto con valori di CII in intorno 1,64. Analizzando la tabella possiamo dedurre che lo stadio dei sani è quello che ha ottenuto le prestazioni migliori. I valori di SSIM e PSNR dimostrano l'ottima fedeltà strutturale ottenuta dalla GAN in confronto alla controparte originale. La metrica che maggiormente caratterizza la qualità e il realismo tra i vari stadi è il FID, con un valore pari a

28,59 per lo stadio sano, seguito dallo stadio di atrofia geografica con un FID di 29,08. A differenza dello stadio intermedio che presenta un valore di 55,89 e lo stadio neovascolare con un valore di 43,12, evidenziando una maggiore difficoltà della GAN nel riprodurre i pattern di questi stadi.

Per la texture analysis, nell'analisi delle features del primo ordine gli errori relativi delle features si attestano mediamente intorno al 5%, fatta eccezione per entropia e skewness che hanno andamenti differenti rispetto alle altre features e ordini di grandezza maggiori, arrivando fino al 162% per l'entropia nello stadio intermedio e 51% per la skewness nello stadio Atrofia Geografica, dovuto alla complessità dell'immagine analizzata e alla sua asimmetria, con alcune regioni che presentano densità di intensità maggiore rispetto alla media complessiva dell'immagine. Analizzando gli andamenti delle features si può notare come generalmente la GAN segua l'andamento delle features calcolate sulle immagini originali con un'inversione dell'andamento principalmente per lo stadio neovascolare. Questo comporta un'immagine strutturalmente simile all'immagine originale, con i sani rappresentati meglio rispetto agli altri stadi grazie alla loro maggiore omogeneità e regolarità strutturale, ma con difficoltà nel riprodurre strutture complesse di piccola scala come le strutture vascolari dello stadio neovascolare. Questo viene confermato anche guardando gli andamenti delle feature del secondo ordine e i rispettivi errori relativi che hanno ottenuto un aumento significativo delle medie degli errori relativi, dovuti ad una maggiore sensibilità delle features GLCM a piccole discrepanze introdotte dal generatore. Questa elevata sensibilità fa sì che negli stadi con maggiore omogeneità vi sia un amplificazione di ogni minima differenza in termini relativi, mentre negli stadi patologici questo ha un impatto minore per via della variabilità intrinseca della struttura retinica.

In conclusione, i risultati ottenuti dimostrano che il modello è in grado di generare immagini globalmente realistiche e coerenti con le originali, come anche evidenziato dai risultati della prima GUI. Rimangono ancora limiti nella riproduzione delle strutture più complesse e dei pattern locali, in particolare negli stadi patologici avanzati, come evidenziato dai risultati della seconda GUI. Queste considerazioni suggeriscono che la GAN sviluppata possa essere efficacemente impiegata per la generazione di immagini OCT sintetiche, utili ad ampliare i dataset attraverso tecniche di data augmentation e a supportare lo sviluppo di modelli di machine learning e deep learning più robusti. Inoltre, tali immagini possono essere utilizzati a scopo didattico.

## **4.2 Sviluppi futuri**

Gli sviluppi futuri di questo lavoro potranno concentrarsi su diversi aspetti:

- Miglioramento dell'estrazione della mappa semantica attraverso un allenamento multistadio della GMM, utilizzando l'intero dataset a disposizione.
- Ampliamento del dataset per aumentare la capacità di generalizzazione della GAN.
- Applicazione di trasformazioni al dataset in modo da testare la robustezza della GAN.
- Validazione delle metriche di tessitura del primo e del secondo ordine sui vari layer della retina.
- Validazione qualitativa da parte di più clinici in modo da avere dati più oggettivi e robusti.

In definitiva, il modello sviluppato conferma la possibilità di produrre immagini sintetiche verosimili e coerenti, offrendo nuove opportunità per la ricerca e la formazione clinica.

# Bibliografia

- [1] National Eye Institute. *How the Eyes Work*. Accessed: 2025-08-16. National Eye Institute. 2022. URL: <https://www.nei.nih.gov/learn-about-eye-health/healthy-vision/how-eyes-work> (cit. alle pp. 1, 2).
- [2] K. H. Nguyen, B. C. Patel e P. Tadi. «Anatomy, Head and Neck: Eye Retina». In: *StatPearls*. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK537226/>. Treasure Island (FL): StatPearls Publishing, 2023 (cit. a p. 1).
- [3] Christopher S. Langlo, Aana Amin e Susanna S. Park. «Optical coherence tomography retinal imaging: Narrative review of technological advancements and clinical applications». In: *Annals of Translational Medicine* 13.2 (2025), p. 17. DOI: 10.21037/atm-24-211. URL: <https://doi.org/10.21037/atm-24-211> (cit. a p. 4).
- [4] Shlomit Schaal, Ahmed ElTanboly, Marwa Ismail, Andy Switala, Ahmed Shalaby, Amir Hadayer, Omar Abdelmegid e Ayman El-Baz. «A novel automated method for the objective quantification of Retinal layers reveals sequential changes that occur in the Normal retina with age». In: *Frontiers in Nanoscience and Nanotechnology Research* 2.5 (2016), pp. 184–189. ISSN: 2397-6527. DOI: 10.15761/FNN.1000133 (cit. a p. 4).
- [5] H. Vyawahare e P. Shinde. «Age-related macular degeneration: Epidemiology, pathophysiology, diagnosis, and treatment». In: *Cureus* 14.9 (2022). DOI: 10.7759/cureus.29527 (cit. a p. 4).
- [6] National Eye Institute. *Age-related macular degeneration*. Accessed: 2025-07-15. U.S. Department of Health & Human Services. 2025. URL: <https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/age-related-macular-degeneration> (cit. a p. 4).
- [7] W. L. Wong, X. Su, X. Li, C. M. G. Cheung, R. Klein, C. Y. Cheng e T. Y. Wong. «Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis». In: *The Lancet Global Health* 2.2 (2014). DOI: 10.1016/S2214-109X(13)70145-1 (cit. a p. 4).

- [8] The Age-Related Eye Disease Study Research Group. «The Age-Related Eye Disease Study (AREDS): Design Implications». In: *Control Clin Trials* 20.6 (1999). Author manuscript available in PMC 2006 June 2, pp. 573–600 (cit. a p. 5).
- [9] Milan Sonka e Michael D. Abramoff. «Quantitative analysis of retinal OCT». In: *Iowa Institute for Biomedical Imaging, The University of Iowa* (2016) (cit. a p. 7).
- [10] Silke Aumann, Sabine Donner, Jörg Fischer, Frank Müller e Josef F. Bille. «Optical Coherence Tomography». In: *High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics*. A cura di Stefan W. Hell e Robert N. Weinreb. 2019, pp. 59–69. DOI: 10.1007/978-3-030-16638-0\_3 (cit. alle pp. 8–10).
- [11] PhysicsOpenLab. *Interferometro di Michelson-Morley*. Consultato il 10 agosto 2025. 2020. URL: <https://physicsopenlab.org/2020/05/16/interferometro-di-michelson-morley/> (cit. a p. 8).
- [12] Wolfgang Drexler, Mengyang Liu, Abhishek Kumar, Tschackad Kamali, Angelika Unterhuber e Rainer A. Leitgeb. «Optical coherence tomography today: Speed, contrast, and multimodality». In: *Journal of Biomedical Optics* 19.7 (2014), p. 071412. DOI: 10.1117/1.JBO.19.7.071412 (cit. a p. 10).
- [13] IBM. *Artificial intelligence (AI)*. Accessed: 2025-08-06. IBM Corporation. 2025. URL: <https://www.ibm.com/it-it/think/topics/artificial-intelligence> (cit. a p. 10).
- [14] A. Al Kuwaiti, K. Nazer, A. Al-Reedy, S. Al-Shehri, A. Al-Muhanna, A. V. Subbarayalu, D. Al Muhanna e F. A. Al-Muhanna. «A review of the role of artificial intelligence in healthcare». In: *Journal of Personalized Medicine* 13.6 (2023), p. 251. DOI: 10.3390/jpm13060251 (cit. a p. 10).
- [15] IBM. *Che cos'è l'AI in medicina?* Accessed: 2025-08-06. IBM Corporation. 2025. URL: <https://www.ibm.com/it-it/topics/artificial-intelligence-medicine> (cit. a p. 10).
- [16] K. Guo, J. Chen, T. Qiu, S. Guo, T. Luo, T. Chen e S. Ren. «MedGAN: An adaptive GAN approach for medical image generation». In: *Computers in Biology and Medicine* 169 (2023), p. 107599. DOI: 10.1016/j.compbiomed.2023.107599 (cit. a p. 11).
- [17] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville e Yoshua Bengio. «Generative adversarial nets». In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 27. 2014 (cit. alle pp. 11, 12).

- [18] Humanativa. *Reti neurali Generative (GAN)*. [Online; accesso effettuato il 11 ottobre 2025]. maggio 2022. URL: <https://humanativaspa.it/reti-neurali-generative-gan/> (cit. a p. 13).
- [19] Data Masters. *Introduzione all'algoritmo di ottimizzazione Adam*. Accessed: 2025-08-14. Data Masters. 2023. URL: <https://datamasters.it/blog/introduzione-allalgoritmo-di-ottimizzazione-adam/> (cit. a p. 14).
- [20] Diederik P. Kingma e Jimmy Lei Ba. «Adam: A Method for Stochastic Optimization». In: *International Conference on Learning Representations (ICLR)*. 2015. URL: <https://arxiv.org/abs/1412.6980> (cit. a p. 14).
- [21] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou e Alexei A. Efros. «Image-to-image translation with conditional adversarial networks». In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, pp. 5967–5976. DOI: 10.1109/CVPR.2017.632 (cit. alle pp. 14, 21).
- [22] H. Adiyaman, Y. E. Varul, T. Bakırman e B. Bayram. «Stripe Error Correction for Landsat-7 Using Deep Learning». In: *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 93 (2024), pp. 51–63. DOI: 10.1007/s41064-024-00306-x (cit. a p. 15).
- [23] H. Sun et al. «HRPGAN: A GAN-based Model to Generate High-resolution Remote Sensing Images». In: *IOP Conference Series: Earth and Environmental Science* 428 (2020), p. 012060. DOI: 10.1088/1755-1315/428/1/012060 (cit. a p. 23).
- [24] S. Saravanan e P. Siva Kumar. «Image Contrast Enhancement Using Histogram Equalization Techniques: Review». In: *International Journal of Advances in Computer Science and Technology* 3.3 (mar. 2014). URL: <http://warse.org/pdfs/2014/ijacst03332014.pdf> (cit. a p. 23).
- [25] *Radiomic Features – First Order Statistics*. <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.firstorder>. Accessed: 2025-10-11. 2025 (cit. a p. 23).
- [26] *PyRadiomics Documentation: Gray Level Co-occurrence Matrix (GLCM) Features*. <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.glcm>. Accessed: 2025-10-11. 2025 (cit. a p. 24).