



**Politecnico
di Torino**

Corso di Laurea Magistrale in Ingegneria Biomedica
Orientamento in strumentazione biomedica
A.a. 2024/2025
Sessione di Laurea di ottobre 2025

Modelli generativi per la sintesi di immagini OCT retiniche relative alla degenerazione maculare legata all'età

Relatori:
Prof. Filippo Molinari

Candidato:
Bussi Daniele

Correlatori:
Reibaldi Michele
Meiburger Kristen Mariko
Salvi Massimo



**Politecnico
di Torino**

In collaborazione con



INDICE

SOMMARIO	7
1 INTRODUZIONE	9
1.1 Degenerazione maculare legata all'età (AMD)	10
1.2 Optical Coherence Tomography(OCT)	15
1.3 Generative adversarial network (GAN)	20
1.3.1 Applicazioni in oftalmologia	23
2 MATERIALI E METODI	25
2.1 Sistema di acquisizione	25
2.2 Dataset utilizzato	26
2.3 Preprocessing	28
2.3.1 Algoritmi utilizzati	32
2.4 Allenamento GAN	44
3.4.1 Immagini Generate dalla GAN	48
3 VALIDAZIONE	51
3.1 GUI PER VALIDAZIONE QUALITATIVA	51
3.1.1 Risultati Validazione Qualitativa	55
3.2 VALIDAZIONE QUANTITATIVA	57
3.2.1 RISULTATI VALIDAZIONE QUANTITATIVA	67
4 CONCLUSIONI	76
4.1 Sviluppi Futuri	78
BIBLIOGRAFIA:	79

INDICE DELLE FIGURE

Figura 1: Schema della regione maculare. Modificato da: Y. Deng et al, 2022, Genes & Diseases	10
Figura 2: Schema della genesi della malattia AMD e dei fattori di rischio che la causano, Modificato da: Y. Deng et al, 2022, Genes & Diseases	11
Figura 3: OCT Occhio Sano	14
Figura 4: OCT Stadio Intermedio	14
Figura 5: OCT Stadio Neovascolare	14
Figura 6: OCT Atrofia Geografica	14
Figura 7: schema del principio di funzionamento delle A-scans e misurazione di spessore della cornea di un coniglio[4]	15
Figura 8: set up dell'interferometro e principio di funzionamento delle tre strategie di processamento segnale OCT [7]	17
Figura 9: grafico del numero di pubblicazioni dal 1997 al 2025 tratto dal sito PubMed	17
Figura 10: dimostrazione dell'efficacia della OCTA. L'immagine OCTA (F) riesce a rilevare una neovascolarizzazione coroidale secondaria che nella angiografia con fluoresceina in fase tardiva (D) non era stata diagnosticata [8]	19
Figura 11: Principio di funzionamento di una GAN per sintetizzare OCT retinici[12].....	20
Figura 12: immagine dello spectralis HRA+OCT (tratto dal sito di spectralis).....	25
Figura 13: esempio di immagine OCT catturata dal Heidelberg Spectralis HRA+OCT	25
Figura 14a/b/c: variabilità d'orientamento dell'immagine	27
Figura 15a/b/c/ variabilità di posizione dell'immagine	27
Figura 16a/b/c/ variabilità di rumore	27
Figura 17: Pipeline di Preprocessing	31
Figura 18: esempio applicazione del filtro gaussiano [19]	32
Figura 19: processo di denoising con l'algoritmo non local means	35
Figura 20: Esempio di scomposizione in gaussiane di una distribuzione [25]	39
Figura 21: Esempio di risultato dell'algoritmo GMM sulle immagini del dataset	39
Figura 22: a/b/c/d: esempio dell'effetto del mean shift all'aumentare del raggio spaziale(sp) [28]	41
Figura 23: confronto tra le diverse tipologie di processing con Otsu [29].....	43
Figura 24: Esempio di un'immagine di train	44
Figura 25: Esempio di un'immagine di Test	44
Figura 26: Immagini generate della GAN dopo l'allenamento per lo stadio Sano	48
Figura 27: Immagini gen. della GAN dopo l'allenamento per lo stadio Intermedio	49
Figura 28: Immagini gen. della GAN dopo l'allenamento per lo stadio Neovascolare	49
Figura 29: Immagini generate della GAN dopo l'allenamento per lo stadio Atrofia Geografica	50
Figura 30: Rappresentazione della GUI, dopo il settaggio dei dataset	51
Figura 31: Avviso che appare all'avvio dell'applicazione	52
Figura 32: Rappresentazioni dei menu a tendina e del contatore delle immagini	52
Figura 33: Rappresentazione parte bassa GUI	52
Figura 34: Foglio generato dalla GUI coi parametri di validazione "Valutazione OCT"	53

Figura 35: GUI per la validazione blind	53
Figura 36: avviso di salvataggio del voto e risconto diretto dell'immagine	54
Figura 37: Foglio generato dalla GUI alla fine del processo di validazione "RealorFake"	54
Figura 38: Risultati ottenuti dalla validazione qualitativa, raggruppati per voto della GUI "Valutazione OCT"	55
Figura 39: Confronto dell'Andamento dell'Energia lungo gli stadi per le immagini reali e false	67
Figura 40: Confronto dell'Andamento dell'Entropia lungo gli stadi per le immagini reali e false	68
Figura 41: Confronto dell'Andamento della MAD lungo gli stadi per le immagini reali e false	68
Figura 42: Confronto dell'Andamento della RMS lungo gli stadi per le immagini reali e false	68
Figura 43: Confronto dell'Andamento della Skewness lungo gli stadi per le immagini reali e false	69
Figura 44: Confronto dell'Andamento della Kurtosis lungo gli stadi per le immagini reali e false	69
Figura 45: Confronto dell'Andamento della Uniformity lungo gli stadi per le immagini reali e false	69
Figura 46: Confronto dell'Andamento del contrasto lungo gli stadi per le immagini reali e false	70
Figura 47: Confronto dell'Andamento della Difference Aerge lungo gli stadi per le immagini reali e false	71
Figura 48: Confronto dell'Andamento della Difference Entropy lungo gli stadi per le immagini reali e false	71
Figura 49: Confronto dell'Andamento della Difference Variance lungo gli stadi per le immagini reali e false	71
Figura 50: Confronto dell'Andamento della Inverse Variance lungo gli stadi per le immagini reali e false	72
Figura 51: Confronto dell'Andamento dell'Autocorrelazione lungo gli stadi per le immagini reali e false	72
Figura 52: Confronto dell'Andamento della Correlazione lungo gli stadi per le immagini reali e false	72
Figura 53: Confronto dell'Andamento della Joint Average lungo gli stadi per le immagini reali e false	73
Figura 54: Confronto dell'Andamento della Joint Energy lungo gli stadi per le immagini reali e false	73
Figura 55: Confronto dell'Andamento della Joint Entropy lungo gli stadi per le immagini reali e false	73
Figura 56: Confronto dell'Andamento del MCC lungo gli stadi per le immagini reali e false	74
Figura 57: Confronto dell'Andamento della Sum Average lungo gli stadi per le immagini reali e false	74
Figura 58: Confronto dell'Andamento della Sum Entropy lungo gli stadi per le immagini reali e false	74

INDICE DELLE TABELLE

Tabella 1: Trattamenti disponibili per stadio Neovascolare e Atrofia Geografica	13
Tabella 2: Numero di pazienti e immagini elaborate per ogni stadio	26
Tabella 3: Caratteristiche della GAN allenata	45
Tabella 4: Risultati ottenuti dalla validazione qualitativa della GUI "Valutazione OCT"	55
Tabella 5: Risultati della validazione qualitativa sulle immagini reali, mediante GUI "RealOrFake"	56
Tabella 6: Risultati della validazione qualitativa sulle immagini sintetiche, mediante GUI "RealOrFake"	56
Tabella 7: Risultati ottenuti dall' analisi tessutale del primo ordine con relative deviazioni standard	67
Tabella 8: Risultati ottenuti dall' analisi tessutale del secondo ordine con relative deviazioni standard	70
Tabella 9: Risultati metriche non Tessutali	75

SOMMARIO

Nei Paesi industrializzati la degenerazione maculare legata all'età (Age-related Macular Degeneration, AMD) rappresenta la principale causa di perdita visiva irreversibile nella popolazione anziana. Questa patologia retinica degenerativa interessa primariamente la regione maculare, compromettendo in maniera significativa la funzione visiva centrale. Attualmente, la tomografia a coerenza ottica (Optical Coherence Tomography, OCT) è considerata il gold standard per la diagnosi precoce e per la caratterizzazione morfologica delle lesioni associate all'AMD in quanto consente un'analisi non invasiva e ad alta risoluzione delle strutture retiniche. La combinazione di questa tecnologia e di algoritmi basati sulle reti neurali (IA) renderebbe meno gravoso il lavoro dei medici nel rilevare e monitorare il decorso della malattia nel tempo.

Il dataset preso in esame comprende immagini OCT relative a pazienti appartenenti ai diversi stadi della malattia (163 pazienti sani, 272 con Stadio Intermedio, 150 con Atrofia Geografica e 154 con stadio Neovascolare). Per ciascun paziente sono state raccolte 49 immagini, per un totale complessivo di 36.211 immagini OCT. Nella tesi presentata è stato analizzato il dataset al fine di allenare una rete generativa avversaria (Generative Adversarial Network o GAN) nella produzione di immagini OCT fittizie della retina aventi le stesse caratteristiche dello stadio associato. Nella prima fase del percorso è stato analizzato il contenuto semantico delle immagini originali in modo da dare alla rete un'informazione generale della struttura dell'occhio, così da poter generare una serie d'immagini coerenti tra loro che possano simulare un occhio

umano reale. Nella seconda parte è stata allenata una GAN con architettura piramidale usando il contenuto semantico analizzato e le immagini originali per un totale di cento epoche.

Le immagini ottenute sono state valutate sia qualitativamente dai clinici che quantitativamente. Nei pazienti sani, i valori ottenuti mostrano un PSNR di 23,19, un SSIM pari a 0,46, un FID di 28,59, errori relativi di texture analysis intorno al 3,7% e un Contrast Improvement Index (CII) pari a 1,6. Nei pazienti con stadio intermedio, i parametri risultano leggermente differenti: il PSNR scende a 22,13, l'SSIM si attesta a 0,44, mentre il FID cresce sensibilmente fino a 55,7. Gli errori relativi di texture analysis rimangono contenuti, circa il 3,5% e il CII risulta pari a 1,65. Nei pazienti con Atrofia Geografica, invece, si rileva un PSNR di 22,36 e un SSIM di 0,43. In questo gruppo il FID risulta pari a 29,08, mentre gli errori relativi di texture analysis si riducono ulteriormente, attestandosi intorno al 2,6%. Il CII, invece, mostra un incremento, raggiungendo il valore di 1,69. Infine, nei pazienti con stadio Neovascolare, i valori riportano un PSNR di 22,62, un SSIM di 0,44, un FID pari a 43,12, con errori relativi di texture analysis leggermente più alti, circa il 3,8% e un CII di 1,66. I risultati confermano la qualità delle immagini prodotte con un aumento generale del contrasto nelle immagini generate e una buona fedeltà strutturale risultando molto simili alle controparti originali. Grazie ai risultati ottenuti le immagini generate potranno essere utilizzate sia per la formazione dei medici specializzandi sia come dataset di supporto (Data Augmentation) per algoritmi di classificazione automatica, contribuendo ad aumentare l'accuratezza nel rilevamento dei diversi stadi di AMD e fornendo un valido supporto alla pratica clinica.

CAPITOLO 1

INTRODUZIONE

In questa tesi si propone un modello per la generazione di immagini sintetiche basato sull'utilizzo delle GAN (Generative Adversarial Networks), con l'obiettivo di produrre immagini retiniche realistiche che rappresentino i diversi stadi della Degenerazione Maculare legata all'età (AMD). Lo scopo principale di questo lavoro è la creazione di un dataset artificiale che possa supportare i medici sia in ambito accademico, sia come strumento utile per ampliare i dataset di training utilizzati nello sviluppo di modelli di classificazione automatica.

La tesi approfondisce gli aspetti teorici che riguardano sia il funzionamento del modello proposto, sia le caratteristiche cliniche e visive della patologia studiata. Successivamente sono illustrati i risultati sperimentali ottenuti e le considerazioni conclusive che emergono dall'analisi dei dati, evidenziando punti di forza, limiti e possibili sviluppi futuri del lavoro.

1.1 Degenerazione maculare legata all'età (AMD)

La degenerazione maculare legata all'età è una malattia che colpisce l'area centrale dell'occhio, denominata macula.

La macula è un'area retinica di forma circolare, con un diametro di 5.5 mm, posta 4,0-5,0 mm temporalmente e 0,5-0,8 mm inferiormente rispetto al centro del disco ottico. Al suo interno è presente la fovea, area costituita da coni posti a stretto contatto tra loro e responsabile della visione centrale [fig. 1], [1].

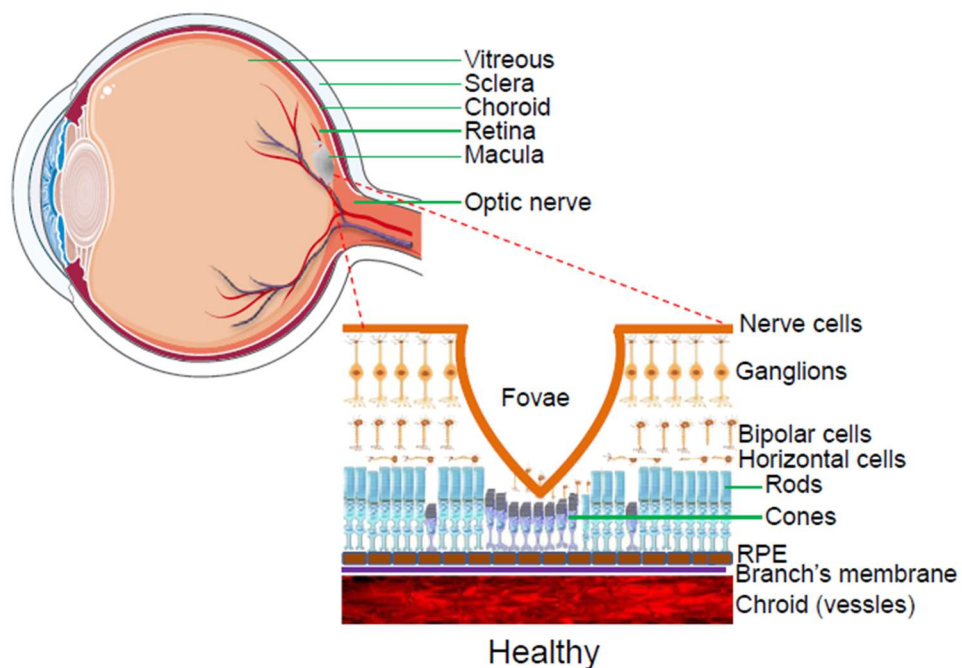


Figura 1: Schema della regione maculare. Modificato da: Y. Deng et al, 2022, Genes & Diseases [1].

La AMD rappresenta una delle principali cause di perdita della vista, soprattutto nei paesi industrializzati, in relazione alla sua stretta correlazione con l'età. Il tasso di prevalenza della malattia è variabile nelle diverse etnie ed è stimato essere del 12.33% nella popolazione europea [1]. Si stima che a livello mondiale la AMD sia presente nell'8.7% della popolazione, pari a 196 milioni di persone affette nel 2020; considerando l'invecchiamento della popolazione mondiale il numero delle persone affette da AMD raggiungerebbe i 288 milioni nel 2040[2].

L'età rappresenta, come già sottolineato, il principale fattore di rischio per la AMD, ma altri fattori giocano un importante ruolo nella genesi della malattia, quali ipertensione arteriosa, l'aterosclerosi, il diabete mellito, il fumo e l'assunzione di alcolici [1].

Anche la genetica svolge un ruolo importante nella eziopatogenesi della malattia, essendo risultati essere maggiormente associati con l'AMD i loci genici CFH e HTRA1. Nella patogenesi della malattia sono implicati molteplici meccanismi: dallo stress ossidativo alle alterazioni del metabolismo lipidico, al ruolo di fattori infiammatori ed immunitari [fig. 2], [1].

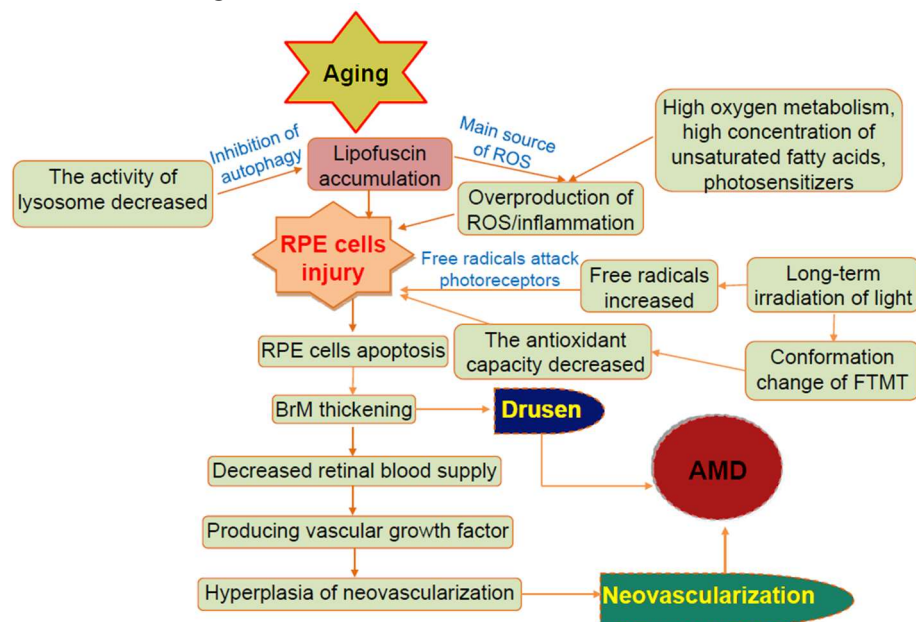


Figura 2: Schema della genesi della malattia AMD e dei fattori di rischio che la causano, Modificato da: Y. Deng et al, 2022, Genes & Diseases [1].

La diagnosi della AMD si avvale di diverse metodiche quali l'imaging ad infrarossi, l'auto fluorescenza del fundus, l'angiografia indocianina verde, l'OCT e l'angiografia OCT

Il decorso della malattia presenta quattro stadi [1]:

- 1. NORMALE DEGENERAZIONE DOVUTA ALL'ETA: in questo stadio** iniziale ci sono piccoli cambiamenti nella macula dovuti all' invecchiamento con comparsa di Drusen (depositi di materiale extracellulare) di piccole dimensioni ($<63\ \mu\text{m}$) e nessuna alterazione dei coni e bastoncelli all'interno della retina.
- 2. STADIO INTERMEDIO (intermediate state):** caratterizzato dalla presenza di più Drusen di grandi dimensioni (diametro da $63\ \mu\text{m}$ a $125\ \mu\text{m}$) o di un unico Drusen con diametro maggiore o uguale a $125\ \mu\text{m}$. Come risultato si ha la nascita di anomalie a livello dell'Epitelio Pigmentato Retinico (EPR)[Fig.3].
- 3. AMD AVANZATA:** questo stadio è caratterizzato dalla contemporanea presenza di più Drusen di diametro superiore a $125\ \mu\text{m}$ e di uno dei due seguenti differenti tipi di degenerazione maculare [Fig4]:
 - a. NEOVASCULARIZZAZIONE:** neoformazione di capillari coroideali che crescono attraverso le lesioni dell'EPR e della membrana di Bruch (MB); tale alterazione può esitare nella Vasculopatia Coroidale Polipoidale (VCP) o nella Neovascolarizzazione Coroidale (NC). Questo stadio viene anche definito Wet AMD in quanto comporta una

perdita di sangue e siero all'interno della retina, creando una regione di liquido più o meno estesa [Fig.5].

b. ATROFIA GEOGRAFICA: cambiamento cromatico dei Drusen verso il bianco-giallastro tra la zona della MB e la zona dell'EPR. Questo stadio è indicato anche come Dry AMD in quanto non sono presenti liquidi. Caratteristici dell'atrofia geografica sono la presenza di una visione distorta simmetrica, la pigmentazione maculare distorta in entrambi gli occhi, la scomparsa del riflesso foveale e la variazione dei poli posteriori[fig.6].

La terapia della AMD prevede innanzi tutto un'azione sui fattori di rischio modificabili della malattia, attraverso la sospensione del fumo, la riduzione dell'assunzione di alcolici, l'introduzione di una dieta mediterranea e l'aumento dell'attività fisica. Esistono inoltre oggi terapie specifiche che prevedano l'utilizzo di fattori di anti crescita dell'endotelio vascolare (VEGF) o di inibitori di fattori del complemento [tab. 1][2].

Trattamenti AMD Neovascolare	Trattamenti AMG Atrofia Geografica
Ranibizumab	Pegcetacoplan (anti C3).
Aflibercept	Avacincaptad pegol (anti C5).
Brolucizumab	Anti-sense oligonucleotide targeting complement factor B.
Faricimab	Complement factor I gene therapy
Bevacizumab (off-label)	-
Biosimilars	-

Tabella 1: Trattamenti disponibili per stadio Neovascolare e Atrofia Geografica

I trattamenti per la AMD neovascolare indicati in tabella sono attualmente disponibili, mentre quelle per l'atrofia geografica sono ancora in fase sperimentale.

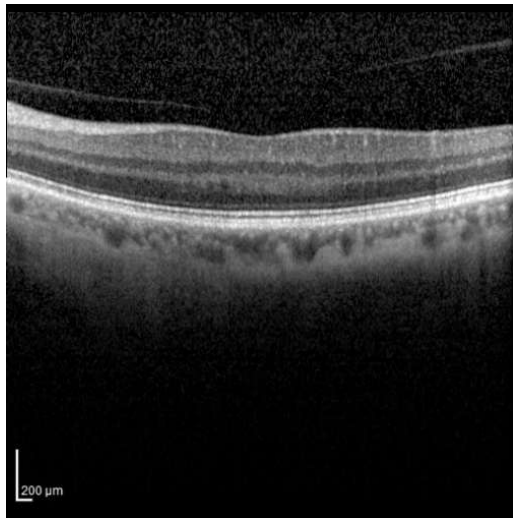


Figura 3: OCT Occhio Sano

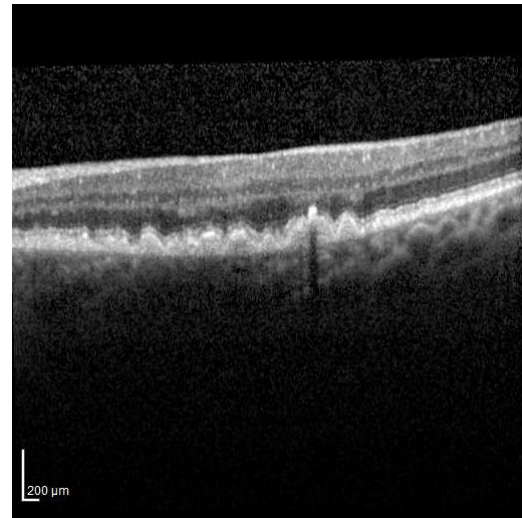


Figura 4: OCT Stadio Intermedio

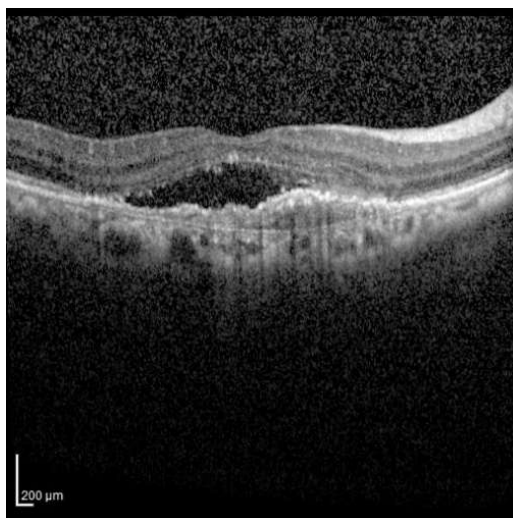


Figura 5: OCT Stadio Neovascolare

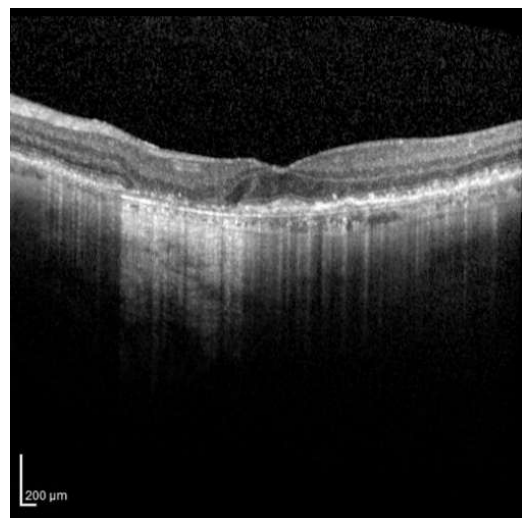


Figura 6: OCT Atrofia Geografica

1.2 Optical Coherence Tomography(OCT)

Lo strumento diagnostico più diffuso per diagnosticare l'AMD e monitorarne la degenerazione nel tempo è l'Optical Coherence Tomography o OCT.

Questa tecnica rappresenta una delle innovazioni tecnologiche più rilevanti in oftalmologia degli ultimi trent'anni.

Le prime pubblicazioni risalgono agli anni '90, anni in cui questa tecnica viene sviluppata ampiamente e, grazie all'applicazione dell'interferometria a bassa coerenza dette A-scans (o interferenza a luce bianca usando un fascio di luce con lunghezza d'onda $\lambda=840$ nm), ha rapidamente rivoluzionato l'imaging retinico [Fig.7], [4].

Questi primi strumenti definiti a dominio temporale (TD-OCT) fornivano immagini sezionali della retina, consentendo di visualizzare la morfologia con una risoluzione innovativa e molto alta rispetto gli standard dell'epoca.

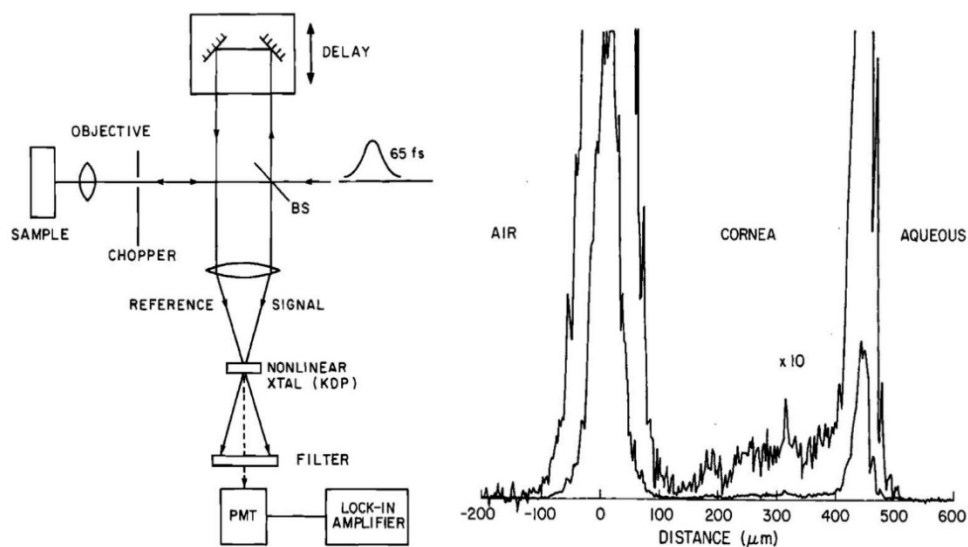


Figura 7: schema del principio di funzionamento delle A-scans e misurazione di spessore della cornea di un coniglio [4].

Nonostante tutti questi vantaggi, rimaneva un grande svantaggio che rendeva L'OCT potente ma non efficiente, ovvero il tempo di acquisizione che si attestava intorno a diversi minuti per una sola scansione 2D. Nei primi anni del 2000 si svilupparono due tecniche di acquisizione per supplire a questo problema. Esse avevano in comune il concetto di raccogliere molto più rapidamente i dati di riflettanza mantenendo lo specchio di riferimento in una posizione fissa e:

1. Per una tecnica si usava una sorgente luminosa che includesse una gamma di lunghezze d'onda in modo da poter essere interrogata in parallelo aggiungendo uno spettrometro al rilevatore.
2. Per l'altra tecnica invece la sorgente luminosa doveva essere costituita da un laser a banda stretta che spazzasse rapidamente e ripetutamente una gamma di lunghezze d'onda insieme a un rilevatore che registrasse il modello di interferenza in più punti temporali all'interno di ciascuna scansione spettrale.

Fu da queste due idee che si crearono nel primo caso le prime OCT che sfruttavano lo spettro (spectral-domain DS-OCT), mentre nel secondo caso le prime OCT che sfruttavano le frequenze (frequency-domain (FD-OCT) o Swept-Source (SS-OCT)) [Fig.8].

Entrambe le tecnologie usano la trasformata di Fourier per calcolare il profilo di riflettanza lungo l'asse Z, per questo esse vengono anche associate col nome di OCT nel dominio di Fourier.

Grazie a questa innovazione. Gli attuali scanner riescono ad acquisire 100.000 scansioni lungo l'asse Z (A-scan) al secondo.

Inoltre, è stata migliorata anche la qualità dell'immagine, con una migliore risoluzione, rendendo possibile l'analisi volumetrica e quantitativa di strutture retiniche e coroideali [7].

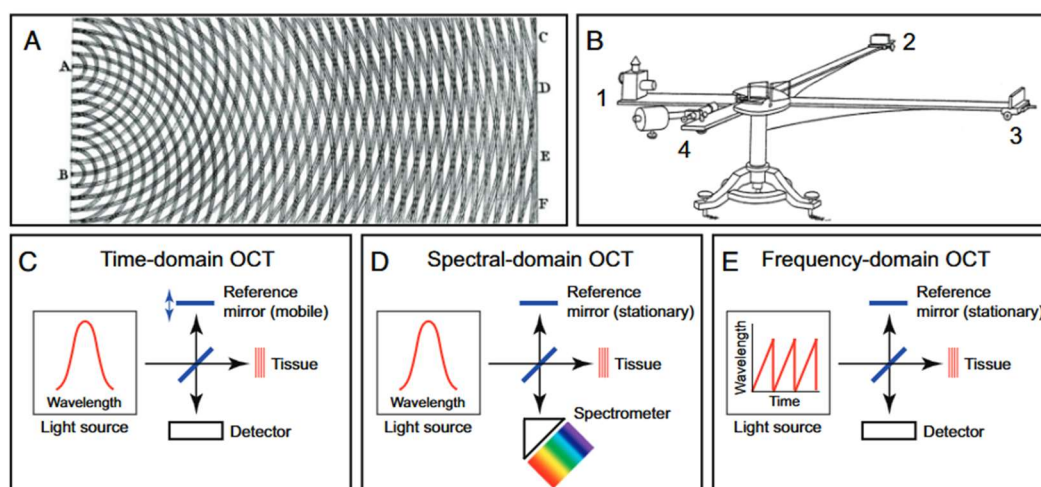


Figura 8: set up dell'interferometro e principio di funzionamento delle tre strategie di processamento segnale OCT [7]

Queste qualità unite alla non invasività e ripetibilità dell'esame hanno reso l'esame OCT un ottimo esame di screening, permettendo di ridurre l'uso di metodologie tradizionali come l'angiografia con fluoresceina (FA), che, pur essendo utile per la classificazione della neovascolarizzazione coroideale, presenta rischi legati all'uso di coloranti [6].

Sintomo della loro efficacia è il numero di pubblicazioni scientifiche che sono state divulgate negli ultimi trent'anni che ha avuto un'importante impennata dagli anni 2005 fino agli anni 2020 in cui ha avuto il suo apice con 3018 pubblicazioni [Fig.9].

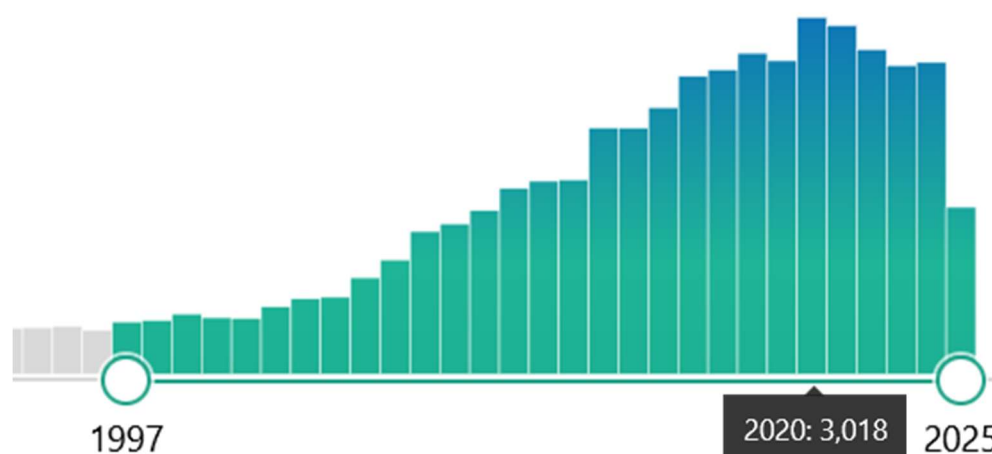


Figura 9: grafico del numero di pubblicazioni dal 1997 al 2025 tratto dal sito PubMed

È stata l'introduzione delle terapie anti-angiogeniche (come l'anti-VEGF o fattore di crescita endoteliale vascolare) a consolidare l'esame OCT come gold standard per la diagnosi e monitoraggio della risposta al trattamento della degenerazione maculare legata all'età (AMD) e nello specifico sullo stadio neovascolare.

Questo risultato è dovuto al ruolo fondamentale che ha avuto l'OCT in quanto ha fornito i biomarcatori visivi indispensabili per lo sviluppo e la valutazione di tali trattamenti: la possibilità di identificare e quantificare i fluidi intraretinali, subretinali e sotto l'epitelio pigmentato ha reso possibile calibrare i protocolli terapeutici, passando da regimi fissi a strategie personalizzate basate sull'imaging OCT. Così facendo si poteva ridurre il numero di iniezioni non necessarie e migliorare la qualità di vita dei pazienti [5].

L'ultima significativa innovazione nell'ambito delle OCT è stata l'introduzione della OCT-Angiografica (OCTA) che ha ulteriormente ampliato le possibilità diagnostiche proposte da questa tecnologia.

L'OCTA consente di visualizzare la microvascolarizzazione retinica e coroidale senza necessità di coloranti, rappresentando metodo valido quanto la FA e angiografia con idocianina verde, ma con molti meno rischi per il paziente.

In relazione all'AMD, l'OCTA permette l'identificazione precoce delle neovascolarizzazioni e il monitoraggio più dettagliato della risposta ai trattamenti rispetto agli standard precedenti di OCT, tuttavia permangono i limiti legati ad artefatti e alla necessità di standardizzazione [Fig.10], [8].

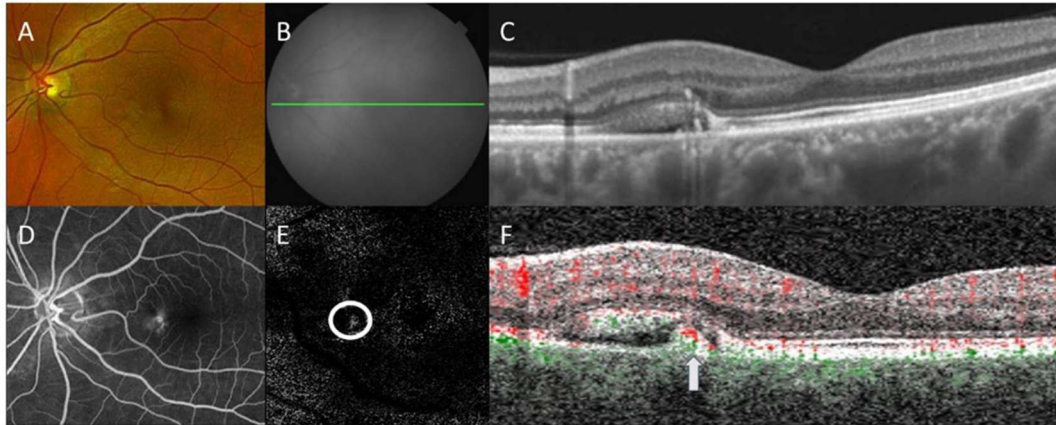


Figura 10: dimostrazione dell'efficacia della OCTA. L'immagine OCTA (F) riesce a rilevare una neovascolarizzazione coroidale secondaria che nella angiografia con fluoresceina in fase tardiva (D) non era stata diagnosticata [8].

Oggi, l'OCT e le sue evoluzioni rappresentano un paradigma di come una tecnologia nata dalla ricerca traslazionale abbia trasformato la pratica clinica. Non solo ha rivoluzionato la gestione della AMD, ma è diventata un modello di riferimento anche in altre patologie come glaucoma e retinopatia diabetica [5], [8].

1.3 Generative adversarial network (GAN)

Negli ultimi anni, l'introduzione del *deep learning* in oftalmologia ha aperto nuove prospettive nella diagnosi, prognosi e gestione delle patologie oculari. In particolare, le Generative Adversarial Networks (GAN) si sono affermate come strumenti promettenti, grazie alla loro capacità di apprendere distribuzioni complesse di dati e generare immagini sintetiche con un'elevata verosimiglianza rispetto a quelle originali. Questo aspetto risulta particolarmente rilevante in oftalmologia, dove la disponibilità di dataset ampi e bilanciati rappresenta una sfida critica [9].

Una GAN è costituita da due reti neurali che si allenano in maniera competitiva: il **generatore** e il **discriminatore** [Fig.11].

- Il **generatore (G)** ha il compito di produrre dati sintetici, a partire da un vettore di rumore casuale o da input condizionati. L'obiettivo è creare campioni che imitino il più possibile i dati reali.
- Il **discriminatore (D)** ha il ruolo di selezionare le immagini che riceve in input e dire se esse sono reali o sintetiche.

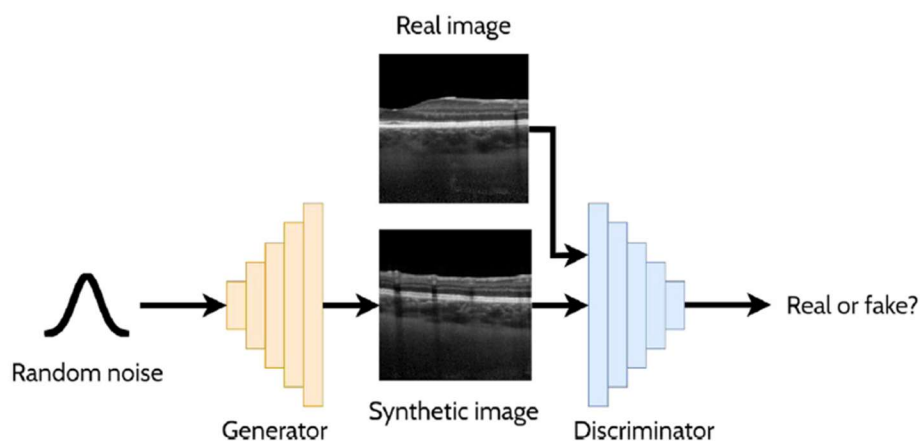


Figura 11: Principio di funzionamento di una GAN per sintetizzare OCT retinici [12].

Durante l'addestramento, il generatore cerca di ingannare il discriminatore producendo dati sempre più realistici, mentre il discriminatore cerca di migliorare la sua capacità di distinzione. Questo equilibrio di forze tra generatore e discriminatore è il concetto di base delle GAN che è espresso dalla funzione obiettivo della GAN classica che è la seguente:

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

In questo equilibrio a somma zero, il generatore (G) cerca di minimizzare la funzione di perdita condivisa mentre il discriminatore (D) cerca di massimizzarla [10], [11].

L'obiettivo è di arrivare ad un equilibrio teorico, con il passare delle epoche di addestramento, dove le immagini sintetiche diventano indistinguibili da quelle reali, ovvero il discriminatore ha una probabilità del 50% di classificare un'immagine come reale o sintetica per ogni immagine che riceve da input.

Questo modello però ha dei limiti che possono essere raggruppati in due gruppi:

1. Limiti di sensibilità:

- a. *Training instabile*: dovuto dalla funzione obiettivo stessa, che introducendo un equilibrio tra due reti neurali, può portare a oscillazioni, divergenze o collasso del training stesso se uno tra generatore o discriminatore diventa più preponderante dell'altro, non lasciandogli la possibilità di apprendere a sua volta, ma venirne sopraffatto.
- b. *Vanishing gradients*: dovuto ad un discriminatore troppo accurato già dalle prime epoche. Se il discriminatore è troppo accurato manderà al

generatore dei gradienti deboli che renderanno difficile l'aggiornamento dei pesi e quindi l'allenamento del generatore stesso.

- c. *Sensibilità agli iperparametri*: proprio per la sua natura strettamente legata in modo critico al learning rate, architettura, inizializzazione e rapporto G/D, piccole variazioni in uno di questi fattori possono portare ad un addestramento fallimentare.

2. Limiti di qualità:

- a. *Mode collapse*: Fenomeno nel quale il generatore produce solo un sottoinsieme ristretto della distribuzione reale producendo immagini ripetitive o con una variabilità limitata e non rappresentativa dell'intera distribuzione dei dati.
- b. *Artefatti visivi*: Immagini generate con anomalie strutturali o pattern innaturali che possono essere un problema importante in ambito medico dove piccoli dettagli sono essenziali.

Numerose varianti della GAN classica sono state proposte per superare questi limiti di stabilità e qualità come le **Wasserstein GAN (WGAN)** in cui viene introdotta la distanza di Wasserstein per ridurre l'instabilità e il vanishing gradients [12], o le **Progressive GAN** che adottano l'addestramento progressivo del generatore (da bassa ad alta risoluzione) per migliorare qualità e stabilità [14], oppure le **Conditional/CycleGAN**, ampiamente adottate per compiti di traduzione tra domini, grazie alla supervisione del generatore tramite un'etichetta o immagine che faccia da target così da ridurre fenomeni di mode collapse e migliorare la coerenza dei risultati[12].

1.3.1 Applicazioni in oftalmologia

Le GAN hanno trovato applicazione in molteplici ambiti legati all'oftalmica come la segmentazione di strutture anatomiche, la sintesi di immagini al fine di fare Data Augmentation, oppure la rimozione di rumore e artefatti, la traduzione tra domini di imaging differenti e la super-risoluzione. Queste metodologie hanno mostrato un potenziale notevole nell'incrementare l'accuratezza dei modelli diagnostici, pur evidenziando limiti come il rischio di introdurre artefatti non fisiologici [13].

Le immagini OCT sintetiche di patologie retiniche non solo riscuotono approvazione da parte degli esperti per via della loro verosimiglianza, ma possono anche essere impiegate per addestrare modelli di classificazione al fine di renderli più accurati. Questi ultimi, nei recenti studi, hanno mantenuto prestazioni comparabili a quelli addestrati esclusivamente su dati reali, confermando l'utilità della sintesi di dati per fronteggiare il problema della scarsità di immagini [15].

Per migliorare sempre di più le prestazioni delle GAN sono state pensate anche architetture a due stadi in grado di accorpare diversi aspetti di un'immagine per produrre immagini retiniche sintetiche realistiche, passando dalla sua segmentazione dei vasi nel primo step e arrivando all'immagine retinica sintetica nel secondo step. Queste immagini sono state poi date in pasto al discriminatore accompagnate dalle rispettive mappe di segmentazione vascolare. Questo approccio ha dimostrato di migliorare significativamente le prestazioni nei compiti di segmentazione, anche in condizioni di *few-shot learning*, evidenziando il valore delle GAN come strumenti per creare dataset di qualità [16].

Ulteriori studi hanno applicato le GAN all'elaborazione di immagini retiniche a bassa qualità, migliorandone il contrasto, introducendo metodologie capaci di preservare le strutture anatomiche rilevanti. Questo processo è definito *image enhancement*. Queste tecniche si sono dimostrate particolarmente utili in scenari clinici reali, dove la qualità delle immagini può risultare compromessa da fattori tecnici o operativi [17].

Oltre alle applicazioni tecniche, bisogna anche evidenziare il ruolo delle GAN nel più ampio contesto dell'intelligenza artificiale generativa in oftalmologia. Oltre a potenziali benefici, emergono questioni etiche e regolatorie, come il rischio di bias introdotti da dataset sbilanciati e rispettive implementazioni sbagliate. Tra le argomentazioni attuali ci sono molte critiche sulla necessità di validazioni cliniche robuste e l'importanza di garantire la trasparenza dei modelli al fine di salvaguardare la salute dei cittadini. Le GAN vengono così inquadrate non solo come strumenti per la ricerca, ma come potenziali protagonisti nel futuro della pratica oftalmologica come aiuto concreto agli specialisti, a patto che vengano superati i limiti ancora presenti in termini di affidabilità, generalizzabilità e sicurezza [18].

In sintesi, la letteratura mostra come le GAN rappresentino un pilastro emergente nell'applicazione del deep learning in oftalmologia. Esse offrono soluzioni concrete a problematiche legate alla scarsità di dati, alla qualità delle immagini e alla variabilità tra domini di acquisizione, con prospettive che spaziano dalla diagnosi assistita alla formazione medica. Tuttavia, la loro adozione clinica su larga scala richiede un ulteriore consolidamento metodologico e un'attenta considerazione delle implicazioni etiche e regolatorie.

CAPITOLO 2

MATERIALI E METODI

2.1 Sistema di acquisizione

Per fare le acquisizioni delle immagini si è utilizzato il Heidelberg Spectralis HRA+OCT device (Heidelberg Engineering, Heidelberg, Germany) [fig. 12] che accoppia alle 49 immagini OCT (Optical coherence tomography), fatte durante l'esame, la loro posizione nello spazio tramite una foto in gray scale del fundus retinico e un puntatore, il quale indica il rispettivo punto di acquisizione dell'immagine rispetto all'intero occhio[fig.13].

Le immagini sono acquisite con una risoluzione 512x496.



Figura 12: immagine dello spectralis HRA+OCT (tratto dal sito di spectralis)

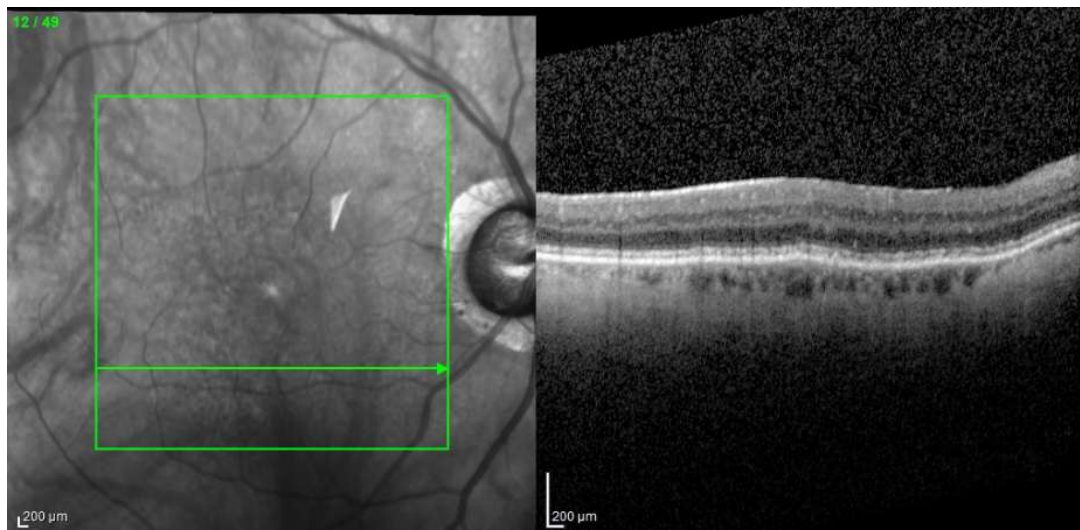


Figura 13: esempio di immagine OCT catturata dal Heidelberg Spectralis HRA+OCT

2.2 Dataset utilizzato

Il dataset utilizzato in questa tesi proviene dalla banca dati dei pazienti in cura al *Department of Ophthalmology, "City of Health and Science" Hospital* di Torino.

I pazienti sono stati divisi in quattro classi dai medici, tre appartenenti ai diversi stadi della patologia e una rappresentata dalle persone sane, usate come confronto. In totale si sono analizzati 739 pazienti per un totale di 36.211 immagini processate, divise nelle seguenti classi [Tab. 2]:

stadio	Numerosità Pazienti	Numerosità immagini
Sani	163	7.987
Intermedio	272	13.328
Atrofia Geografica	150	7350
Neovascolare	154	7546
Totale	739	36.211

Tabella 2: Numero di pazienti e immagini elaborate per ogni stadio

Il dataset utilizzato non contiene dati sensibili e riconducibili all'identità del paziente analizzato, in quanto è stato mandato dall'ospedale già con un ID progressivo e la classe di appartenenza. Esso contiene una forte variabilità sia nell'orientazione [Fig.14 a/b/c], sia nel posizionamento dell'immagine [Fig.15 a/b/c] che nella distribuzione del rumore sullo sfondo [Fig.16 a/b/c]. In questo caso, la variabilità osservata può essere ricondotta sia a quella inter-soggetto, sia a quella intra-soggetto. Infatti, da un lato le immagini provengono da pazienti diversi, che possono reagire in maniera differente al macchinario (per esempio alcuni mantengono la calma mentre altri risultano più agitati) o avere fisiologie diverse; dall'altro lato, ci

sono immagini dello stesso paziente acquisite in momenti differenti, in cui la progressione della malattia può determinare una risposta diversa allo stesso tipo di esame.

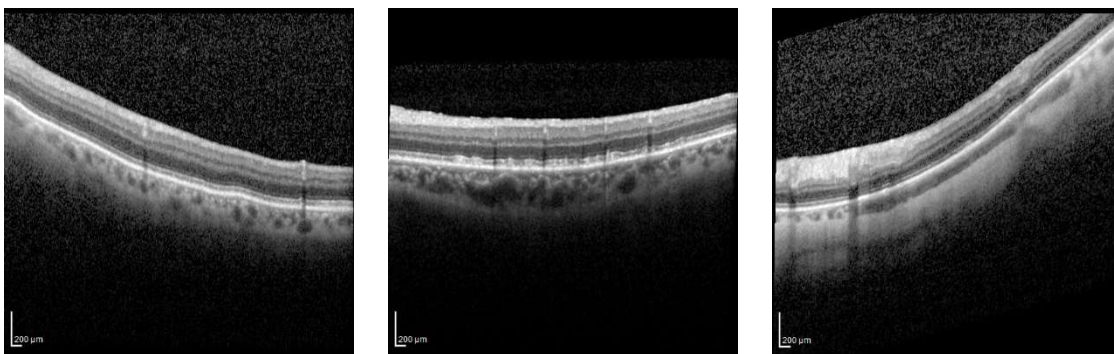


Figura 14a/b/c: variabilità d'orientamento dell'immagine

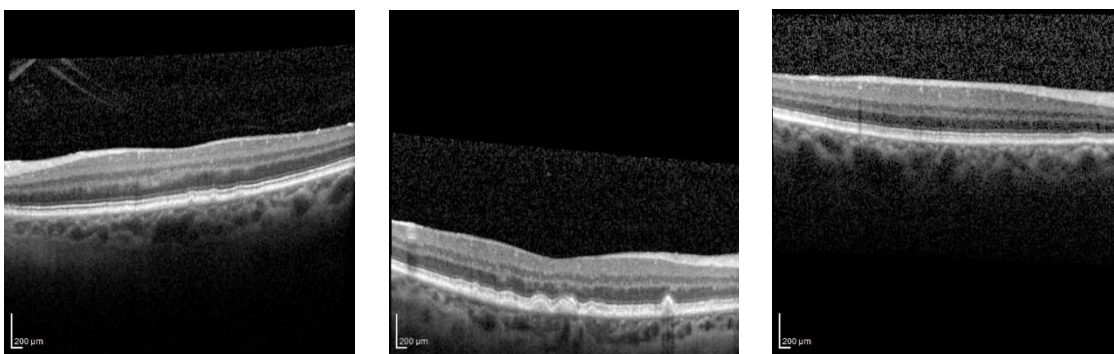


Figura 15a/b/c/ variabilità di posizione dell'immagine

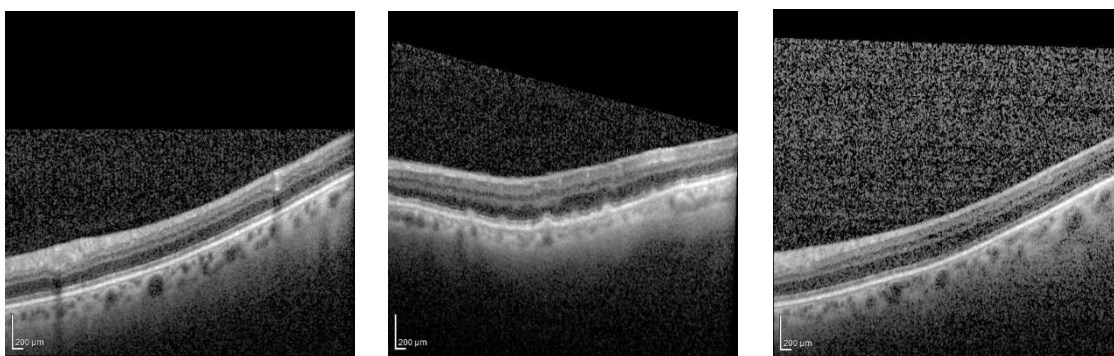


Figura 16a/b/c/ variabilità di rumore

2.3 Preprocessing

Il primo passo è stato quello di prendere tutte le immagini contenute nella cartella di ogni paziente ed elaborarle al fine di isolare esclusivamente la regione di interesse relativa all'OCT (Optical Coherence Tomography), in particolare l'immagine della parte destra, attraverso un processo di ritaglio mirato e, successivamente, è stato applicato dello zeropadding per garantire che tutte le immagini avessero una dimensione quadrata uniforme di 512x512 pixel. Operazione necessaria per standardizzare tutte le immagini prima dei successivi processi di filtraggio e segmentazione, facilitando così l'elaborazione automatizzata e riducendo le possibili variazioni dovute a dimensioni differenti.

Una volta ottenute le immagini preliminari, è stato applicato un processo di filtraggio per ridurre il rumore di tipo "sale e pepe", particolarmente marcato nelle immagini OCT. In particolare, sono stati testati e applicati diversi filtri in cascata per garantire una pulizia ottimale dell'immagine:

- **Filtro Gaussiano**, con sigma pari a 0.5, per attenuare il rumore senza compromettere eccessivamente i dettagli strutturali degli strati retinici.
- **fastNlMeansDenoising**, che utilizza l'algoritmo Non-Local Means Denoising. Questo metodo sostituisce il valore di ciascun pixel con la media pesata dei pixel contenuti in patch simili nelle vicinanze. Il filtro è stato applicato con un'intensità di 7 su patch di dimensioni 7x7, considerando un'area di ricerca attorno al pixel di 21x21. L'utilizzo di questo approccio permette di preservare i bordi e i dettagli sottili, garantendo al contempo una riduzione efficace del rumore.

Una volta ottenute le immagini filtrate, sono state create in parallelo due versioni principali per la successiva fase di segmentazione:

1. **Segmentazione tramite Gaussian Mixture Model (GMM):** in questa fase, gli strati retinici sono stati segmentati in scala di grigi utilizzando sette livelli di intensità. Successivamente, i valori dei livelli di grigio sono stati rimappati in modo coerente verso valori più chiari (43, 50, 85, 128, 170, 213, 255). Questo processo consente di aumentare il contrasto tra gli strati retinici e il background, oltre a mantenere una coerenza tra le immagini dello stesso occhio, associando a ciascun strato retinico un livello di grigio univoco e facilmente riconoscibile in tutte le immagini.
2. **Identificazione della Region Of Interest (ROI):** attraverso l'applicazione in cascata del filtro **Mean Shift** e l'**Otsu thresholding**, è stata creata una maschera binaria in cui i valori 1 identificano la retina e i valori 0 il background circostante. Questo passaggio permette di isolare con precisione l'area della retina dal resto dell'immagine, enfatizzando le strutture di interesse e riducendo ulteriormente il rumore presente nello sfondo.

Successivamente, le segmentazioni ottenute tramite GMM e le maschere binarie sono state combinate per ottenere la **segmentazione semantica della retina**. Il risultato finale è un'immagine in cui il background è completamente nero, mentre gli strati retinici sono rappresentati da sette livelli di grigio distinti e coerenti tra le immagini. Questo approccio permette non solo di ridurre significativamente il rumore di tipo "sale e pepe" presente soprattutto nello sfondo, ma anche di

aumentare il contrasto tra la retina e il background, facilitando l'analisi visiva e computazionale delle strutture retiniche.

In sintesi, il filtraggio, la segmentazione e la rimappatura dei livelli di grigio consentono di ottenere dataset di immagini OCT altamente standardizzati e privi di rumore, con una segmentazione semantica accurata della retina che può essere utilizzata per l'analisi automatizzata o addestramento di algoritmi di intelligenza artificiale. Questo flusso di lavoro garantisce un equilibrio ottimale tra qualità dell'immagine e conservazione dei dettagli anatomici [Fig. 17].

Pipeline Preprocessing

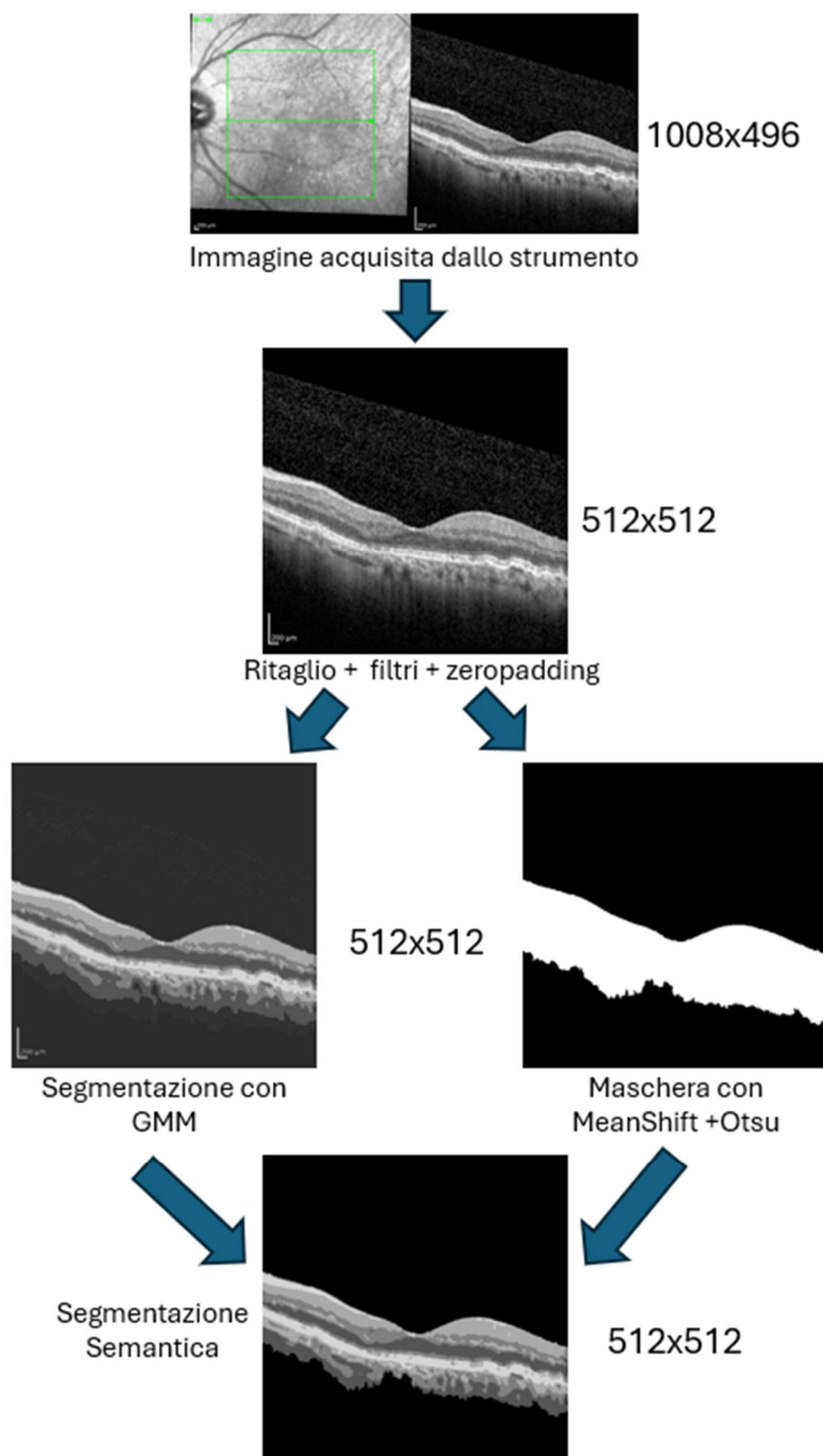


Figura 17: Pipeline di Preprocessing

2.3.1 Algoritmi utilizzati

-Filtro Gaussiano

Questo filtro è uno dei filtri più diffusi nell'ambito dell'imaging. Esso consiste in un filtro passa-basso non uniforme utilizzato per ridurre il rumore tramite l'attenuazione delle alte frequenze, quindi riducendo la distorsione dei segnali più bassi e più alti con un effetto di appiattimento dell'immagine [Fig.18].

il Kernel è definito in questo modo.

$$G(x, y, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right)$$

Nel caso di questa tesi σ_x e σ_y hanno lo stesso valore e rappresentano la deviazione standard, responsabile della larghezza della distribuzione Gaussiana voluta nella rispettiva direzione [19],[20].

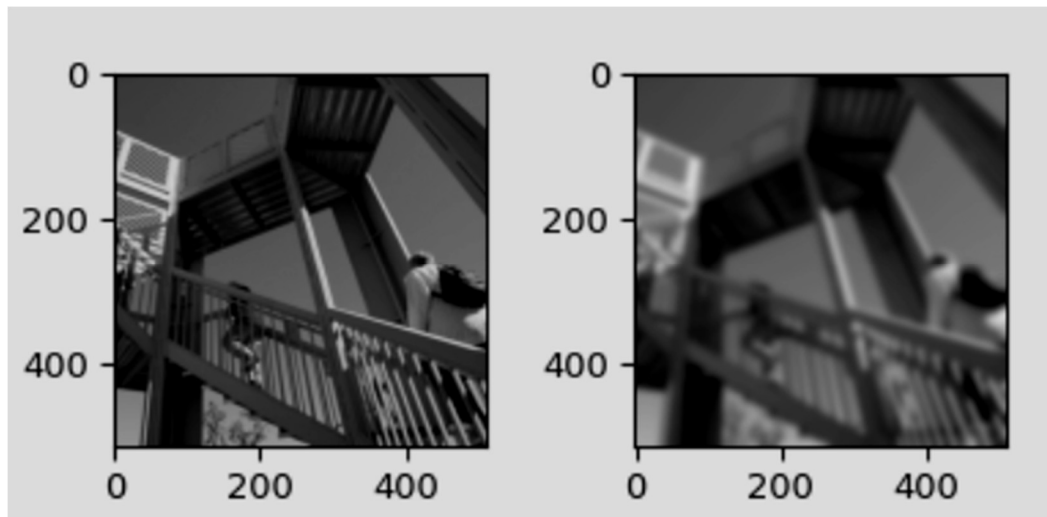


Figura 18: esempio applicazione del filtro gaussiano [19]

-Fast NL Means Denoising

Questo algoritmo si basa sul concetto dei Non Local Means (NLM). Questo metodo si basa sul principio di sostituire il colore di un pixel con una media dei colori di pixel simili in una vasta porzione di dell'immagine.

La variante fast introduce la scelta della dimensione massima di scansione così da non dover scansionare tutta l'immagine ma solo una porzione di essa per ogni pixel.

La rimozione del rumore del pixel viene effettuata calcolando il colore medio dei pixel più simili all'interno della finestra di scansione secondo la seguente formula:

$$NLu(p) = \frac{1}{C(p)} \int f(d(B(p), B(q))) u(q) dq,$$

Dove $d(B(p), B(q))$ è una distanza euclidea tra patch di immagini centrate rispettivamente in p e q , f è una funzione decrescente e $C(p)$ è il fattore di normalizzazione.

Questo algoritmo sul fondamento che nel calcolare la distanza euclidea $d(B(p), B(q))$, tutti i pixel nella patch $B(p)$ hanno la stessa importanza e quindi il peso $f(d(B(p), B(q)))$ può essere utilizzato per rimuovere il rumore da tutti i pixel nel patch $B(p)$ e non solo da p .

Per calcolare la distanza euclidea di un'immagine $u = (u_1, u_2, u_3)$ e un certo pixel p si usano le formule

$$\hat{u}_i(p) = \frac{1}{C(p)} \sum_{q \in B(p,r)} u_i(q) w(p, q), \quad C(p) = \sum_{q \in B(p,r)} w(p, q)$$

Nel caso in esame le immagini sono in scala di grigio quindi avremo una sola componente u mentre la componente $B(p, r)$

indica una regione centrata in p di dimensioni $(2r + 1) \times (2r + 1)$ nel caso in esame è stata usata una finestra di scansione di 21×21 .

I pesi $w(p, q)$ a loro volta dipendono dalla distanza euclidea al quadrato d^2 definita come:

$$d^2 = d^2(B(p, f), B(q, f))$$

Delle due patches centrate in p e q .

Esplicitando l'ultima formula si ottiene:

$$d^2(B(p, f), B(q, f)) = \frac{1}{3(2f + 1)^2} \sum_{i=1}^3 \sum_{j \in B(0, f)} (u_i(p + j) - u_i(q + j))^2$$

Da questa formula si può notare come il valore di ogni pixel viene riscritto come media dei pixel più simili.

L'algoritmo usa un kernel esponenziale per il calcolo dei pesi definito nel modo seguente:

$$w(p, q) = e^{-\frac{\max(d^2 - 2\sigma^2, 0.0)}{h^2}}$$

Dove σ definisce la deviazione standard e h^2 è il parametro di filtraggio che viene chiesto in input (nel caso di questa tesi è settato a sette per un filtraggio alto).

Inoltre, se il valore della distanza al quadrato è più alta di σ^2 il valore del peso viene settato a 1 in modo da non far diminuire eccessivamente il valore del peso.

Infine, Il peso del pixel di riferimento p nella media è impostato al massimo dei pesi nel vicinato $B(p, r)$ per non creare un'eccessiva ponderazione del denoising.

Applicando la procedura di calcolo della media sopra descritta, otteniamo un valore di pixel p con attenuamento del rumore per ogni pixel [21], [22], [Fig.19].

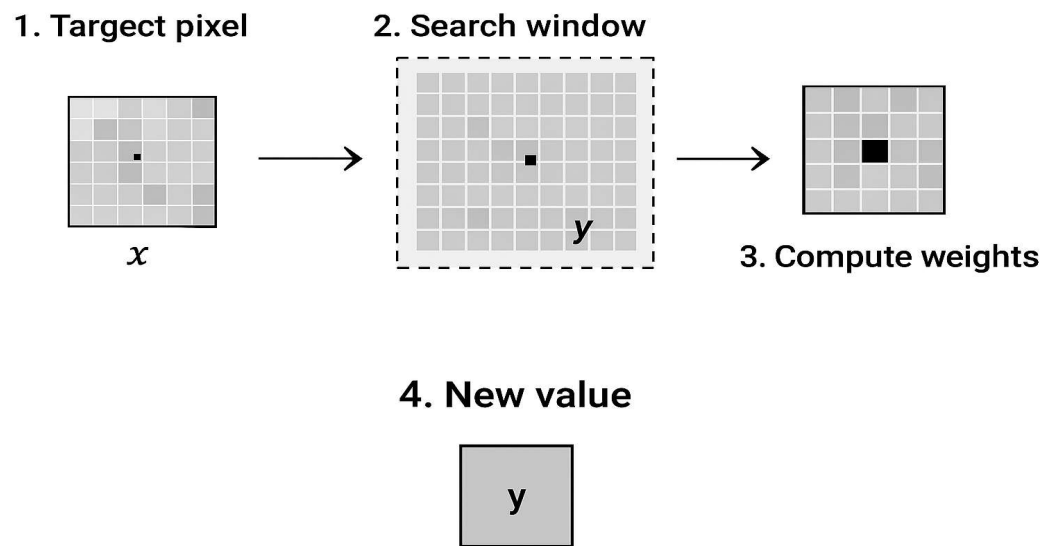


Figura 19: processo di denoising con l'algoritmo non-local means

-Gaussian Mixture Model (GMM)

Questo algoritmo è un modello di densità che rappresenta la distribuzione dei dati come somma pesata di K distribuzioni Gaussiane caratterizzate da tre parametri:

- **Peso di mixing (π_k)** calcolati in modo che la loro sommatoria sia 1 e che siano tutti positivi.
- **Media (μ_k)**
- **Matrice di covarianza (Σ_k)**

La densità di probabilità viene così definita:

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x} | \mu_k, \Sigma_k)$$

Gli algoritmi GMM sono principalmente utilizzati per fare modellazione di segnali, stima di densità o, come nel caso di questa tesi, clustering perché riescono ad approssimare una distribuzione complessa ad una somma di gaussiane [23][25].

Per stimare i parametri di K l'algoritmo GMM si basa sul metodo di massima verosimiglianza (MLE) che cerca di massimizzare la seguente formula:

$$\ln p(\mathbf{X} | \pi, \mu, \Sigma) = \sum_{n=1}^N \ln \left(\sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}^{(n)} | \mu_k, \Sigma_k) \right)$$

Per via della somma interna è molto difficile effettuare un'ottimizzazione diretta per cui per massimizzare questa espressione ricorre all'algoritmo EM (Expectation-Maximization)

che sfrutta il concetto di variabili latenti che indicano quale componente ha generato ogni punto.

Il principio di funzionamento di questo algoritmo consiste nell'inserire variabili latenti $z^{(n)}$ in modo da portare la formula precedente ad avere questa forma:

$$\begin{aligned}\ell(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \ln p(\mathbf{X}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^N \ln p(\mathbf{x}^{(n)}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \sum_{n=1}^N \ln \sum_{z^{(n)}=1}^K p(\mathbf{x}^{(n)}|z^{(n)}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) p(z^{(n)}|\boldsymbol{\pi})\end{aligned}$$

Per poter stimare i valori di $z^{(n)}$ l'algoritmo esegue tre fasi e alterna l'esecuzione delle ultime due fino ad arrivare a convergenza [23]:

- **Inizializzazione:** fase in cui l'algoritmo da dei valori arbitrari di π_k, μ_k, Σ_k da cui partire
- **E-step (Expectation):** in questa fase l'algoritmo calcola la *responsibilities* (γ_k) su z , ovvero la probabilità a posteriori, che ci dice quanto pensiamo abbia contribuito ogni Gaussiana K a generare quel punto.

$$\begin{aligned}\gamma_k = p(z = k|\mathbf{x}) &= \frac{p(z = k)p(\mathbf{x}|z = k)}{p(\mathbf{x})} \\ &= \frac{p(z = k)p(\mathbf{x}|z = k)}{\sum_{j=1}^K p(z = j)p(\mathbf{x}|z = j)} \\ &= \frac{\pi_k \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x}|\mu_j, \Sigma_j)}\end{aligned}$$

Una volta ottenuto il parametro γ_k si implementa nella formula della verosimiglianza attesa ottenendo la formula finale.

$$\begin{aligned}
& \mathbb{E}_{P(z^{(i)}|\mathbf{x}^{(i)})} \left[\sum_i \log(P(\mathbf{x}^{(i)}, z^{(i)}|\Theta)) \right] \\
&= \sum_i \sum_k \left(\gamma_k^{(i)} \log(P(z^i = k|\Theta)) + \log(P(\mathbf{x}^{(i)}|z^{(i)} = k, \Theta)) \right) \\
&= \sum_i \sum_k \gamma_k^{(i)} \left(\log(\pi_k) + \log(\mathcal{N}(\mathbf{x}^{(i)}; \mu_k, \Sigma_k)) \right) \\
&= \sum_k \sum_i \gamma_k^{(i)} \log(\pi_k) + \sum_k \sum_i \gamma_k^{(i)} \log(\mathcal{N}(\mathbf{x}^{(i)}; \mu_k, \Sigma_k))
\end{aligned}$$

- **M-step(maximization):** una volta scoperta la verosimiglianza attesa si aggiornano i parametri π_k, μ_k, Σ_k di ogni gaussiana K , usando i $\gamma_k^{(i)}$ ottenuti, per fare in modo da massimizzare la probabilità che essa generi i parametri di cui è ora responsabile.

I nuovi parametri così ottenuti saranno:

$$\begin{aligned}
\mu_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} \mathbf{x}^{(n)} \\
\Sigma_k &= \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} (\mathbf{x}^{(n)} - \mu_k)(\mathbf{x}^{(n)} - \mu_k)^T \\
\pi_k &= \frac{N_k}{N} \quad \text{with} \quad N_k = \sum_{n=1}^N \gamma_k^{(n)}
\end{aligned}$$

L'algoritmo ripete gli step E ed M fino a che non arriva a convergenza individuando a che gaussiana appartiene ogni punto [23], [24], [Fig.20].

Una volta individuate le gaussiane applica una tonalità di colore a tutti i punti appartenenti a quella gaussiana, così da creare dei cluster [Fig.21].

Questo algoritmo è spesso usato in combinazione con altri metodi di denoising al fine di amplificare l'efficacia del filtraggio [26].

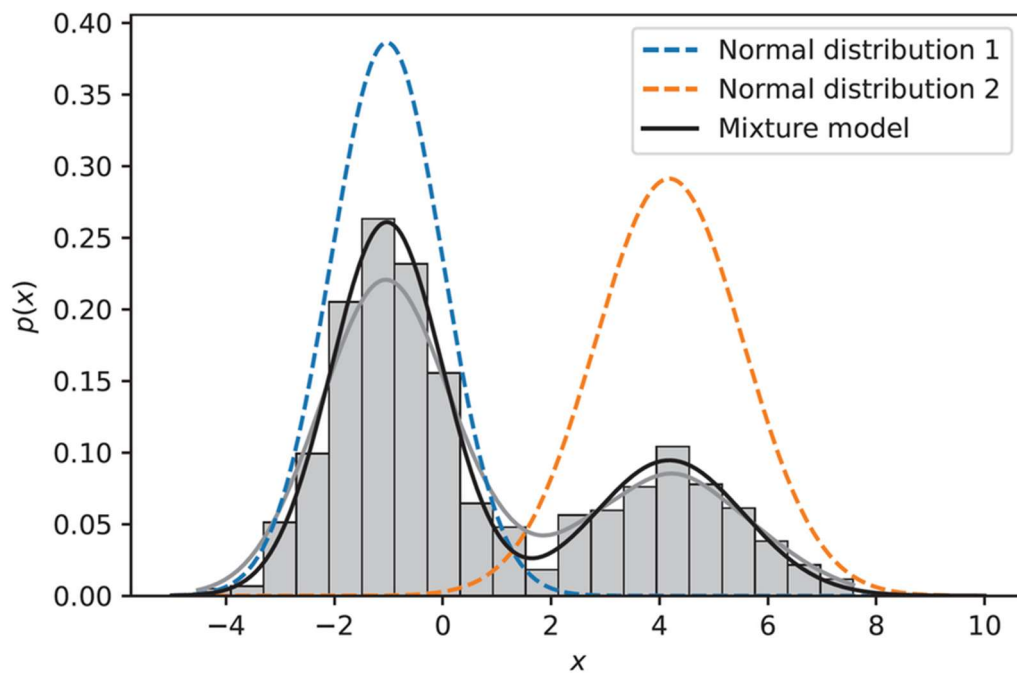


Figura 20: Esempio di scomposizione in gaussiane di una distribuzione [25]

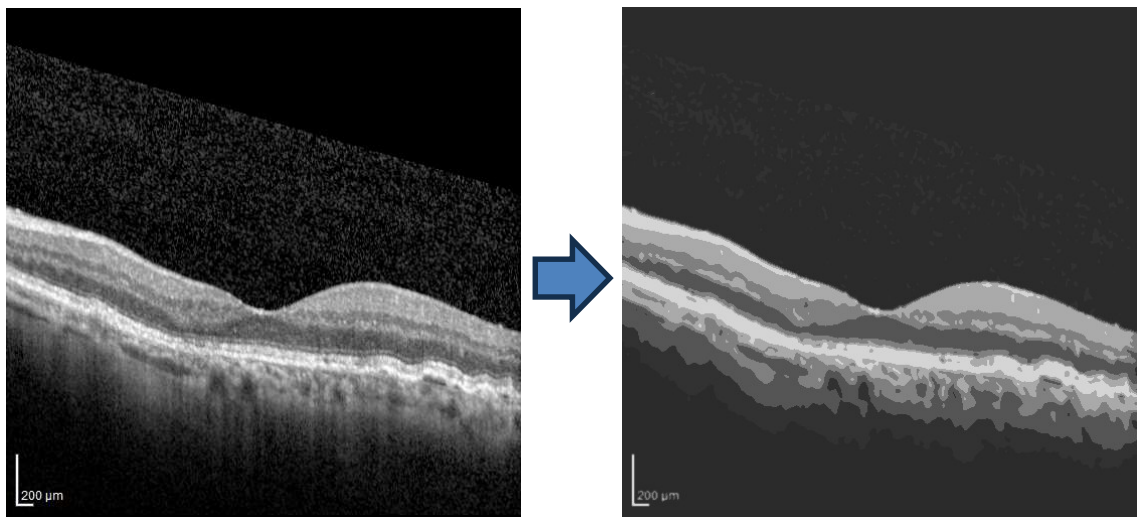


Figura 21: Esempio di risultato dell'algoritmo GMM sulle immagini del dataset

-Filtro Mean shift

È una funzione iterativa di ricerca di moda, in particolare cerca i massimi di densità in uno spazio di dati definito dai parametri sia sul piano dello spazio (PIXEL) sia sul piano delle intensità (piano dei colori).

In particolare, per ogni pixel (X, Y) dell'immagine di input, la funzione esegue iterazioni meanshift, ovvero considera il vicinato del pixel (X, Y) nell'iperspazio congiunto spazio-colore secondo questi parametri:

$$(x, y) : X - \mathbf{sp} \leq x \leq X + \mathbf{sp}, Y - \mathbf{sp} \leq y \leq Y + \mathbf{sp}$$

$$||(R, G, B) - (r, g, b)|| \leq \mathbf{sr}$$

Dove sp sta per *spatial radius* (raggio spaziale) e definisce la finestra nel piano dello spazio e sr sta per *color window radius* (raggio colore) e definisce la soglia di similarità nello spazio RGB, mentre (R, G, B) e (r, g, b) sono rispettivamente i valori delle tre componenti di colore associate al pixel di riferimento e al pixel candidato (nel caso in esame solo la prima componente ha un valore, tutte le altre hanno 0).

Una volta trovati tutti i vicini vengono calcolati il valore spaziale medio (X', Y') e il vettore cromatico medio (R', G', B') , che fungono da centro del vicinato nell'iterazione successiva:

$$(X, Y) (X', Y'), (R, G, B) (R', G', B')$$

Al termine di tutte le iterazioni viene associato al valore del pixel iniziale (che è quello di riferimento), il valore trovato all'ultima iterazione, ovvero la media dei colori trovata all'iterazione finale [27], [Fig.22 a/b/c/d].

$$I(X, Y) < -(R^*, G^*, B^*)$$

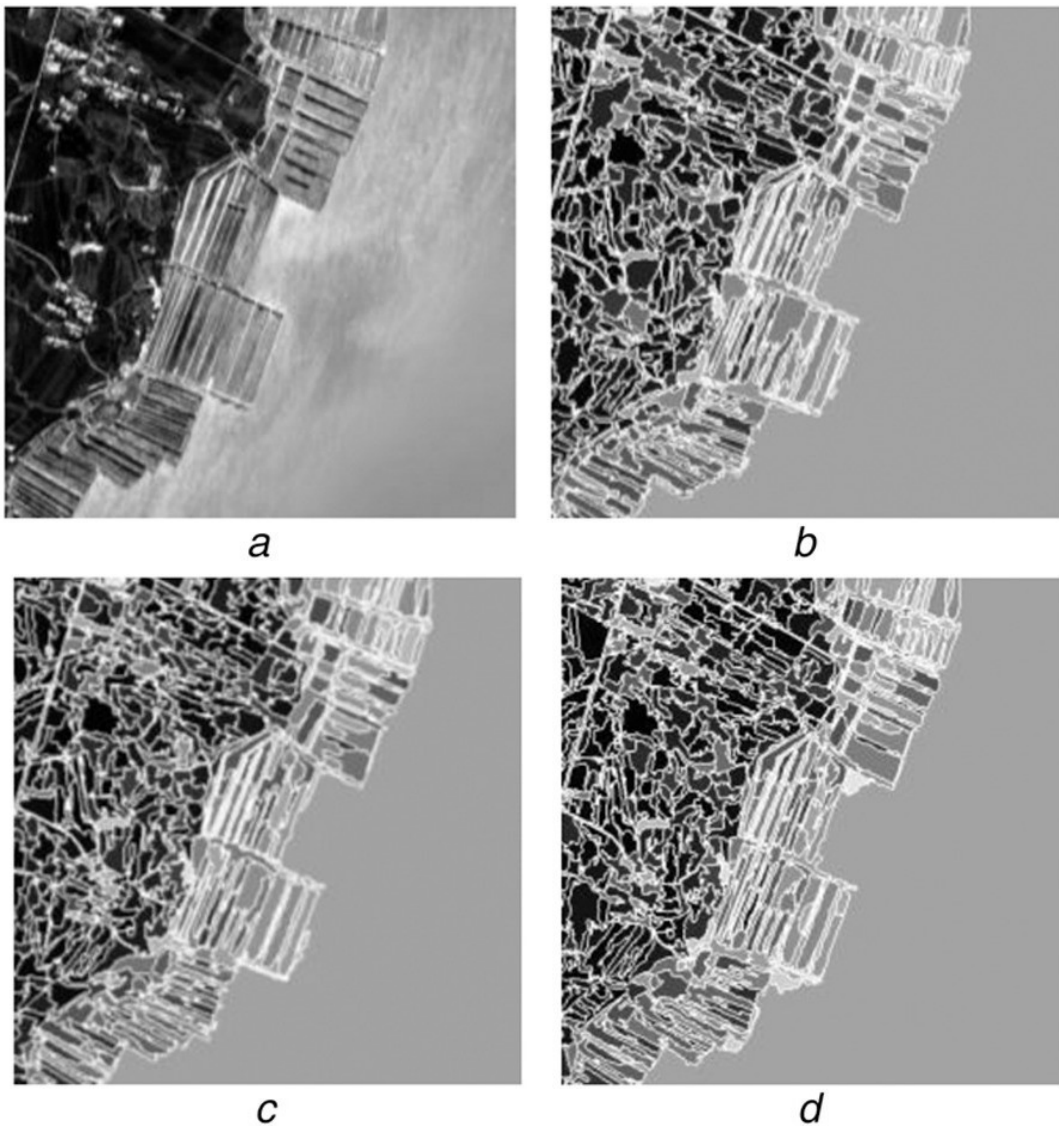


Figura 22 a/b/c/d: esempio dell'effetto del mean shift all'aumentare del raggio spaziale(sp) [28]

-Otsu thresholding

Questo algoritmo viene usato per effettuare la segmentazione binaria di immagini in scala di grigi, in modo da separare lo sfondo (settato a 0) dal soggetto che si vuole enfatizzare (settato a 1) [Fig.23].

Per trovare la soglia ottimale in grado di determinare queste due classi l'algoritmo cerca di minimizzare la varianza intraclassa ponderata, data dalla relazione:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t)$$

Dove:

$$\bullet \quad q_1(t) = \sum_{i=1}^t P(i) \quad \& \quad q_2(t) = \sum_{i=t+1}^I P(i)$$

Sono rispettivamente le probabilità delle due classi

$$\bullet \quad \sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad \& \quad \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)}$$

Sono le varianze delle rispettive classi

$$\bullet \quad \mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad \& \quad \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)}$$

Sono rispettivamente le medie pesate delle classi [29],[30].

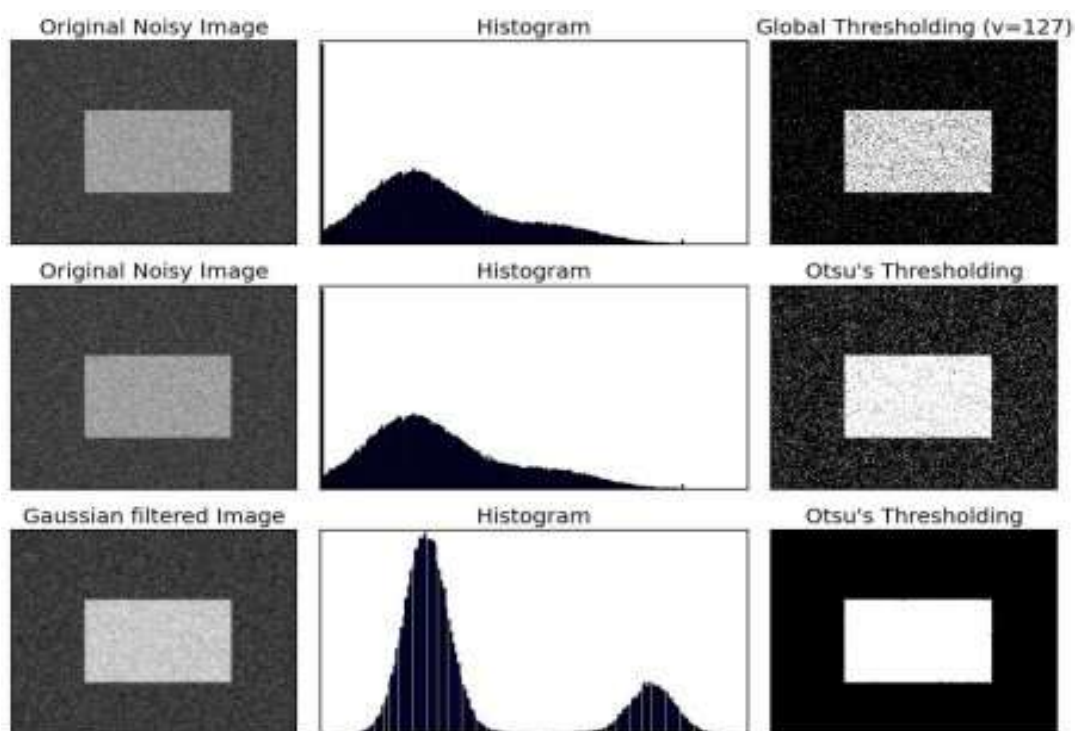


Figura 23: confronto tra le diverse tipologie di processing con Otsu [29]

2.4 Allenamento GAN

Al fine di allenare la GAN, ogni stadio è stato allenato separatamente con una divisione in training, contenente l'80% delle immagini del dataset, e test, contenente il restante 20%. Per comporre le immagini di train, sono state unite le immagini di segmentazione, a sinistra, con quelle originali, a destra [Fig. 24]. Mentre le immagini di test, invece, sono composte unendo le immagini di segmentazione, a sinistra, con un quadrato nero delle stesse dimensioni della segmentazione [Fig. 25].

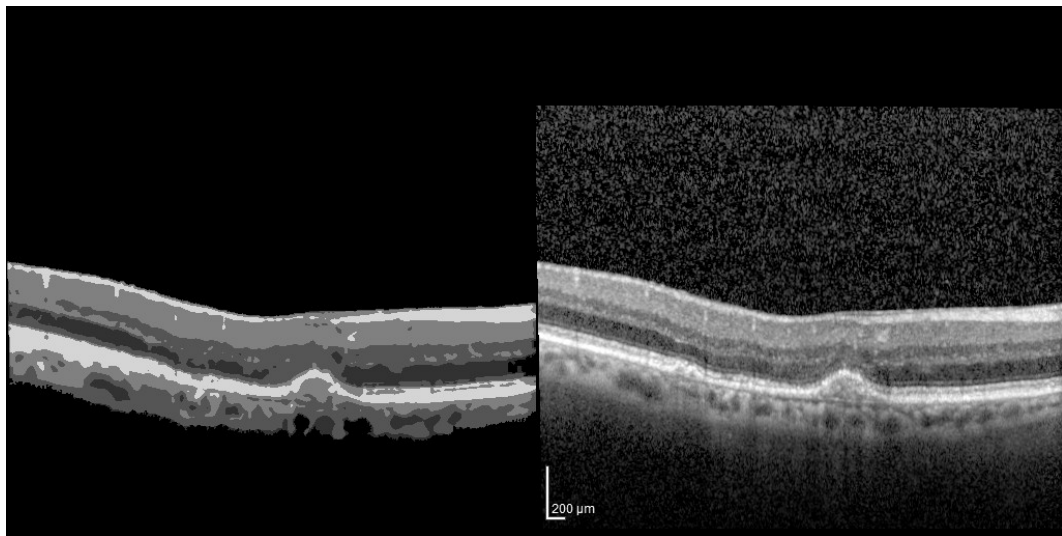


Figura 24: Esempio di un'immagine di train

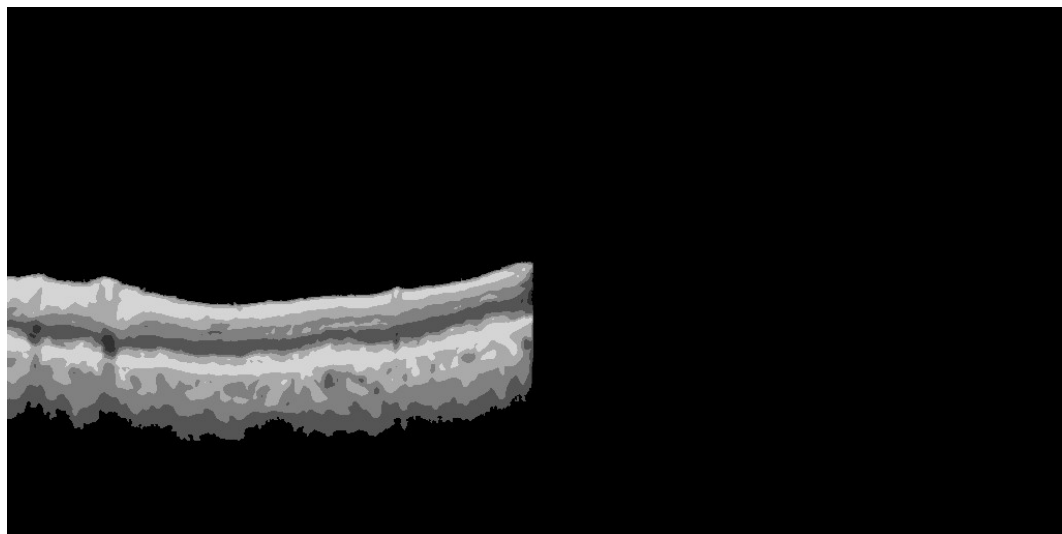


Figura 25: Esempio di un'immagine di Test

La GAN allenata in questa tesi è di tipo piramidale avente le seguenti caratteristiche [Tab.3]:

Componente	Parametro / Configurazione
Generatore	U-net 256x256
Discriminatore	PatchGAN, patch size 70×70
Input/Output	1 canale, dimensione 512×512
Batch size	8
Funzione di perdita	LSGAN + loss percettiva VGG + loss multilivello (L_2, L_3, L_4) + 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 a LR costante + 100 con LR lineare decrescente)

Tabella 3: Caratteristiche della GAN allenata

La rete descritta in tabella è una GAN sviluppata per la generazione di immagini in scala di grigi con dimensione 512×512 pixel. Si tratta di un'architettura ispirata al modello Pix2Pix, basata quindi su una Conditional GAN, in cui il generatore impara a produrre un'immagine di output coerente a partire da un'immagine di input corrispondente. L'obiettivo è quello di ottenere risultati visivamente realistici e strutturalmente fedeli all'immagine originale, combinando l'apprendimento supervisionato con il principio di competizione tipico delle GAN.

- Il **generatore** della rete è una U-Net 256×256 , una struttura encoder-decoder con connessioni di salto (*skip connections*) tra i livelli simmetrici dell'encoder e del decoder. Questa architettura consente di catturare simultaneamente informazioni globali e dettagli locali: l'encoder comprime progressivamente l'immagine riducendone la risoluzione e aumentando la profondità delle

feature, mentre il decoder ricostruisce l'immagine a piena risoluzione sfruttando le informazioni provenienti dall'encoder attraverso le connessioni skip. In questo modo la rete è in grado di mantenere i dettagli spaziali fini e le strutture presenti nell'immagine di partenza, evitando di perdere informazioni importanti durante il processo di codifica. Questa rete è stata usata in quanto particolarmente adatta in contesti in cui è necessario preservare la geometria e i contorni.

- Il **discriminatore** è invece di tipo PatchGAN con dimensione di patch pari a 70×70 . A differenza dei discriminatori tradizionali, che valutano l'intera immagine come reale o generata, il PatchGAN analizza piccole porzioni locali (le patch) e ne valuta la plausibilità. Questa scelta si traduce in un controllo più preciso della qualità dei dettagli delle texture, poiché il discriminatore impara a riconoscere le caratteristiche locali tipiche delle immagini reali, come bordi, transizioni di intensità e pattern strutturali [31].

Applicare un discriminatore di questo tipo migliora le caratteristiche del generatore che è incentivato a produrre immagini non solo globalmente coerenti, ma anche localmente realistiche.

- La **funzione di perdita** complessiva integra diversi fattori che agiscono su livelli complementari. La componente principale è la LSGAN loss (Least Squares GAN), che sostituisce la classica funzione di perdita logaritmica delle GAN con una formulazione ai minimi quadrati. Questo accorgimento contribuisce a rendere l'addestramento più stabile e a migliorare la qualità dei gradienti [32]. A questa si aggiunge una loss percettiva basata sulla rete VGG, che confronta le immagini generate e reali a livello di feature interne piuttosto

che di singoli pixel, permettendo di ottimizzare la qualità visiva percepita [33]. A tutto questo si aggiunge inoltre l'implementazione di una loss multilivello (L_2, L_3, L_4) in modo da assicurare coerenza strutturale tra le immagini su diversi gradi di astrazione, e due termini di regolarizzazione aggiuntivi: uno basato sull'operatore Sobel, utile a preservare i bordi e i contorni, e uno di regolarizzazione convoluzionale, che limita la crescita dei pesi e contribuisce a migliorare la generalizzazione del modello [34].

L'addestramento della rete avviene su batch di dimensione 8, con normalizzazione dei dati tramite Batch Normalization e inizializzazione dei pesi secondo una distribuzione normale [35]. Per l'ottimizzazione è stato scelto l'algoritmo Adam, configurato con un learning rate pari a 0.0002 e un parametro β_1 impostato a 0.5, comunemente utilizzati per garantire un buon equilibrio tra stabilità e velocità di convergenza. Durante il processo di training, la GAN è stata allenata per 200 epoche complessive: con un learning rate costante per le prime 100, mentre per le successive 100 vi è un decremento linearmente, così da favorire una convergenza graduale e ridurre il rischio di oscillazioni nella fase finale.

In sintesi, la GAN proposta combina una U-Net come generatore e un PatchGAN come discriminatore, integrando una funzione di perdita composita che tiene conto sia della fedeltà percettiva sia della coerenza strutturale.

3.4.1 Immagini Generate dalla GAN

Come risultato dell'allenamento, la GAN ha prodotto delle immagini verosimili, con alta verosimiglianza nella regione della retina, ma molto più bassa nel riprodurre il rumore circostante [Fig. 26-29].

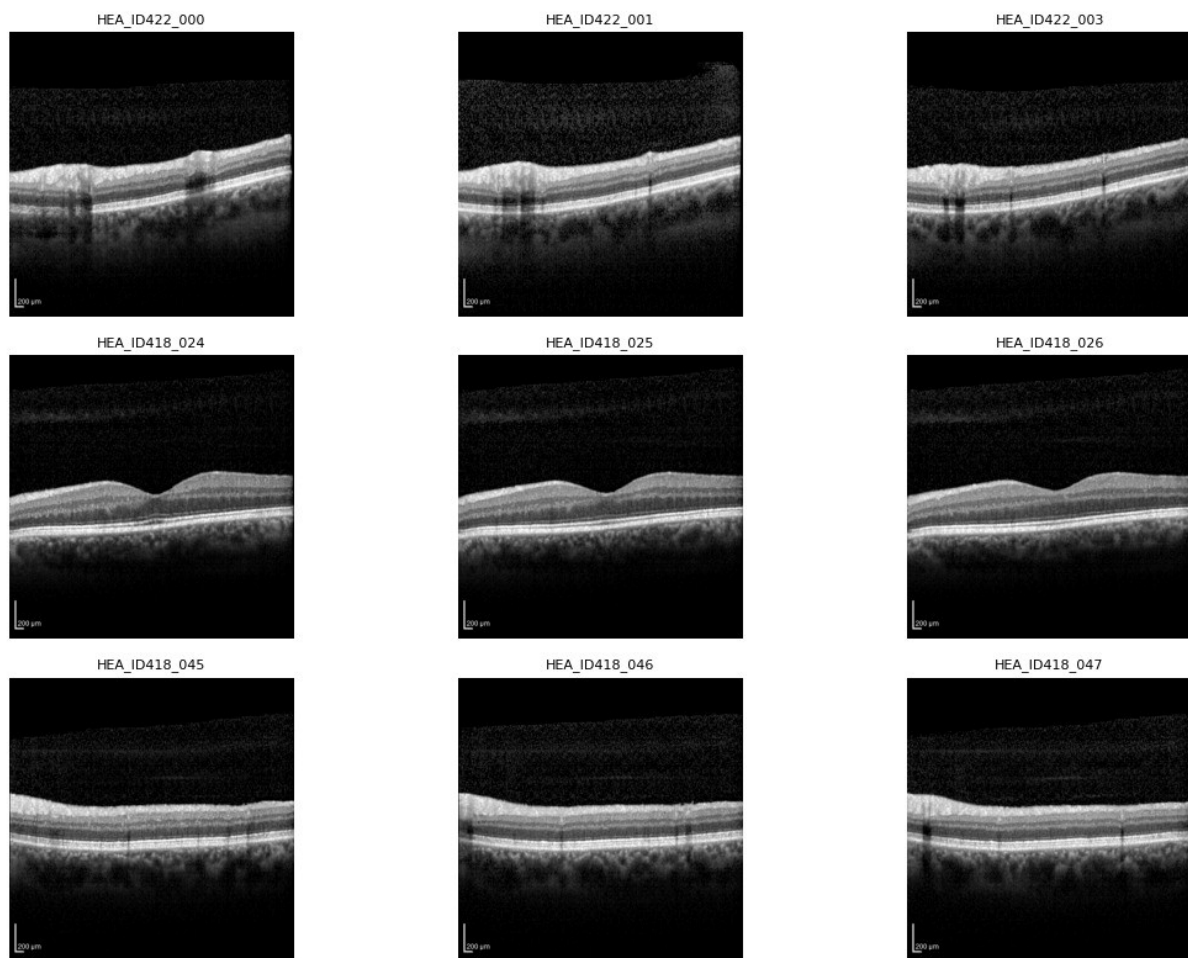


Figura 26: Immagini generate della GAN dopo l'allenamento per lo stadio Sano

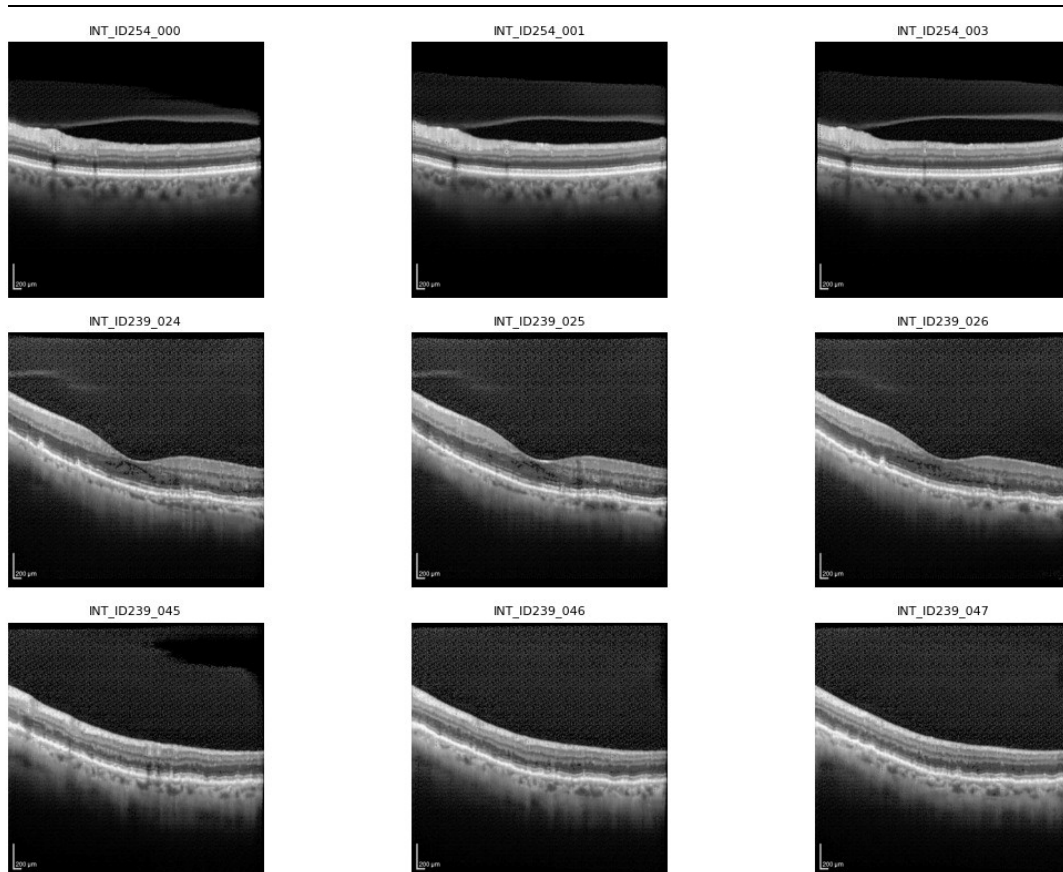


Figura 27: Immagini gen. della GAN dopo l'allenamento per lo stadio Intermedio

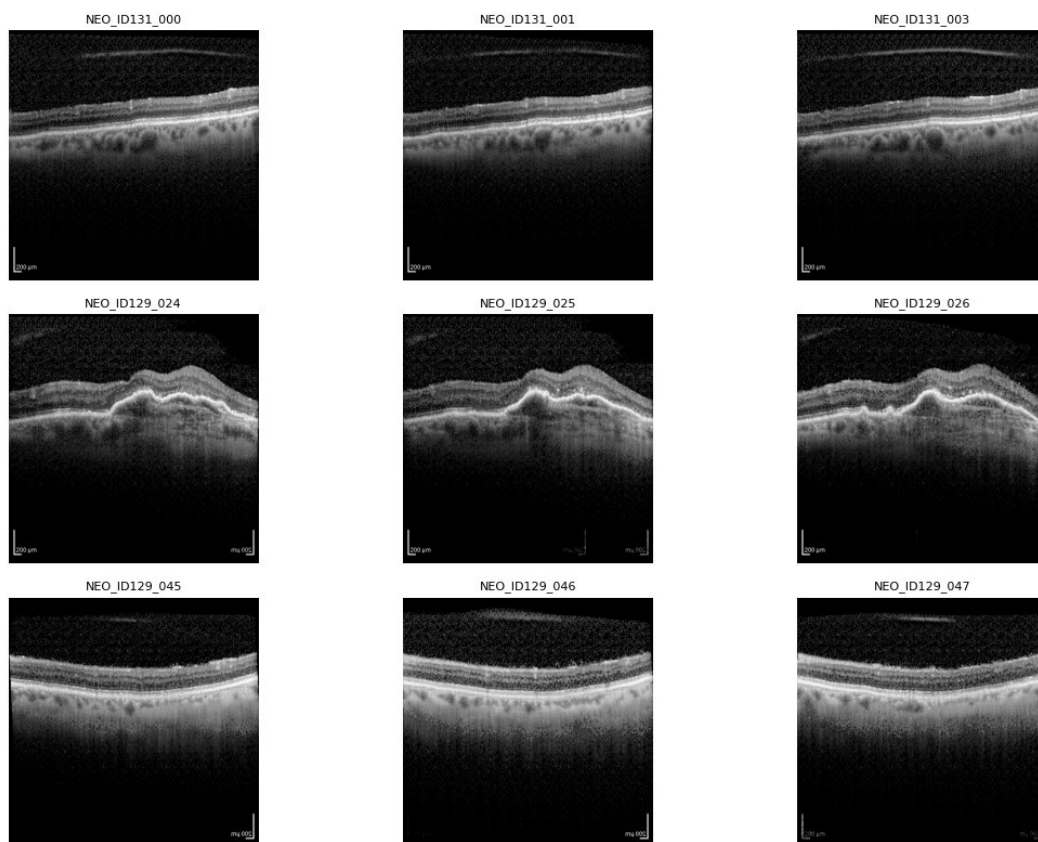


Figura 28: Immagini gen. della GAN dopo l'allenamento per lo stadio Neovascolare

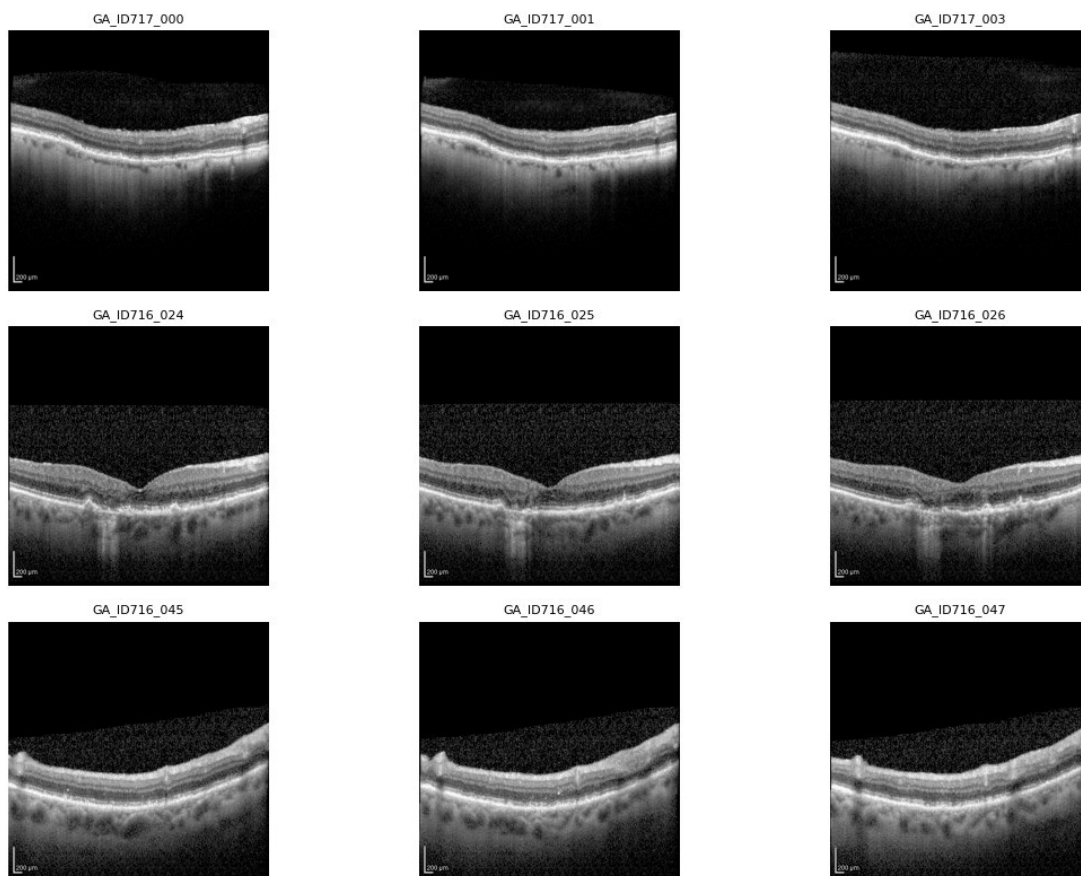


Figura 29: Immagini generate della GAN dopo l'allenamento per lo stadio Atrofia Geografica

Capitolo 3

VALIDAZIONE

Il processo di validazione è stato eseguito sia tramite il calcolo di metriche rilevanti quali *contrast improvement index (CII)*, *Frechet Inception Distance (FID)*, *Structural similarity index measure (SSIM)*, *Peak Signal-to-Noise Ratio (PSNR)* e *analisi tissutale del primo e secondo ordine*, sia da tramite un riscontro diretto dei medici, fornendogli due GUI per la validazione qualitativa delle immagini.

3.1 GUI PER VALIDAZIONE QUALITATIVA

Le GUI sono state create mediante il Tool “MATLAB App Designer” e successivamente fornite ai medici.

Esse permettono di validare le immagini secondo due criteri:

- La prima GUI, chiamata **Valutazione OCT** si presenta nel seguente modo [Fig. 30]:

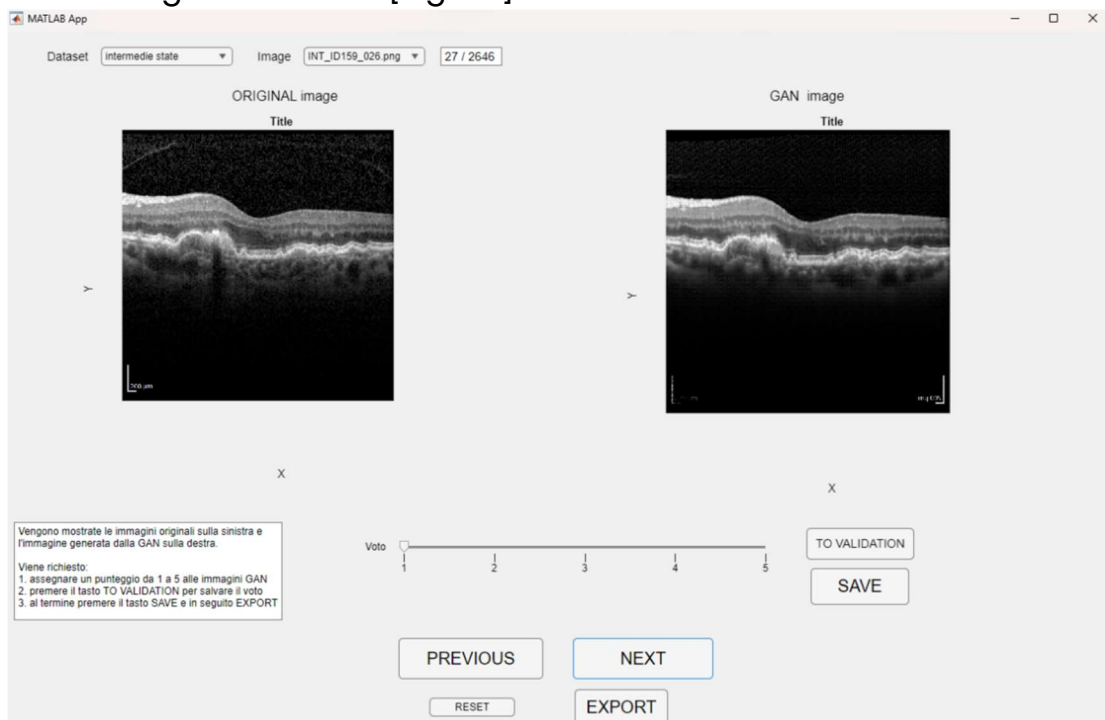


Figura 30: Rappresentazione della GUI, dopo il settaggio dei dataset

appena si apre l'applicazione appare un avviso che chiede di selezionare le cartelle dei dataset [Fig. 31]:

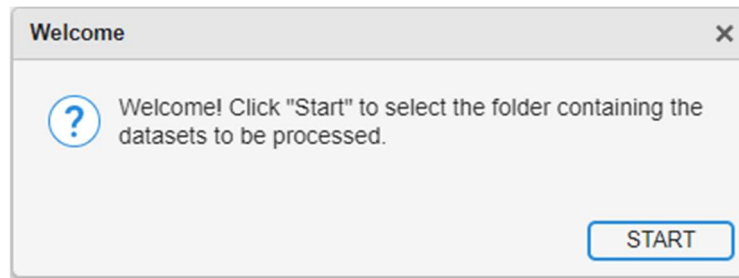


Figura 31: Avviso che appare all'avvio dell'applicazione

Una volta selezionate saranno visibili sulla GUI, a sinistra l'immagine originale, mentre a destra la controparte sintetizzata dalla GAN. Nella parte alta della GUI ci sono due menù a tendina che rendono disponibile la navigazione attraverso gli stadi e attraverso le immagini di quello stadio affiancati da un contatore che indica quale immagine si sta analizzando e il numero d'immagini presente nel dataset [Fig. 32].

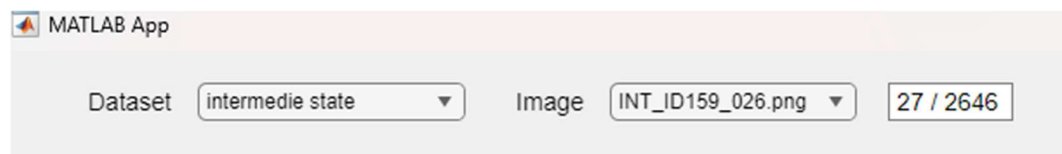


Figura 32: Rappresentazioni dei menu a tendina e del contatore delle immagini

Nella parte bassa, a sinistra vi è una legenda dove poter vedere tutti i passaggi da eseguire per la validazione, al centro uno slider con la possibilità di selezionare un voto da 1 a 5 in base al tasso di verosimiglianza più i bottoni per navigare lungo le immagini avanti e in dietro o per resettare la GUI in caso di problemi ed in fine, per esportare il file excel una volta finita la validazione, a destra vi sono due tasti per il salvataggio dei voti di ogni immagine [Fig. 33].



Figura 33: Rappresentazione parte bassa GUI

Cliccando il bottone export, l'app genera un file Excel con i risultati dei voti dati che riporta il nome dell'immagine, lo stadio a cui appartiene e il voto assegnato [Fig.34].

	A	B	C	D	E	F
1	name	dataset	voto			
2	GA_ID596_004	geographic atrophy	2			
3	EAR_ID455_000	intermediate state	3			
4	HEA_ID297_000	healthy	2			
5	HEA_ID297_001	healthy	4			
6	NEO_ID011_000	neovascular state	3			
7	HEA_ID297_002	healthy	2			
8	HEA_ID297_003	healthy	3			
9	HEA_ID297_004	healthy	5			
10	NEO_ID011_002	neovascular state	3			

Figura 34: Foglio generato dalla GUI coi parametri di validazione “Valutazione OCT”

- La seconda GUI, chiamata **RealOrFake** invece consiste in un piccolo gioco in cui si vede una sola immagine e bisogna dire se è reale o falsa, indicandone anche il grado di certezza, essa è strutturata allo stesso modo della precedente GUI, ma con un'immagine sola e l'impossibilità di scegliere le immagini dal menu a tendina in alto, ma solo lo stadio [Fig. 35].

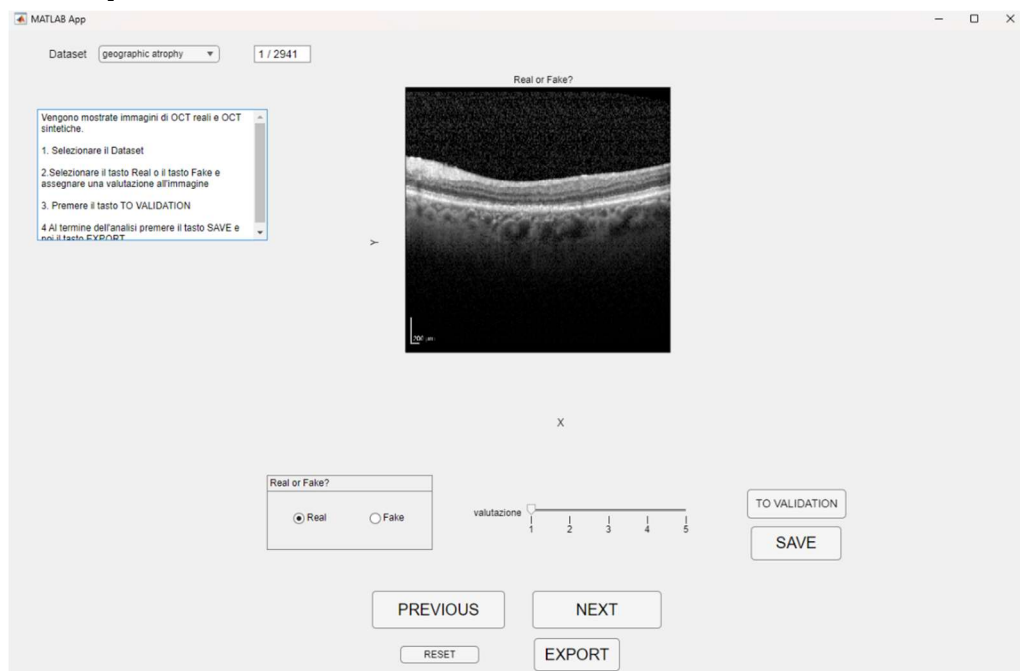


Figura 35: GUI per la validazione blind

A differenza della GUI precedente, quando si preme il tasto "TO VALIDATION" appare un avviso che attesta il salvataggio del voto e visualizza anche il nome dell'immagine così da poter fare avere un riscontro immediato e capire se si ha risposto in modo giusto o no [Fig.36]. Se nel nome è compreso "fake_B" vorrà dire che è generata, se non è presente vuol dire che è quella originale.

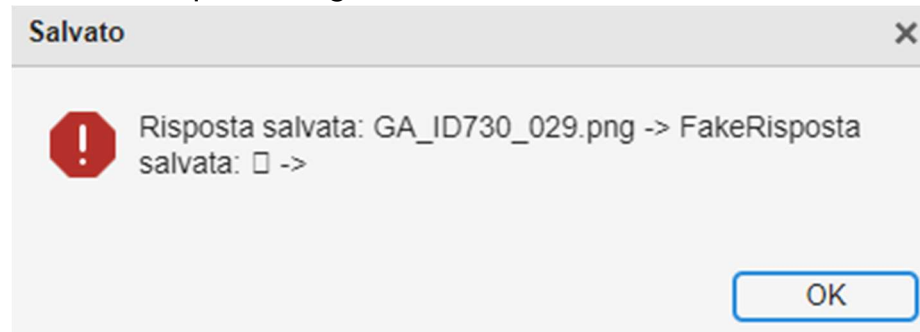


Figura 36: avviso di salvataggio del voto e riscontro diretto dell'immagine

Anche in questo caso, appena si è premuto il tasto export viene generato un file di Excel contenente il nome, la risposta data e il voto assegnato ad essa [Fig.37].

	A	B	C	D	E
1	ImageName	UserResponse	score		
2	GA_ID643_041.png	Fake	4		
3	GA_ID706_035_fake_B.png	Real	3		
4	GA_ID721_015_fake_B.png	Real	3		
5	INT_ID237_019_fake_B.png	Fake	5		
6	GA_ID605_009.png	Real	5		
7	GA_ID692_045.png	Real	5		
8	GA_ID652_014_fake_B.png	Fake	4		
9	GA_ID616_045_fake_B.png	Fake	2		
10	GA_ID642_016.png	Fake	4		
11	NEO_ID011_046_fake_B.png	Fake	5		
12	GA_ID730_029.png	Fake	5		

Figura 37: Foglio generato dalla GUI alla fine del processo di validazione "RealOrFake"

3.1.1 Risultati Validazione Qualitativa

I risultati ottenuti dalla validazione dei medici sono i seguenti per la GUI “Valutazione OCT”, dove si rappresenta la percentuale di voto per stadio [tab. 4] e l’andamento delle percentuali di voti per ogni stadio [Fig.38].

Stadio \ voto	1	2	3	4	5
Sani	0,51%	10,10%	1,14%	1,33%	86,92%
Stadio intermedio	3,28%	16,96%	11,99%	53,86%	13,90%
Atrofia geografica	0,07%	23,87%	1,58%	6,81%	67,68%
Stadio Neovascolare	0,68%	16,51%	9,00%	31,79%	42,02%

Tabella 4: Risultati ottenuti dalla validazione qualitativa della GUI “Valutazione OCT”

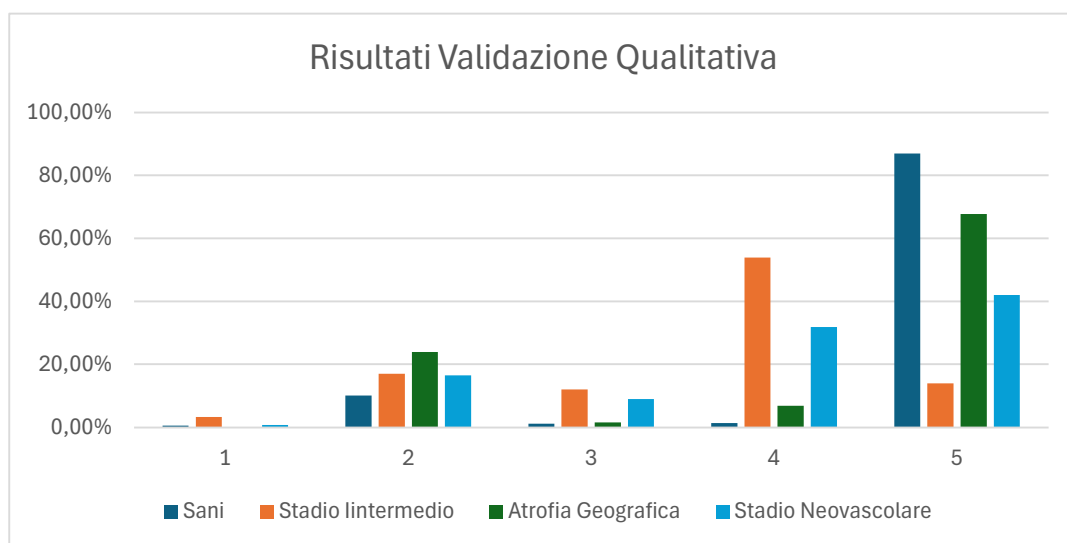


Figura 38: Risultati ottenuti dalla validazione qualitativa, raggruppati per voto della GUI “Valutazione OCT”

Di seguito vengono rappresentati i risultati ottenuti per la GUI "RealOrFake" in cui vengono riportati i valori di selezione per le immagini reali [Tab.5] e per le immagini sintetiche [Tab.6].

Risposta \ voto	3	4	5	Σ	%
Reale	4	3	25	32	33%
Falsa	5	1	58	64	67%
Σ	9	4	83	96	-
%	9%	4%	87%	-	-

Tabella 5: Risultati della validazione qualitativa sulle immagini reali, mediante GUI "RealOrFake"

Risposta \ voto	2	3	4	5	Σ	%
Reale	3	6	4	40	53	70%
Falsa	2	2	1	18	23	30%
Σ	5	8	5	58	76	-
%	6%	11%	6%	77%	-	-

Tabella 6: Risultati della validazione qualitativa sulle immagini sintetiche, mediante GUI "RealOrFake"

3.2 VALIDAZIONE QUANTITATIVA

Per la validazione qualitativa delle immagini sintetiche sono state adottate diverse metriche, al fine di analizzare in maniera approfondita i molteplici aspetti visivi e strutturali delle immagini. Nel presente paragrafo verranno illustrate nel dettaglio le metriche impiegate e presentati i relativi risultati.

-Contrast Improvement index (CII)

Questa metrica è considerata un importante parametro di riferimento che aiuta a confrontare le prestazioni delle tecniche di image enhancement. Essa permette di calcolare il miglioramento del contrasto attraverso la formula:

$$CII = \frac{C_{Proposed}}{C_{Original}}$$

Dove C rappresenta il valore medio del contrasto locale misurato in una finestra di 3x3 e $C_{proposed}$ rappresenta il valore medio di contrasto nell'immagine generata mentre $C_{original}$ rappresenta il valore medio di contrasto riferito all'immagine originale [36].

Un'immagine con un contrasto migliorato rende più semplice l'individuazione di patologie di piccole dimensioni, indice di stadi iniziali, e poterle trattare tempestivamente così da aumentare l'efficacia del trattamento [36].

-Fréchet inception distance (FID)

Questa metrica è diventato uno standard per la valutazione della qualità delle immagini sintetiche mediche, grazie alla sua capacità di simulare il giudizio umano. Essa utilizza la distanza di Fréchet tra due gaussiane multivariate costituite a partire da rappresentazioni estratte da una rete pre-allenata [37].

In questa tesi è stata usata la libreria di Pytorch "*TorchMetrics*" per calcolare il FID usando il comando "*FrechetInceptionDistance()*". Essa utilizza la seguente formula per calcolare il FID:

$$FID = \|\mu - \mu_w\|^2 + tr(\Sigma + \Sigma_w - 2(\Sigma\Sigma_w)^{\frac{1}{2}})$$

Ovvero la distanza tra le due distribuzioni normali multivariate delle feature, calcolate sulle immagini originali $\mathcal{N}(\mu, \Sigma)$ e calcolate sulle immagini sintetiche $\mathcal{N}(\mu_w, \Sigma_w)$ [38].

Nonostante il FID sia un buon parametro per attestare la qualità di un'immagine sintetica, recenti studi hanno dimostrato che esso è anche affetto da forti BIAS e non è in grado di rilevare se una GAN ha memorizzato il set di addestramento [39].

-Structural Similarity Index Measure (SSIM)

Questa metrica è principalmente usata in medicina per calcolare la similarità percepita. Essa si concentra maggiormente sulla trama dell'immagine, analizzando il degrado della qualità sulle strutture visibili [40]. In questa tesi si è utilizzata la funzione *StructuralSimilarityIndexMeasure()* della libreria *TorchMetrics* che implementa la seguente formula [41]:

$$SSIM = \frac{(2\sigma_{nc} + c_2)(2\mu_n\mu_c + c_1)}{(\mu_n^2 + \mu_c^2 + c_1)(\sigma_n^2 + \sigma_c^2 + c_2)}$$

Dove:

- σ_n e σ_c sono rispettivamente le deviazioni standard delle coppie di patch d'immagini "rumore (n) e pulite (c)" che nel caso di questa tesi si associa a n le immagini originali e a c le immagini generate.
- σ_{nc} rappresenta la covarianza tra le due patch d'immagini
- μ_n e μ_c rappresentano invece i valori assoluti delle rispettive patch analizzate
- c_1 e c_2 sono delle costanti stabilizzatrici quando le somme degli altri contributi si avvicina molto allo 0, definite in questo modo:

$$C_1 = (k_1L)^2, \quad C_2 = (k_2L)^2$$

- Con k_1 e k_2 sono delle costanti di piccole dimensioni, per *TorchMetrics* sono predefinite rispettivamente a 0.01 e 0.03.
- L rappresenta il range dinamico del valore di pixel, nel caso di questa tesi 255 in quanto si analizzano immagini in 8 bit [40],[41],[42].

- *Peak Signal-to-Noise Ratio (PSNR)*

Questa metrica è spesso usata insieme all'SSIM e valuta la qualità di un'immagine ricostruita confrontandola con una di riferimento, fornendo una misura quantitativa della fedeltà e della somiglianza tra le due immagini [20].

In questa tesi è stata applicata la funzione *PeakSignalNoiseRatio()* dalla libreria *TorchMetrics* che implementa la seguente formula [43]:

$$\text{PSNR}(I, J) = 10 * \log_{10} \left(\frac{\max(I)^2}{\text{MSE}(I, J)} \right)$$

Dove:

- I, J rappresentano le due immagini prese in esame, rispettivamente l'originale e la ricostruita
- $\max(I)^2$ rappresenta il valore di pixel massimo rappresentabile dall'immagine al quadrato (nel caso in esame in questa tesi $\max(I) = 255$).
- $\text{MSE}(I, J)$ rappresenta l'errore quadratico medio tra il valore di pixel dell'immagine originale e il valore di pixel dell'immagine ricostruita espresso come:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (I_i - J_i)^2$$

- Con I_i e J_i , i rispettivi valori dello stesso pixel, preso dall'immagine originale e da quella sintetica [20].

Questa funzione restituisce o il valore medio di tutti i PSNR calcolati o una matrice, con le stesse dimensioni dell'immagine, contenente tutti i valori di PSNR calcolati nei rispettivi pixel [43].

- *Analisi tessutale radiomica del primo e secondo ordine*

L'analisi tessutale (o texture analysis) è una tecnica di radiomica che consente di estrarre informazioni quantitative dai dati di imaging medico per descrivere le caratteristiche strutturali e spaziali dei tessuti attraverso l'analisi della distribuzione dei pixel e la relazione tra loro. L'obiettivo è identificare pattern invisibili all'occhio umano ma significativi da un punto di vista diagnostico [44]. Per effettuare l'analisi tessutale delle immagini originali e sintetiche è stata utilizzata la libreria "Pyradiomics" che contiene all'interno un'ampia gamma di features estraibili tramite analisi tessutale, si è fatta una selezione delle features da estrarre in base all'utilità nel campo della tesi [45][46].

Le features del primo ordine descrivono la distribuzione delle intensità dei voxel all'interno della regione d'interesse, definita da una maschera binaria (ROI). Assumendo X come insieme di voxel N_p contenuti nella ROI, $P(i)$ l'istogramma del primo ordine con N_g livelli discreti di grigio, dove esso rappresenta il numero di bins diversi da zero, ma equidistanti da esso di una distanza pari al parametro binWidth e $p(i)$ l'istogramma normalizzato del primo ordine definito come $\frac{P(i)}{N_p}$; sono state implementate le seguenti features [40]:

- **ENERGIA**

$$energy = \sum_{i=1}^{N_p} (X(i) + c)^2$$

L'energia è la misura della grandezza dei valori dei voxel in un'immagine. Un valore maggiore implica una maggiore somma dei quadrati di questi valori.

- **ENTROPIA**

$$entropy = - \sum_{i=1}^{N_g} p(i) \log_2 (p(i) + \epsilon)$$

specifica l'incertezza/casualità dei valori dell'immagine. Misura la quantità media di informazioni necessarie per codificare i valori dell'immagine.

- **MEAN ABSOLUTE DEVIATION**

$$MAD = \frac{1}{N_p} \sum_{i=1}^{N_p} |\mathbf{X}(i) - \bar{X}|$$

La mean absolute deviation è la distanza media di tutti i valori di intensità dal valore medio dell'array dell'immagine.

- **ROOT MEAN SQUARED**

$$RMS = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) + c)^2}$$

RMS è la radice quadrata della media di tutti i valori di intensità al quadrato. È un'altra misura della grandezza dei valori dell'immagine.

- **SKENWNESS**

$$skewness = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2} \right)^3}$$

misura l'asimmetria della distribuzione dei valori rispetto al valore medio. A seconda di dove si allunga la coda e si concentra la massa della distribuzione, questo valore può essere positivo o negativo.

- **KURTOSIS**

$$kurtosis = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^4}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2 \right)^2}$$

La curtosi è una misura del "picco" della distribuzione dei valori nella ROI dell'immagine. Una curtosi più alta implica che la massa della distribuzione è concentrata verso le code piuttosto che verso la media. Una curtosi più bassa implica il contrario: la massa della distribuzione è concentrata verso un picco vicino al valore medio.

- **UNIFORMITY**

$$uniformity = \sum_{i=1}^{N_g} p(i)^2$$

L'uniformità è una misura della somma dei quadrati di ciascun valore di intensità. Si tratta di una misura dell'omogeneità della matrice di immagini, dove una maggiore uniformità implica una maggiore omogeneità o una minore gamma di valori di intensità discreti.

Le features del secondo ordine sono stati calcolati con il metodo GLCM implementato in “*Pyradiomics*”, usando la matrice di co-occorrenza a livelli di grigio. Questa matrice descrive la funzione di probabilità congiunta del secondo ordine di una regione dell'immagine, vincolata dalla maschera applicata, definita come $P(i, j | \delta, \theta)$. $L'(i, j)$ esimo elemento di questa matrice rappresenta il numero di volte in cui la combinazione dei livelli i e j si verifica in due pixel dell'immagine, separati da una distanza di δ (definita dal voxel centrale secondo la norma infinito) pixel lungo l'angolo θ . Le features implementate sono le seguenti [45]:

- **CONTRASTO**

$$contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j)$$

Il contrasto è una misura della variazione locale dell'intensità, che favorisce i valori lontani dalla diagonale. Un valore maggiore è correlato a una maggiore disparità nei valori di intensità tra voxel vicini.

- **DIFFERENCE AVERAGE**

$$difference\ average = \sum_{k=0}^{N_g-1} k p_{x-y}(k)$$

La media delle differenze misura il rapporto tra le occorrenze di coppie con valori di intensità simili e le occorrenze di coppie con valori di intensità diversi.

- **DIFFERENCE ENTROPY**

$$difference\ entropy = \sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2 (p_{x-y}(k) + \epsilon)$$

La difference entropy è una misura della casualità/variabilità delle differenze di valore dell'intensità del vicinato.

- **DIFFERENCE VARIANCE**

$$difference\ variance = \sum_{k=0}^{N_g-1} (k - DA)^2 p_{x-y}(k)$$

Misura l'eterogeneità e attribuisce un peso maggiore alle coppie di livelli di intensità che si discostano maggiormente dalla media.

- **INVERSE VARIANCE**

$$inverse\ variance = \sum_{k=1}^{N_g-1} \frac{p_{x-y}(k)}{k^2}$$

Dire qualcosa

- **AUTOCORRELAZIONE**

$$autocorrelation = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) ij$$

misura l'entità della finezza e della grossolanità della texture.

- **CORRELAZIONE**

$$correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) ij - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)}$$

È un valore compreso tra 0 (non correlato) e 1 (perfettamente correlato) che mostra la dipendenza lineare dei valori del livello di grigio dai rispettivi voxel nel GLCM.

- **JOINT AVERAGE**

$$joint\ average = \mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) i$$

Restituisce l'intensità media dei livelli di grigio della distribuzione i .

- **JOINT ENERGY**

$$joint\ energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i, j))^2$$

Misura i modelli omogenei nell'immagine. Un'energia maggiore implica che nell'immagine sono presenti più coppie di valori di intensità che si avvicinano tra loro a frequenze più elevate.

- **JOINT ENTROPY**

$$joint\ entropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2 (p(i, j) + \epsilon)$$

Misura la casualità/variabilità dei valori di intensità del vicinato.

- **MAXIMAL CORRELATION COEFFICIENT (MCC)**

$$MCC = \sqrt{\text{second largest eigenvalue of } Q}$$

$$Q(i, j) = \sum_{k=0}^{N_g} \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}$$

Misura la complessità della texture ed è compresa tra 0 e 1.

- **SUM AVERAGE**

$$sum\ average = \sum_{k=2}^{2N_g} p_{x+y}(k)k$$

misura il rapporto tra le occorrenze di coppie con valori di intensità inferiori e le occorrenze di coppie con valori di intensità superiori.

- **SUM ENTROPY**

$$sum\ entropy = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2 (p_{x+y}(k) + \epsilon)$$

È la somma delle differenze dei valori di intensità del vicinato [45].

3.2.1 RISULTATI VALIDAZIONE QUANTITATIVA

Vengono presentati i risultati ottenuti dall'analisi tessutale del primo ordine espressi come errori relativi [Tab. 7] e i relativi andamenti [Fig. 39-45]:

Stadio	Sa	Int	Ga	Neo
Energy	0.033±0.03	0.047±0.03	0.038±0.03	0.041±0.03
Entropy	1.12±2.31	1.62±3.78	0.76±1.44	1.05±2.31
MAD	0.03±0.03	0.03±0.03	0.03±0.02	0.03±0.02
RMS	0.02±2.2	0.02±0.02	0.02±0.02	0.02±0.02
Skewness	0.41±2.2	0.41±1.59	0.51±1.64	0.44±1.49
Kurtosis	0.04±0.04	0.05±0.04	0.04±0.03	0.06±0.05
Uniformity	0.07±0.56	0.02±0.32	0.002±0.02	0.04±0.43

Tabella 7: Risultati ottenuti dall' analisi tessutale del primo ordine con relative deviazioni standard

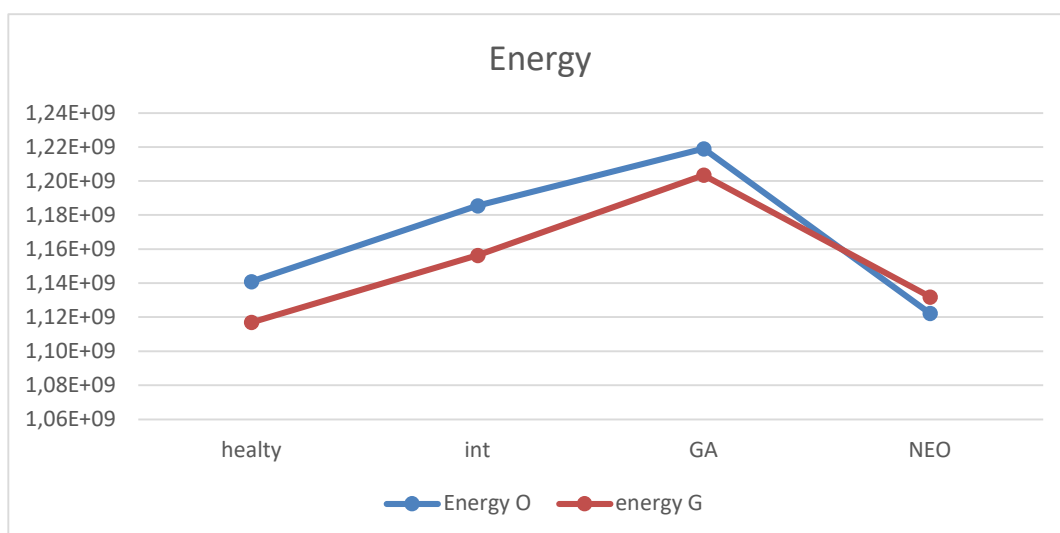


Figura 39: Confronto dell'Andamento dell'Energia lungo gli stadi per le immagini reali e false

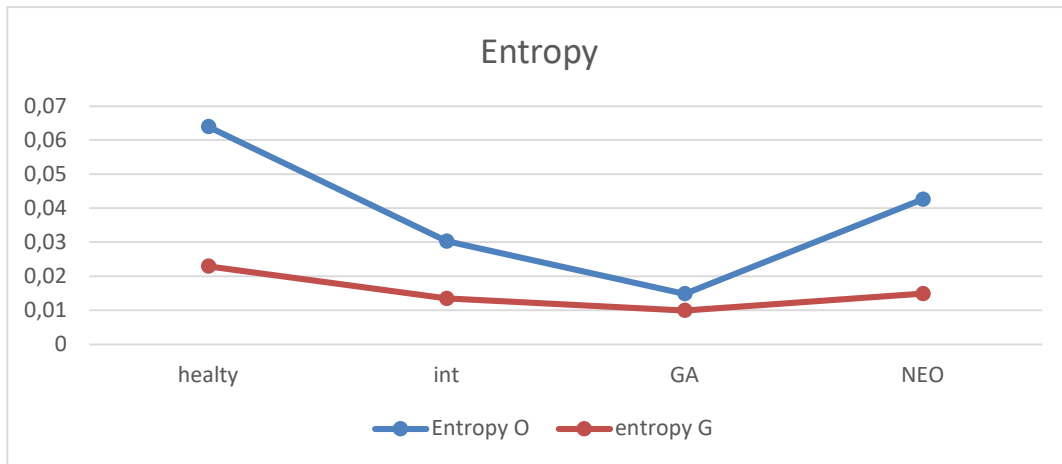


Figura 40: Confronto dell'Andamento dell'Entropia lungo gli stadi per le immagini reali e false

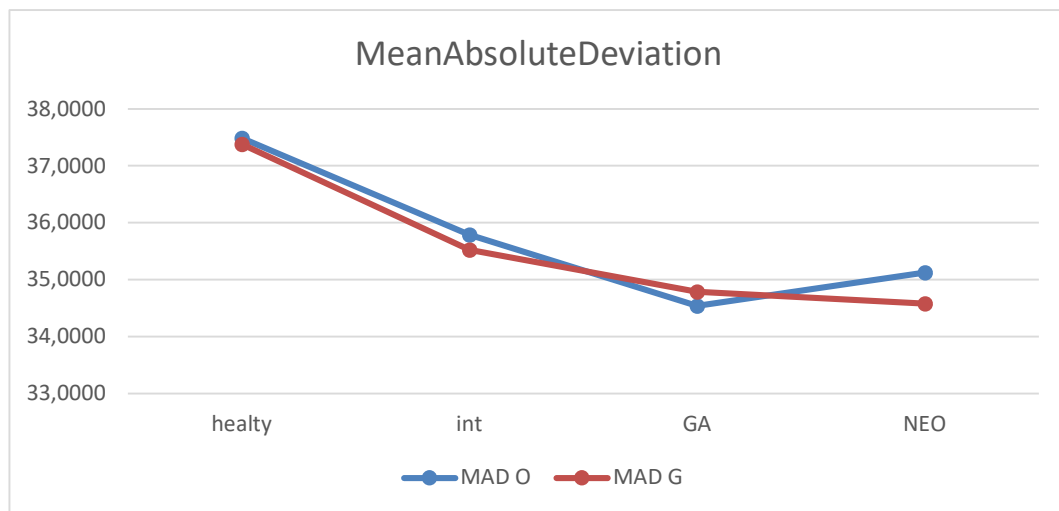


Figura 41: Confronto dell'Andamento della MAD lungo gli stadi per le immagini reali e false

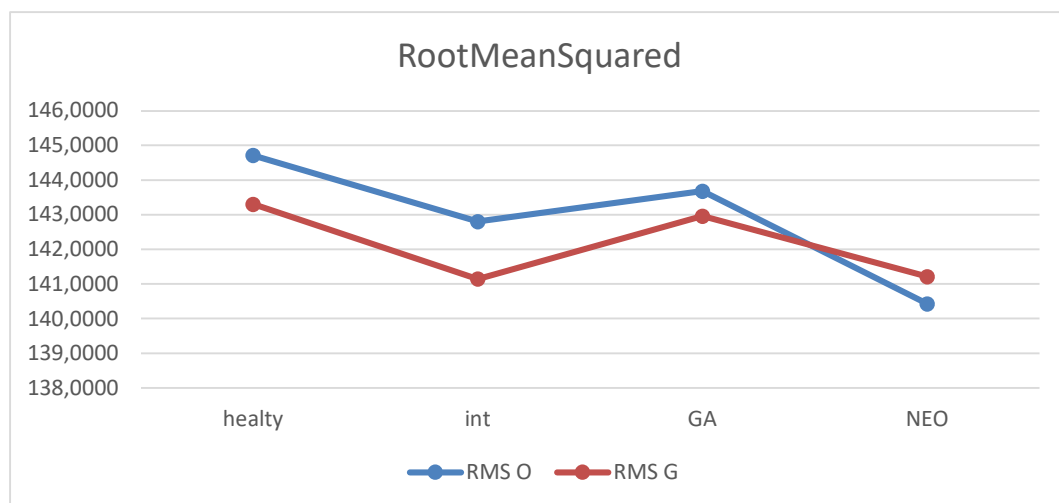


Figura 42: Confronto dell'Andamento della RMS lungo gli stadi per le immagini reali e false

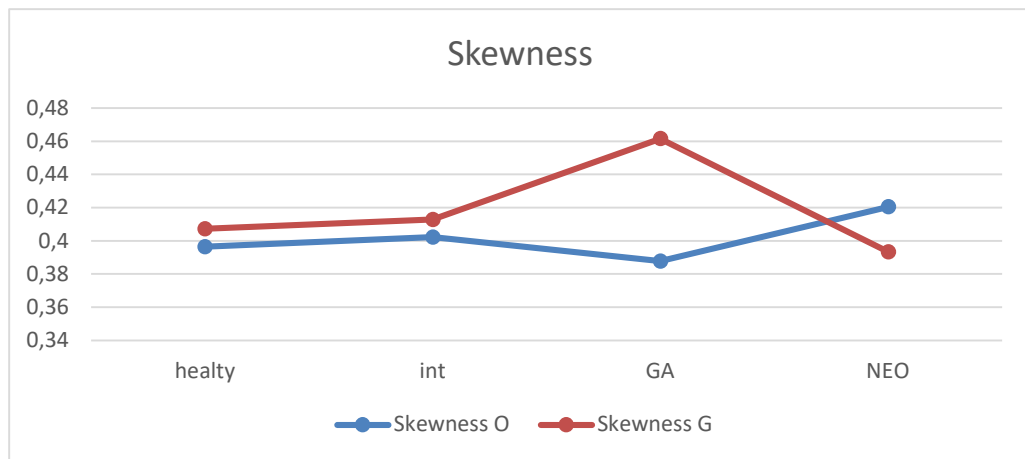


Figura 43: Confronto dell'Andamento della Skewness lungo gli stadi per le immagini reali e false

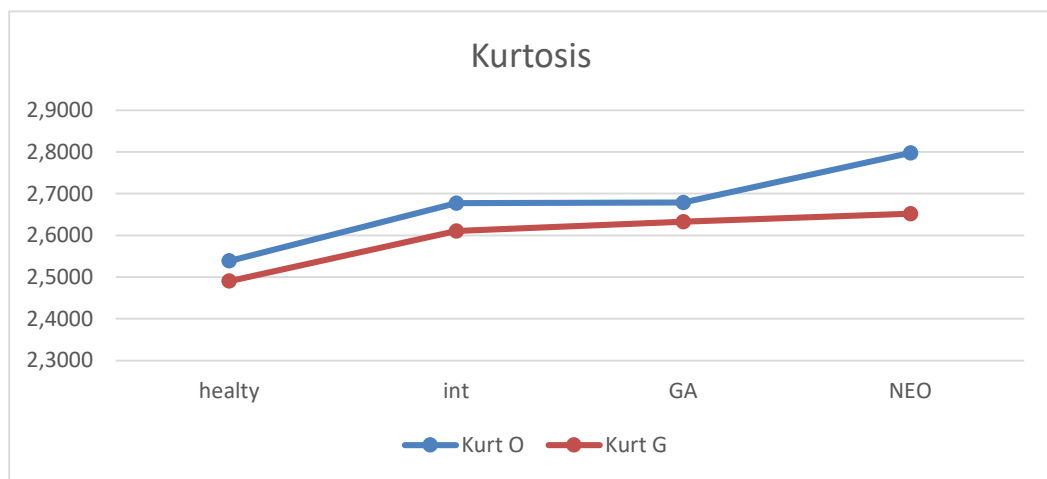


Figura 44: Confronto dell'Andamento della Kurtosis lungo gli stadi per le immagini reali e false

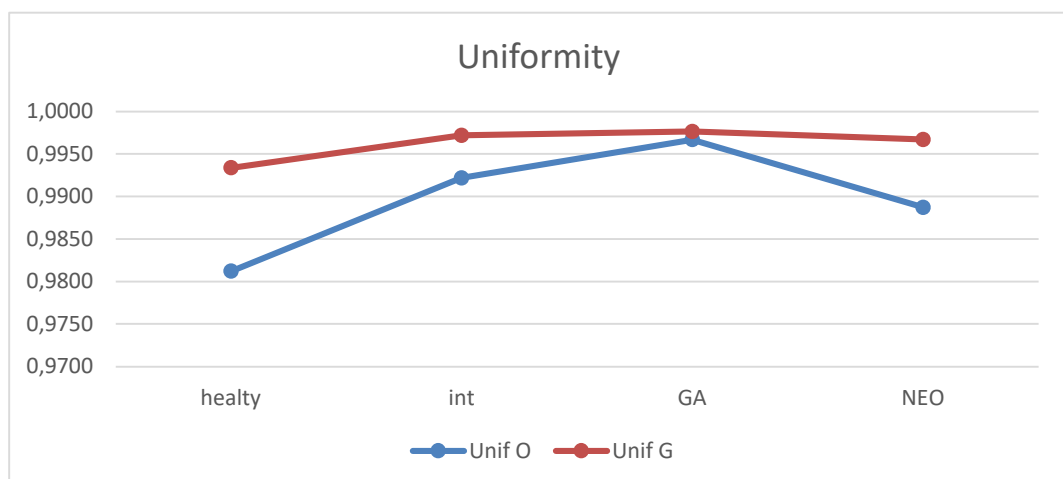


Figura 45: Confronto dell'Andamento della Uniformity lungo gli stadi per le immagini reali e false

Vengono presentati i risultati ottenuti dall'analisi tessutale del secondo ordine espressi come errori relativi [Tab. 8] con i relativi andamenti [Fig. 46-58]:

Stadio	Sa	Int	Ga	Neo
Contrast	1,31±2.99	0.86±4.46	0.86±1.46	1.20±2.65
Difference Average	1.31±2.99	0.86±4.46	0.86±1.46	1.20±2.65
Difference Entropy	1.14±2.79	0.77±3.75	0.77±1.49	1.02±2.14
Difference Variance	1.308±2.99	0.86±4.45	0.86±1.46	1.20±2.65
Inverse Variance	1.31±2.99	0.86±4.46	0.86±1.46	1.20±2.65
Autocorrelation	0.01±0.09	0.004±0.06	0.003±0.11	0.07±0.07
Correlation	0.48±0.81	0.78±0.84	0.79±1.98	0.59±1.13
Joint Average	0.008±0.07	0.001±0.04	0.001±0.03	0.004±0.05
Joint Energy	0.157±1.31	0.004±0.75	0.004±0.3	0.09±1.01
Joint Entropy	1.208±2.9	0.80±4.05	0.80±1.7	1.1±2.49
MCC	0.491±0.83	0.790±0.86	0.79±1.99	0.61±1.24
Sum Average	0.009±0.07	0.001±0.04	0.001±0.03	0.005±0.05
Sum Entropy	1.23±3.12	0.8±3.97	0.8±1.68	1.1±2.46

Tabella 8: Risultati ottenuti dall' analisi tessutale del secondo ordine con relative deviazioni standard

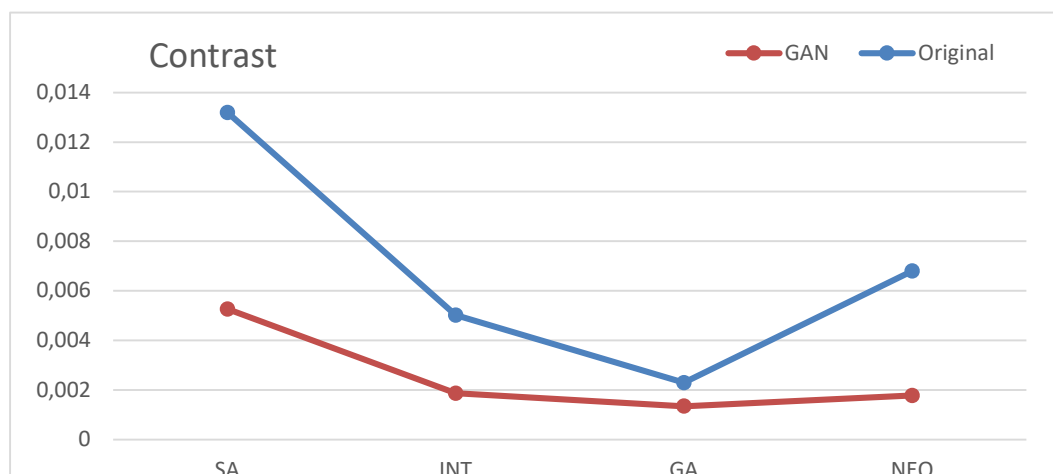


Figura 46: Confronto dell'Andamento del contrasto lungo gli stadi per le immagini reali e false

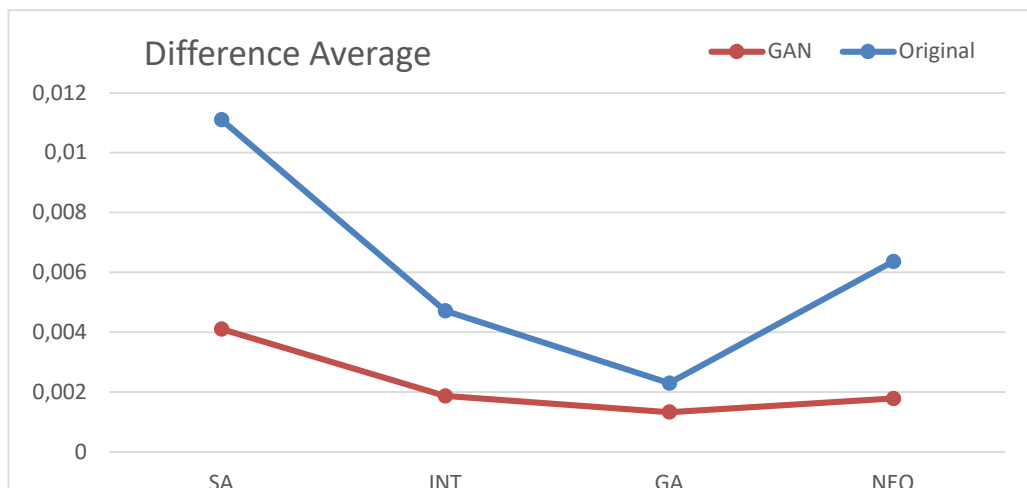


Figura 47: Confronto dell'Andamento della Difference Average lungo gli stadi per le immagini reali e false

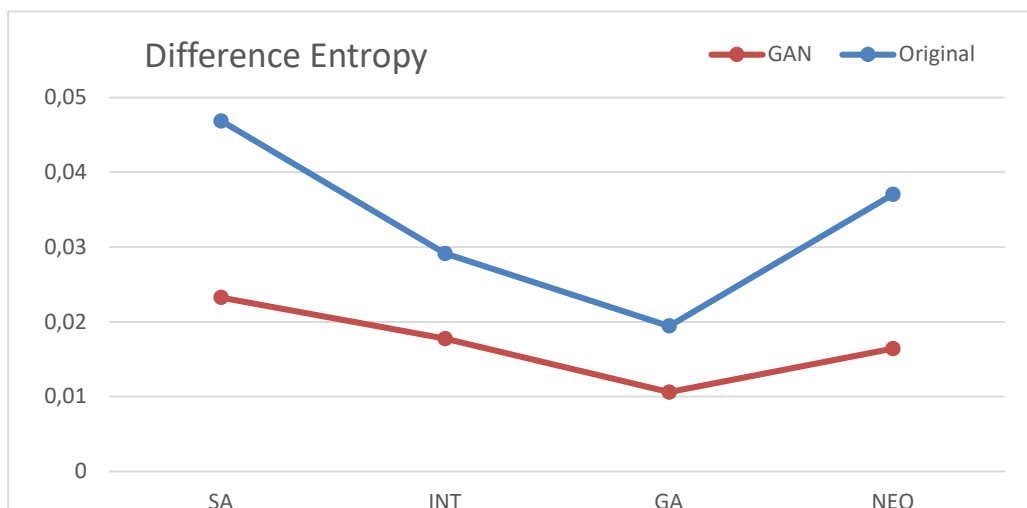


Figura 48: Confronto dell'Andamento della Difference Entropy lungo gli stadi per le immagini reali e false

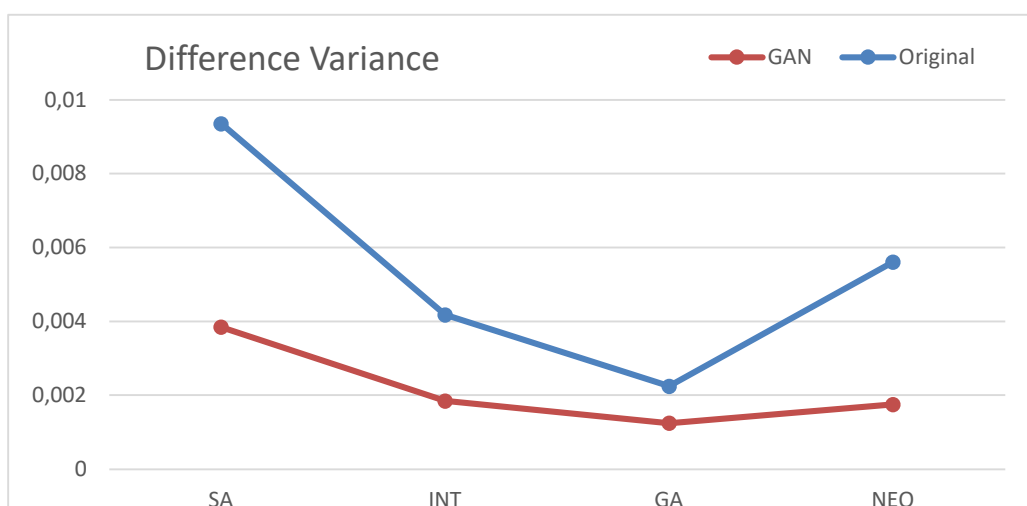


Figura 49: Confronto dell'Andamento della Difference Variance lungo gli stadi per le immagini reali e false

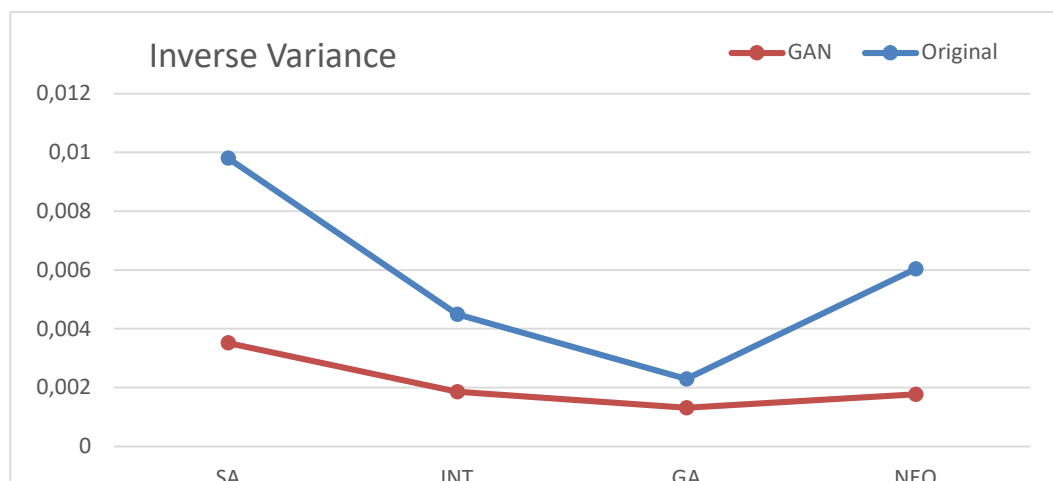


Figura 50: Confronto dell'Andamento della Inverse Variance lungo gli stadi per le immagini reali e false

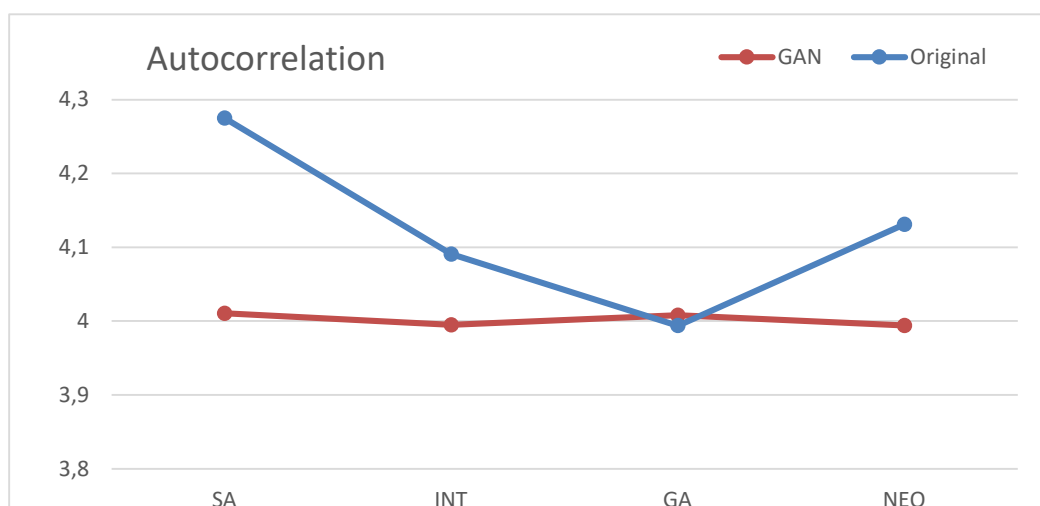


Figura 51: Confronto dell'Andamento dell'Autocorrelazione lungo gli stadi per le immagini reali e false

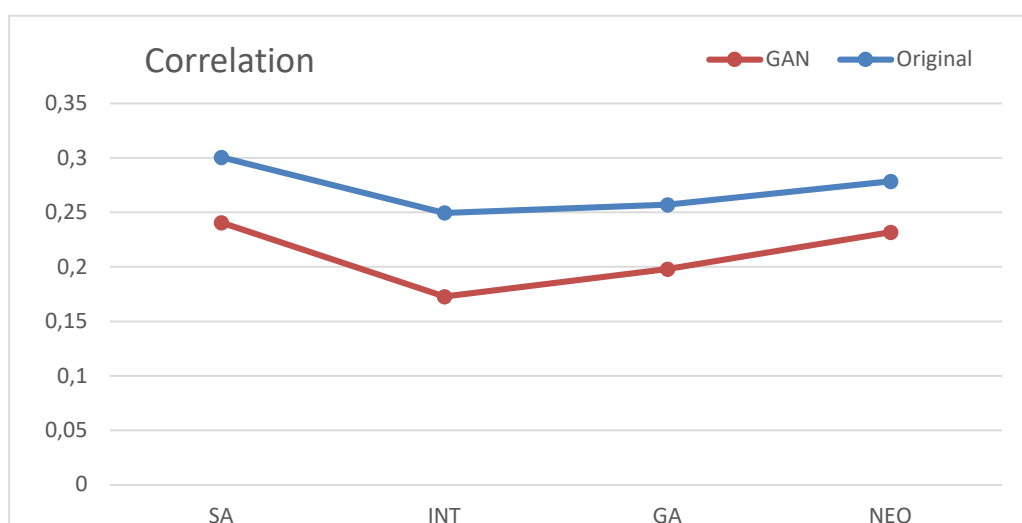


Figura 52: Confronto dell'Andamento della Correlazione lungo gli stadi per le immagini reali e false

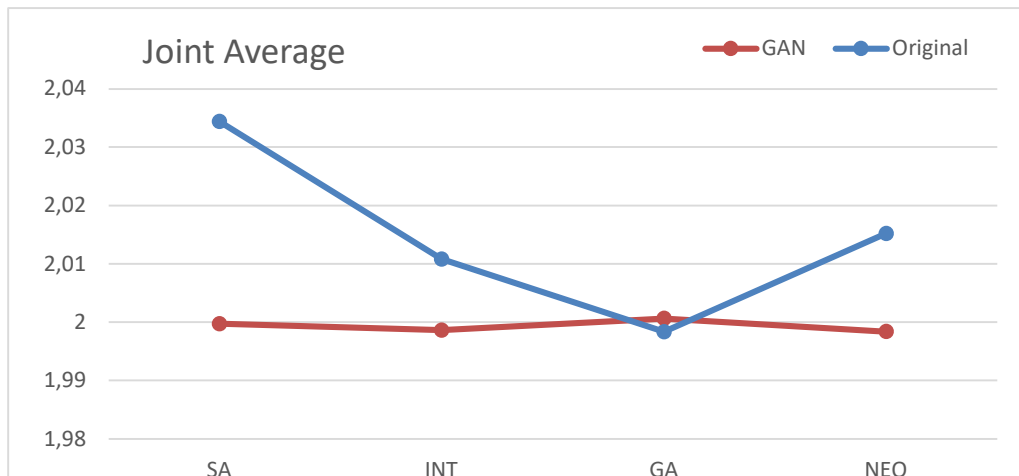


Figura 53: Confronto dell'Andamento della Joint Average lungo gli stadi per le immagini reali e false

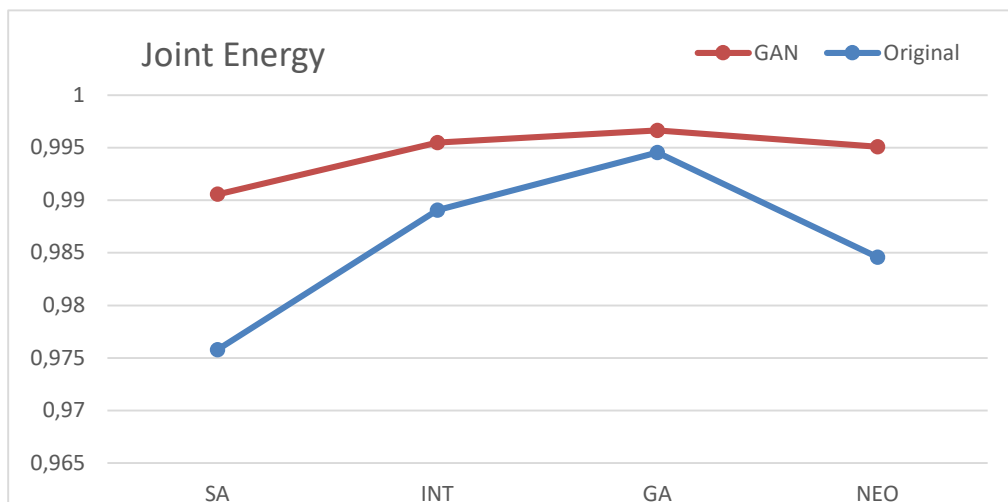


Figura 54: Confronto dell'Andamento della Joint Energy lungo gli stadi per le immagini reali e false

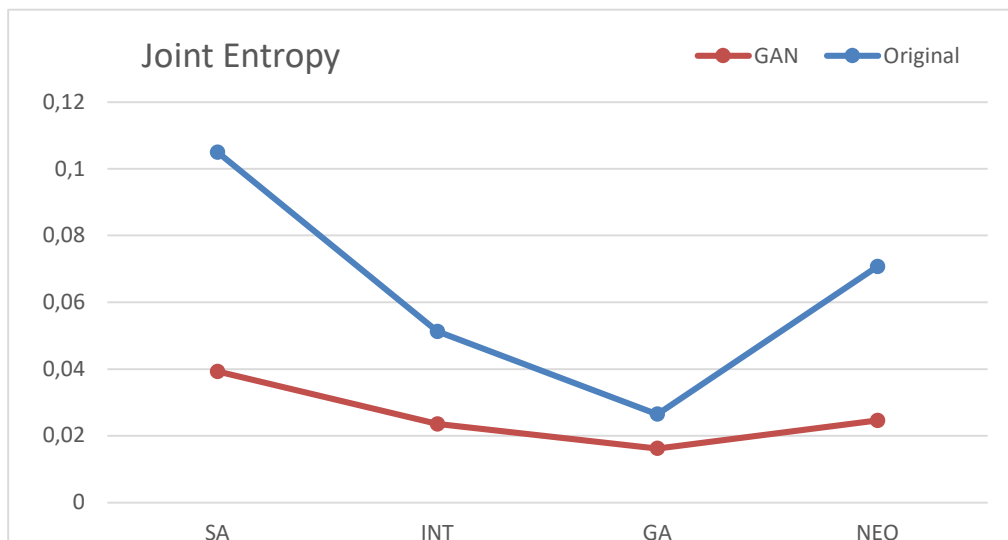


Figura 55: Confronto dell'Andamento della Joint Entropy lungo gli stadi per le immagini reali e false

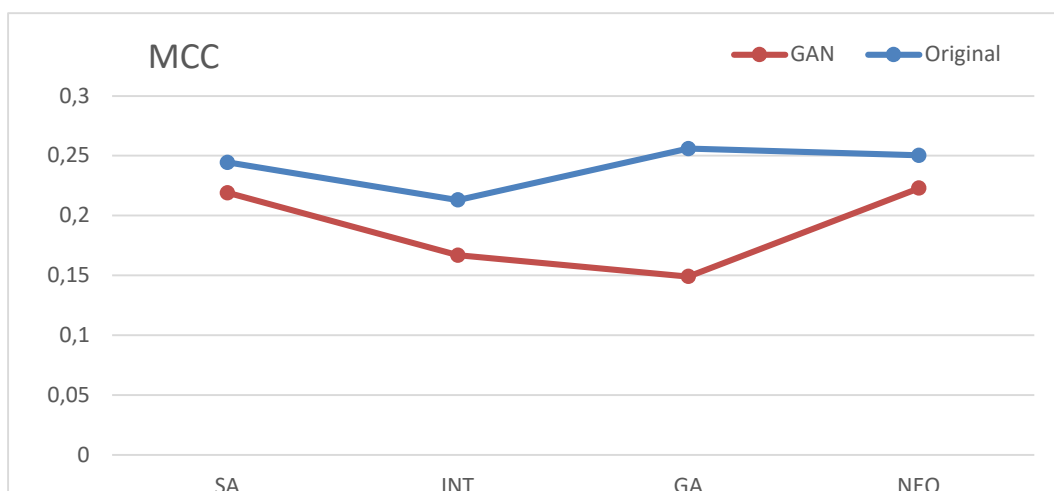


Figura 56: Confronto dell'Andamento del MCC lungo gli stadi per le immagini reali e false

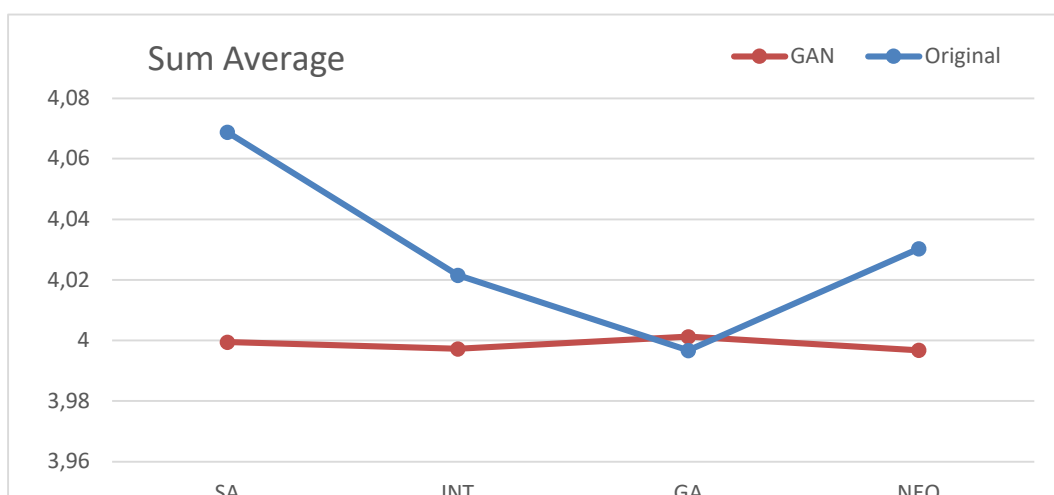


Figura 57: Confronto dell'Andamento della Sum Average lungo gli stadi per le immagini reali e false

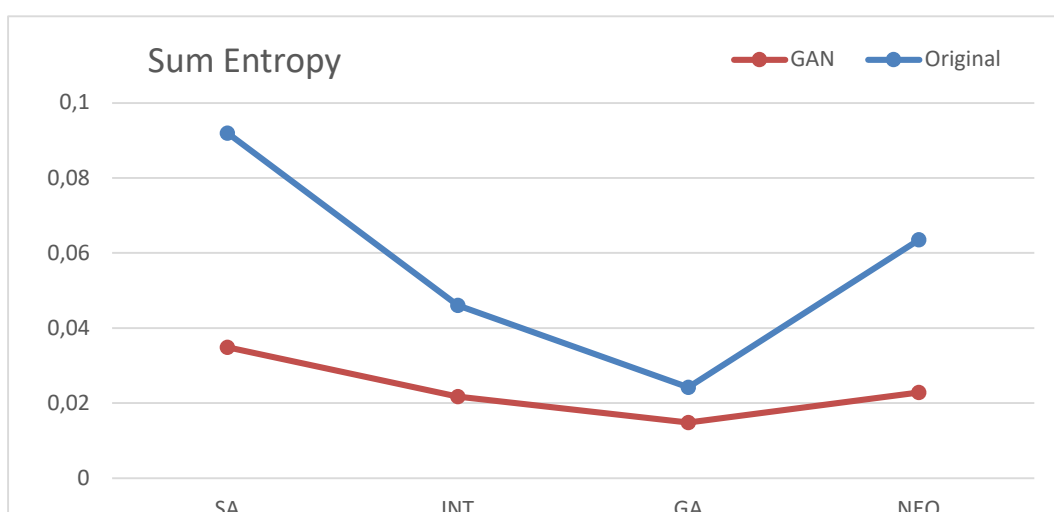


Figura 58: Confronto dell'Andamento della Sum Entropy lungo gli stadi per le immagini reali e false

Si presentano i risultati delle metriche non relative all' analisi tessutale [Tab. 9]:

Stadio	Sa	Int	Ga	Neo
CII	1,59±0,27	1,65±0,24	1,66±0,25	1,69±0,22
FID	28,59	55,89	43,12	29,08
SSIM	0,46±0,1	0,44±0,09	0,44±0,09	0,43±0,08
PSNR	23,19±2,73	22,13±2,39	22,62±2,33	22,36±2,14

Tabella 9:Risultati metriche non Tessutali

Capitolo 4

CONCLUSIONI

Dall'analisi qualitativa della GUI *"Valutazione OCT"*, 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 GUI *"RealOrFake"* 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 confrontata 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 per l'Atrofia Geografica. Ciò è dovuto alla complessità dell'immagine analizzata e alla sua asimmetria, con alcune regioni che presentano densità di intensità molto maggiori 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 negli errori delle features del secondo ordine 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. L'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, confermato anche dai relativi andamenti.

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 avvalorano l'ipotesi di utilizzo di questa GAN per generare immagini sintetiche allo scopo di fare Data augmentation nello sviluppo di nuovi modelli di classificazione o a scopo didattico per la formazione di medici specializzandi.

4.1 Sviluppi Futuri

Nonostante i risultati ottenuti, per gli sviluppi futuri di questo lavoro, ci sono degli aspetti che andrebbero meglio investigati:

- Miglioramento dell'estrazione della mappa semantica attraverso un allenamento multistadio della GMM, utilizzando l'intero dataset a disposizione.
- Ampliamento del dataset per aumentare le 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 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] Y. Deng, L. Qiao, M. Du, C. Qu, L. Wan, J. Li, and L. Huang, "Age-related macular degeneration: Epidemiology, genetics, pathophysiology, diagnosis, and targeted therapy," *Genes & Diseases*, vol. 9, no. 1, pp. 62–79, 2021.
- [2] R. H. Guymer and T. G. Campbell, "Age-related macular degeneration," *The Lancet*, vol. 401, no. 10386, pp. 1459–1472, 2023, doi: 10.1016/S0140-6736(23)00407-2
- [3] D. Thomas and G. Duguid, "Optical coherence tomography—a review of the principles and contemporary uses in retinal investigation," *Eye*, vol. 18, no. 6, pp. 561–570, 2004, doi: 10.1038/sj.eye.6700729
- [4] J. G. Fujimoto and E. A. Swanson, "The development, commercialization, and impact of optical coherence tomography," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 9, pp. OCT1–OCT13, Jul. 2016. doi: 10.1167/iovs.16-19963.
- [5] P. J. Rosenfeld, J. G. Fujimoto, and E. A. Swanson, "Optical coherence tomography and the development of antiangiogenic therapies in neovascular age-related macular degeneration," *Investigative Ophthalmology & Visual Science*, vol. 57, no. 9, pp. OCT1–OCT13, Jul. 2016. doi: 10.1167/iovs.16-19964
- [6] Medical Advisory Secretariat, "Optical coherence tomography for age-related macular degeneration and diabetic macular edema: an evidence-based analysis," *Ontario Health Technology Assessment Series*, vol. 9, no. 13, pp. 1–21, 2009. PMID: 23074517.
- [7] J. G. Fujimoto et al., "Seeing is believing: The development of optical coherence tomography," *Proceedings of the National Academy of Sciences*, vol. 120, no. 32, e2311129120, Aug. 2023. doi: 10.1073/pnas.2311129120.
- [8] L. M. Heiferman, A. Rahimy, and D. Sarraf, "OCT and OCT angiography update: Clinical application to age-related macular degeneration and diabetic retinopathy," *Retina*, vol. 43, no. 2, pp. 179–193, Feb. 2023. doi: 10.1097/IAE.0000000000003437.
- [9] A. You, J. Kim, and S. K. Kang, "Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey," *Eye and Vision*, vol. 9, no. 44, pp. 1–14, 2022. doi: 10.1186/s40662-022-00277-3.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, 2014, pp. 2672–2680.
- [11] I. Goodfellow, "NIPS 2016 Tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2016. [Online]. Available: <https://arxiv.org/abs/1701.00160>
- [12] J. Kugelman, D. Alonso-Caneiro, S. A. Read, and M. J. Collins, "A review of generative adversarial network applications in optical coherence tomography image analysis," *Journal of Optometry*, vol. 15, Suppl. 1, pp. S1–S11, Jan. 2022, doi: 10.1016/j.optom.2022.09.004.
- [13] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," *arXiv preprint arXiv:1701.07875*, 2017. [Online]. Available: <https://arxiv.org/abs/1701.07875>
- [14] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive Growing of GANs for Improved Quality, Stability, and Variation," in *International Conference on Learning Representations (ICLR)*, 2018. [Online]. Available: <https://arxiv.org/abs/1710.10196>

-
- [15] C. Zheng, Y. Xu, D. K. Wong, and J. Liu, "Assessment of Generative Adversarial Networks Model for Synthetic Optical Coherence Tomography Images of Retinal Disorders," *arXiv preprint arXiv:1910.09748*, 2019. [Online]. Available: <https://arxiv.org/abs/1910.09748>
- [16] A. Andreini, G. Lenzini, and M. Bertini, "A Two Stage GAN for High Resolution Retinal Image Generation and Segmentation," *arXiv preprint arXiv:1907.12296*, 2019. [Online]. Available: <https://arxiv.org/abs/1907.12296>
- [17] W. Zhu, Z. Wang, Z. Huang, and others, "Optimal Transport Guided Unsupervised Learning for Enhancing Low-Quality Retinal Images," *arXiv preprint arXiv:2302.02991*, 2023. [Online]. Available: <https://arxiv.org/abs/2302.02991>
- [18] E. Waisberg, J. Ong, S. A. Kamran, M. Masalkhi, P. Paladugu, N. Zaman, A. G. Lee, and A. Tavakkoli, "Generative artificial intelligence in ophthalmology," *Survey of Ophthalmology*, vol. 70, no. 1, pp. 1–11, 2025, doi: 10.1016/j.survophthal.2024.04.009.
- [19] SciPy Developers, "gaussian_filter — SciPy v1.16.2 Manual", SciPy v1.16.2, Multidimensional Gaussian filter, `scipy.ndimage`, SciPy. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.ndimage.gaussian_filter.html [Accessed: 17-Sep-2025].
- [20] A. D., P. A., S. K. S., and A. K., "Development of Adaptive Gaussian Filter Based Denoising as an Image Enhancement Technique," *Mathematical Modelling of Engineering Problems*, vol. 11, no. 10, pp. 2715–2725, Oct. 2024, doi: [10.18280/mmep.111013](https://doi.org/10.18280/mmep.111013)
- [21] OpenCV Authority, "Denoising: `denoise_TVL1` function," *OpenCV Documentation*, version 4.x, available: https://docs.opencv.org/4.x/d1/d79/group__photo__denoise.html#ga4c6b0031f56ea3f98f768881279ffe93
- [22] A. Buades, B. Coll, e J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, pp. 208–212, 13 Set. 2011, doi: 10.5201/ipol.2011.bcm_nlm.
- [23] University of Toronto, "CSC411 Lectures 15–16: Gaussian Mixture Model & EM," University of Toronto, [Online]. Available: https://www.cs.toronto.edu/~jlucas/teaching/csc411/lectures/lec15_16_handout.pdf [Accessed: 17-Sep-2025]
- [24] M. Bonakdarpour, *Introduction to EM: Gaussian Mixture Models*, FiveMinuteStats, 22-Jan-2016. [Online]. Available: https://stephens999.github.io/fiveMinuteStats/intro_to_em.html [Accessed: 17-Sep-2025].
- [25] R. Yehoshua, "Gaussian Mixture Models (GMMs): From Theory to Implementation," *Towards Data Science*, 09-Jul-2024. [Online]. Available: <https://towardsdatascience.com/gaussian-mixture-models-gmms-from-theory-to-implementation-4406c7fe9847/> [Accessed: 17-Sep-2025]
- [26] Z. Fang, H. Rabbani, A. Hojjatoleslami, and M. Sonka, "Optical coherence tomography image denoising using Gaussianization transform," *Journal of Biomedical Optics*, vol. 22, no. 8, pp. 086011–1–086011–12, Aug. 2017, doi: 10.1117/1.JBO.22.8.086011.
- [27] OpenCV Library, "pyrMeanShiftFiltering — Image Filtering," *OpenCV Documentation*, Version 3.4. [Online]. Available: https://docs.opencv.org/3.4/d4/d86/group__imgproc__filter.html#ga9fabdce9543bd602445f5db3827e4cc0 [Accessed: 23-Sep-2025]
- [28] Jia-Xiang Zhou et al., *Improved fast mean shift algorithm for remote sensing image segmentation* —IET Image Processing, 2015. doi: 10.1049/iet-ipr.2014.0393.
-

-
- [29] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [30] OpenCV, "Image Thresholding," OpenCV Documentation, 2025. [Online]. Available: https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html [Accessed: 25-Sep-2025].
- [31] P. Isola, J. Zhu, T. Zhou e A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *arXiv preprint arXiv:1611.07004*, 2016.
- [32] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, e S. P. Smolley, "Least Squares Generative Adversarial Networks," *arXiv preprint arXiv:1611.04076*, 2017.
- [33] K. Simonyan e A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv: 1409.1556*, 2014.
- [34] G. Rani, A. Misra, V. S. Dhaka, E. Zumpano, e E. Vocaturo, "Spatial feature and resolution maximization GAN for bone suppression in chest radiographs," *Computer Methods and Programs in Biomedicine*, vol. 224, p. 107024, 2022, doi: 10.1016/j.cmpb.2022.107024.
- [35] S. Ioffe e C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv preprint arXiv:1502.03167*, 2015. Disponibile su: <https://arxiv.org/pdf/1502.03167>
- [36] M. S. N. Devi and S. Santhi, "Improved OCT Image Enhancement using CLAHE," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 11, pp. 1351–1355, Sept. 2019, doi: 10.35940/ijitee.J9668.0981119.
- [37] M. Woodland, J. Wood, B. M. Anderson, S. Kundu, E. Lin, E. Koay, B. Odisio, C. Chung, C. Kang, A. M. Venkatesan, S. Yedururi, B. De, Y.-M. Lin, A. B. Patel e K. K. Brock, "Evaluating the Performance of StyleGAN2-ADA on Medical Images," *arXiv preprint arXiv:2210.03786*, 7 ott. 2022.
- [38] "Frechet Inception Distance (FID) — PyTorch-Metrics 1.8.2 documentation," *Lightning AI / TorchMetrics*, [Online]. Available: https://lightning.ai/docs/torchmetrics/stable/image/frechet_inception_distance.html
- [39] Y. Skandarani, P.-M. Jodoin, A. Lalonde, "GANs for Medical Image Synthesis: An Empirical Study," *arXiv preprint arXiv:2105.05318*, 2021.
- [40] H. Ahmed, Q. Zhang, R. Donnan, e A. Alomainy, "Denoising of Optical Coherence Tomography Images in Ophthalmology Using Deep Learning: A Systematic Review," *J. Imaging*, vol. 10, n. 4, art. 86, pp. –, Apr. 2024, doi: 10.3390/jimaging10040086.
- [41] TorchMetrics Contributors, "Structural Similarity Index Measure (SSIM)," *TorchMetrics Documentation*, v1.8.2, Oct. 2025. [Online]. Available: https://lightning.ai/docs/torchmetrics/stable/image/structural_similarity.html
- [42] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*. *IEEE Transactions on Image Processing*, vol. 13, no.4, pp. 600–612, (2004).
- [43] "Peak Signal-to-Noise Ratio (PSNR) — PyTorch-Metrics 1.8.2 documentation," *Lightning AI*, [Online]. Available: https://lightning.ai/docs/torchmetrics/stable/image/peak_signal_noise_ratio.html
- [44] F. Peng, K. Jiang, G. Zhu, J. Li, Y. Zhang e Y. Wang, "Magnetic Resonance Texture Analysis in Myocardial Infarction," *Frontiers in Cardiovascular Medicine*, vol. 8, art. 724271, 2021. [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8581163/pdf/fcvm-08-724271.pdf>
- [45] Pyradiomics, "Radiomic Features," *PyRadiomics Documentation*, versione 3.1.0, [Online]. Available: <https://pyradiomics.readthedocs.io/en/latest/features.html>
- [46] van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillon-Robin, J. C., Pieper, S., Aerts, H. J. W. L. (2017). *Computational Radiomics System to Decode the Radiographic Phenotype*. *Cancer Research*, 77(21), e104–e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339> <https://doi.org/10.1158/0008-5472.CAN-17-0339>
-