

Candidato Fabrizio Ravelli
Matricola 177085

Retinal Disease Classification and Localisation through Convolutional Neural Networks



A.A. 2022/2023

Corso di
Elaborazione di Bioimmagini

Prof. Mario Cesarelli
Prof. Francesco Mercaldo

.summary()

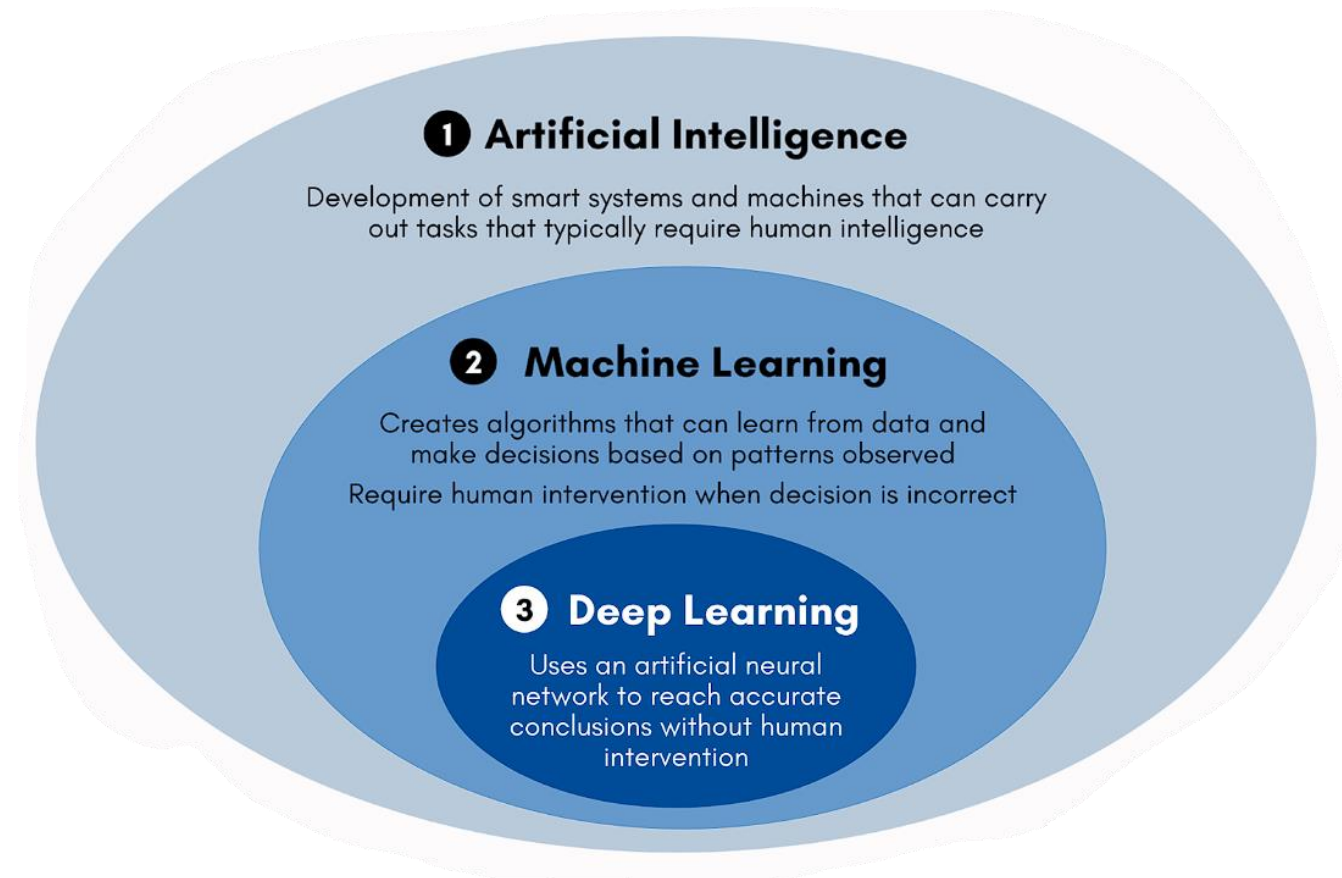
- Deep Learning e Reti Neurali Convoluzionali
- Patologie retiniche
- Stato dell'arte
- Metodologia
- Analisi sperimentale
- Visual Explainability
- Conclusioni e sviluppi futuri

Deep Learning e Reti Neurali Convoluzionali

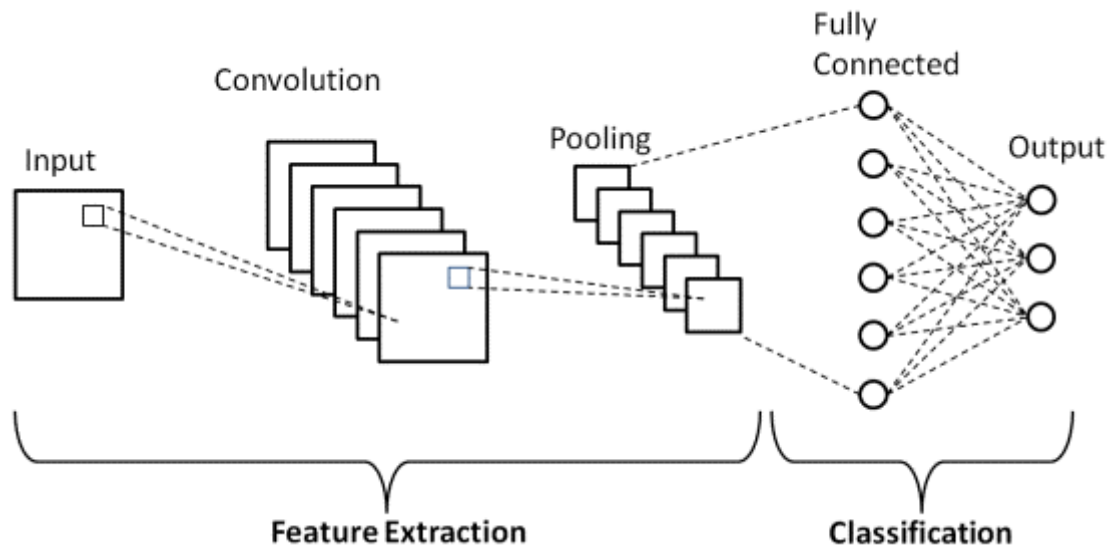
Il **Deep Learning** (DL), tradotto come apprendimento profondo, è una branca dell'Intelligenza artificiale, più propriamente del Machine Learning (ML).

La differenza principale tra Machine Learning e Deep Learning consiste nella fase di feature extraction.

Tra gli «strumenti» offerti dal Deep Learning troviamo le Reti Neurali Profonde (DNNs), le reti neurali ricorrenti (RNNs) e le **reti neurali convoluzionali** (CNNs).



Deep Learning e Reti Neurali Convoluzionali



Le Reti Neurali Convoluzionali (CNNs) sono lo strumento migliore per task di classificazione di immagini, object detection e segmentazione.

Nella loro versione più semplice sono composte da due blocchi sequenziali:

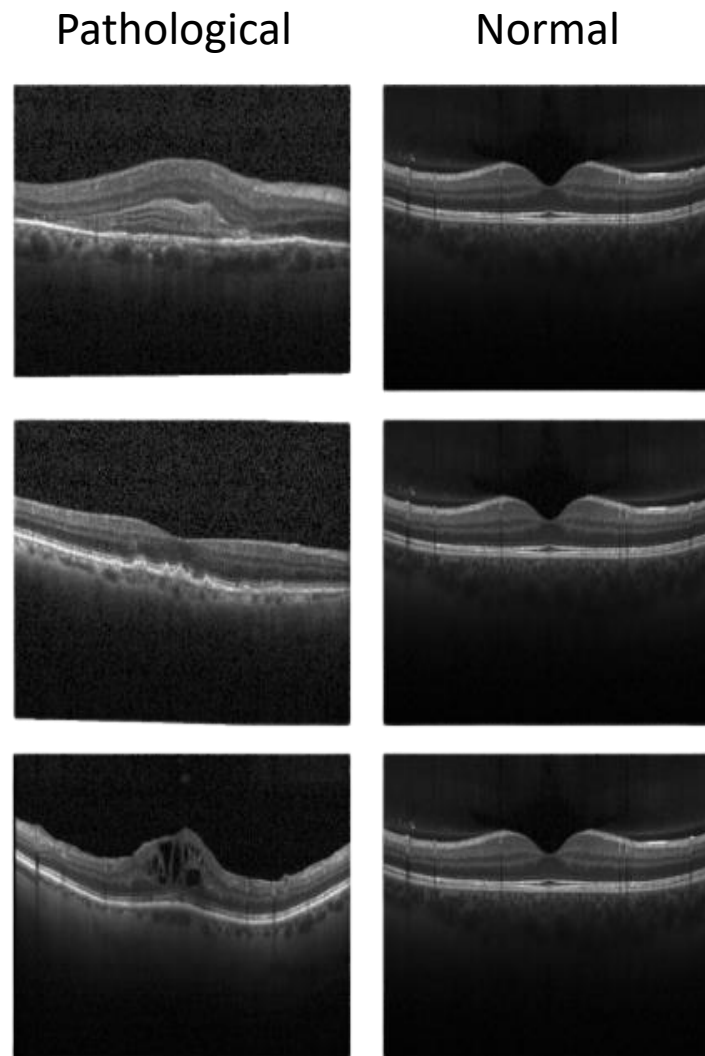
- **Blocco convolutivo** : Formato da layer di convoluzione e pooling, si occupa della parte di feature extraction.
- **Blocco di classificazione** : Formato da layer densamente connessi, si occupa del task di classificazione.

Le architetture possono essere rese più complicate tramite l'aggiunta di layer di regolarizzazione, normalizzazione, ecc...

Patologie retiniche

Le patologie che sono state oggetto di studio in questo lavoro sono:

- Neovascolarizzazione coroidale (**CNV**) : La CNV fa parte dello spettro di complicanze dovute alla degenerazione maculare senile (AMD) e consiste in un crescita anomala di vasi al di sotto della membrana di Bruch. L'emorragia da CNV è molto pericolosa per l'acuità visiva
- Drusen retinali (**DRUSEN**) : Piccoli depositi di materiale proteo-lipidico che si vanno a formare al di sotto della retina. Si suddividono in rigidi e morbidi. Tipicamente si formano tra i 40 e i 60 anni e sono correlati alla degenerazione tissutale dovuta all'età.
- Edema maculare diabetico (**DME**) : Patologia oculare che si presenta insieme alla retinopatia diabetica (DR) e quindi al diabete. La DR provoca fuoriuscite di liquido dai vasi che si vanno ad accumulare sulla macula e ne causa gonfiore.



Stato dell'arte

Sono stati condotti **diversi studi** sulla possibilità di diagnosi delle patologie descritte tramite reti neurali. Per confrontare i risultati ottenuti durante la sperimentazione sono stati selezionati 6 studi che utilizzano il medesimo dataset.

Author	Focus	Localisation	Dataset	Results
Kermany et al. [1]	Retinal diseases and chest X-ray classification.	yes	Kaggle website OCT dataset	99.6% Accuracy
Sunija et al. [2]	Retinal diseases classification.	yes	Kaggle website OCT dataset	99.69% Accuracy
Chen et al. [3]	Retinal diseases classification.	no	Kaggle website OCT dataset	99.19% Accuracy
Tsuji et al. [4]	Retinal diseases classification.	yes	Kaggle website OCT dataset	99.6% Accuracy
Ara et al. [5]	Retinal diseases classification.	no	Kaggle website OCT dataset	98.97% Recall
Mittal et al. [6]	Retinal diseases classification.	no	Kaggle website OCT dataset	98.73% Accuracy

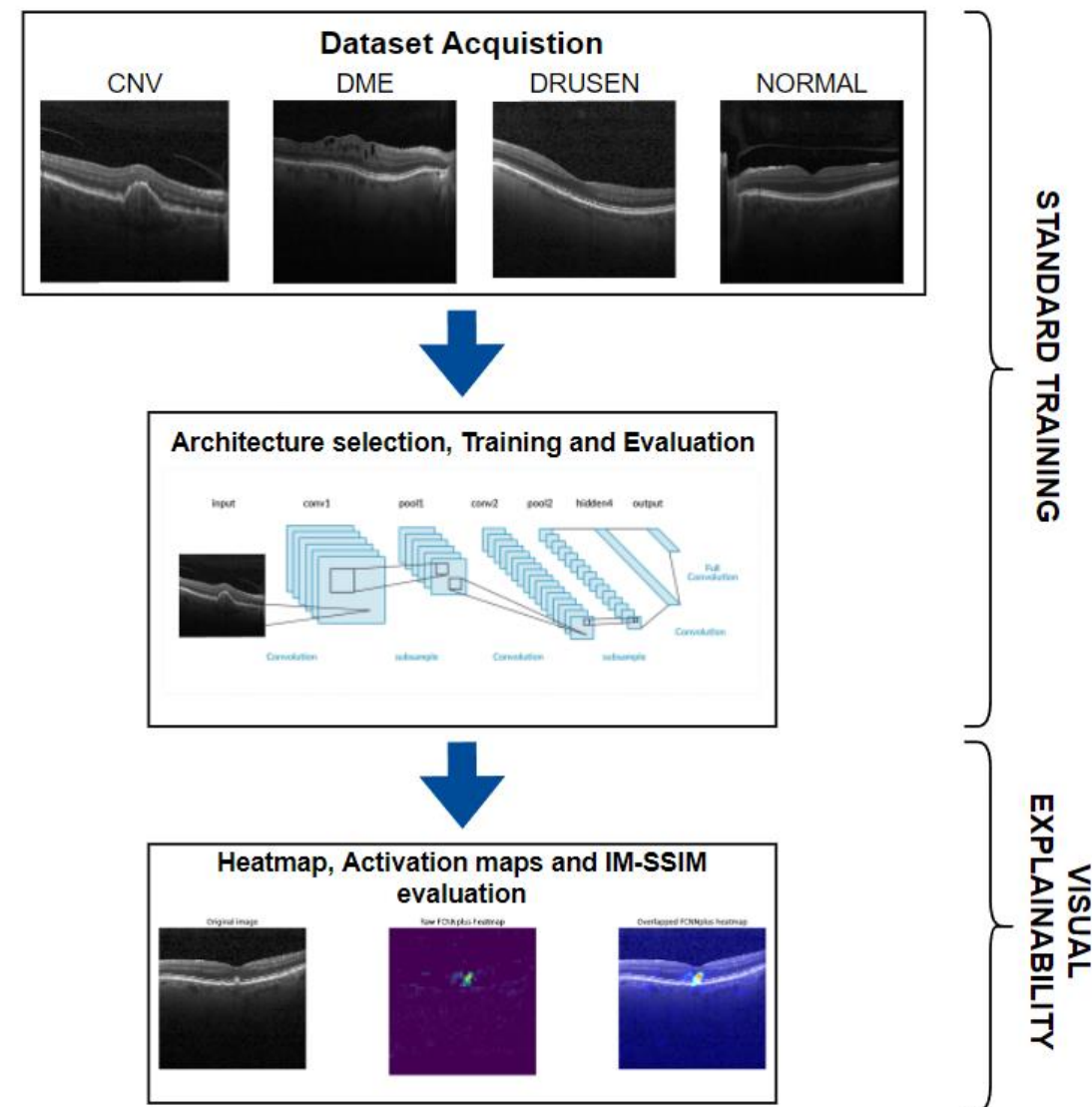
Metodologia

La metodologia proposta vede due passaggi fondamentali :

- **Addestramento standard** : Acquisizione dati, verifica, selezione delle architetture, addestramento e test.
- **Visual Explainability** : Generazione heatmap tramite algoritmo ScoreCAM, visualizzazione dei filtri e delle attivazioni intermedie, stima della robustezza del modello tramite IM-SSIM.

Il codice Python utilizzato è disponibile al link

<https://github.com/reFraw/Ocular-Disease-Recognition>



Metodologia

Folder	Class	# Samples	Total	Percentage
Training	CNV	25310	61284	80
	DME	8788		
	DRUSEN	6256		
	NORMAL	20930		
Validation	CNV	3163	7659	10
	DME	1098		
	DRUSEN	782		
	NORMAL	2616		
Testing	CNV	3165	7664	10
	DME	1099		
	DRUSEN	783		
	NORMAL	2617		

Il dataset utilizzato è disponibile sul sito web Kaggle al link:

<https://www.kaggle.com/datasets/paultimothymooney/kermany2018>

Le immagini sono state ottenute tramite tomografia a coerenza ottica (OCT).

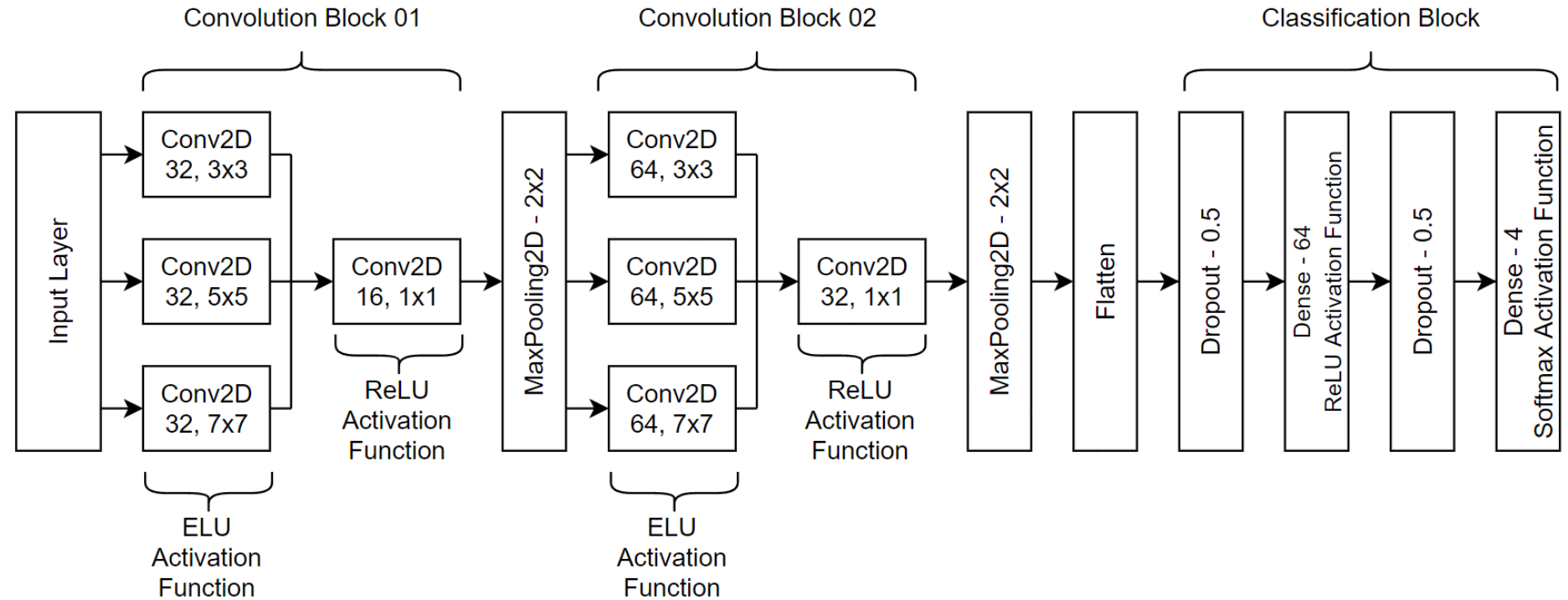
Analizzando i campioni del dataset tramite la chiave SHA-256 è stato riscontrato del data leakage tra le cartelle di training e testing, oltre che diversi campioni ripetuti.

Rimuovendo i duplicati ed effettuando uno splitting 80/10/10 si ottiene un dataset di 76607 campioni suddivisi come mostrato in tabella.

Metodologia

Per la fase di addestramento e test sono state utilizzate un totale di 7 architetture:

- FCNNplus
- STANDARD_CNN
- VGG16
- VGG19
- EfficientNetB0
- InceptionV3
- Xception



Architettura FCNNplus

Metodologia

L'architettura proposta sfrutta due blocchi convoluzionali, denominati **F-Block**, strutturati nel seguente modo:

- **Primo passaggio** : Tre convoluzioni simultanee e indipendenti caratterizzate da dimensioni dei kernel di 3x3, 5x5 e 7x7. I layer convolutivi sfruttano la funzione di attivazione ELU (Exponential Linear Unit) così da mantenere anche le attivazioni di rete negative.
- **Secondo passaggio** : Gli output delle precedenti convoluzioni vengono concatenati e sottoposti ad ulteriore convoluzione con dimensione del kernel pari a 1x1 e $\#filter_concat > \#filter_1x1$. Questa operazione, sfruttando la funzione ReLU (Rectified Linear Unit) permette di ridurre la dimensionalità dei dati conservando al tempo stesso le feature maps migliori.

$$ELU(x) = \begin{cases} x, & x \geq 0 \\ \alpha(e^x - 1), & x < 0 \end{cases} \quad ReLU(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Metodologia

Per fornire Visual Explainability ai modelli ottenuti sono state applicate diverse tecniche:

1. Applicazione preliminare dell'algoritmo GradCAM ai modelli addestrati.
2. Applicazione dell'algoritmo ScoreCAM ai due modelli più performanti.
3. Stima della robustezza del modello tramite IM-SSIM.
4. Valutazione delle attivazioni intermedie e dei filtri convolutivi.

Analisi sperimentale

Per permettere un confronto oggettivo sulle performance quantitative dei modelli ottenuti si è deciso di utilizzare la stessa combinazione di iperparametri per ogni rete:

- Dimensione immagine 100x100 pixel
- Codifica RGB
- Batch size 128
- Learning rate 0.0003

L'addestramento delle reti è stato svolto per un totale di 25 epoche.

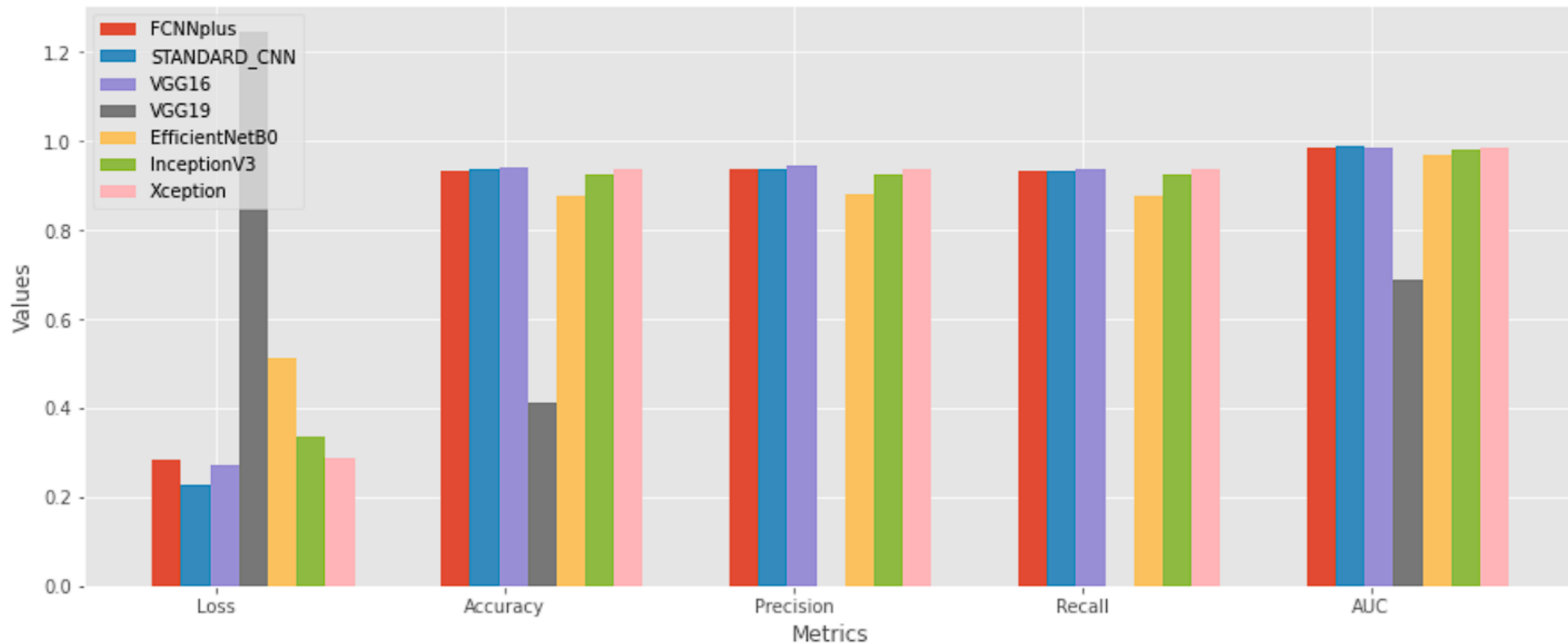
Analisi sperimentale

In tabella sono riportati i risultati ottenuti in fase di test.

Model	Loss	Accuracy	Precision	Recall	AUC
FCNNplus	0.2827	0.9335	0.9347	0.9324	0.9851
STANDARD_CNN	0.2264	0.9350	0.9379	0.9323	0.9893
VGG16	0.2696	0.9419	0.9451	0.9382	0.9860
VGG19	1.2438	0.4130	0.0000	0.0000	0.6884
EfficientNetB0	0.5121	0.8770	0.8803	0.8743	0.9679
InceptionV3	0.3375	0.9245	0.9251	0.9241	0.9800
Xception	0.2884	0.9353	0.9362	0.9346	0.9842

Ad eccezione delle reti VGG19 ed EfficientNetB0, tutti i modelli riescono ad ottenere risultati al di sopra del 90% su tutte le metriche di riferimento.

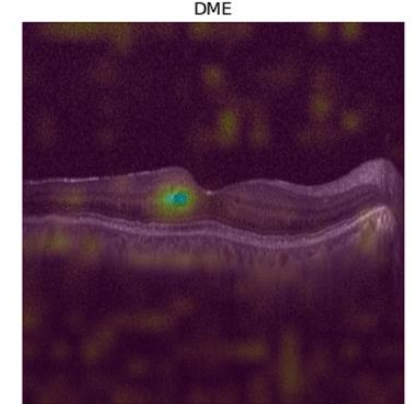
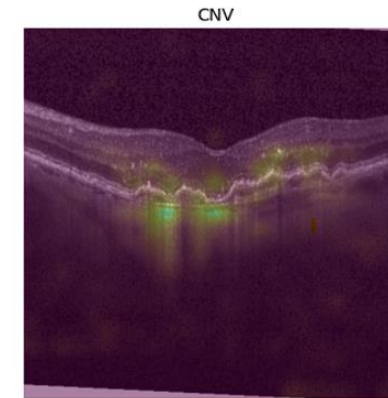
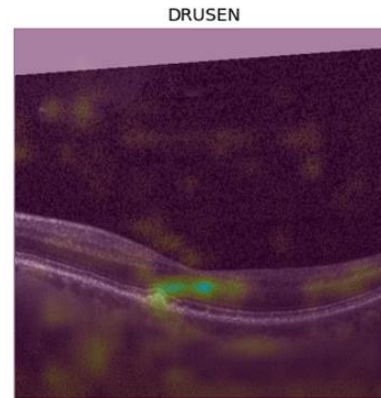
Analisi sperimentale



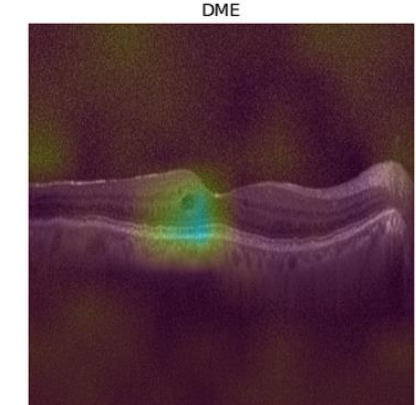
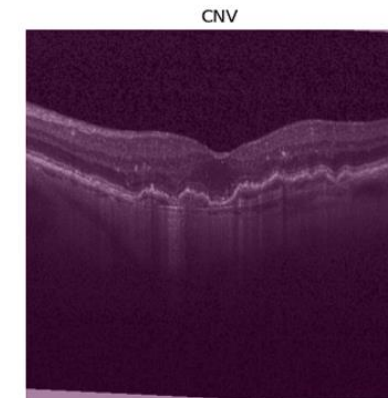
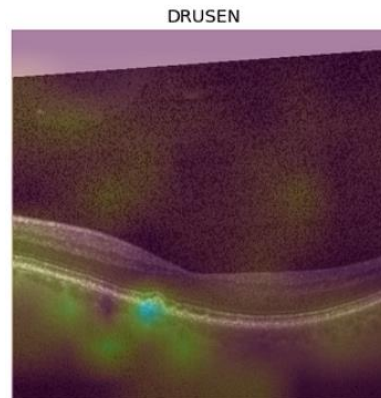
Visual Explainability

Dall'analisi delle heatmap generate tramite l'algoritmo GradCAM è possibile notare che solo i modelli FCNNplus e STANDARD_CNN offrono il giusto compromesso tra performance quantitative e qualitative.

FCNNplus

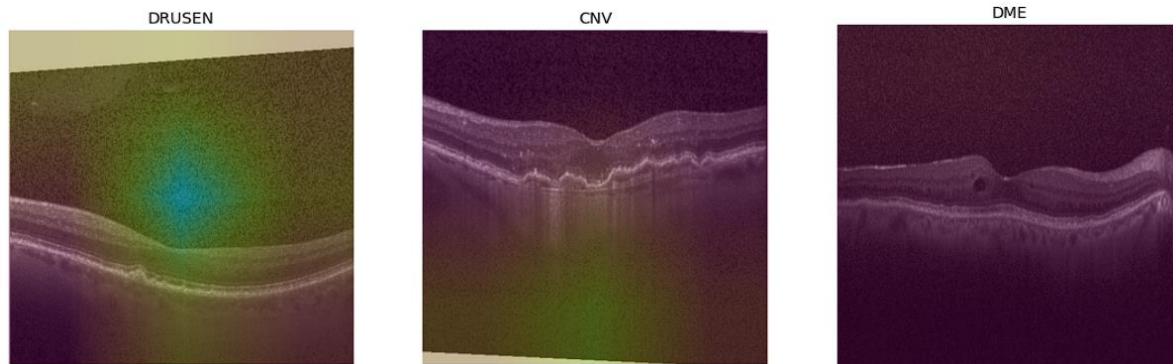


STANDARD_CNN

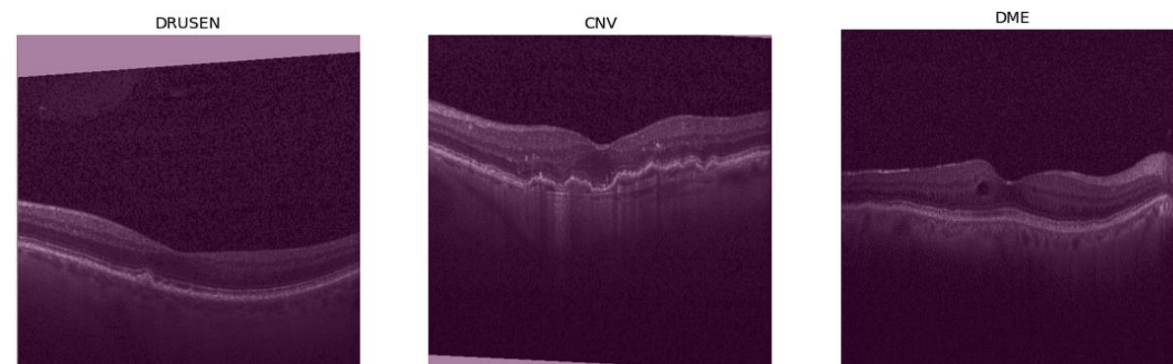


Visual Explainability

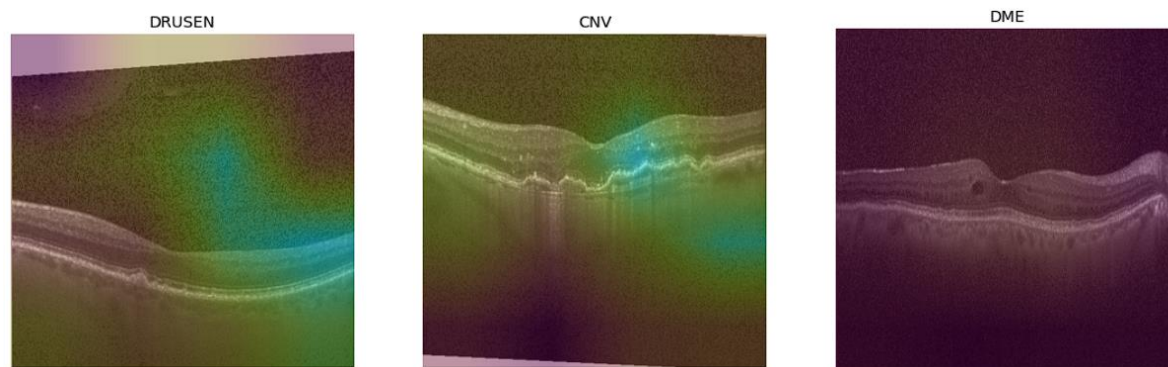
VGG16



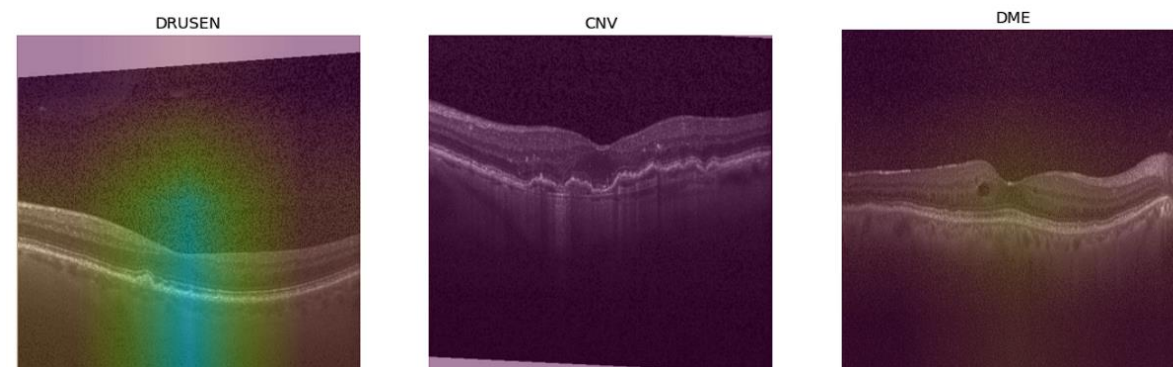
VGG19 / InceptionV3



EfficientNetB0

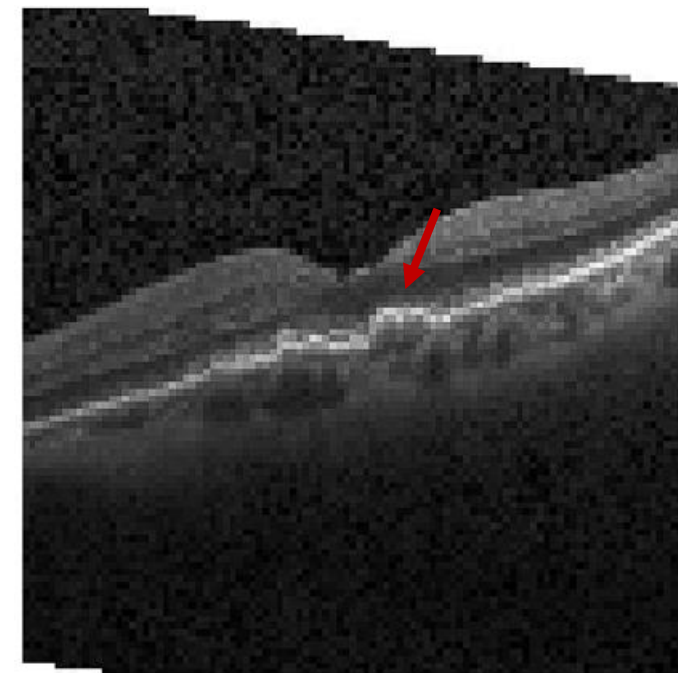
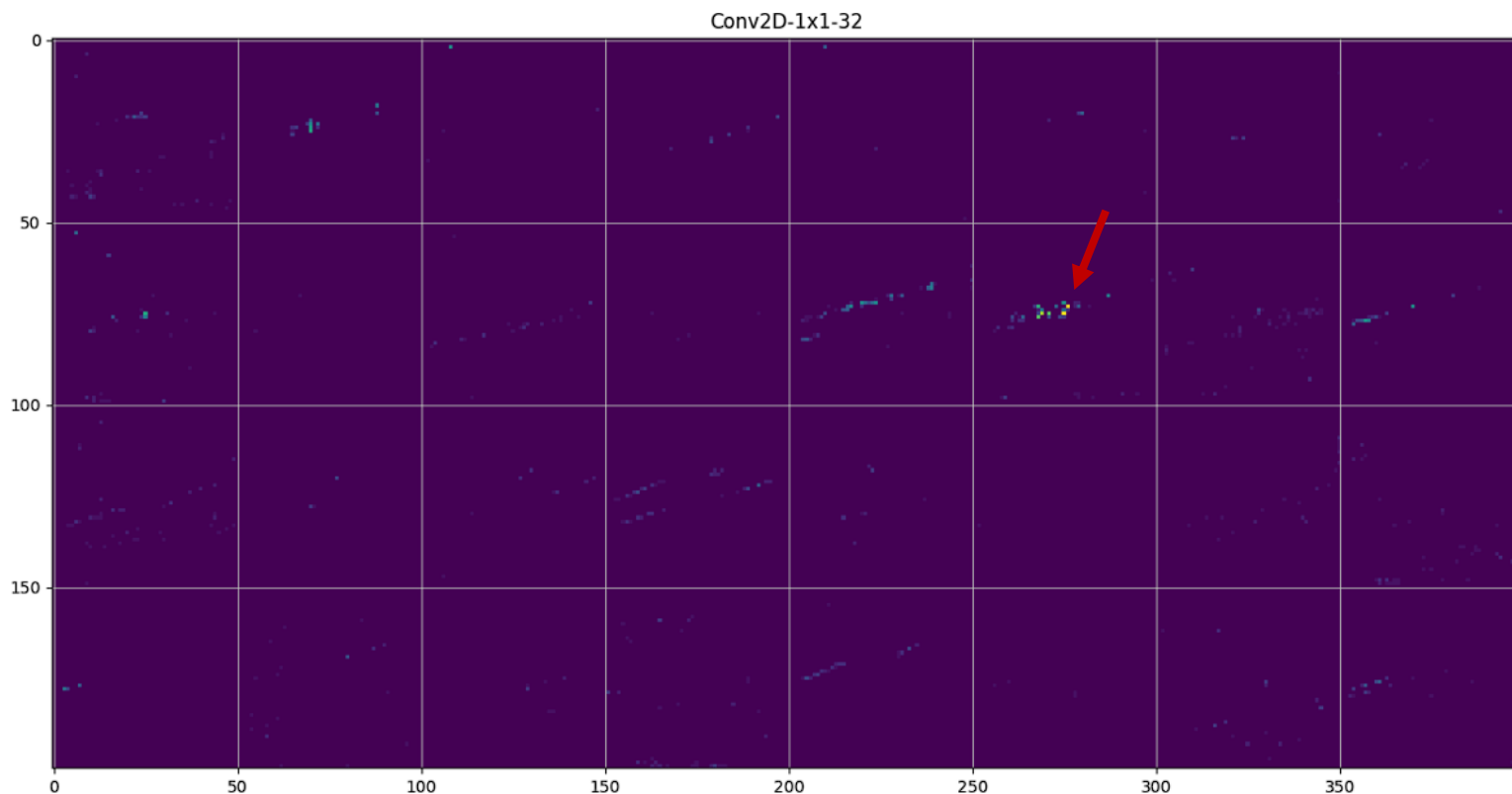


Xception



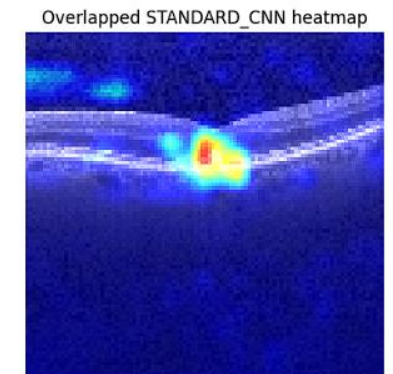
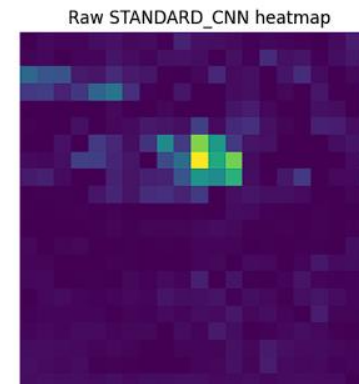
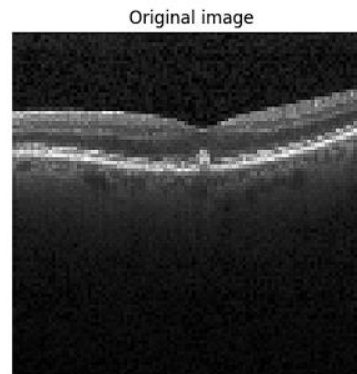
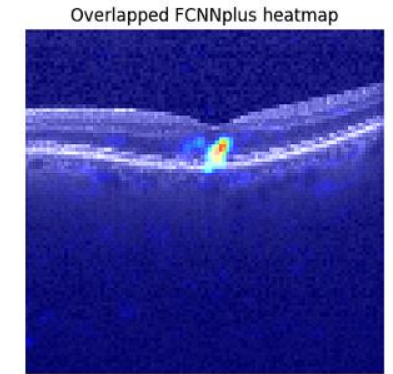
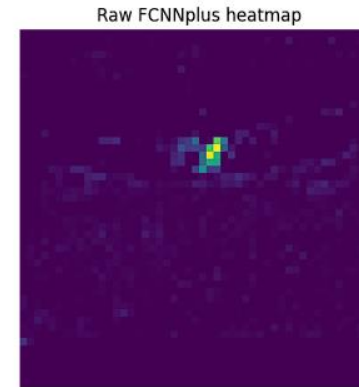
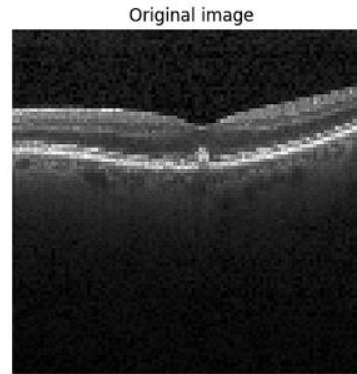
Visual Explainability

L'analisi delle attivazioni intermedie mostra come la rete FCNNplus, che produce heatmap più localizzate e ben definite rispetto al modello STANDARD_CNN, effettivamente vada a valutare le porzioni di immagini rappresentative dello stato patologico.



Visual Explainability

Per fornire ulteriore robustezza al modello FCNNplus si è deciso di confrontare le heatmap, generate con l'algoritmo ScoreCAM, con le medesime ottenute dal modello STANDARD_CNN.

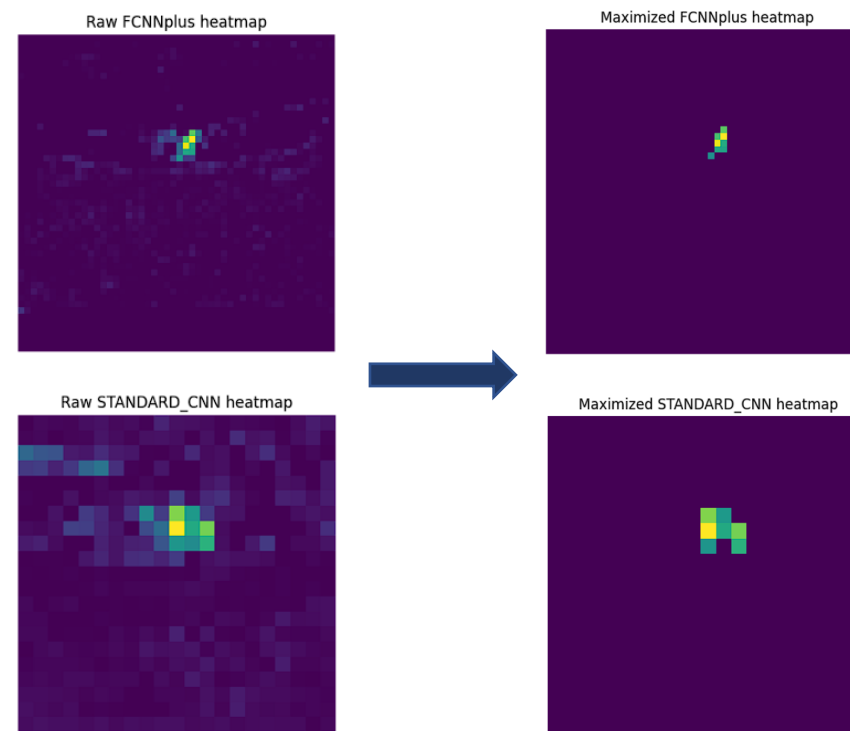


Visual Explainability

Le heatmap generate dai due modelli sulla medesima immagine sono state confrontate tramite la metrica **IM-SSIM** (Inter Model Structural Similarity Index).

La metrica IM-SSIM può assumere valori nell'intervallo $[-1, 1]$, dove 1 indica che le due immagini sono perfettamente uguali.

Per ridurre al minimo il contributo del rumore, e delle aree meno significative delle heatmap, sono stati conservati solo i pixel con valore normalizzato pari o superiore a 0.5.



Class	Mean	Std
CNV	0.8260	0.0621
DME	0.8308	0.0756
DRUSEN	0.8516	0.0678

Conclusioni e Sviluppi futuri

Dall'analisi dei risultati ottenuti sono emerse le seguenti osservazioni:

- La diagnosi di patologie retiniche attraverso modelli CNN-based è attuabile sfruttando le tecnologie attuali.
- Il modello FCNNplus è in grado di distinguere le diverse patologie e identificarle correttamente tramite diversi algoritmi CAM quali GradCAM e ScoreCAM.
- **I risultati ottenuti nei lavori presenti in letteratura, che utilizzano lo stesso dataset, non fanno cenno della presenza di data leakage all'interno dello stesso, pertanto i risultati non possono essere considerati attendibili.**

Ulteriori sviluppi possono focalizzarsi sull'implementazioni di nuove architetture, ottimizzazione degli iperparametri tramite strumenti come KerasTuner e applicazione di tecniche di preprocessing per evidenziare ulteriormente le regioni patologiche.

Bibliografia

- [1] Kermany, Daniel S., et al. "Identifying medical diagnoses and treatable diseases by image-based deep learning." *Cell* 172.5 (2018): 1122-1131.
- [2] Sunija, A. P., et al. "Octnet: A lightweight cnn for retinal disease classification from optical coherence tomography images." *Computer methods and programs in biomedicine* 200 (2021): 105877.
- [3] Chen, Yao-Mei, et al. "Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning." *BMC bioinformatics* 22.5 (2021): 1-16.
- [4] Tsuji, Takumasa, et al. "Classification of optical coherence tomography images using a capsule network." *BMC ophthalmology* 20.1 (2020): 1-9.
- [5] Ara, Rouhollah Kian, et al. "Fast and efficient method for optical coherence tomography images classification using deep learning approach." *Sensors* 22.13 (2022): 4675.
- [6] Mittal, Praveen. "Retinal disease classification using convolutional neural networks algorithm." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12.11 (2021): 5681-5689.

