



# Deep Learning

Classificazione binaria di melanomi con reti  
neurali convoluzionali

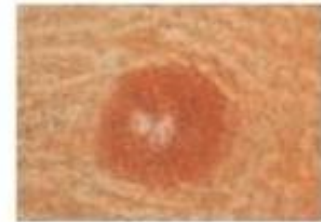
Andrea Moretti

# Il problema

Il cancro alla pelle è uno dei più diffusi tipi di cancro al mondo e la sua **diagnosi** si basa prevalentemente sull'**esame visivo** e successivamente si effettua un'analisi dermoscópica

La classificazione automatica delle lesioni cutanee tramite immagini rappresenta un compito complesso a causa della variabilità sottile e dettagliata nell'aspetto delle lesioni stesse

Benign



Malignant



# Il melanoma

È il tipo di cancro alla pelle più mortale

Dati dell'American Cancer Society (AMS) rispetto alla popolazione americana:

|                     | Casi    | Morti      |
|---------------------|---------|------------|
| Ogni anno (America) | 105`000 | 8`500 (8%) |

Nel 2017 i morti erano 10`000, evidenziando una diminuzione della mortalità, ma il numero di diagnosi di melanoma dal 2015 al 2025 è cresciuto del 42%

# L'uso di AI in medicina

Negli ultimi anni sono stati condotti numerosi studi sull'applicazione dell'intelligenza artificiale in ambito oncologico per la classificazione del melanoma, evidenziando come modello di CNN siano in grado di raggiungere **performance** diagnostiche **comparabili** a quelle dei **dermatologi**

L'individuazione precoce risulta fondamentale, poiché il tasso di sopravvivenza è stimato essere al 99% se diagnosticato nelle fasi iniziali, ma si riduce drasticamente a circa il 14% quando la malattia viene rilevata negli stadi più avanzati

# Dataset utilizzato

Il dataset che ho utilizzato è l'ISIC Challenge 2018, che contiene circa 10`000 immagini di cui solo 1824 rappresentano melanomi (poco più del 18%)

Per questo motivo ho utilizzato una **loss pesata** per bilanciare il contributo per classe

Il **train**, **validation** e **test** sono poi stati suddivisi secondo **80/10/10** e mantengono lo stesso sbilanciamento tra classi

# Approccio

In questo progetto ho condotto un'analisi comparativa delle performance di diversi modelli di CNN per la classificazione di melanomi (binary classification):

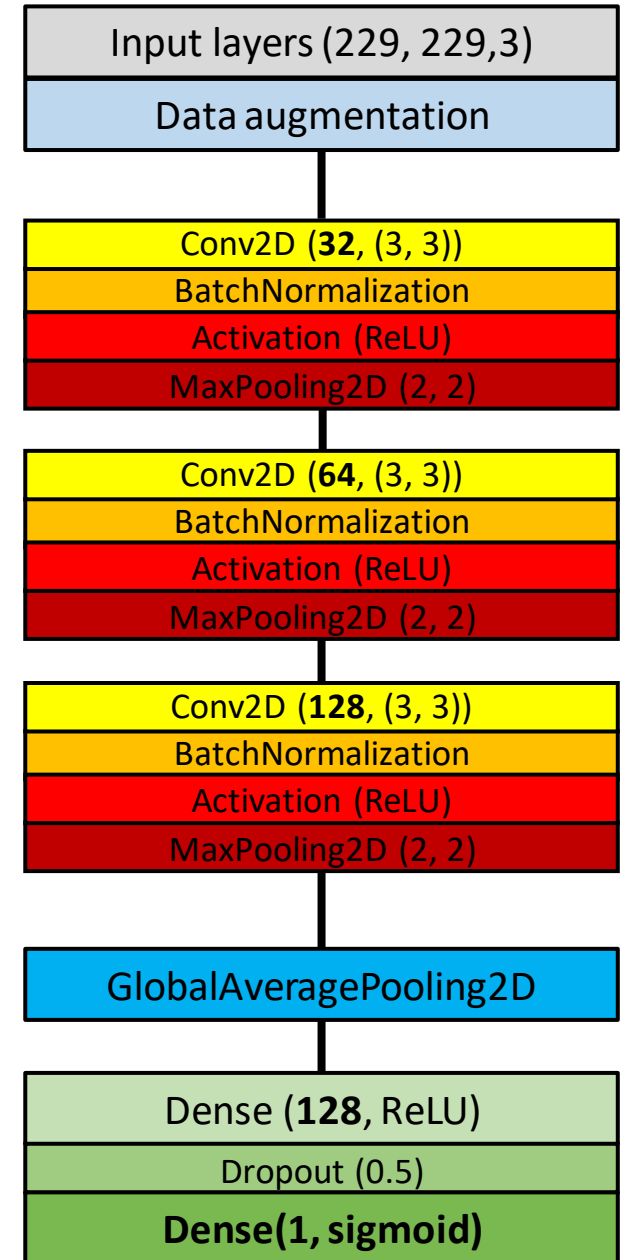
- un modello naïve sviluppato **da zero**
- un modello basato su transfer learning con **InceptionV3**
- un modello basato su transfer learning con **EfficientNet-B3**.

# Modello sviluppato da zero

Il modello è stato allenato esclusivamente sul dataset che ho utilizzato

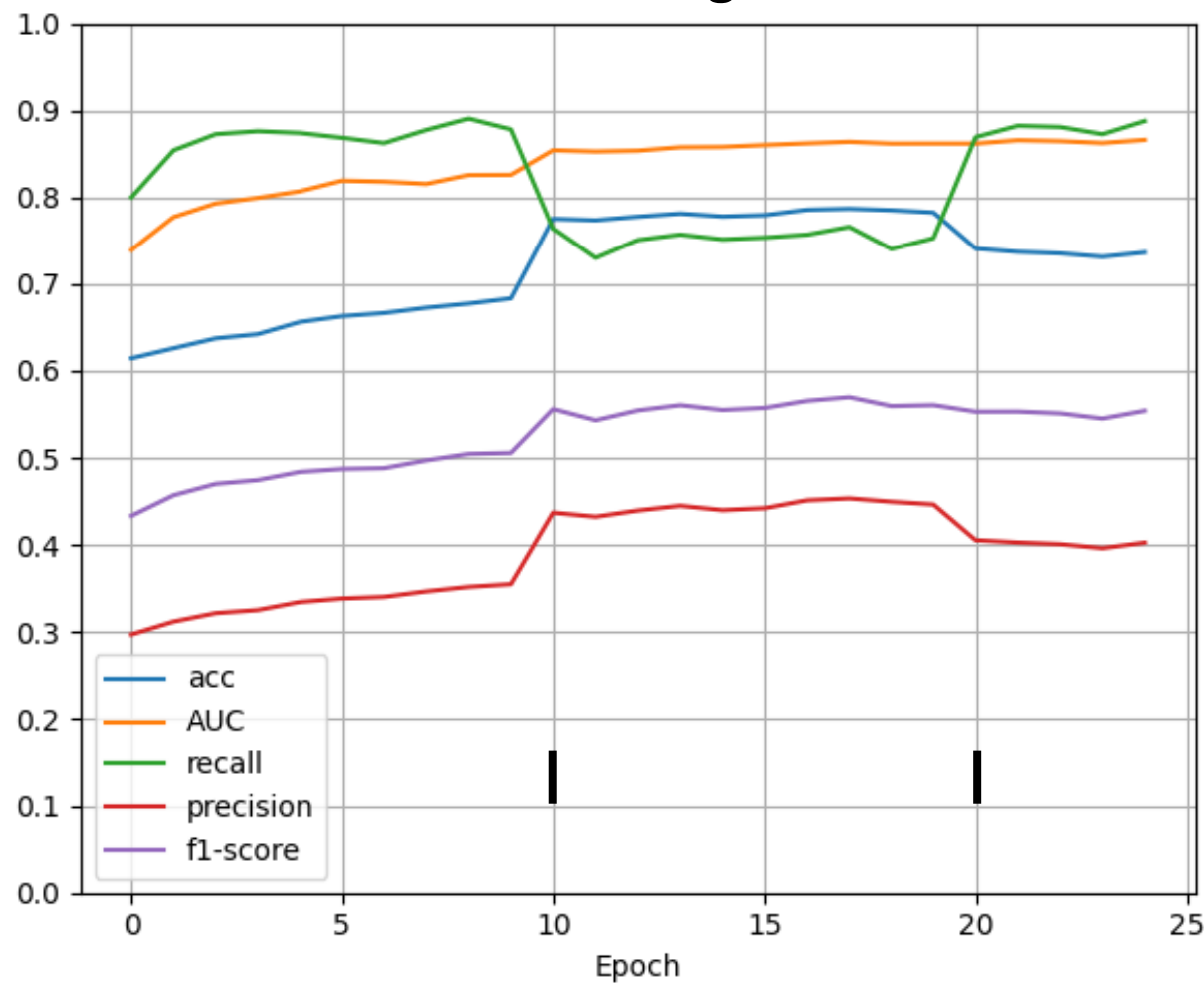
Le immagini utilizzate hanno una dimensione standard di 299×299 RGB

|                      |   |
|----------------------|---|
| <b>Class weight</b>  | 0 → 0.614      1 → 2.702                                      |
| <b>Loss</b>          | binary crossentropy   |
| <b>Optimizer</b>     | Adam  |
| <b>Batch size</b>    | 32  |
| <b>Learning rate</b> | 1e-4 (0-10 epoch)<br>5e-5 (10-20 epoch)<br>1e-5 (20-25 epoch) |

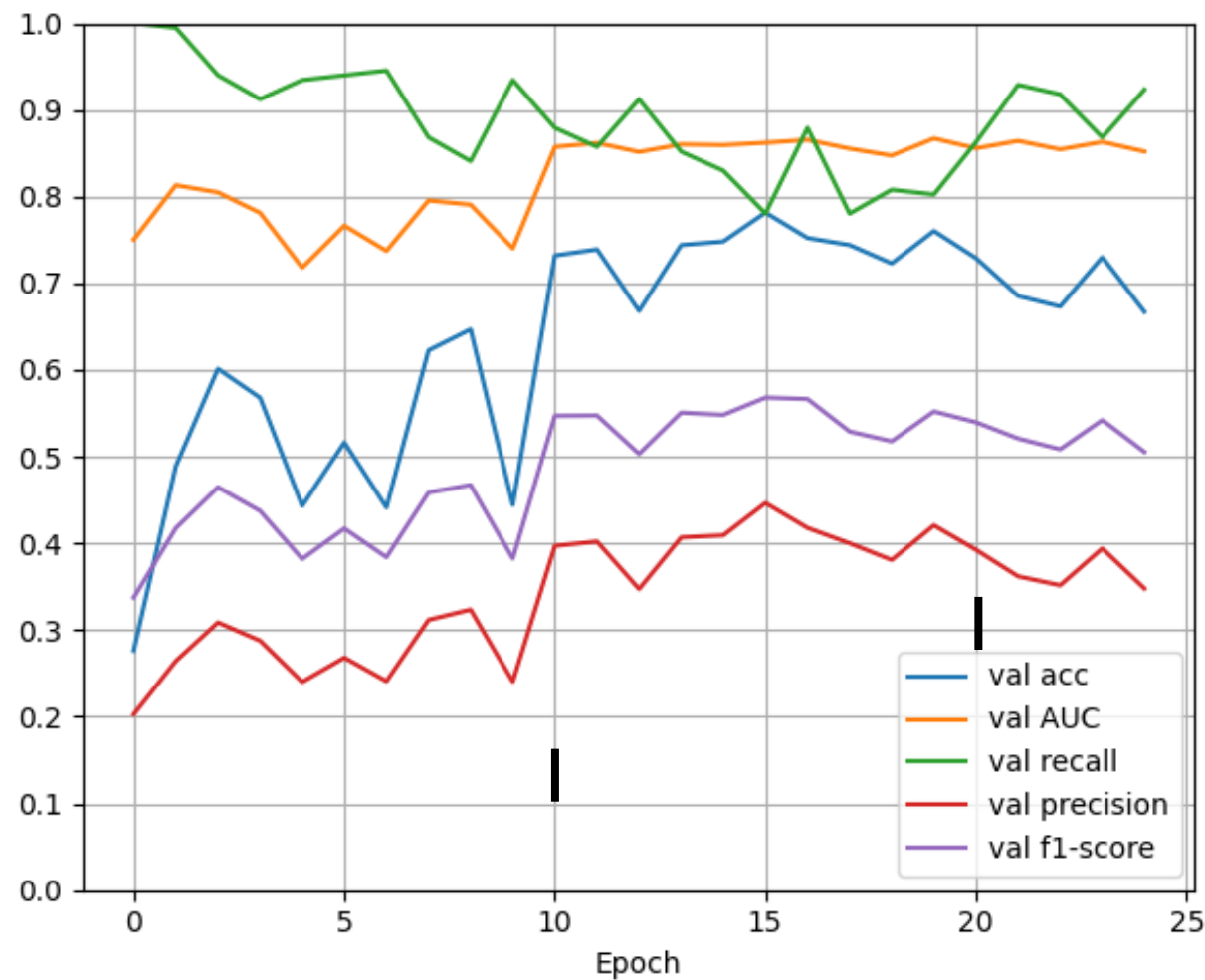


# Modello sviluppato da zero

## Training



## Validation





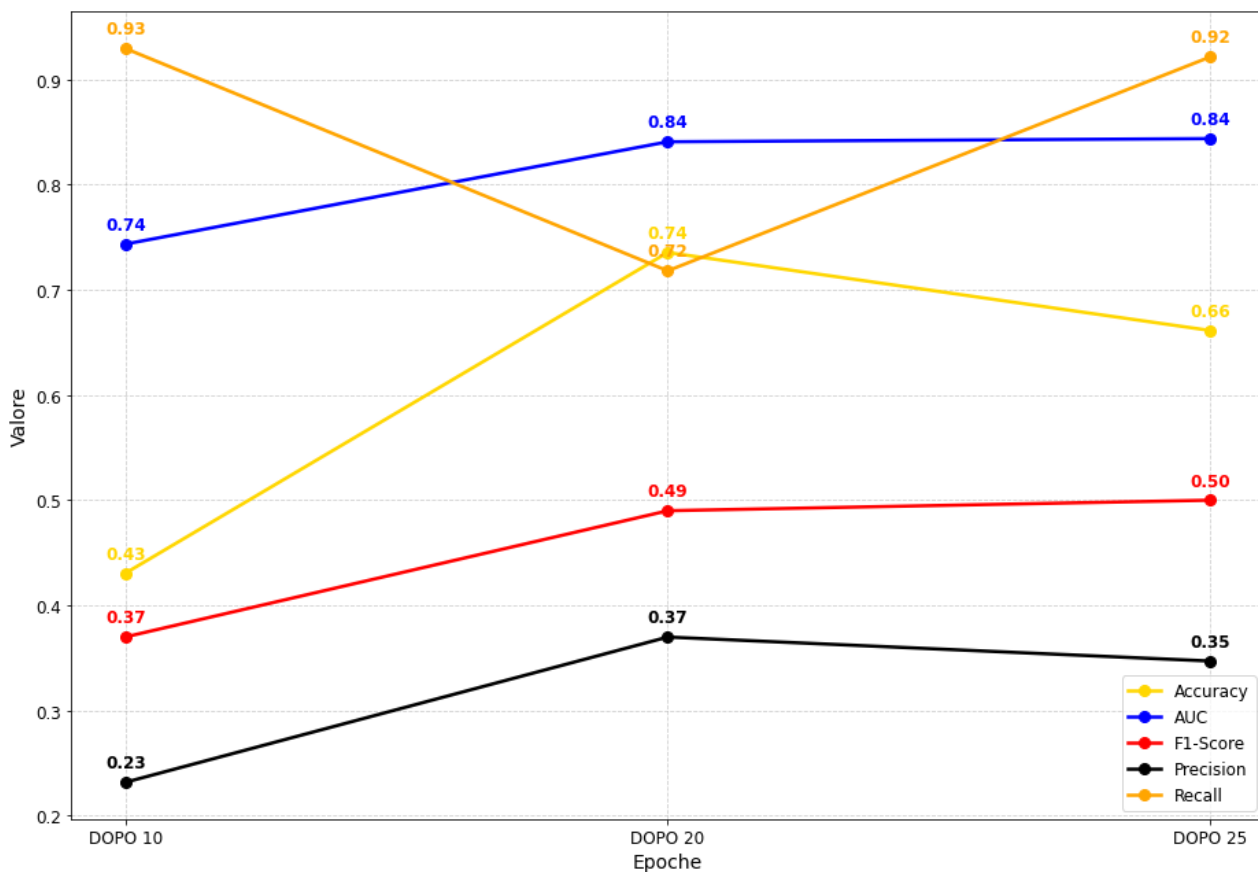
# Modello sviluppato da zero

1

2

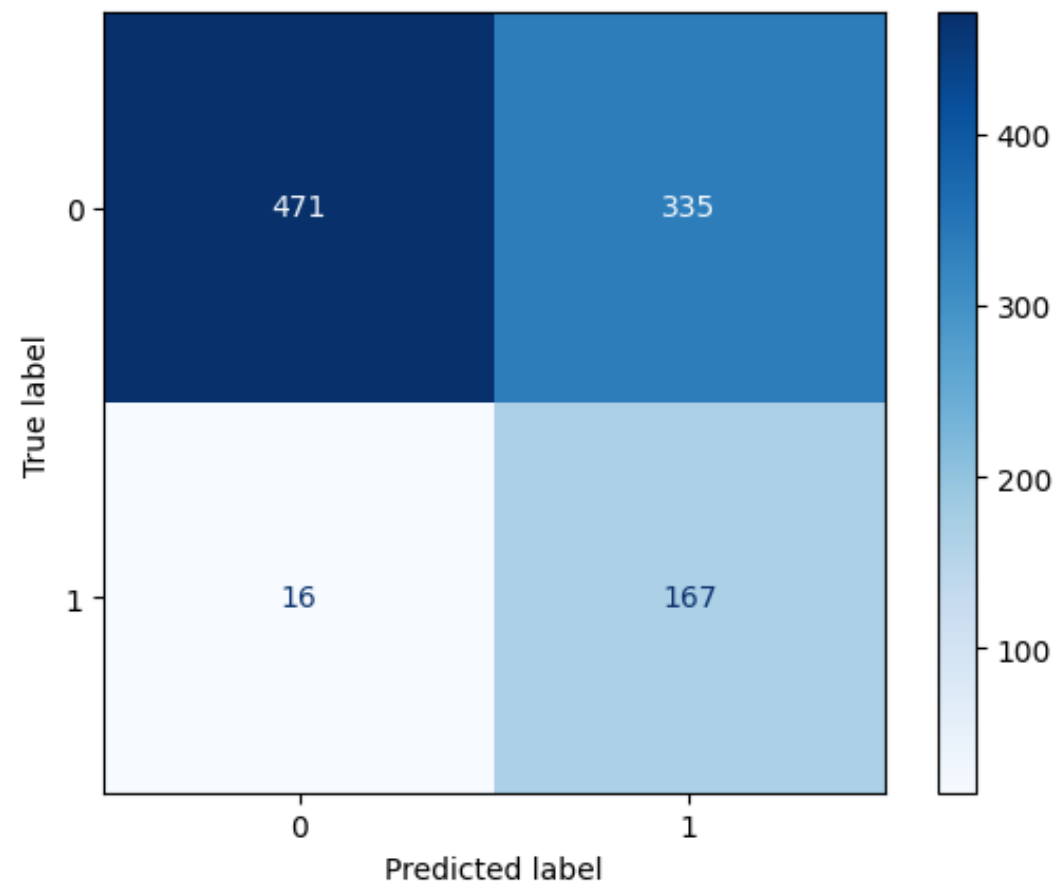
3

## Evaluation sul test set



Media valori predizioni su test set:

## Confusion matrix



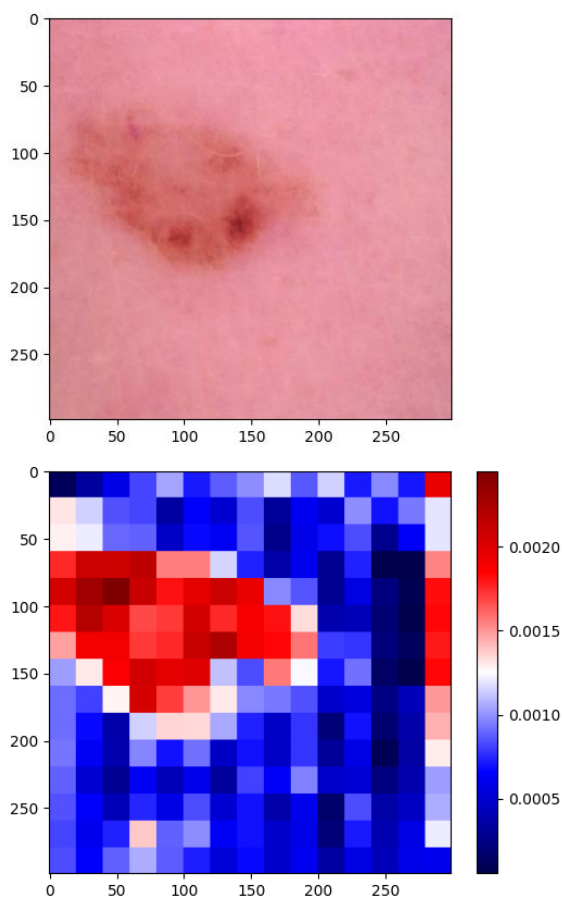
**Label 0: 0.3751**

**Label 1: 0.7515**

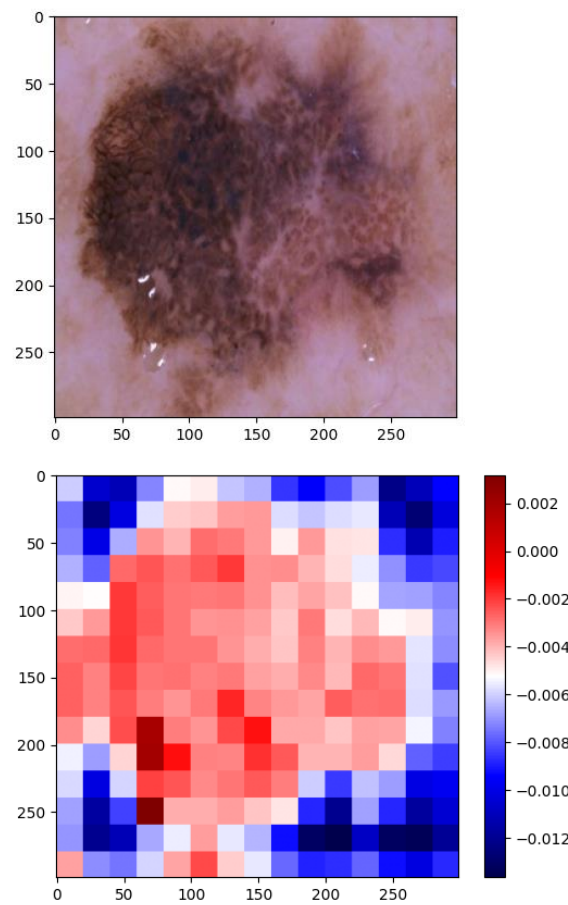
# Modello sviluppato da zero

`patch_size = 20`

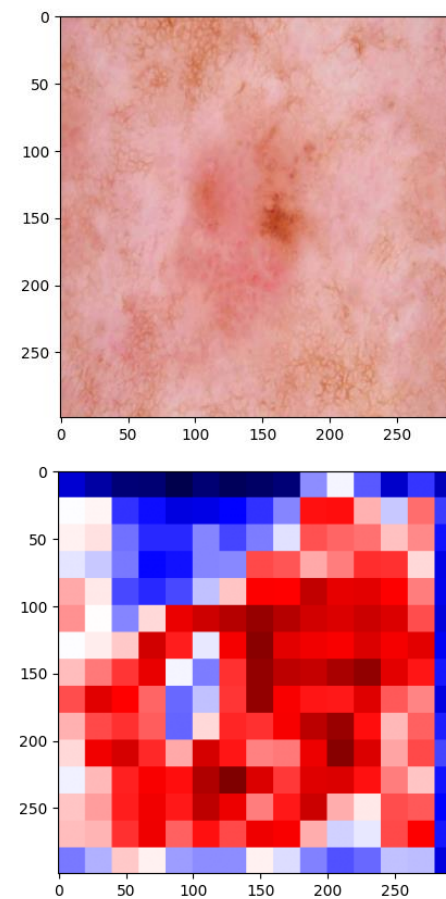
Prediction: 0.0082  
Label: 0



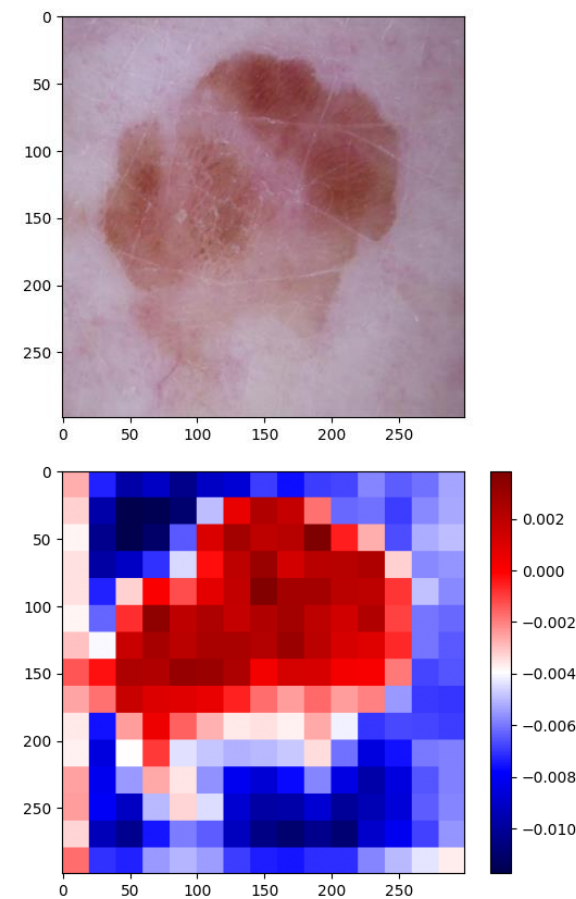
Prediction: 0.7914  
Label: 1



Prediction: 0.9153  
Label: 1



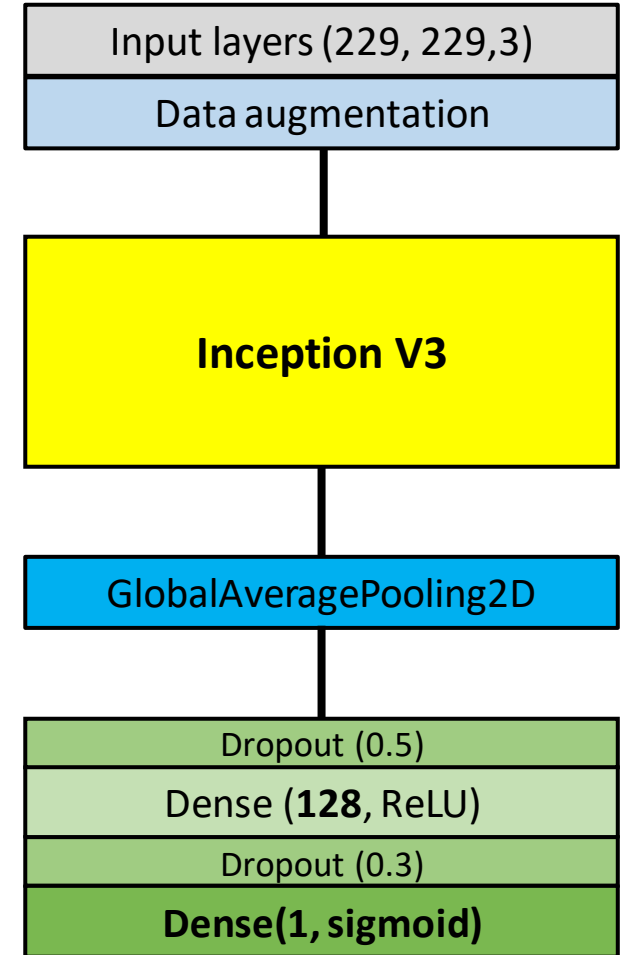
Prediction: 0.7363  
Label: 0



# Modello basato su InceptionV3

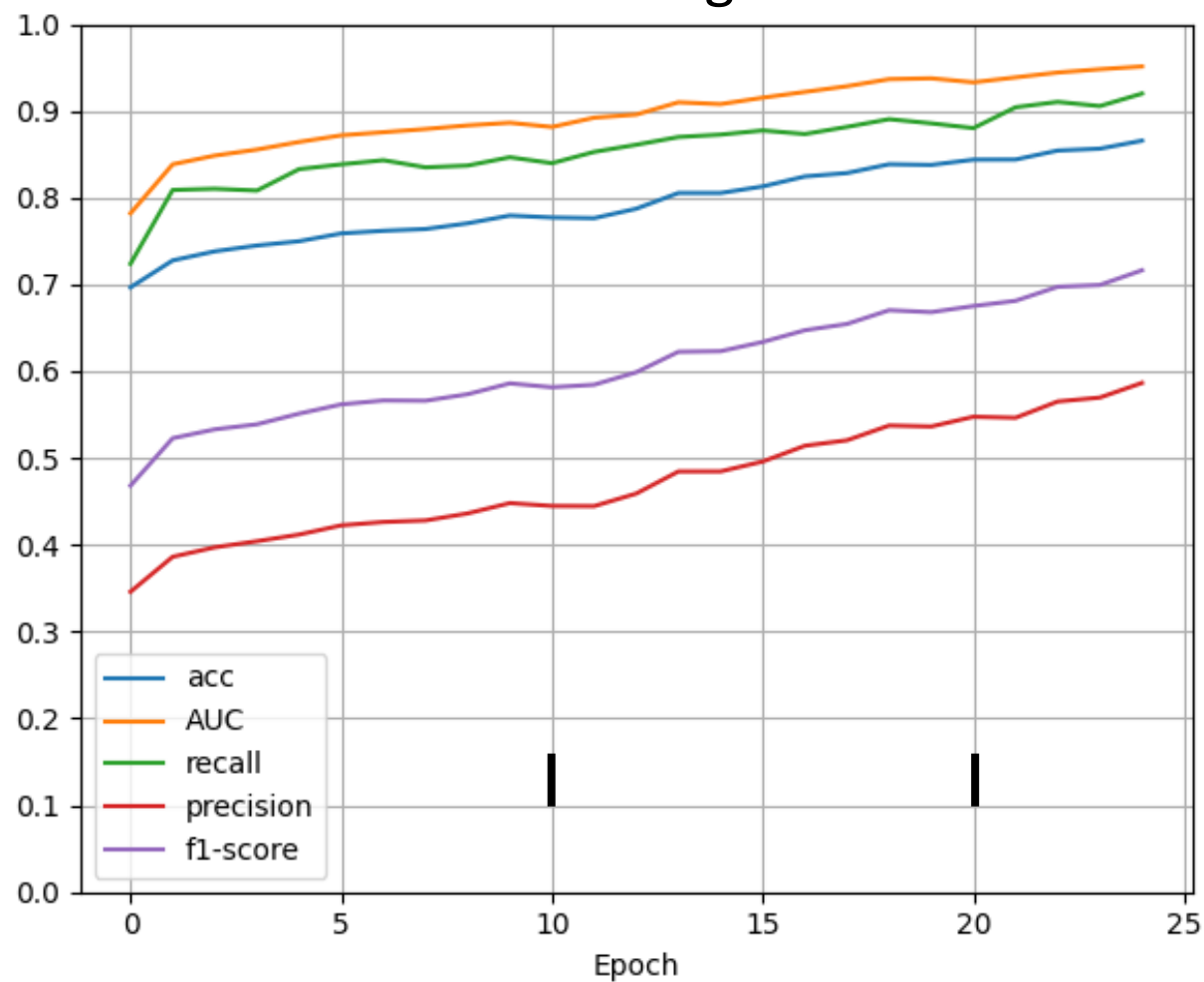
Tramite transfer learning utilizzo Inception V3 con i pesi addestrati su ImageNet e sbloccando inizialmente gli ultimi 25 layers, poi gli ultimi 45 e infine gli ultimi 65

|                      |   |
|----------------------|---|
| <b>Class weight</b>  | 0 → 0.614      1 → 2.702                                      |
| <b>Loss</b>          | binary crossentropy   |
| <b>Optimizer</b>     | Adam  |
| <b>Batch size</b>    | 32  |
| <b>Learning rate</b> | 1e-4 (0-10 epoch)<br>5e-5 (10-20 epoch)<br>1e-5 (20-25 epoch) |

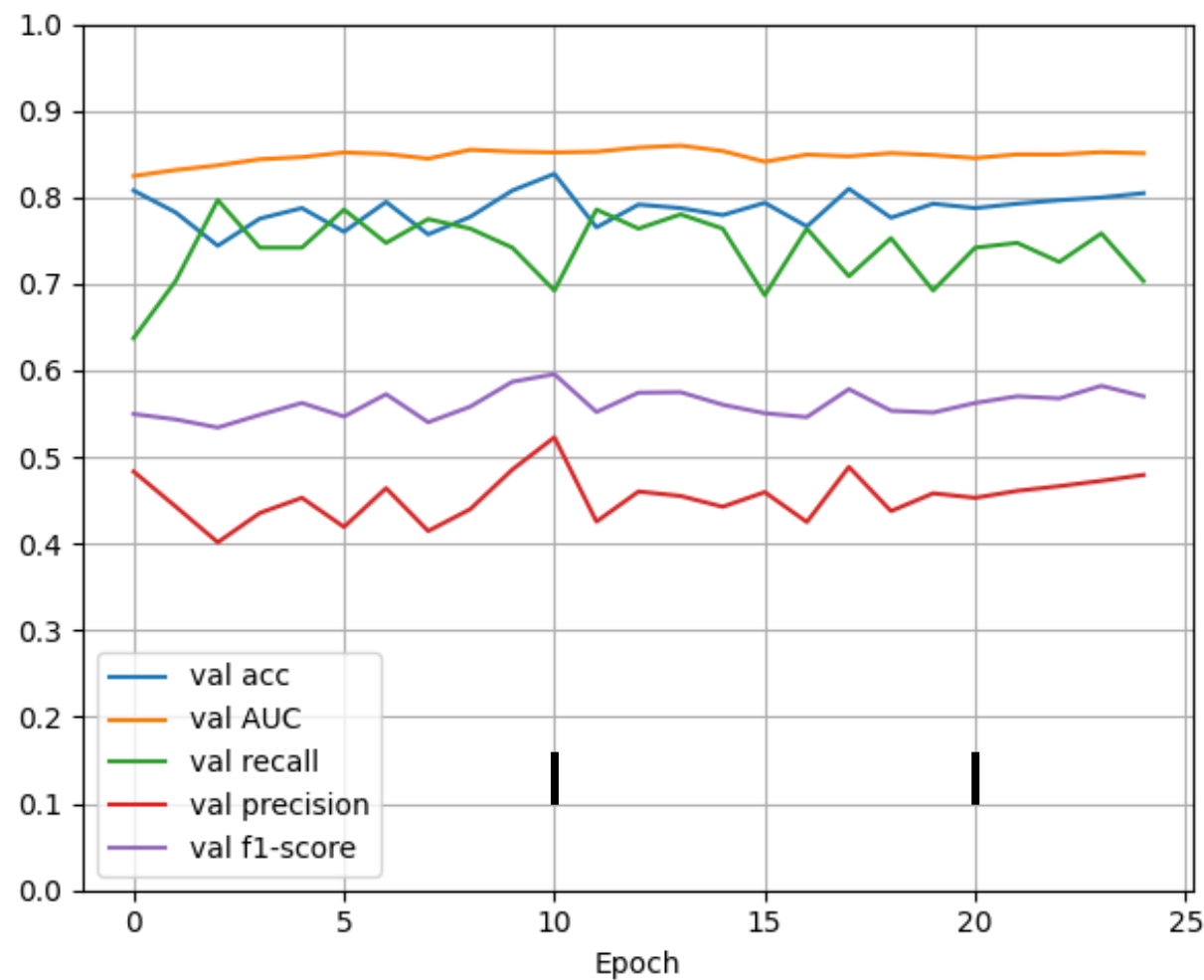


# Modello basato su InceptionV3

## Training



## Validation



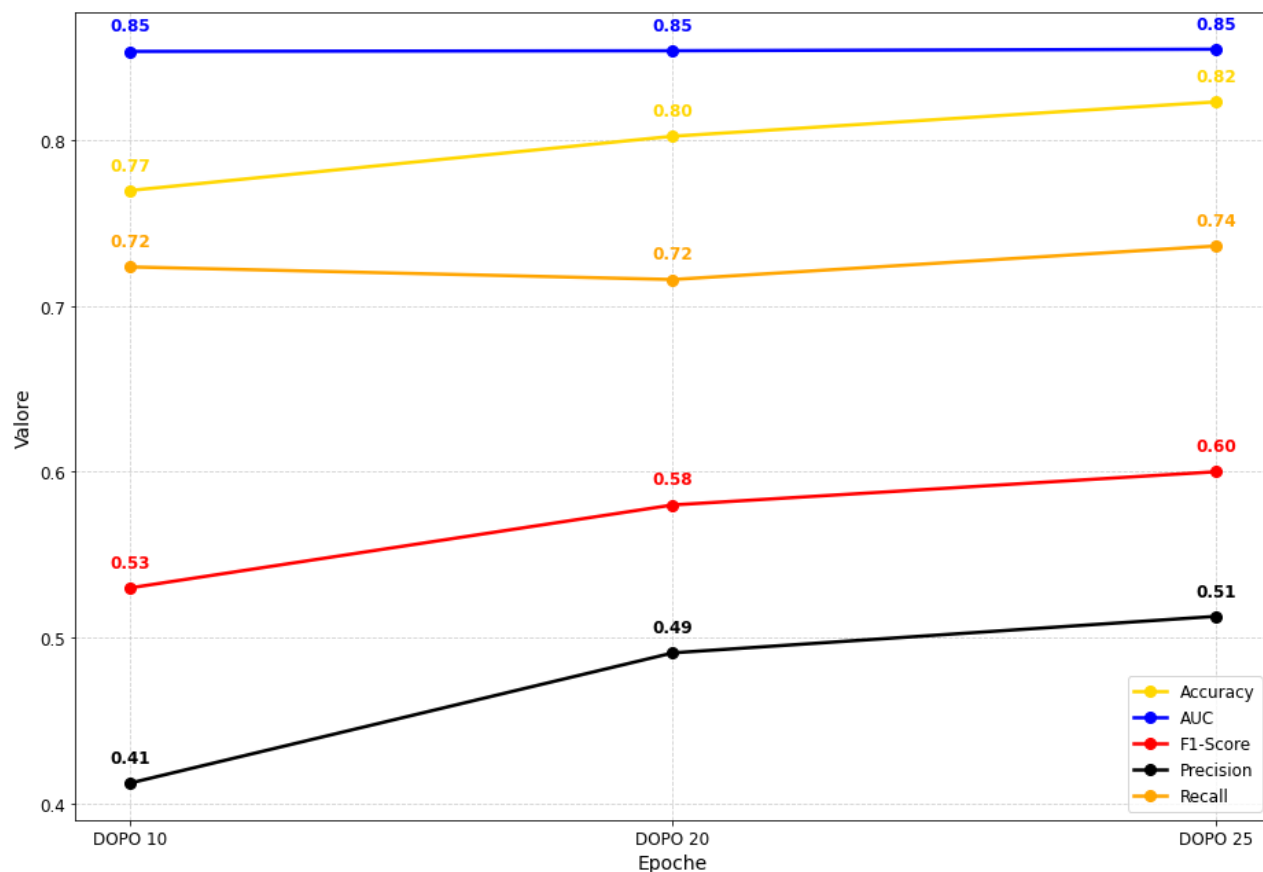
# Modello basato su InceptionV3

1

2

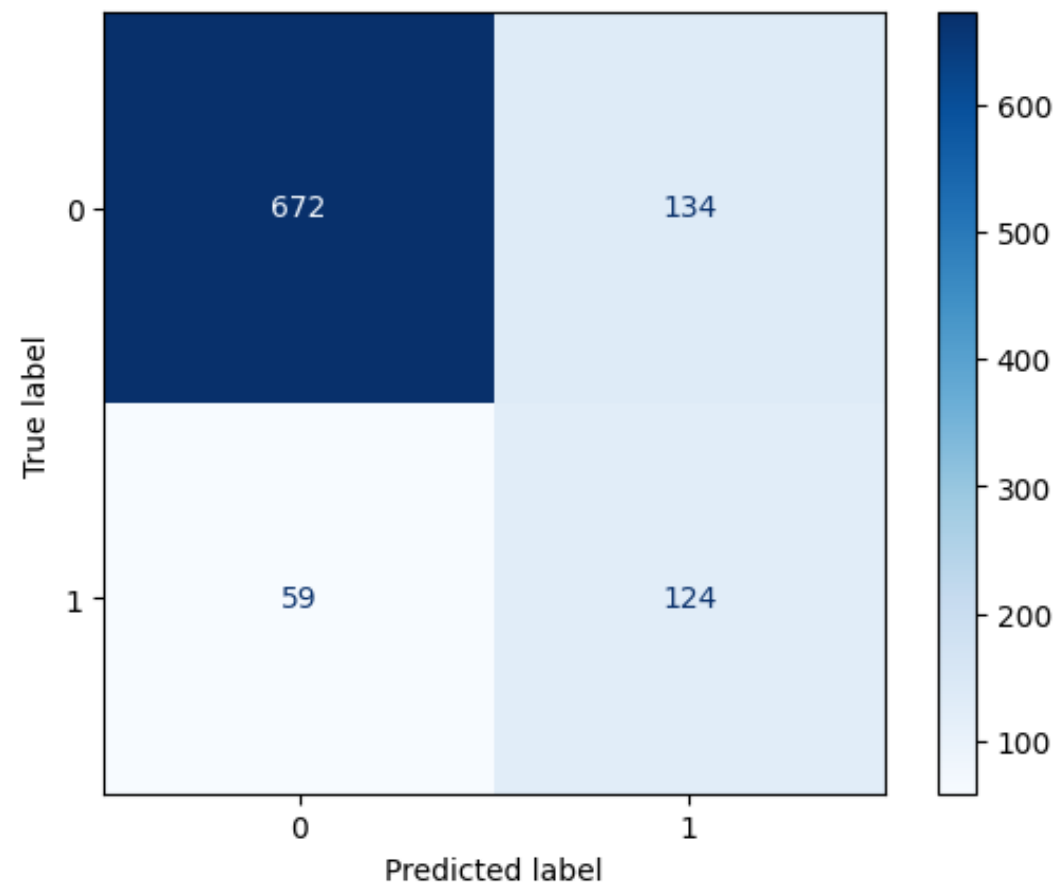
3

Evaluation sul test set



Media valori predizioni su test set:

Confusion matrix



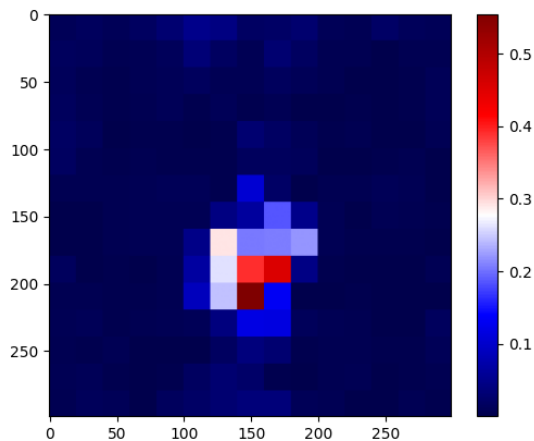
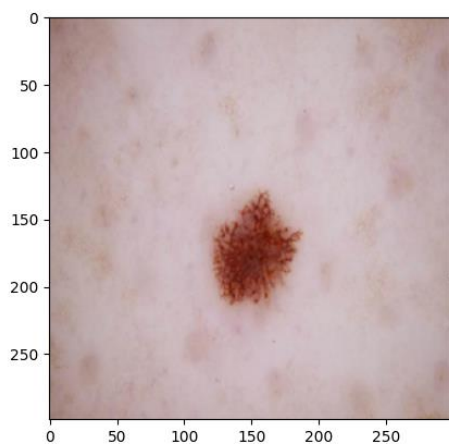
**Label 0: 0.1868**

**Label 1: 0.6530**

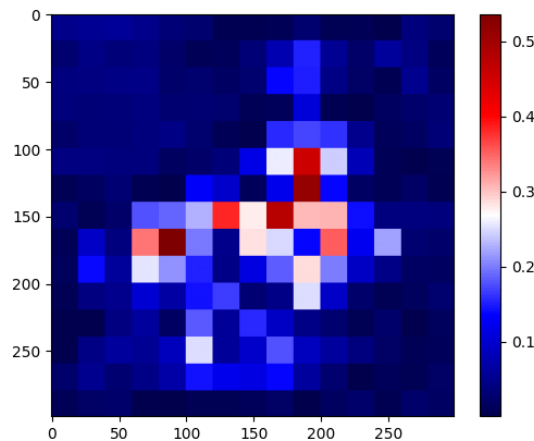
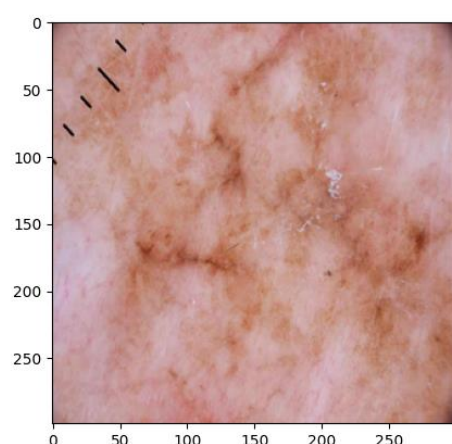
# Modello basato su InceptionV3

`patch_size = 20`

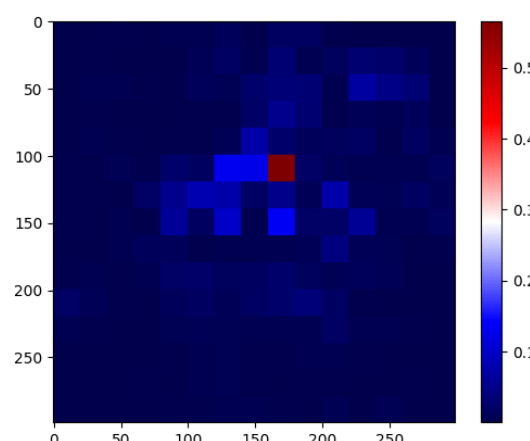
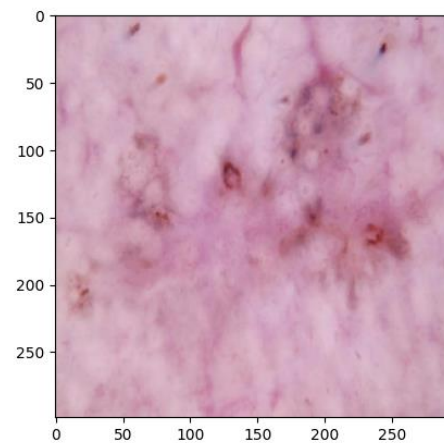
Prediction: 0.0125  
Label: 0



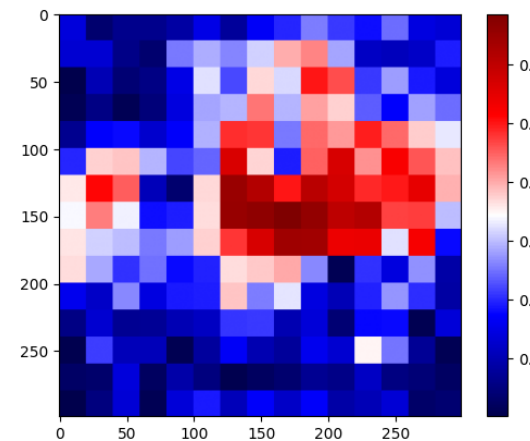
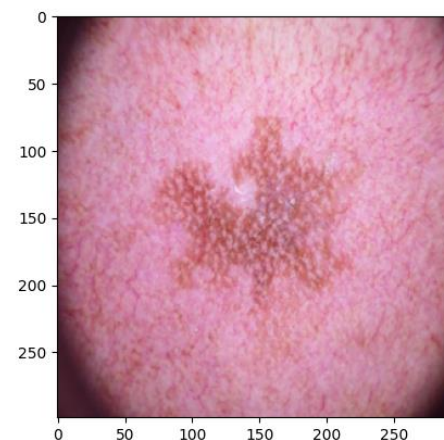
Prediction: 0.9351  
Label: 1



Prediction: 0.9975  
Label: 1



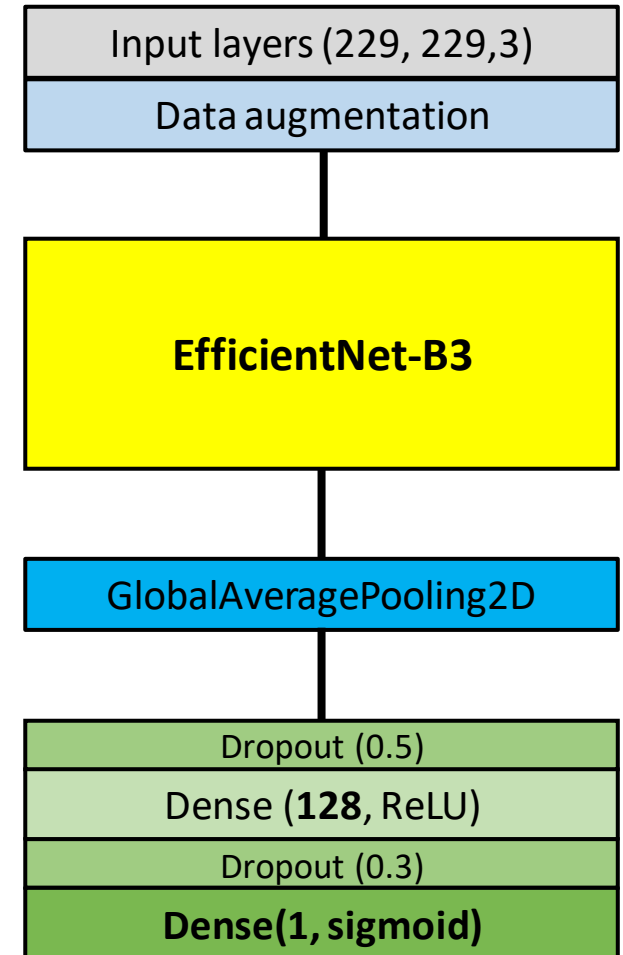
Prediction: 0.7695  
Label: 0



# Modello basato su EfficientNet-B3

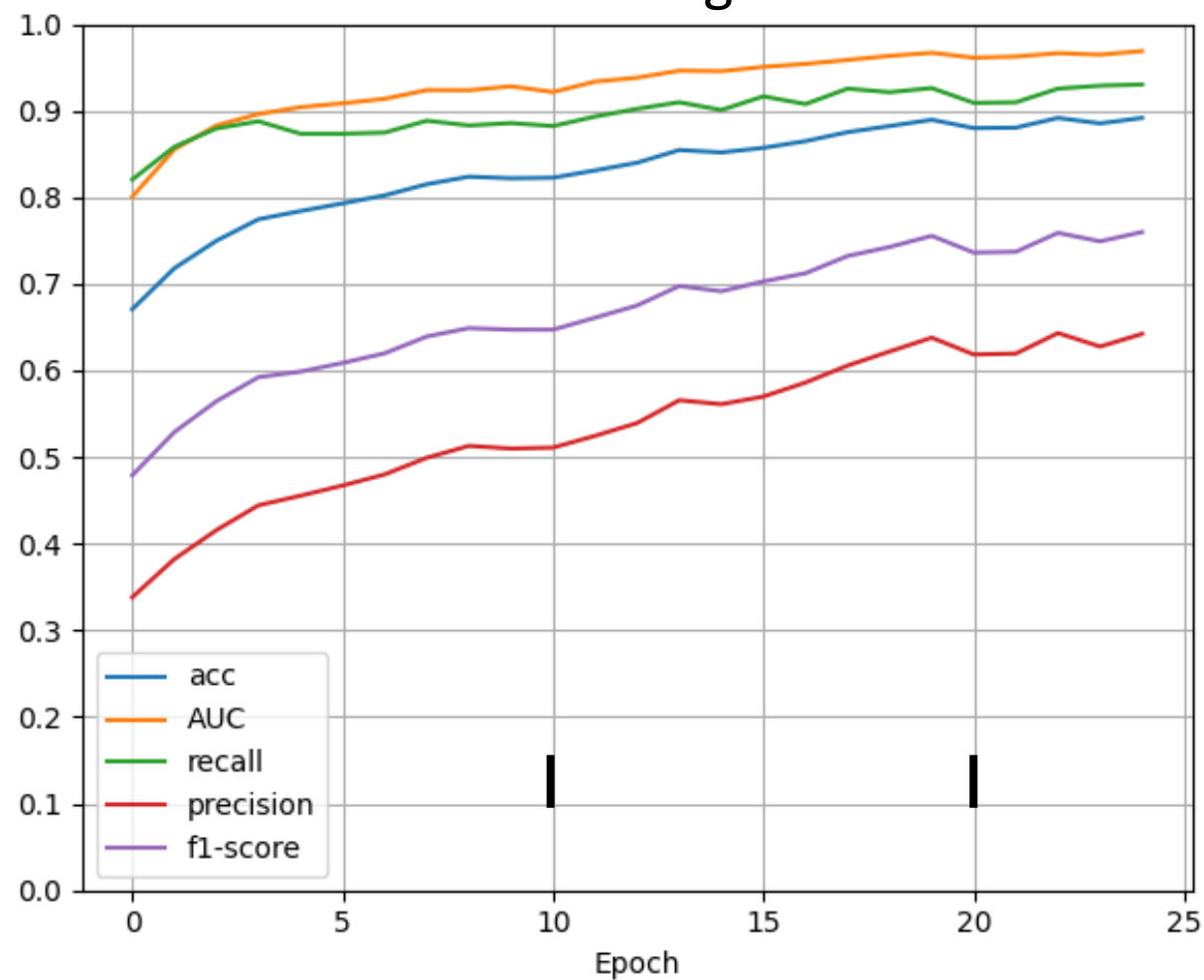
In maniera simile utilizzo EfficientNet-B3 con i pesi addestrati su ImageNet e sbloccando inizialmente gli ultimi 30 layers, poi gli ultimi 60 e infine gli ultimi 80

|                      |   |
|----------------------|---|
| <b>Class weight</b>  | 0 → 0.614      1 → 2.702                                      |
| <b>Loss</b>          | binary crossentropy   |
| <b>Optimizer</b>     | Adam  |
| <b>Batch size</b>    | 32  |
| <b>Learning rate</b> | 1e-4 (0-10 epoch)<br>5e-5 (10-20 epoch)<br>1e-5 (20-25 epoch) |

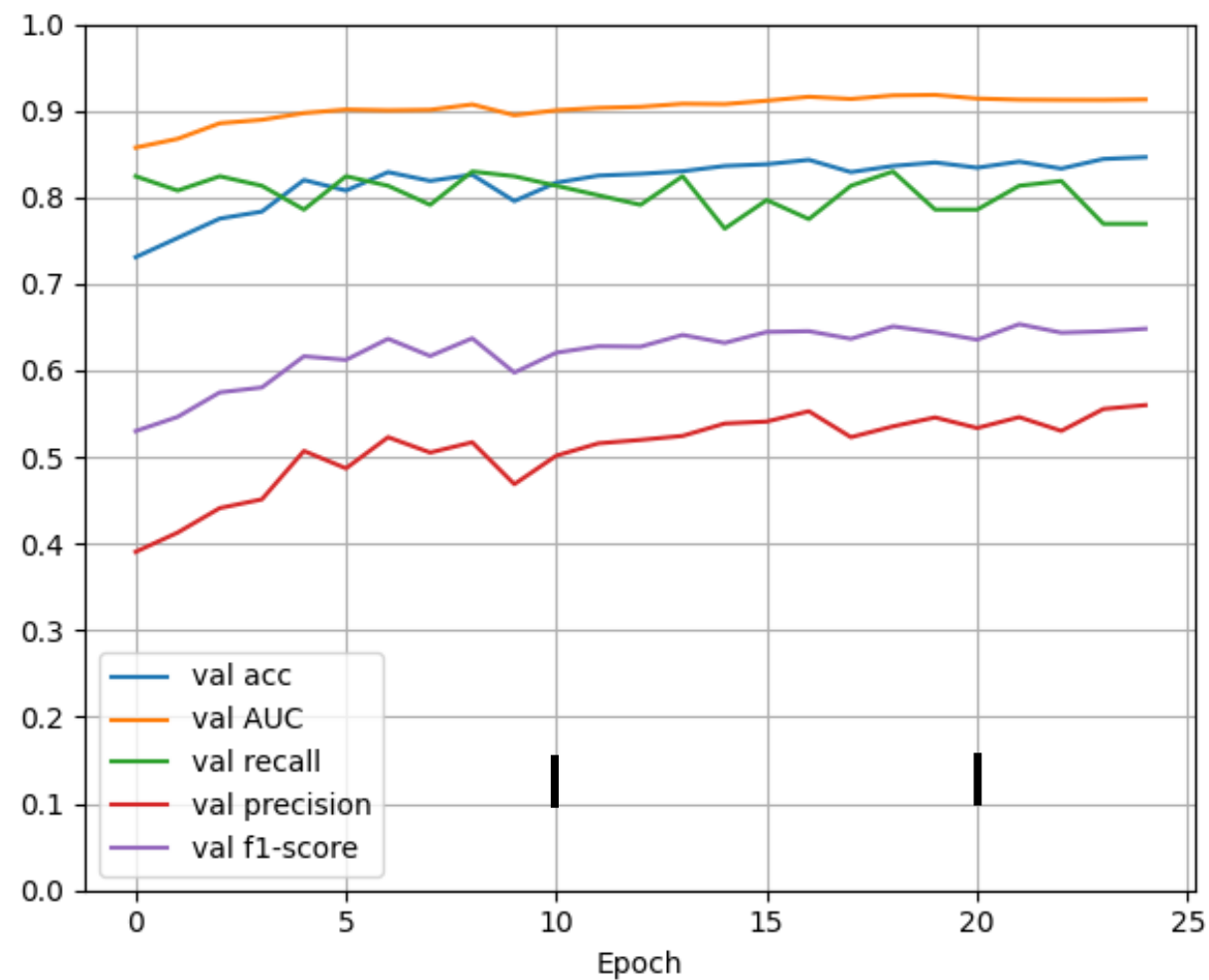


# Modello basato su EfficientNet-B3

## Training



## Validation





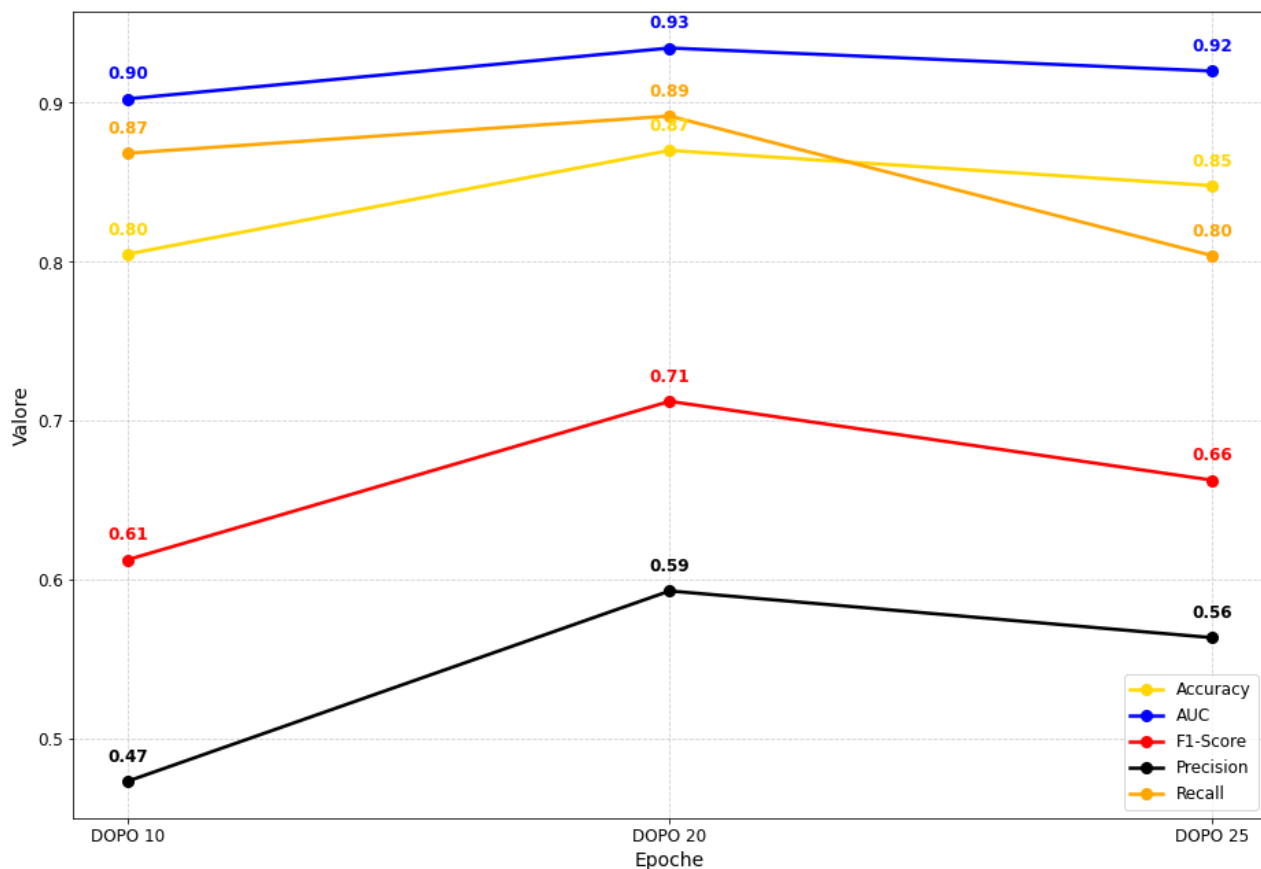
# Modello basato su EfficientNet-B3

1

2

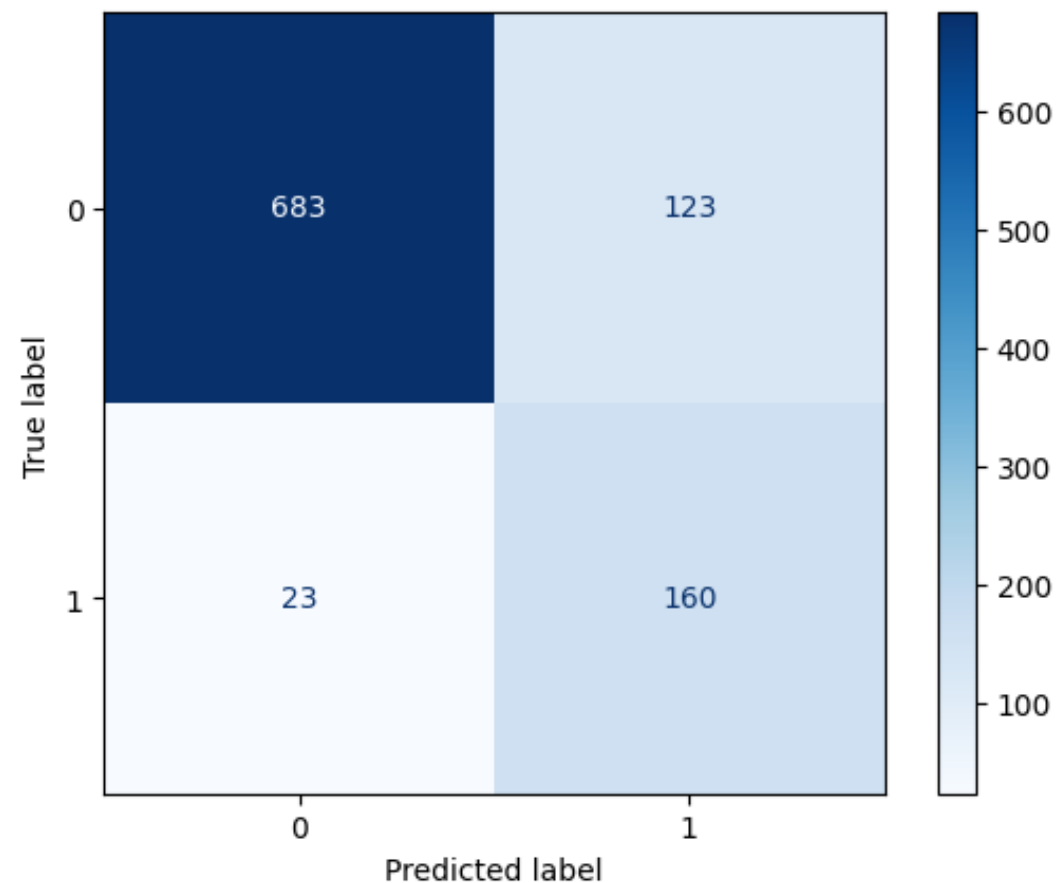
3

Evaluation sul test set



Media valori predizioni su test set:

Confusion matrix

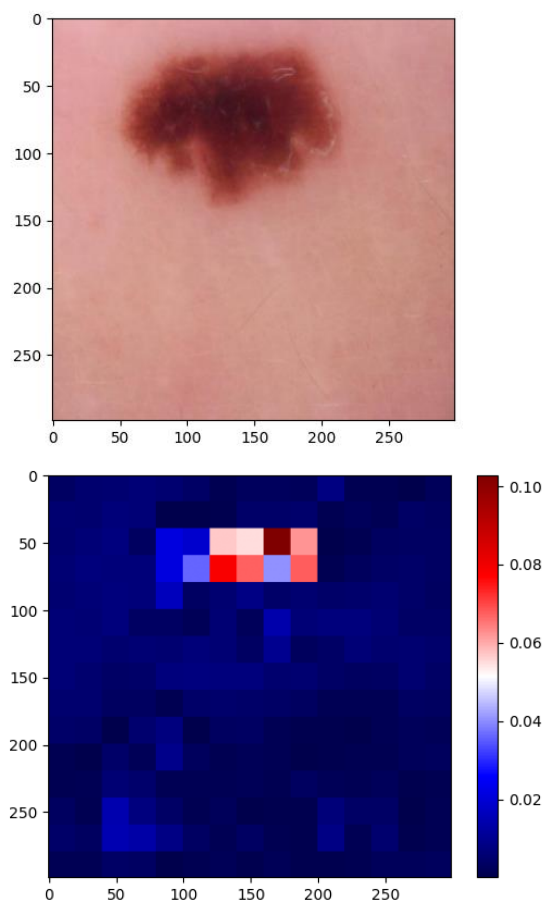


**Label 0: 0.1730**

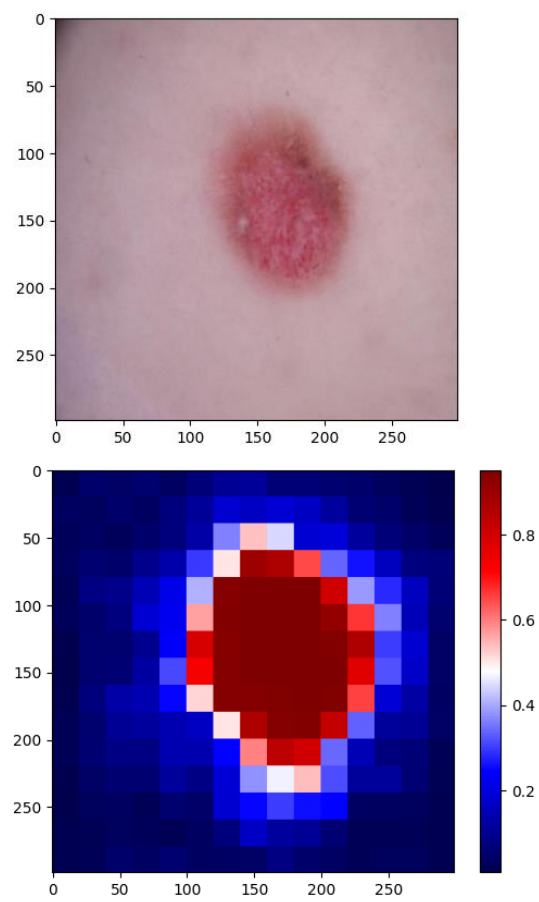
**Label 1: 0.8041**

# Modello basato su EfficientNet-B3 `patch_size = 20`

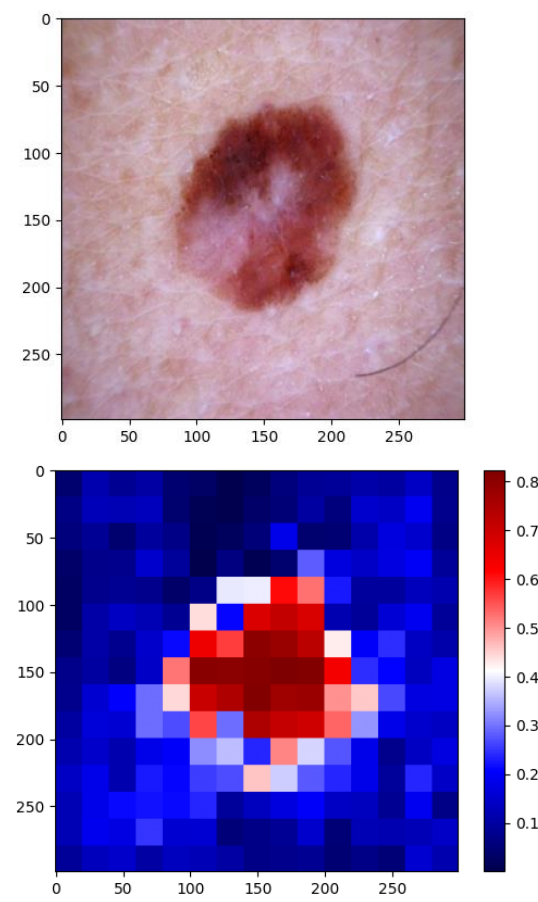
Prediction: 0.0002  
Label: 0



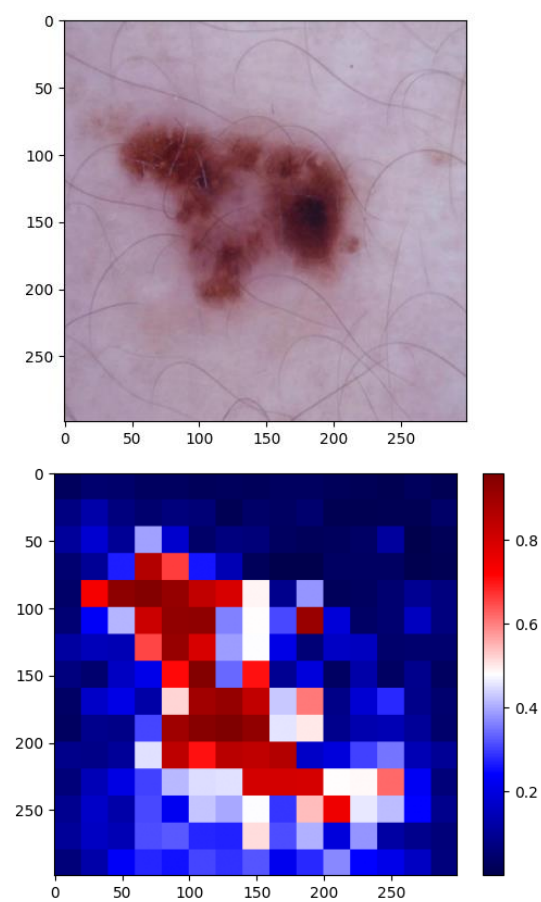
Prediction: 0.9521  
Label: 1



Prediction: 0.8363  
Label: 1



Prediction: 0.9625  
Label: 0



# Conclusioni

Nonostante i limiti dovuti allo sbilanciamento del dataset i modelli hanno prodotto risultati discreti, in particolare i modelli basati su transfer learning hanno superato significativamente il modello sviluppato da zero, raggiungendo performance superiori con una buona affidabilità nel riconoscimento di melanomi

Questi modelli possono essere impiegati come supporto per una prima valutazione di lesioni dermatologiche, anche se non rappresentano uno strumento diagnostico autonomo, poiché le loro capacità predittive non garantiscono ancora un livello di precisione sufficiente