

Università degli Studi di Napoli Federico II



**Dipartimento di Ingegneria Elettrica
delle Tecnologie dell'Informazione**



Elaborazione di Segnali Multimediali

Docente : Luisa Verdoliva

Gruppo 13

**Pacifico Catapano - Marcello Donisi
Lorenzo Discolo - Salvatore Giugliano**

ESM



Pacifico Catapano



Lorenzo Discolo

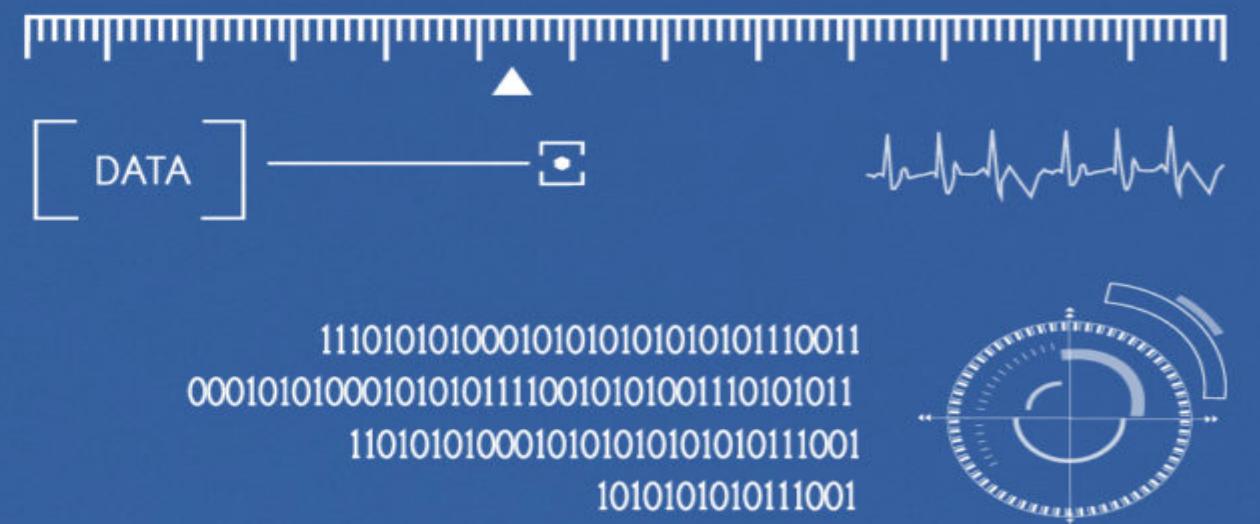


Marcello Donisi



Salvatore Giugliano

Riconoscimento degli attributi facciali

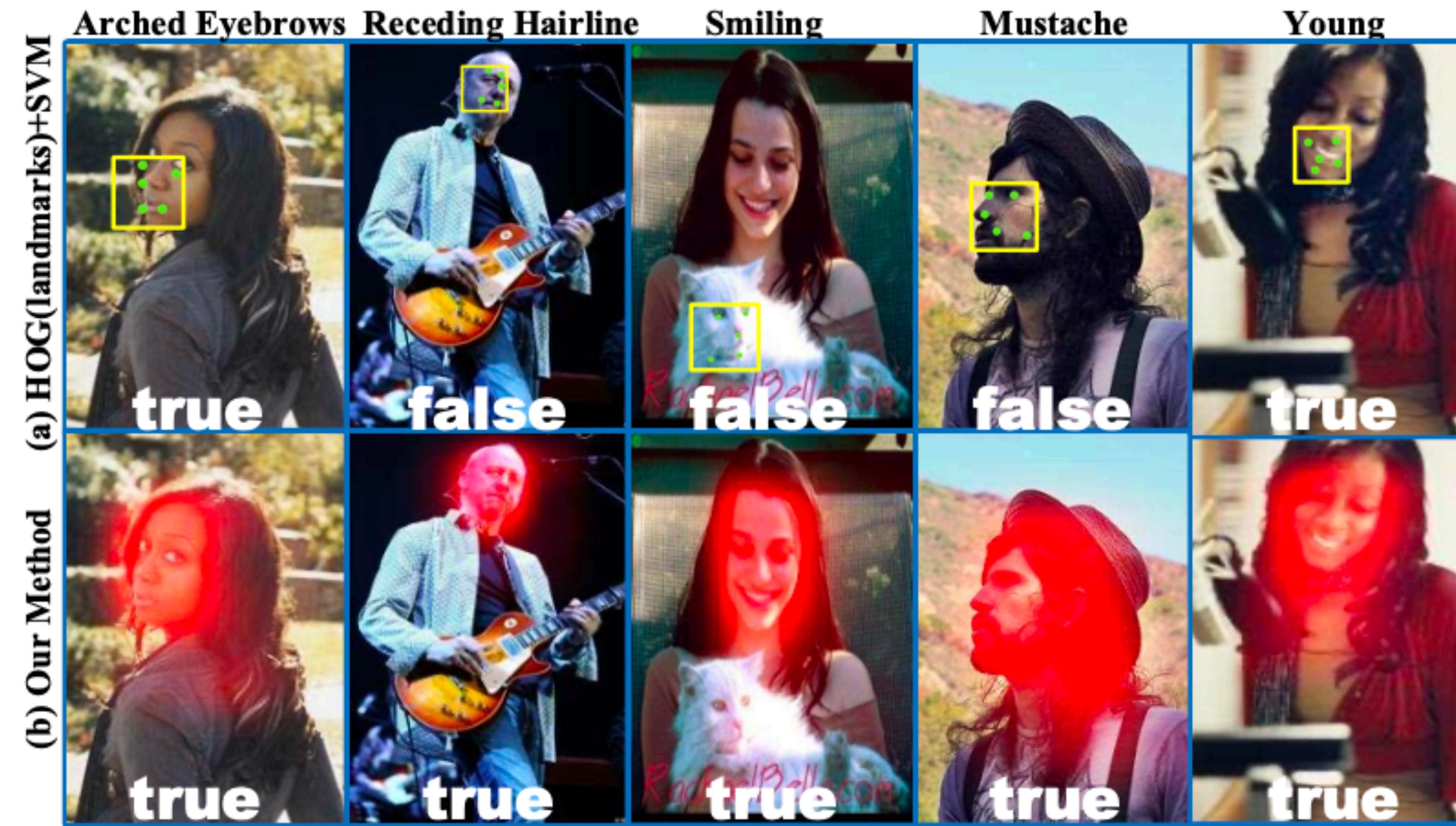


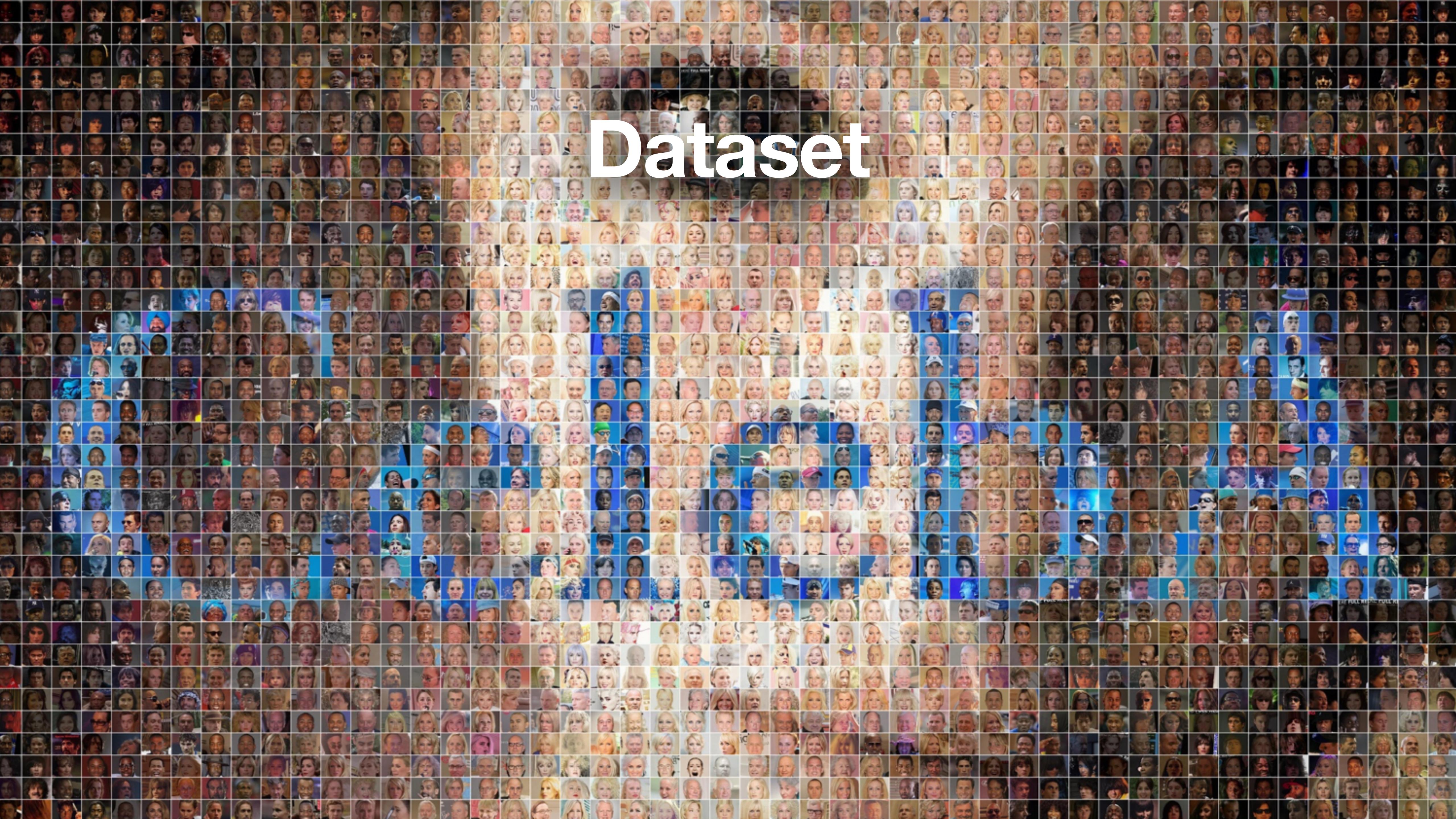
NO: ONE PERSON
GENDER: MAN
AGE GROUP: YOUNG MAN
ETHNICITY: CAUCASIAN
HUMAN BODY PART: HUMAN FACE
TIME: 167 S
DETECTION: 63621 POINTS
POS (X/Y/Z): 1322 / 856 / 21

[Link per il Notebook Colab](#)

Deep Learning

Per Attributi Facciali

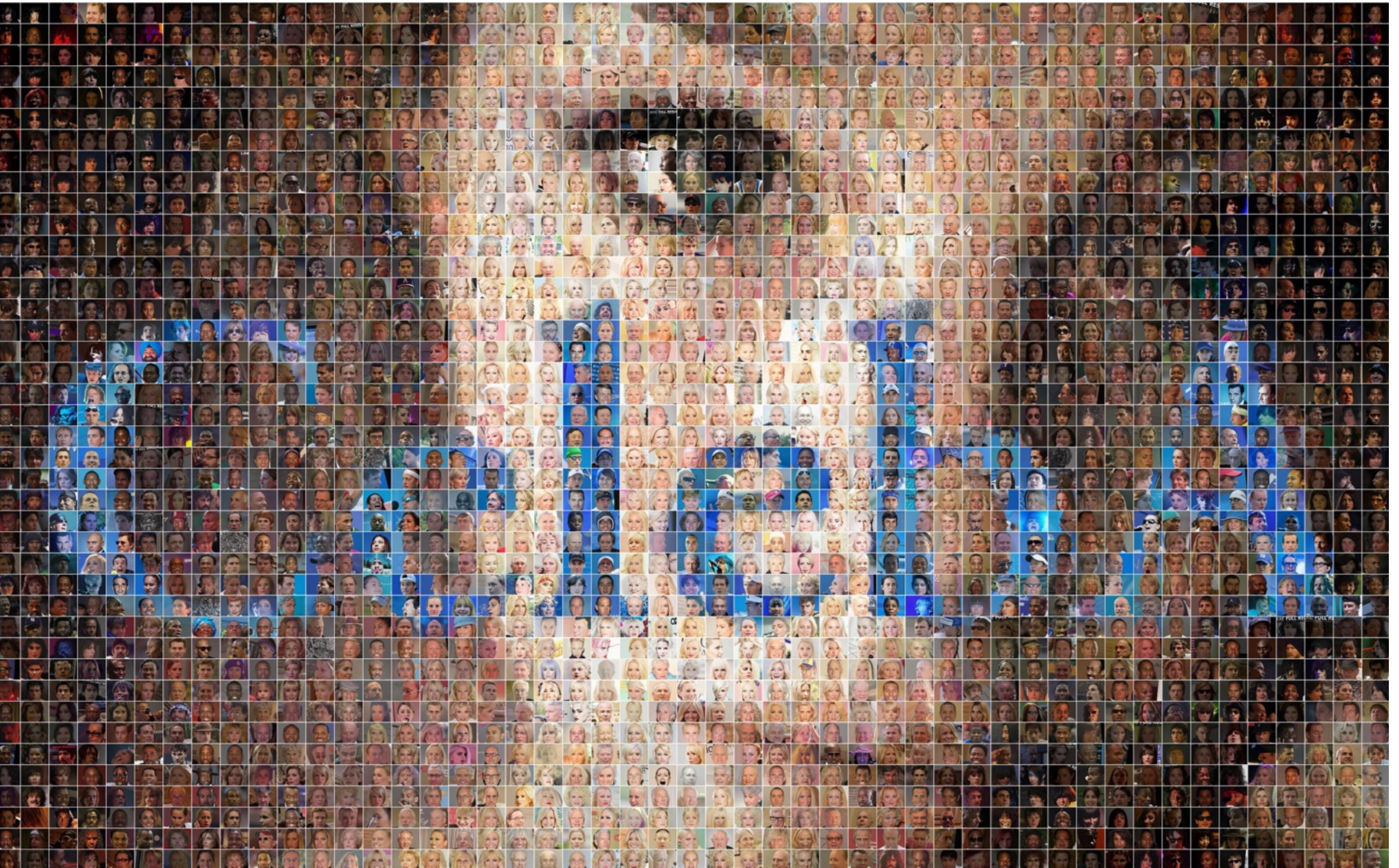




Dataset

Dataset

Celeb A



<http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

- Composta da 202599 immagini
- Ogni immagine ha 40 attributi binari classificati
- Ogni immagine è allineata e tagliata
- Ogni immagine ha dimensioni (178 x 218)

[Link per codice](#)

Dataset

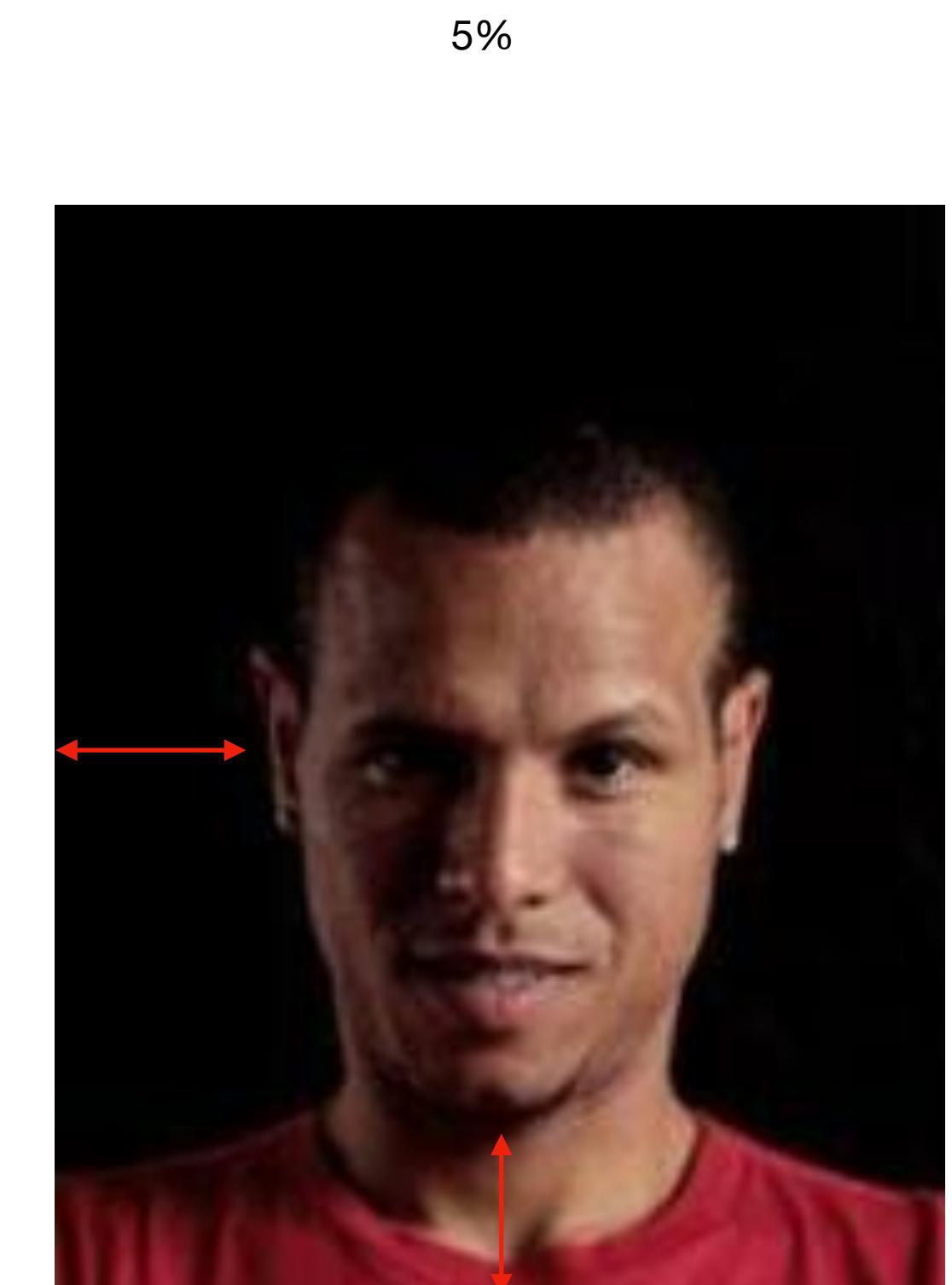
Preparazioni di Immagini



Rotazione

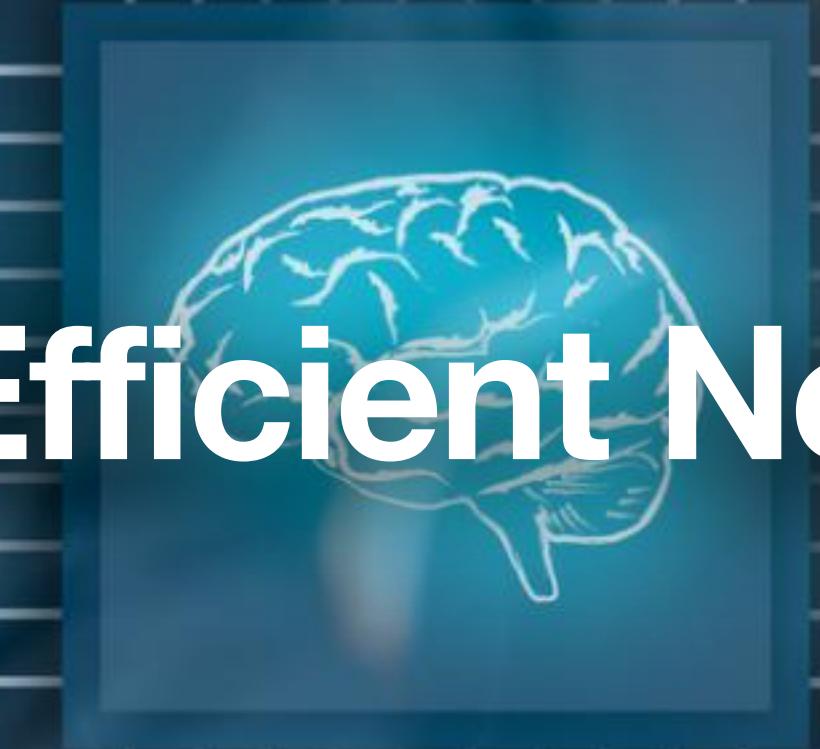


Ridimensionamento



Traslazione

[Link per codice](#)

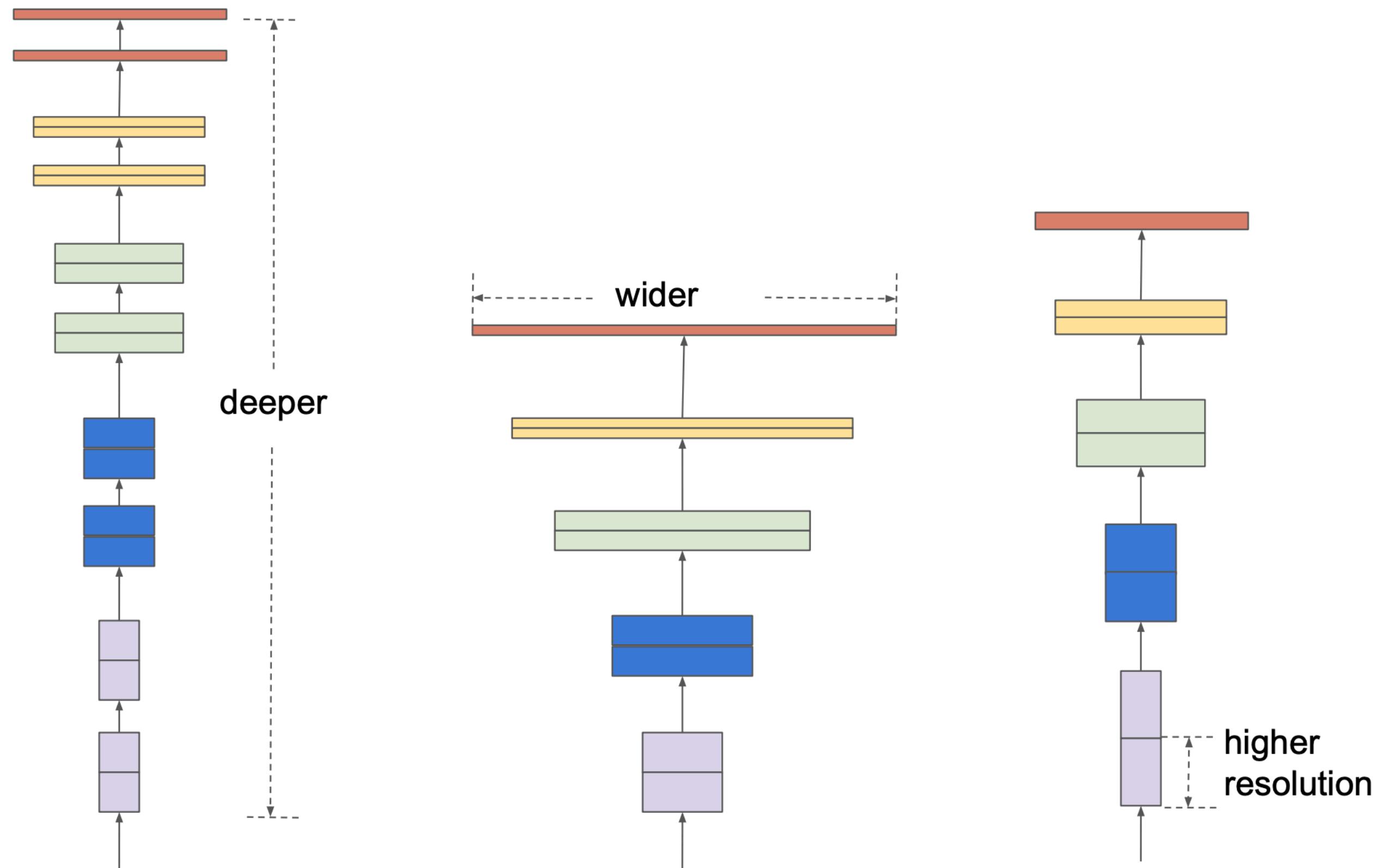


Efficient Net

Efficient Net

Caratteristiche

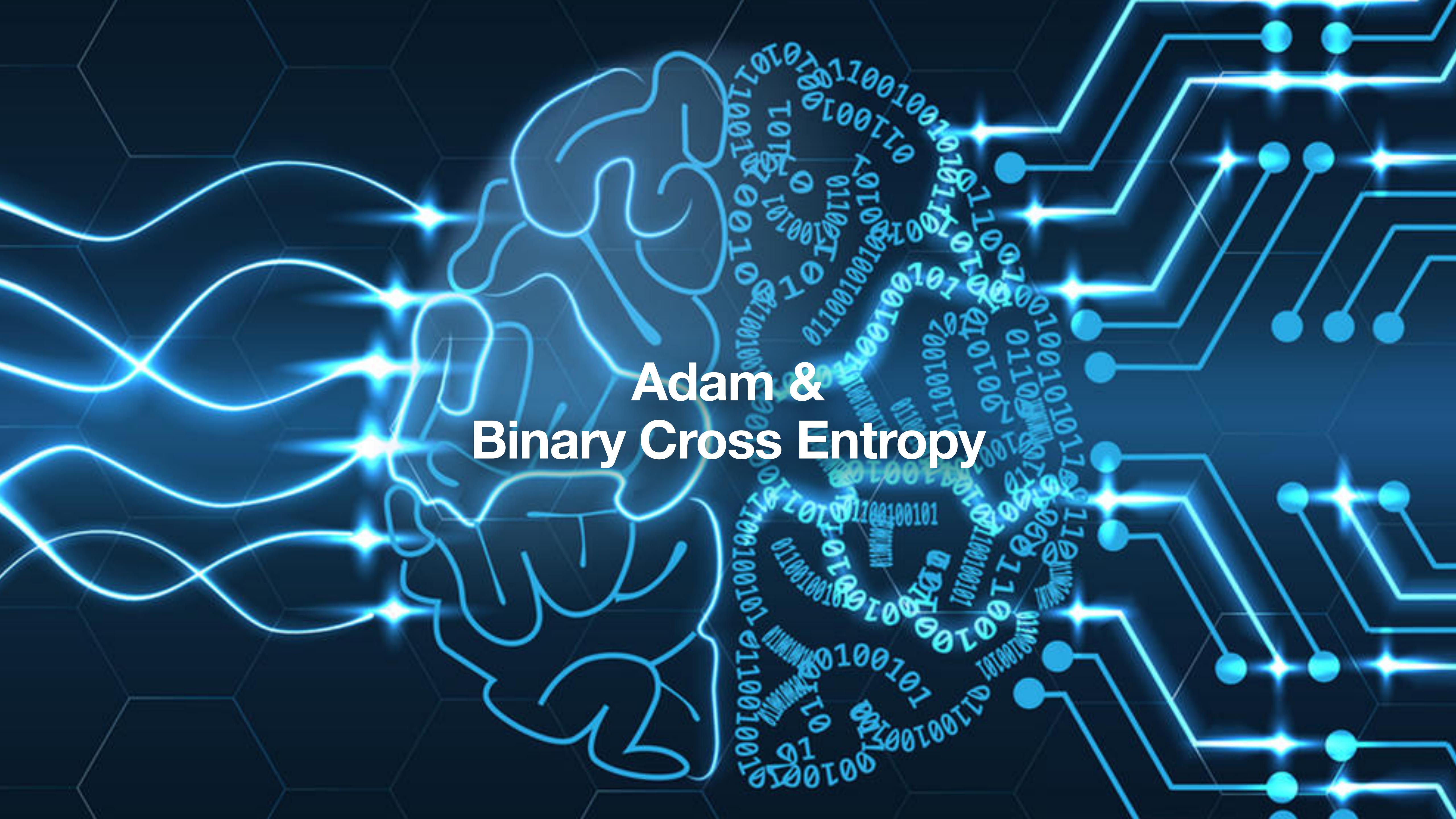
- Profondità
- Larghezza
- Risoluzione



Efficient Net

Prestazioni

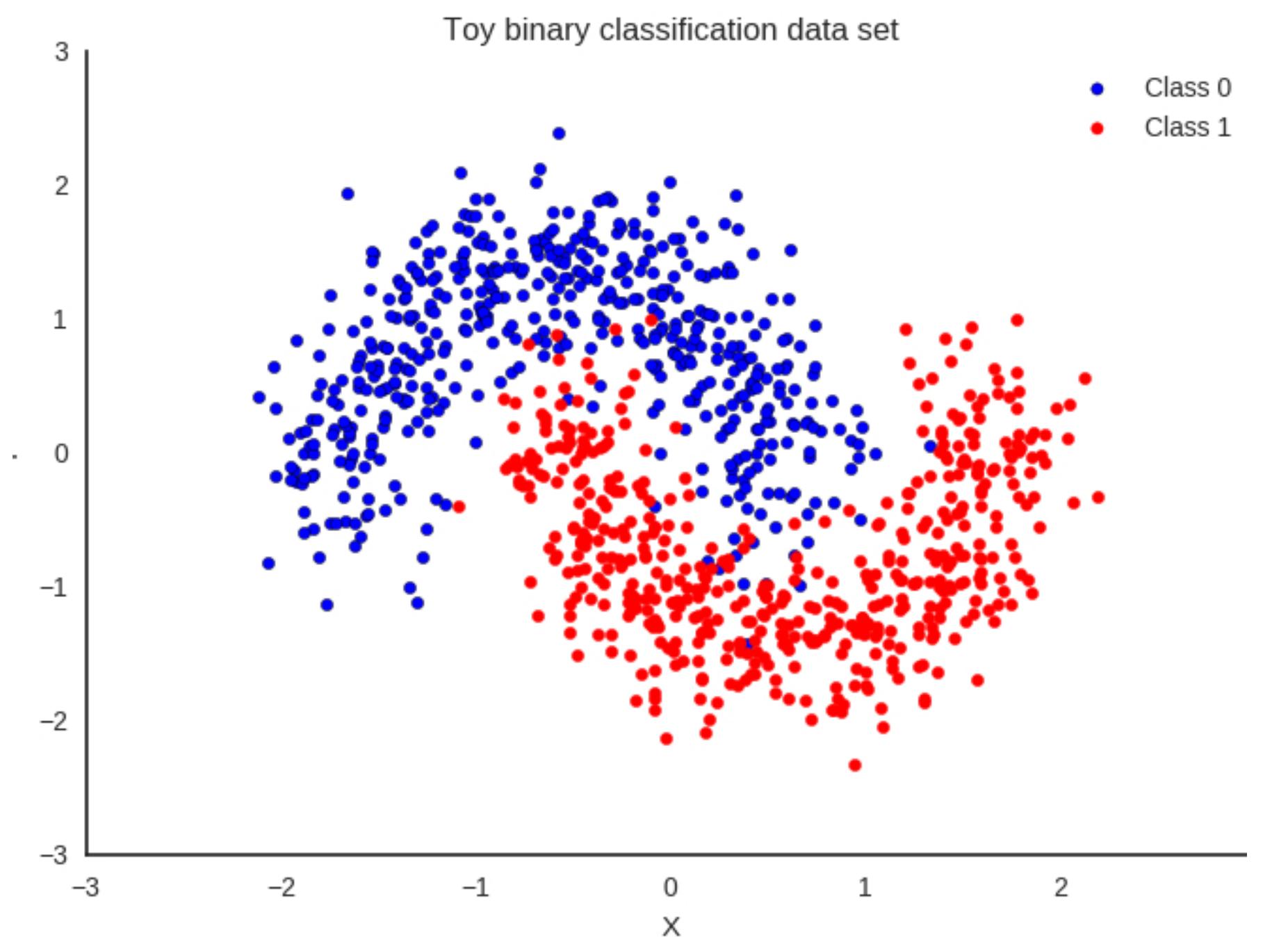
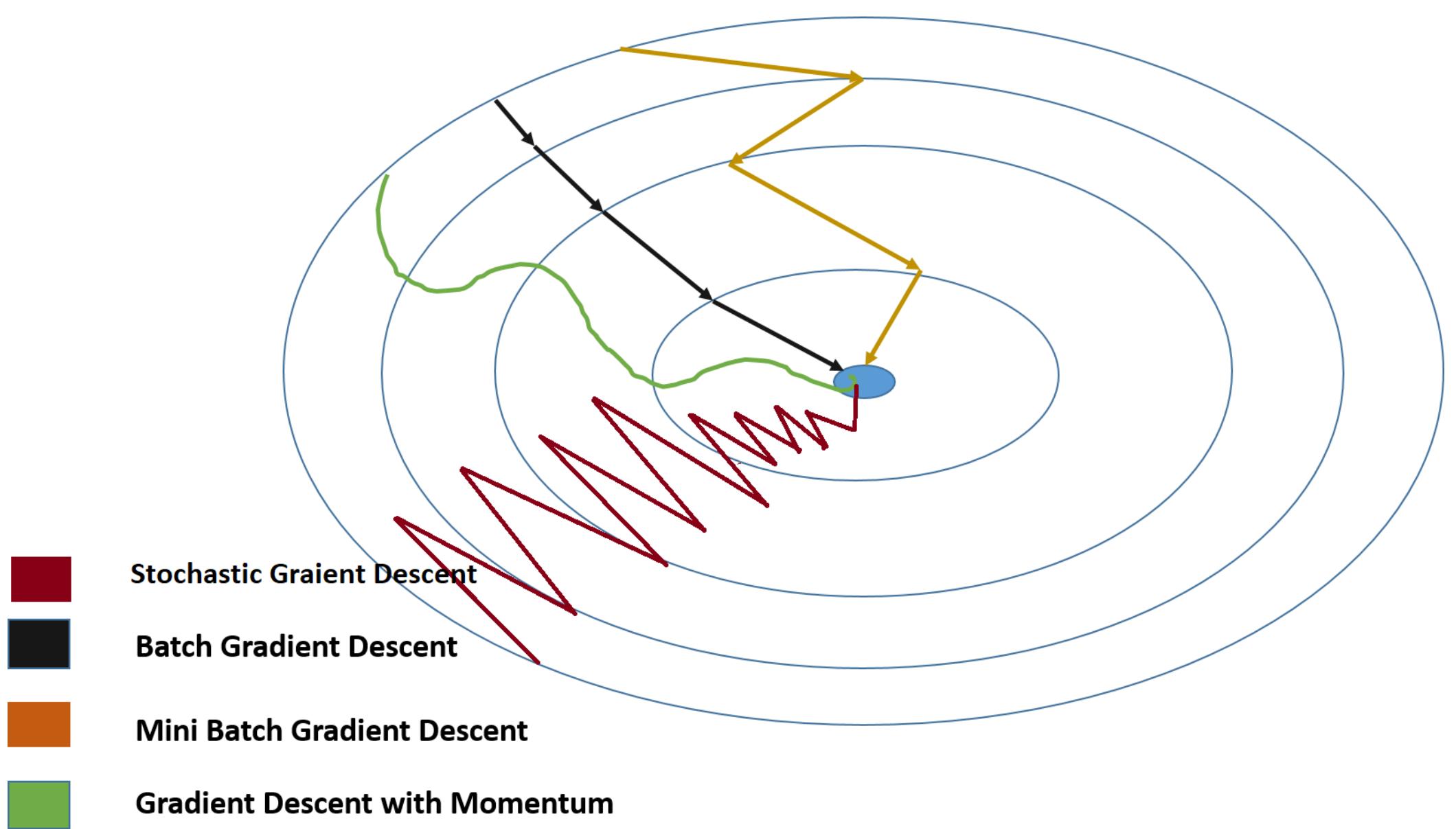
Model	Top-1 Acc.	Top-5 Acc.	#Params
EfficientNet-B0	77.1%	93.3%	5.3M
ResNet-50 (He et al., 2016)	76.0%	93.0%	26M
DenseNet-169 (Huang et al., 2017)	76.2%	93.2%	14M
EfficientNet-B1	79.1%	94.4%	7.8M
ResNet-152 (He et al., 2016)	77.8%	93.8%	60M
DenseNet-264 (Huang et al., 2017)	77.9%	93.9%	34M
Inception-v3 (Szegedy et al., 2016)	78.8%	94.4%	24M
Xception (Chollet, 2017)	79.0%	94.5%	23M
EfficientNet-B2	80.1%	94.9%	9.2M
Inception-v4 (Szegedy et al., 2017)	80.0%	95.0%	48M
Inception-resnet-v2 (Szegedy et al., 2017)	80.1%	95.1%	56M
EfficientNet-B3	81.6%	95.7%	12M
ResNeXt-101 (Xie et al., 2017)	80.9%	95.6%	84M
PolyNet (Zhang et al., 2017)	81.3%	95.8%	92M



Adam & Binary Cross Entropy

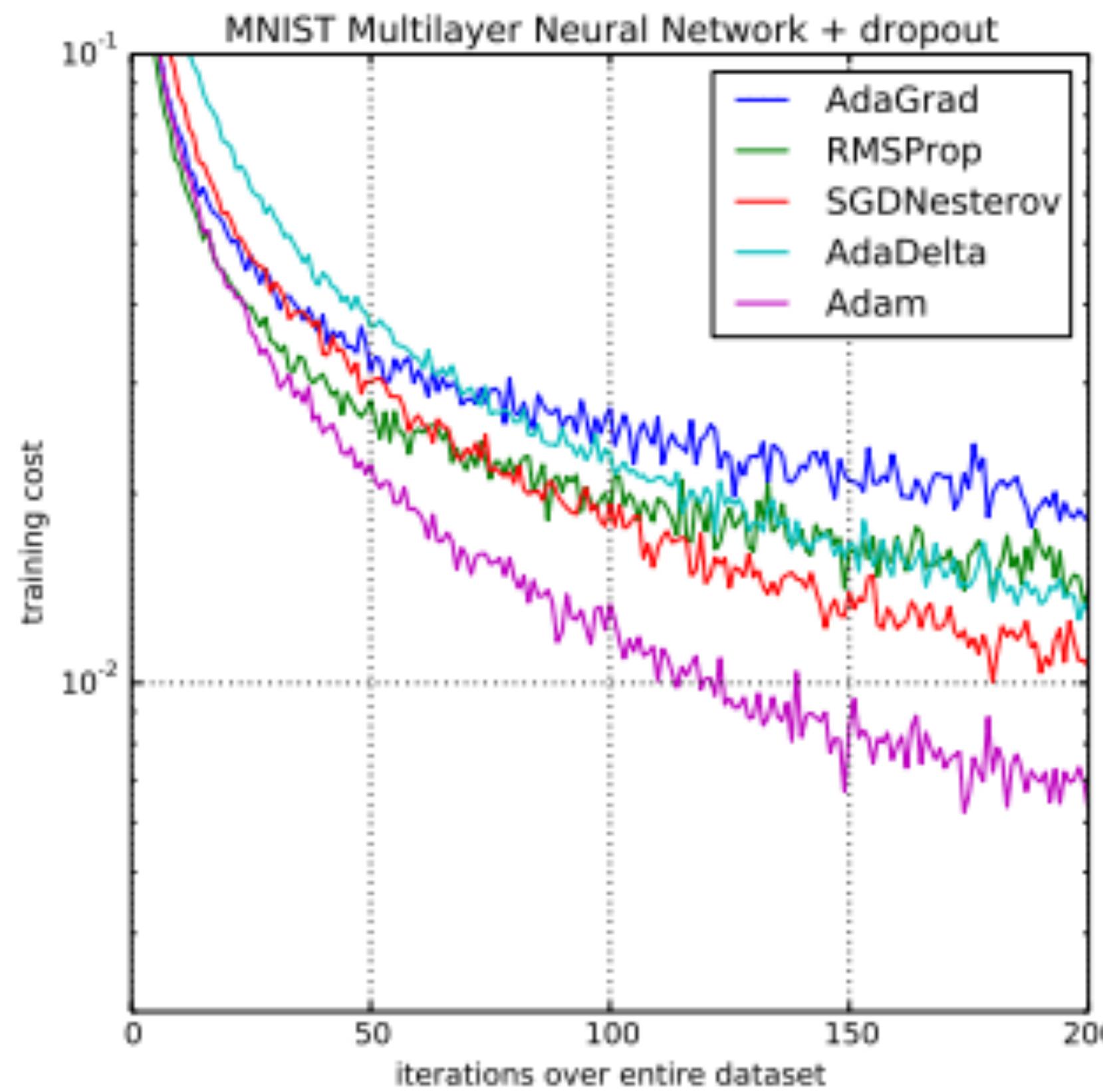
Adam & Binary Cross Entropy

Training cost



Adam & Binary Cross Entropy

Training cost



Combinazione di vantaggi di altre due varianti di discesa stocastica del gradiente:

- Adaptive Gradient Algorithm (AdaGrad)
- Root Mean Square Propagation (RMSProp)

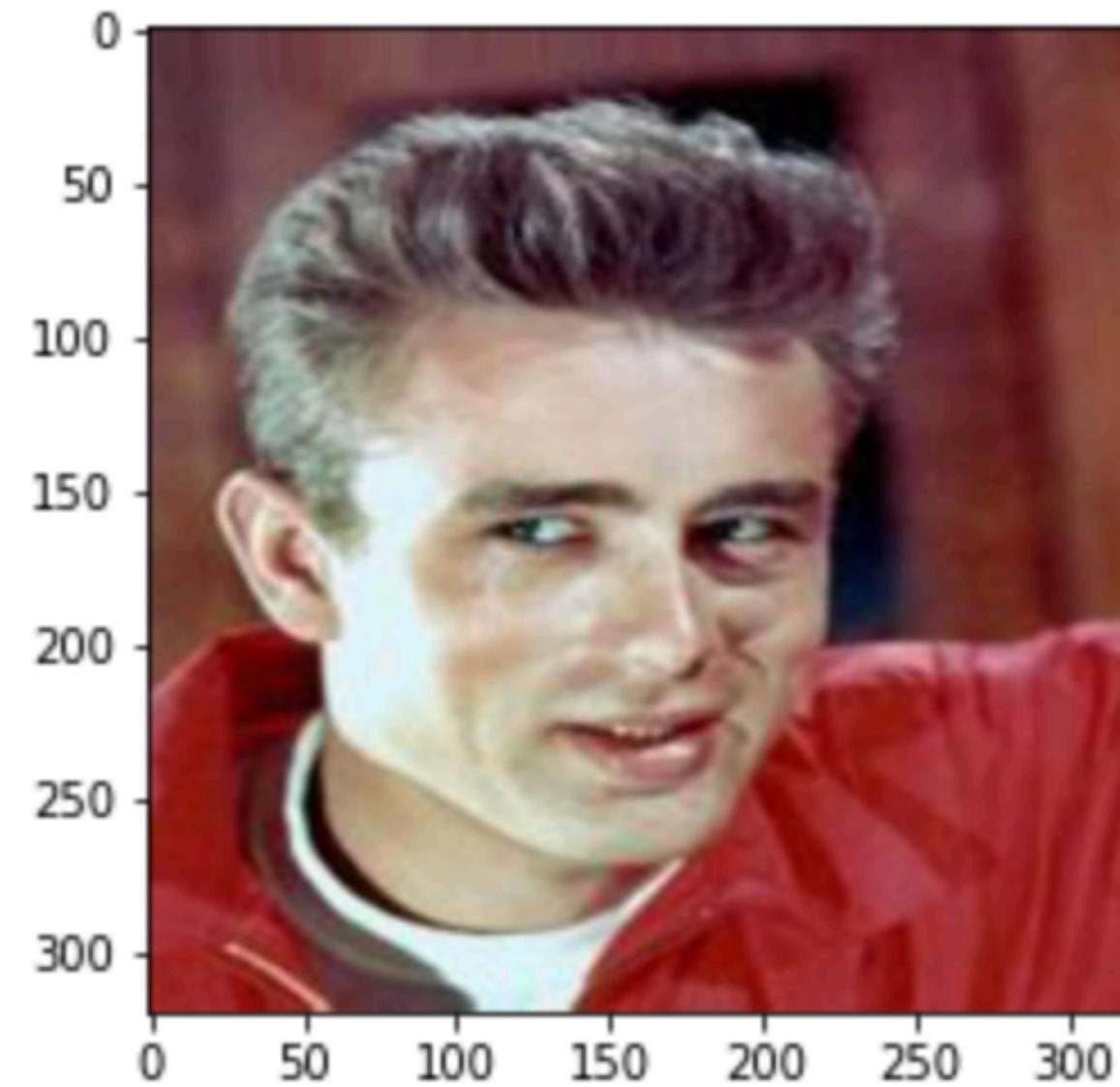
$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad \hat{m}_t = \frac{m_t}{(1 - \beta_1^t)}$$
$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad \hat{v}_t = \frac{v_t}{(1 - \beta_2^t)}$$
$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

The background of the slide features a complex, abstract digital pattern. It consists of numerous small, glowing blue particles that form intricate, flowing lines and waves across the entire frame. These lines vary in intensity, creating a sense of depth and motion. In the upper right quadrant, there is a brighter, more concentrated area where the lines converge and radiate outwards, resembling a starburst or a beam of light. The overall effect is futuristic and dynamic, suggesting concepts like data flow, information, or energy.

Valutazione prestazioni

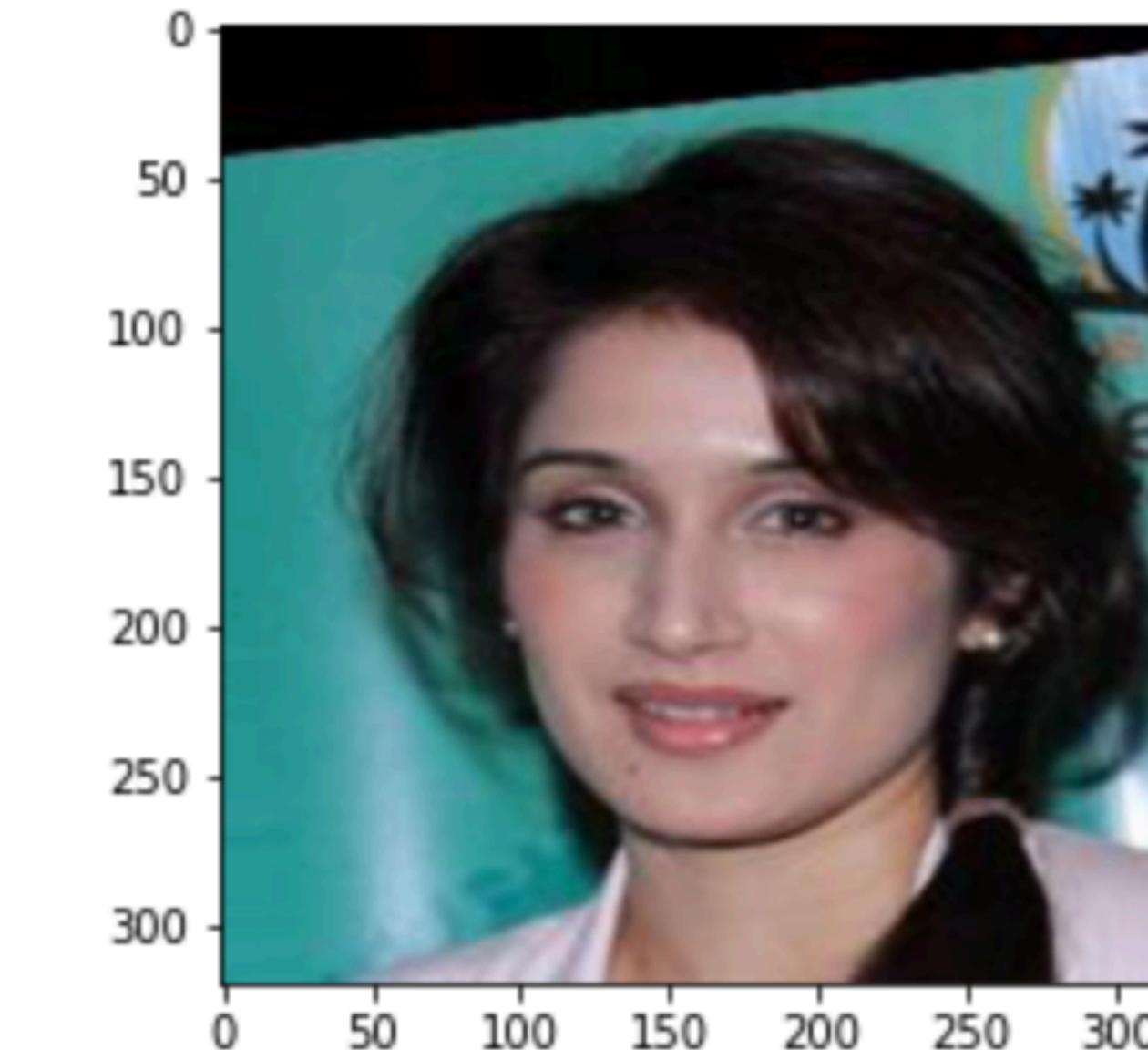
Valutazione prestazioni

Risultati ottenuti



Attributi riconosciuti nell'immagine:

Bags_Under_Eyes : 63.66 %
Male : 99.36 %
Mouth_Slightly_Open : 80.4 %
No_Beard : 92.07 %
Young : 73.92 %

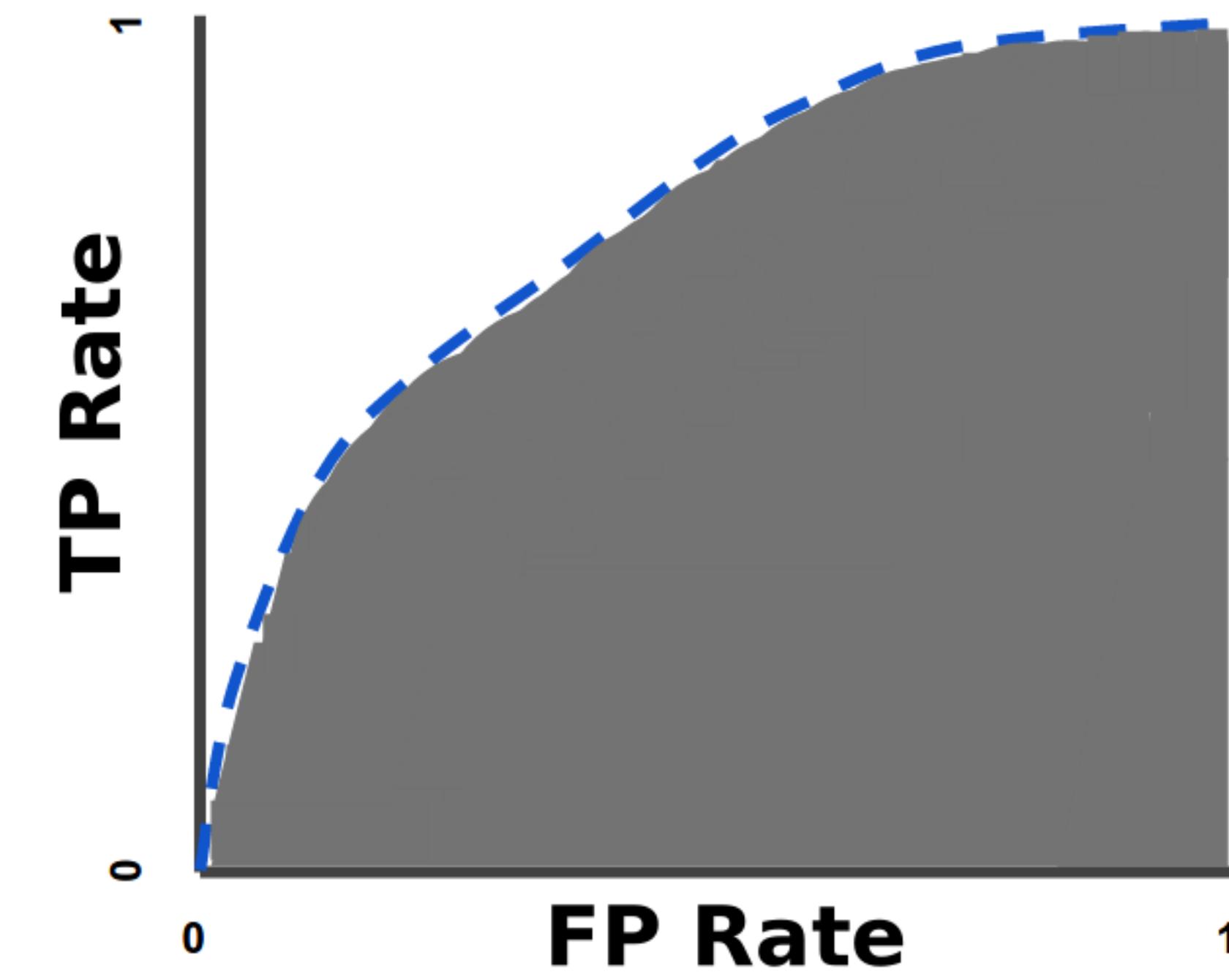
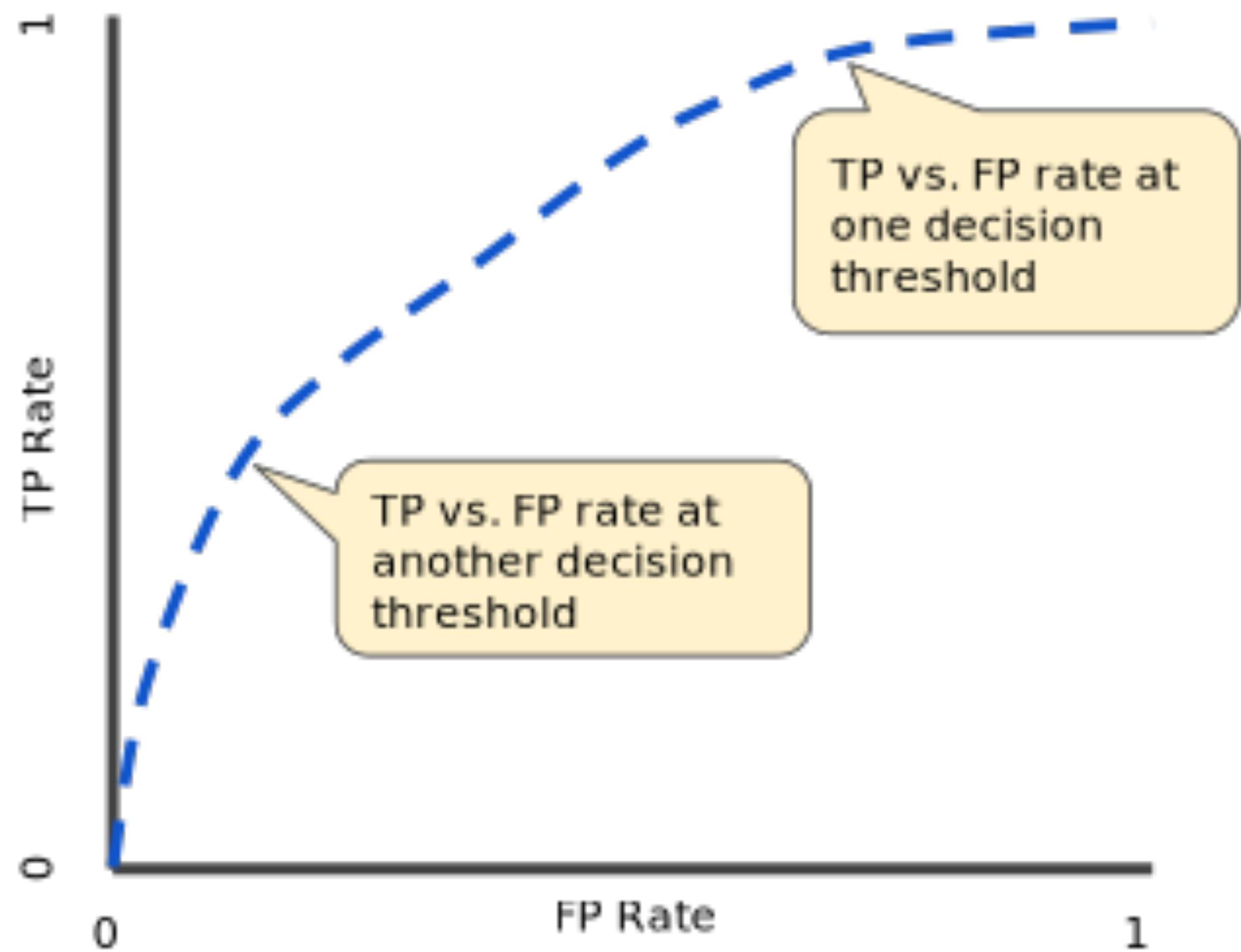


Attributi riconosciuti nell'immagine:

Attractive : 95.88 %
Heavy_Makeup : 92.59 %
Mouth_Slightly_Open : 99.53 %
No_Beard : 99.99 %
Smiling : 89.69 %
Straight_Hair : 62.38 %
Wearing_Earrings : 80.49 %
Wearing_Lipstick : 98.2 %
Young : 96.55 %

Valutazione prestazioni

Area Under Roc Curve



Valutazione prestazioni

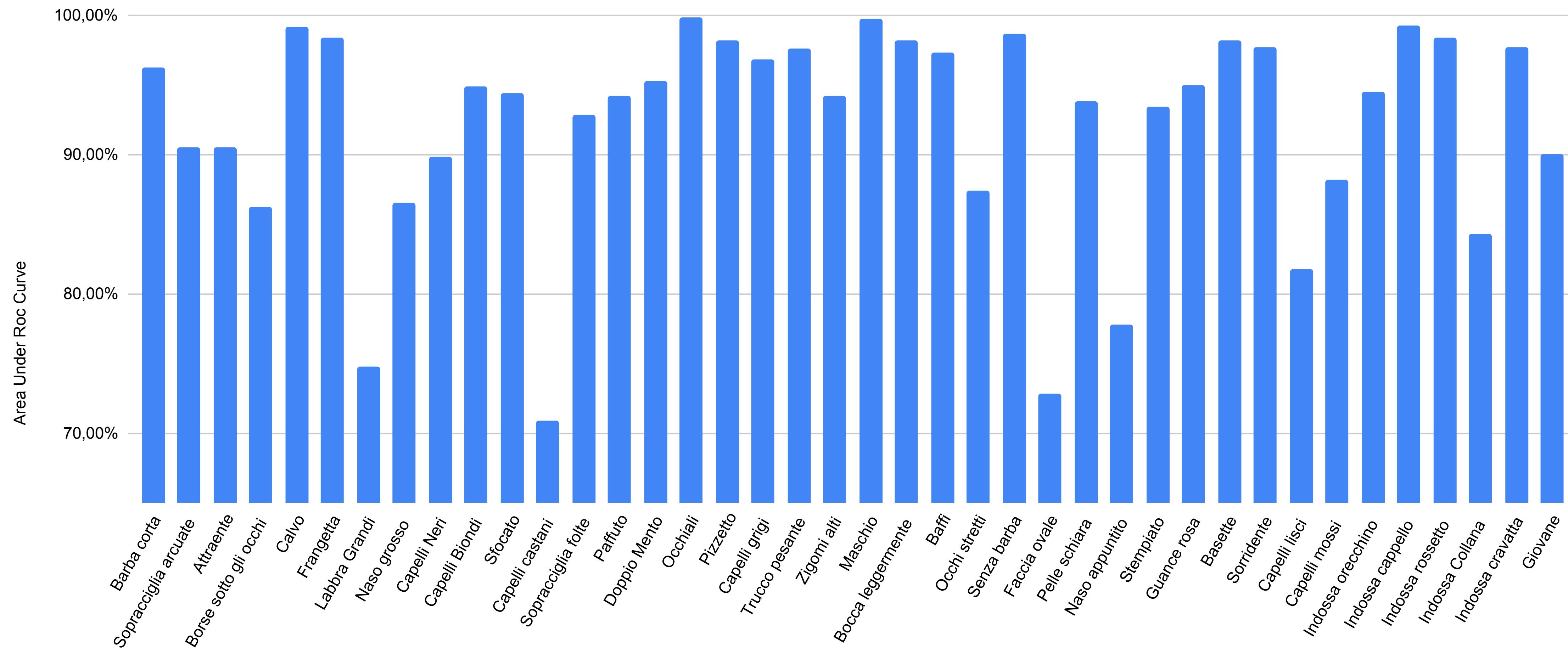
Prestazioni migliori riscontrate

- Learning Rate = 0,001
- Numero di epochhe = 15
- I primi 37 layer della rete (10% dei layer totali) non sono stati allenati
- Tempo impiegato: 4 ore e mezza
- Binary Accuracy media training = 90,18%
- Binary Accuracy media validation = 88,72%



Valutazione prestazioni

Prestazioni per singoli attributi



Grazie per l'attenzione!