

Digital Signal and Image Management

Anna Mattioli - 826381
Beatrice Barzaghi - 831829
Guglielmo Muoio - 826029



Obiettivi del progetto

PROCESSING MONO-DIMENSIONALE

Riconoscere, a partire da un file audio, l'identità di chi sta parlando

BINARY MULTI-OUTPUT CLASSIFICATION

- Augmentation
- MFCC
- B-bot vocal assistant



PROCESSING BI-DIMENSIONALE

Riconoscere, a partire da una foto del volto, l'identità della persona

BINARY MULTI-OUTPUT CLASSIFICATION

- Face detection (MTCNN)
- VGG-Face
- Transfer learning



RETRIEVAL

Trovare all'interno di un dataset di persone famose, i 10 volti che somigliano di più a quello del punto precedente

KD-TREE

- Labeled Faces in the Wild
- VGG-Face
- Celebrity check



Processing di segnali mono-dimensionali

MAIN TASK

Riconoscere, a partire da un file audio, l'identità di chi sta parlando



708 REGISTRAZIONI

- Lunghezza 3 secondi
- Formato .wav
- Nessuna restrizione sulle frasi pronunciate
- Dataset eterogeneo



EXPLORATIVE TASK

Riconoscere l'invocazione dell' assistente vocale
B-Bot



Augmentation

Manipolazione delle tracce audio di training attraverso la variazione di:

- speed - velocità
- noise - rumore di fondo
- pitch - frequenza



Audio features extractors

- MFCC*
- Standard deviation
- Absolute average
- Energy
- Zero-crossing rate



Support Val. accuracy

Anna	23	<div style="width: 96%;"></div> 96%
Beatrice	24	<div style="width: 97%;"></div> 97%
Guglielmo	21	<div style="width: 100%;"></div> 100%
B-Bot	37	<div style="width: 89%;"></div> 89%



Implementazioni

- Multi-class classification
- Binary multi-output classification

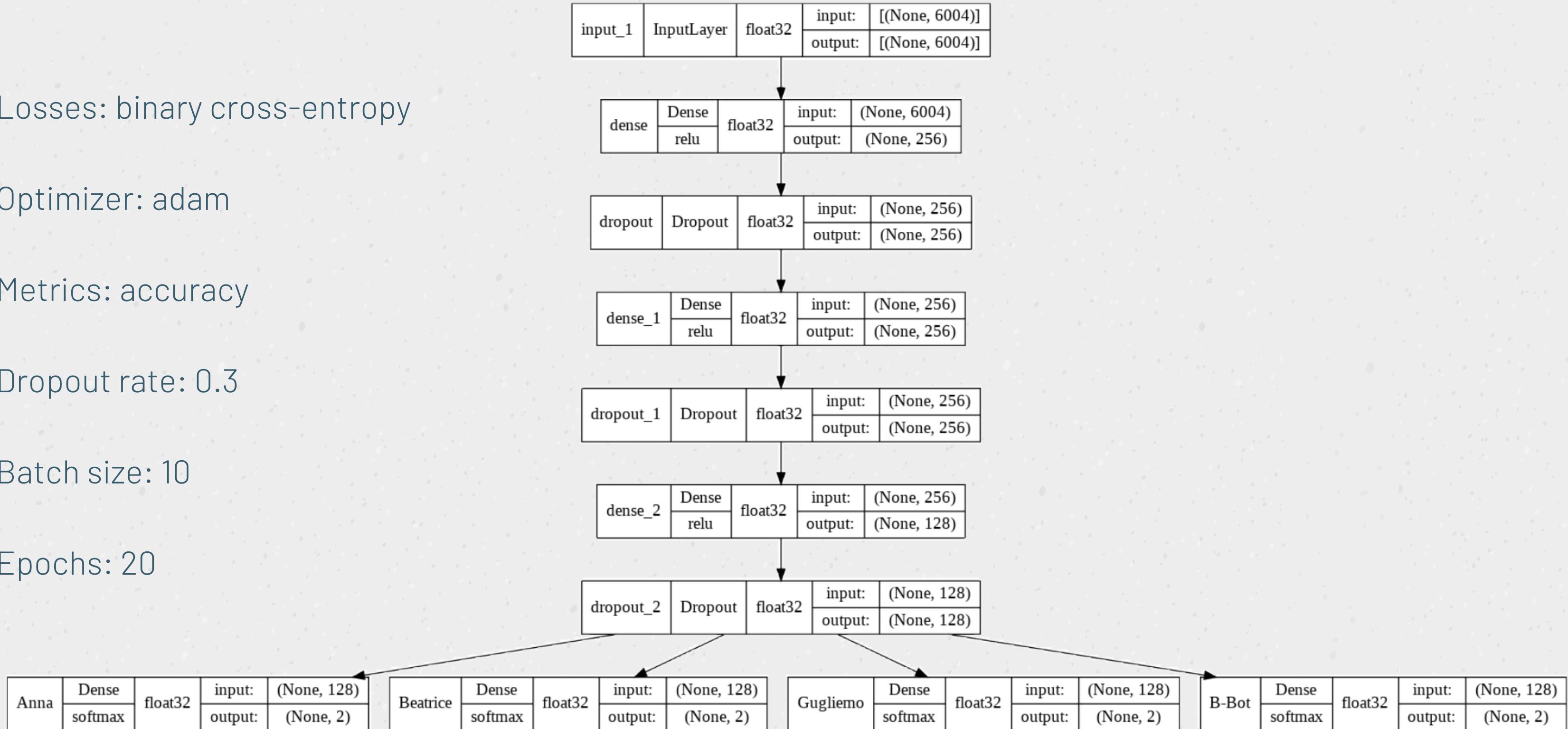
Splitting

- Training: 2548
- Validation: 71

*Mel-Frequency Cepstral Coefficients

Modello di classificazione

- Losses: binary cross-entropy
- Optimizer: adam
- Metrics: accuracy
- Dropout rate: 0.3
- Batch size: 10
- Epochs: 20



Processing di segnali bi-dimensionali

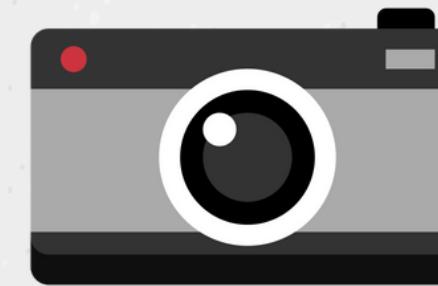
MAIN TASK

Riconoscere, a partire da una foto del volto, l'identità della persona



450 FOTO

- Volto frontale
- Dataset eterogeneo
- Condizioni di illuminazione variabile
- Arco temporale di 5 anni



Pre-processing

- Face detection: acquisizione del volto della persona in foto (MTCNN)
- Resize: ridimensionamento dei volti ad uno standard 224x224 pixels
- Augmentation: flip, rotation, contrast



Transfer learning

Utilizzo del modello VGG-Face, architettura pre-trainata per il riconoscimento facciale

- Fine tuning
- Feature extraction



Implementazioni

- Multi-class classification
- Binary multi-output classification

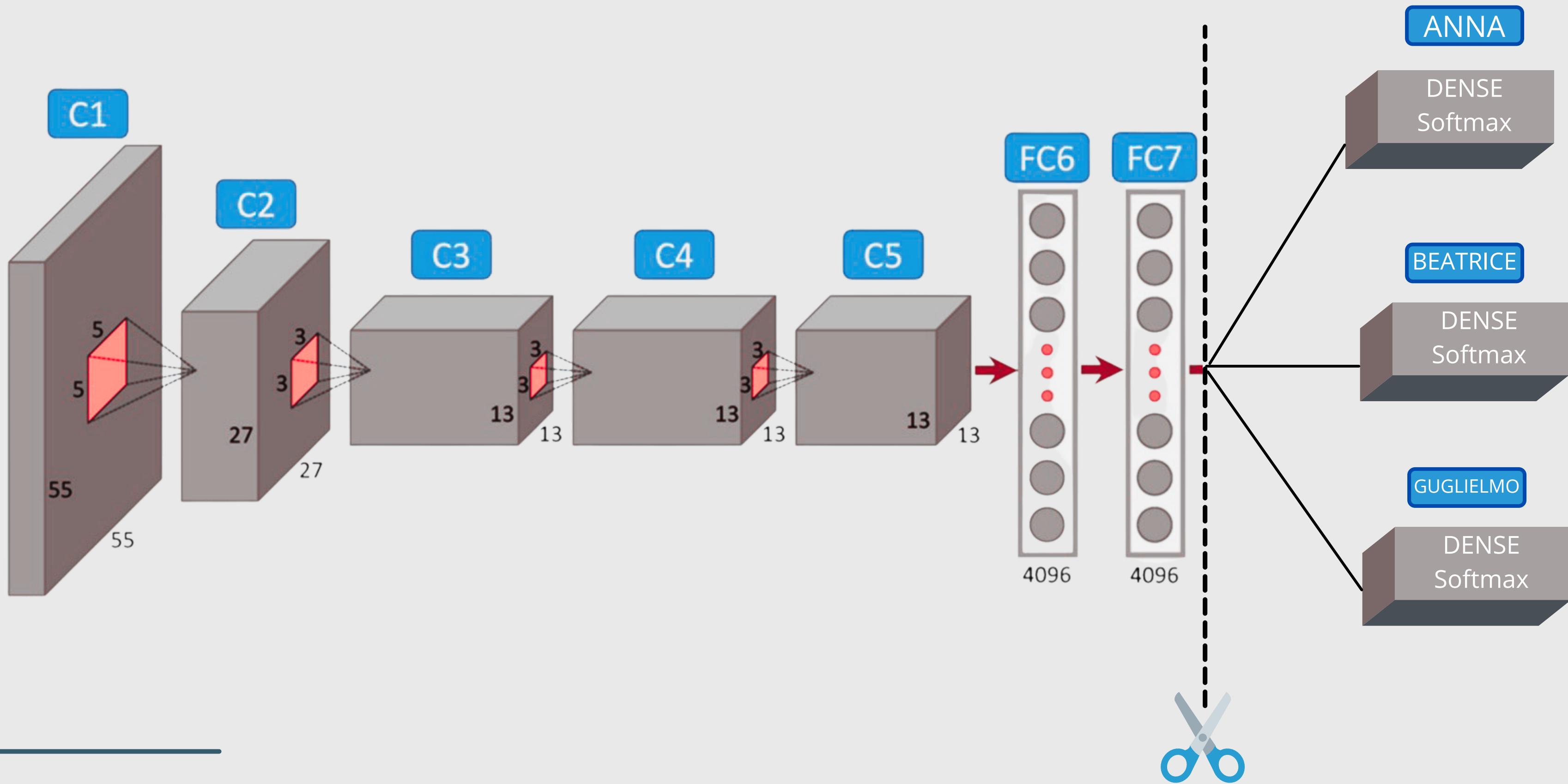
Splitting

- Training: 396
- Validation: 45

	Support	Val. accuracy
Anna	11	<div style="width: 100%;"><div style="width: 100%;">100%</div></div>
Beatrice	18	<div style="width: 100%;"><div style="width: 100%;">100%</div></div>
Guglielmo	14	<div style="width: 100%;"><div style="width: 100%;">100%</div></div>



Modello di classificazione



Retrieval



MAIN TASK

Trovare all'interno di un dataset di persone famose, i 10 volti che somigliano di più a quello del punto precedente



CELEBRITY DATASET*

Labeled Faces in the Wild

- 13233 foto
- 5749 persone

Per lo svolgimento del task

- Vip con almeno 5 foto
- massimo 20 foto per vip



Pre-processing

- Face detection: acquisizione del volto della persona famosa
- Resize: ridimensionamento dei volti ad uno standard 224x224 pixels



Features extraction

Estrazione delle features tramite VGGFace

- Group faces(query)
- Celebrity faces



Implementazioni

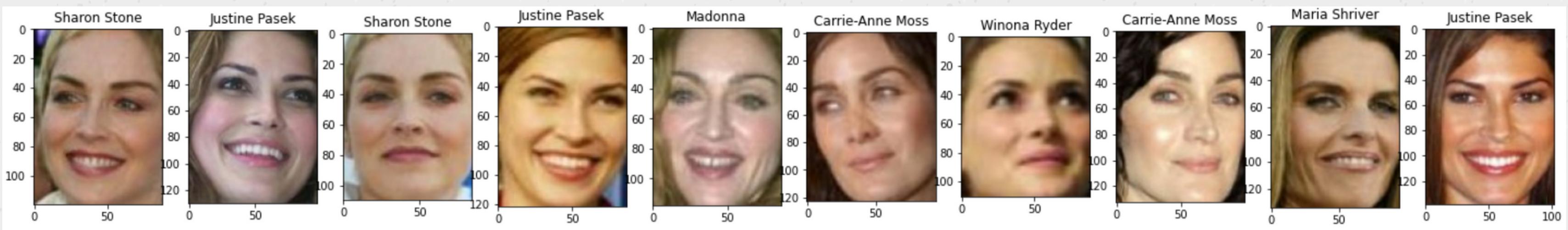
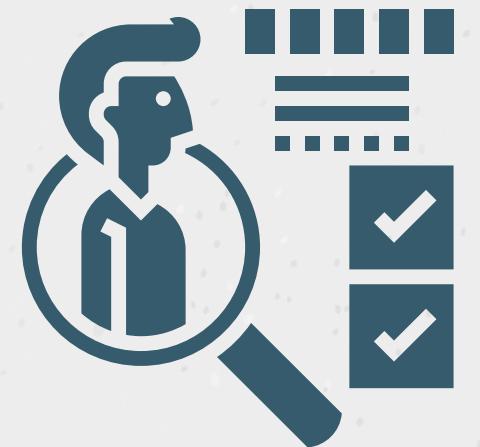
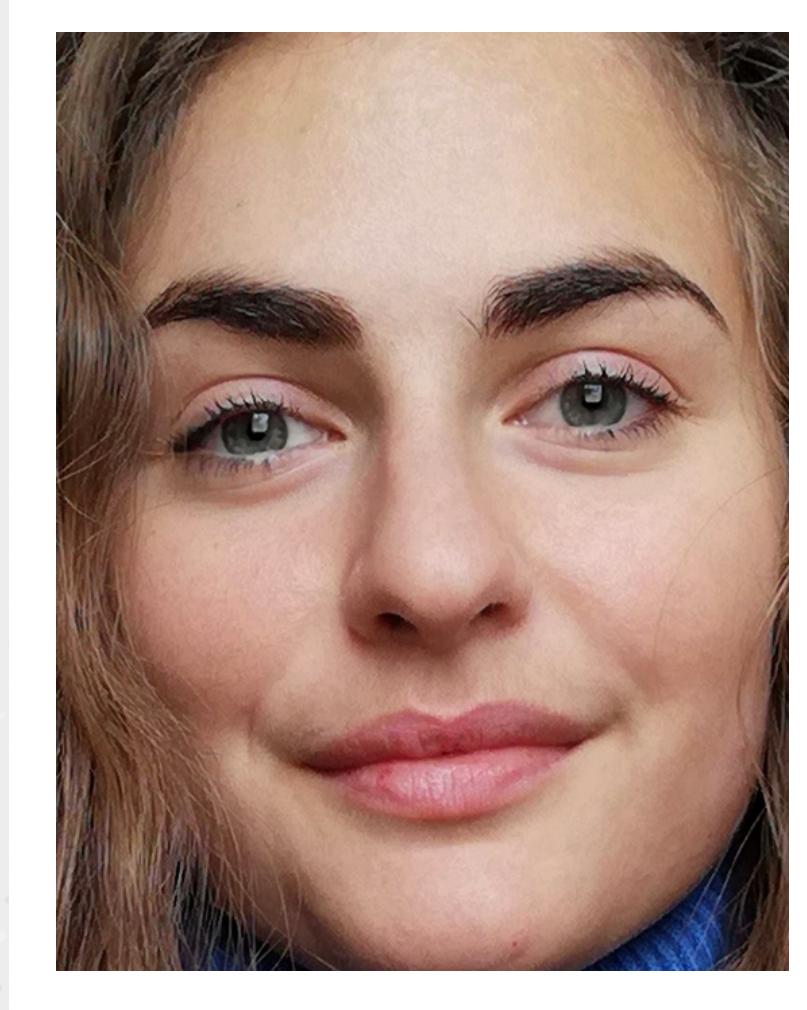
- KDTree: albero di ricerca, vengono estratti i 10 nearest neighbors dell'immagine di input

Criteri di valutazione

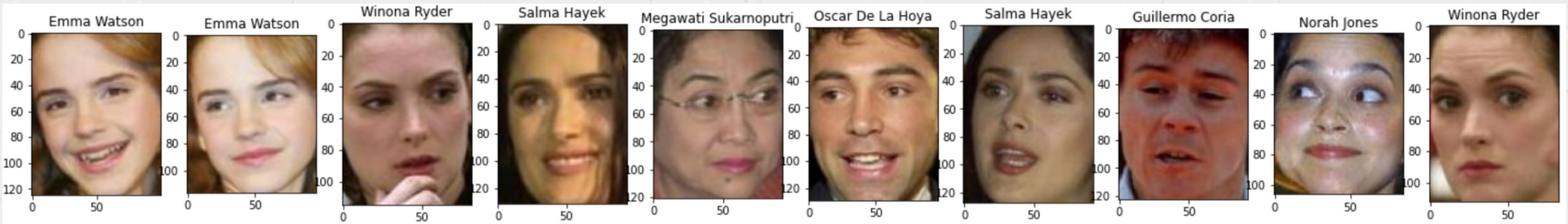
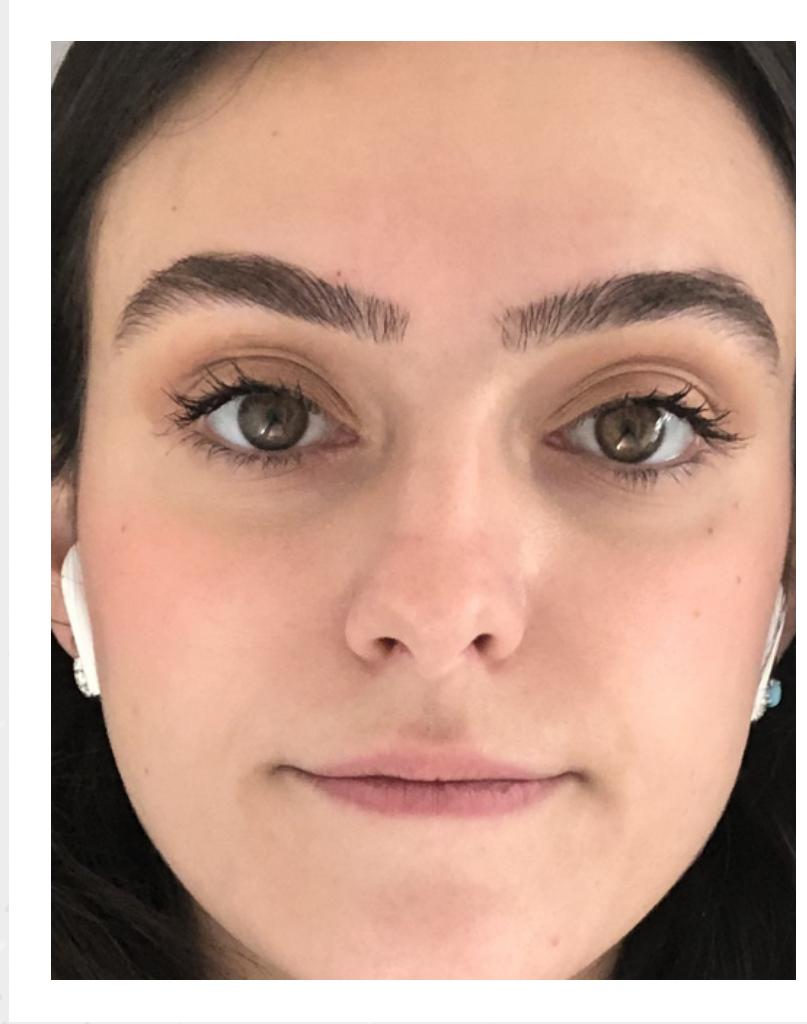
- Riscontro visuale
- Celebrity P@10



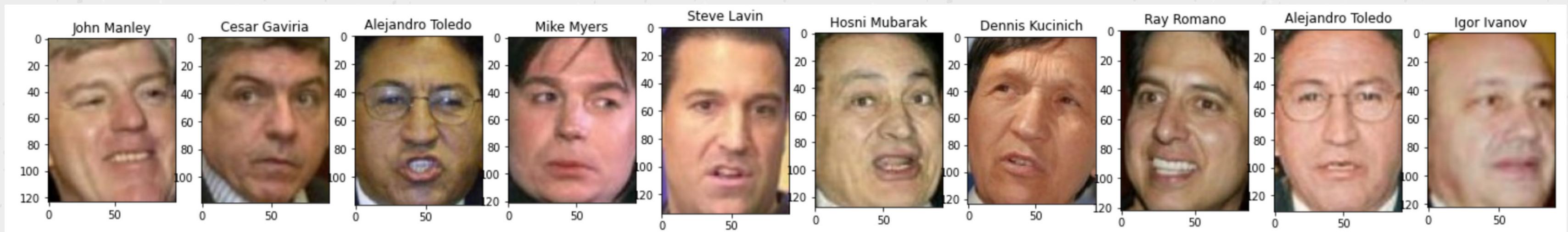
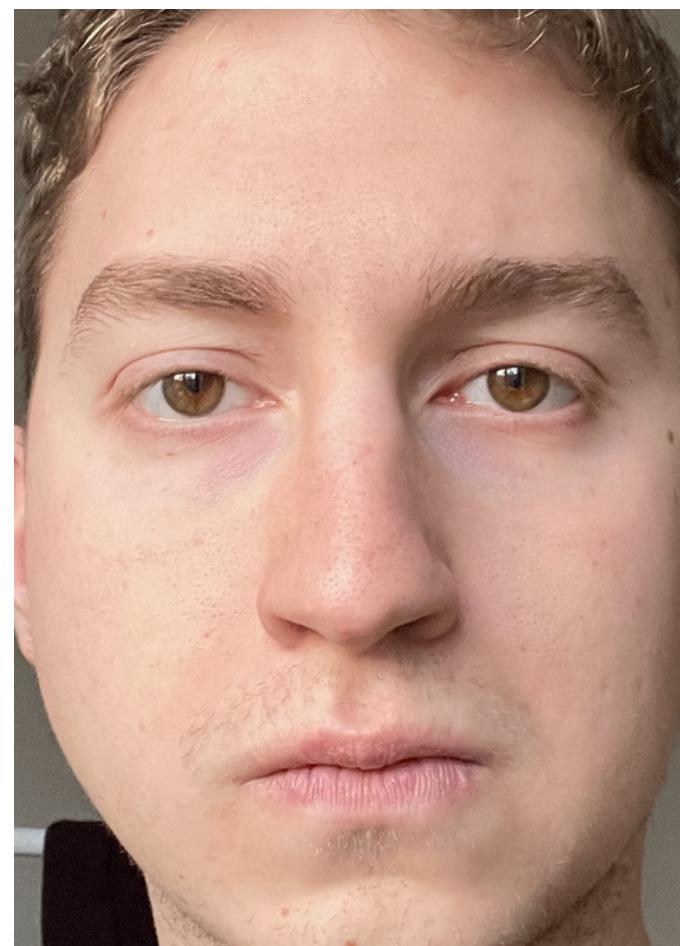
Anna



Beatrice



Guglielmo

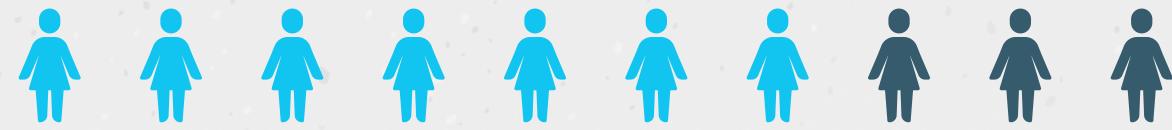


Celebrity

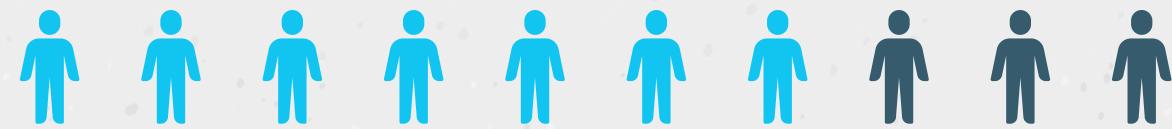
- Angelina Jolie
- Lance Armstrong
- Jennifer Aniston
- Michael Schumacher
- Hugh Grant
- Britney Spears
- George W. Bush
- Serena Williams
- Tiger Woods
- Jiang Zemin

P@10

(7/10)



(7/10)



(8/10)



(5/10)



(9/10)



(9/10)



(10/10)



(5/10)



(9/10)



(10/10)

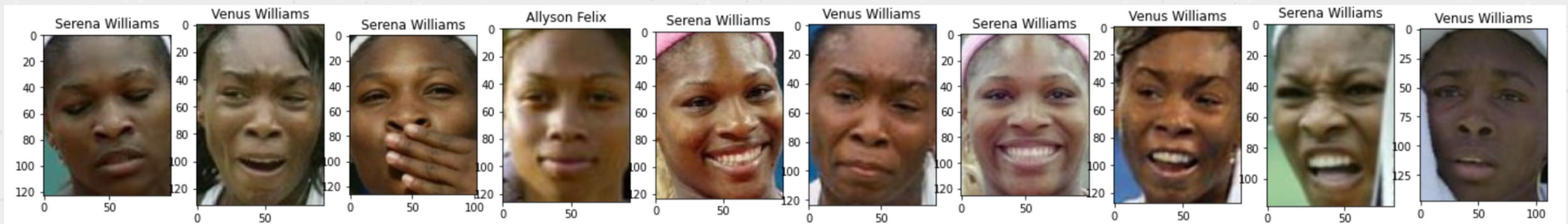


Media: 7.9

Serena Williams



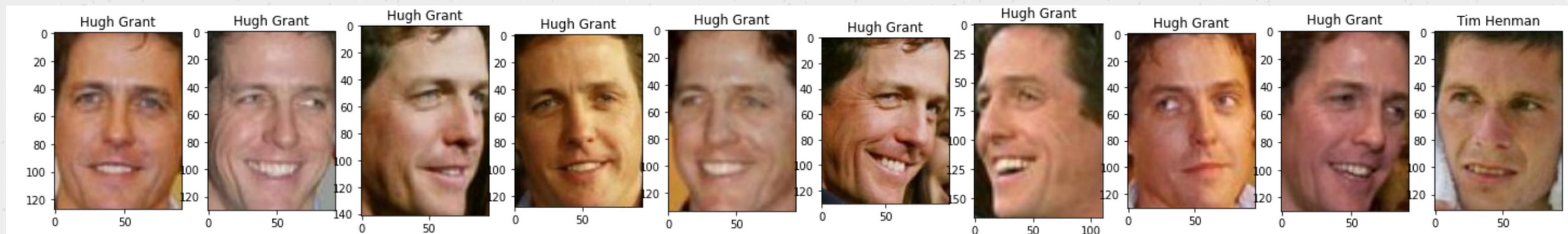
5/10



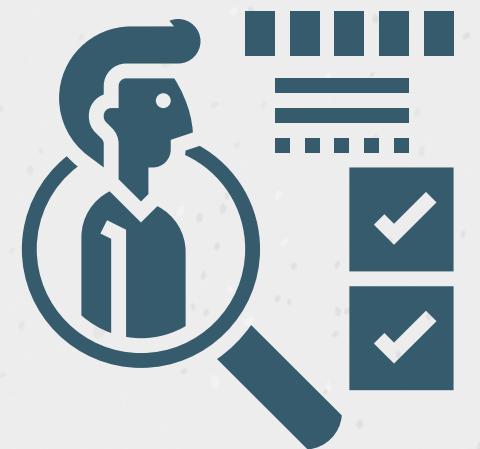
Hugh Grant



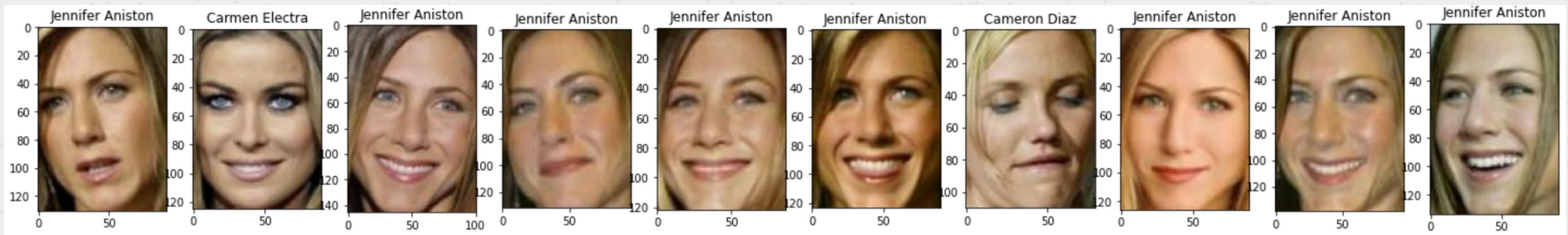
9/10



Jennifer Aniston



8/10



Grazie per
l'attenzione