



Università degli Studi di Salerno
Dipartimento di Informatica



Gender Classification using a Spider Web Method

Progetto Fondamenti di Visione Artificiale e Biometria

a.a. 2019/2020 - Prof. M. Nappi - Dott.sse C. Bisogni e P. Barra

- 1. Introduzione al problema
- 2. Modelli utilizzati
- 3. Dataset
- 4. Architettura
- 5. Risultati ottenuti
- 6. Conclusioni



1

Introduzione al problema



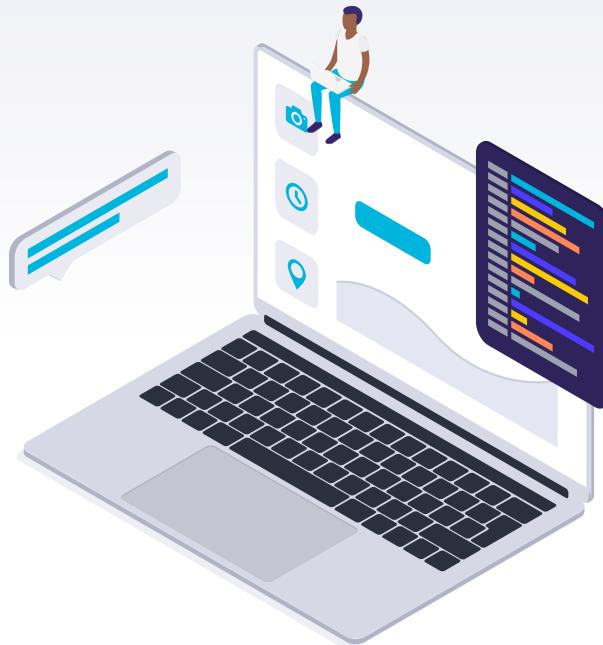
Classificazione del genere umano

- ▶ Il genere è una biometria molto utilizzata, specialmente per dare supporto ad altre biometrie di riconoscimento.
- ▶ Determinare il genere di un volto non chiaro, è un compito impegnativo per i computer a causa della diversità e delle variazioni del volto umano.
- ▶ L'obiettivo è stimare il genere umano andando principalmente ad analizzare la struttura ossea del soggetto, in modo da trascendere da implicazioni di carattere etico legate alla gender classification.



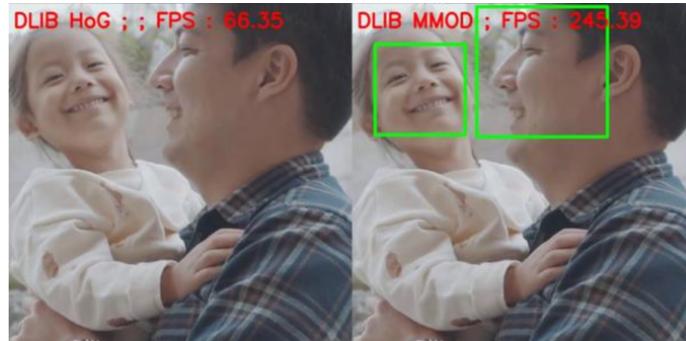
2

Modelli utilizzati



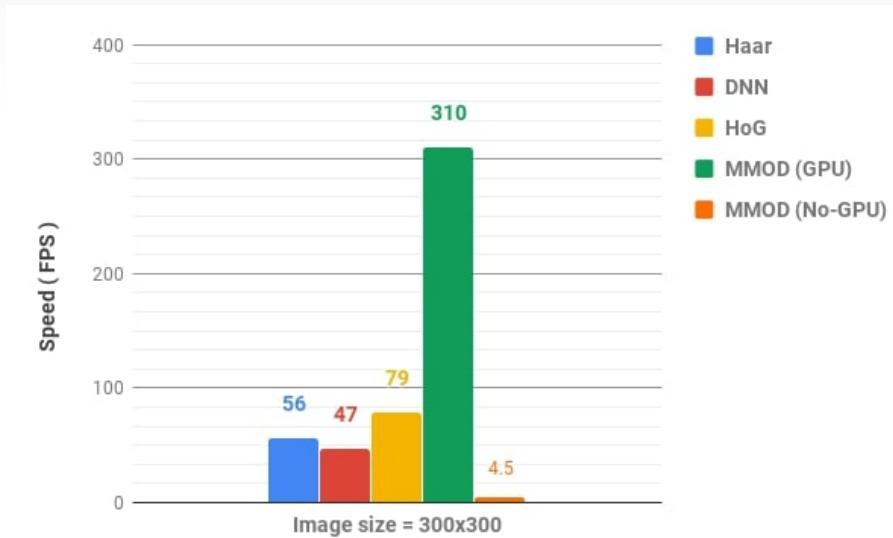
Face detection

- Invece di utilizzare un rilevatore di volti basato su HOG + SVM, si è deciso di utilizzare il rilevatore MMOD basato su una rete neurale.



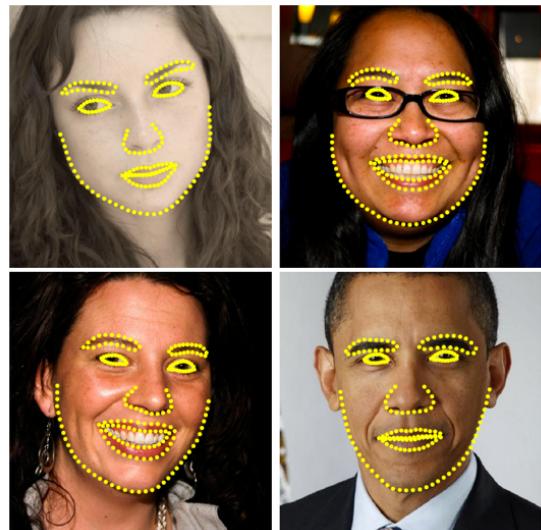
Face detection

- ▶ Utilizzo delle massime prestazioni della GPU (Nvidia Titan Xp)



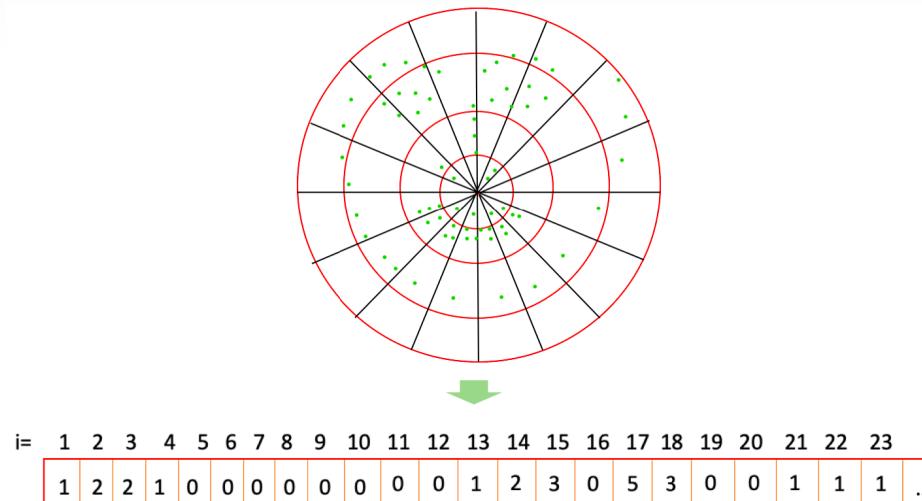
Landmark prediction

- ▶ Il predittore prende come input un'immagine del volto ed emette le posizioni di 68 *facial landmarks* P_i espresse come coppie di coordinate cartesiane (x_i, y_i) con $i = 1, 2, \dots, 68$
- ▶ Il rilevamento dei punti si basa su un insieme di alberi di regressione.



Spider Web Model

- ▶ Modello a forma di ragnatela che applicato ai 68 landmark facciali restituisce un array caratteristico di tali punti individuati sul volto.



Spider Web Model

- ▶ La configurazione della ragnatela utilizzata in questo progetto è codificata come *4C_4S_inv4*, poichè è formata da quattro cerchi e ogni quarto è diviso in quattro sezioni.
- ▶ Il raggio per ogni cerchio (*inv*) è calcolato nel modo seguente
 - ▶ $4/10 * R$ (prima circonferenza)
 - ▶ $7/10 * R$ (seconda circonferenza)
 - ▶ $9/10 * R$ (terza circonferenza)
 - ▶ R (quarta circonferenza)

Rete Neurale

- ▶ Prima rete neurale (ANN1) con numero totale di parametri 19329

Layer	Output Shape	#Parametri	Funzione Attivazione
Dense1	128	8320	Relu
Dense2	64	8256	Relu
Dense3	32	2080	Relu
Dense4	16	528	Relu
Dense5	8	136	Relu
Dense6	1	9	Sigmoid

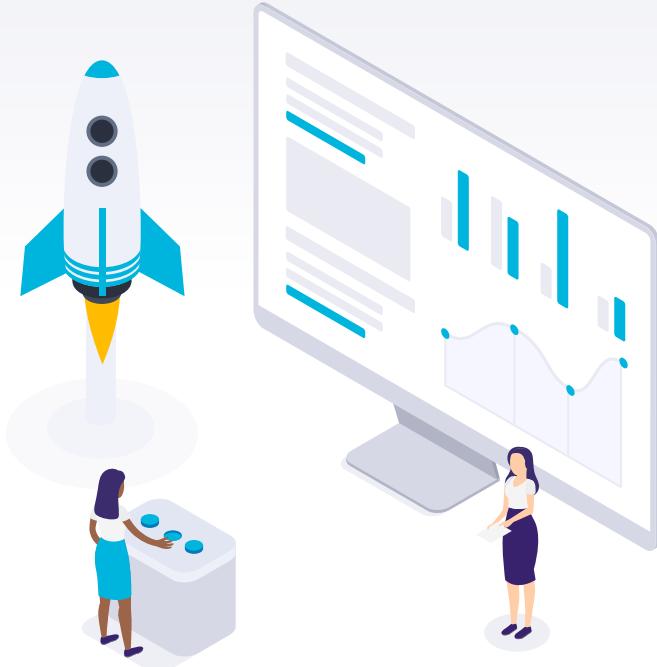
Rete Neurale

- ▶ Seconda rete neurale (ANN2) con numero totale di parametri 7969

Layer	Output Shape	#Parametri	Funzione Attivazione
Dense1	64	4160	Relu
Dense2	32	2080	Relu
Dense3	32	1056	Relu
Dense4	16	528	Relu
Dense5	8	136	Relu
Dense6	1	9	Sigmoid

3

Dataset



CelebA

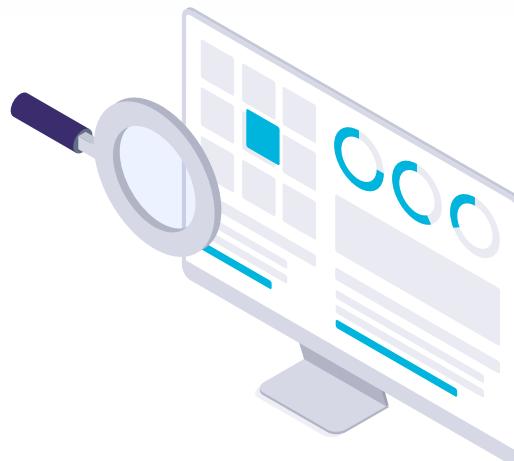
- ▶ Set di dati di attributi di facce su larga scala con oltre 200.000 immagini di celebrità, ognuna con 40 annotazioni di attributi

202.509

numero di immagini di volti

10.177

numero di identità



WheDataset

- ▶ Set di dati di attributi di volti di piccole dimensioni rispetto a CelebA, con annotazioni sul genere, età, occhiali, posa delle testa ecc...

152.917

numero di immagini di volti

6.139

numero di identità



CelebA + WheDataset

- ▶ Unione dei due dataset (non-bilanciati) e tramite un'operazione di shuffle è stato creato un nuovo dataset con:

302.785

numero di immagini di volti

16.316

numero di identità

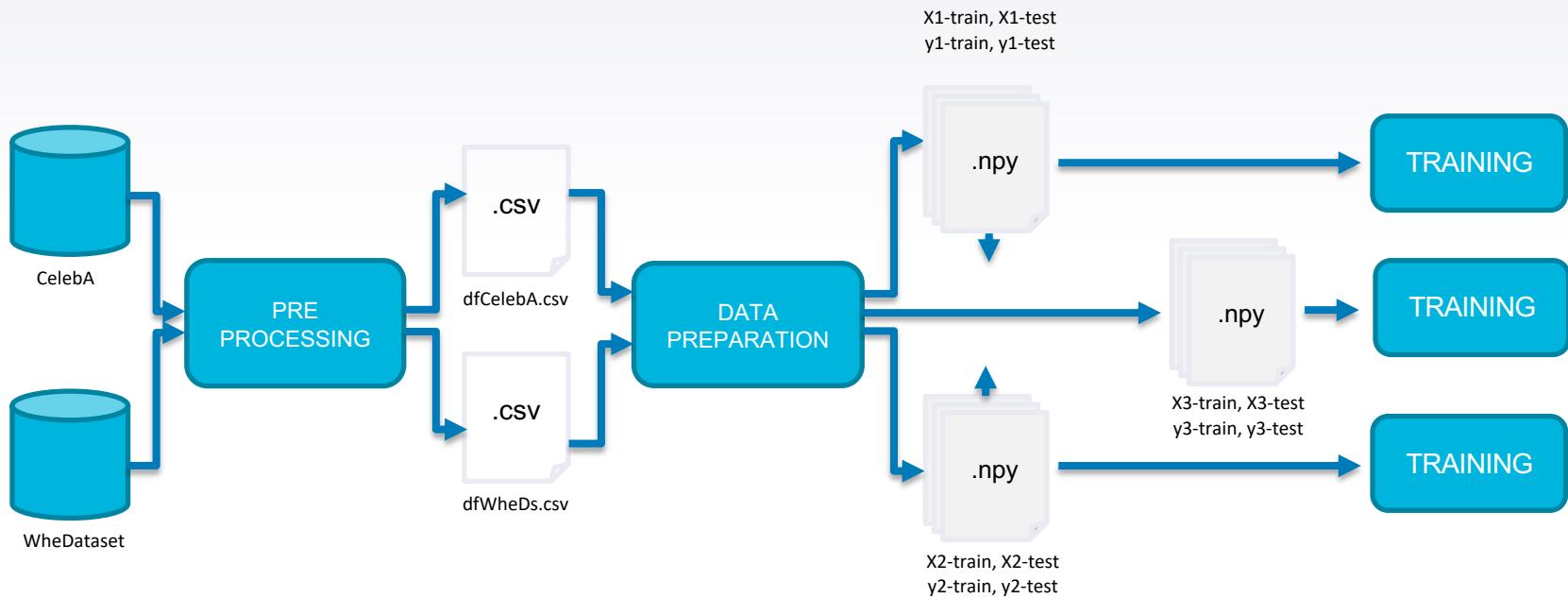


4

Architettura

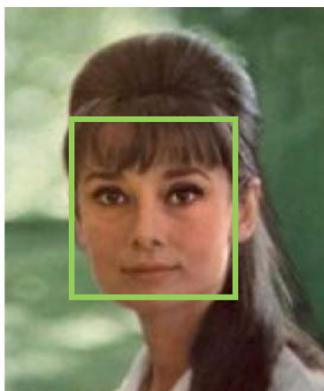


Architettura



Pre processing

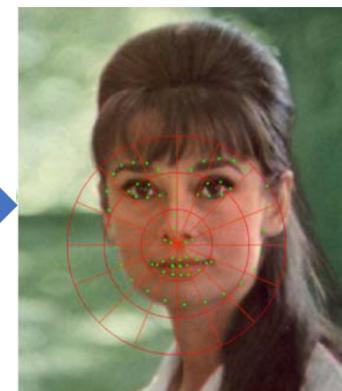
- L'obiettivo principale di questa fase è quello di: estrarre da ogni immagine presente nei due dataset i volti dei soggetti tramite la face detection, predire i 68 landmark facciali e determinare il vettore caratteristica di ogni immagine tramite lo Spider Web Method.



FACE DETECTION



LANDMARK PREDICTION

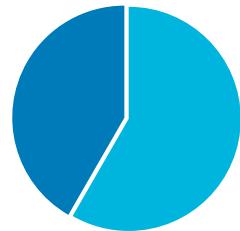


SPIDER WEB MODEL

Data Preparation

- ▶ Prima di effettuare la fase di training sui dati è stata effettuata un'operazione di bilanciamento e normalizzazione dei dati.

CelebA

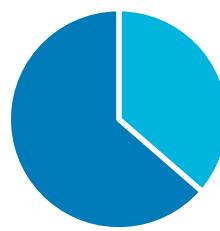


■ Donne ■ Uomini

84.434 uomini

118.165 donne

Whe Dataset

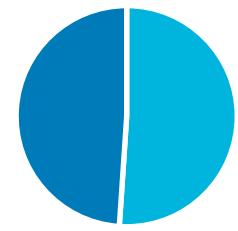


■ Donne ■ Uomini

63.735 uomini

36.451 donne

CelebA + Whe Dataset



■ Donne ■ Uomini

148.169 uomini

154.616 donne

5

Risultati ottenuti

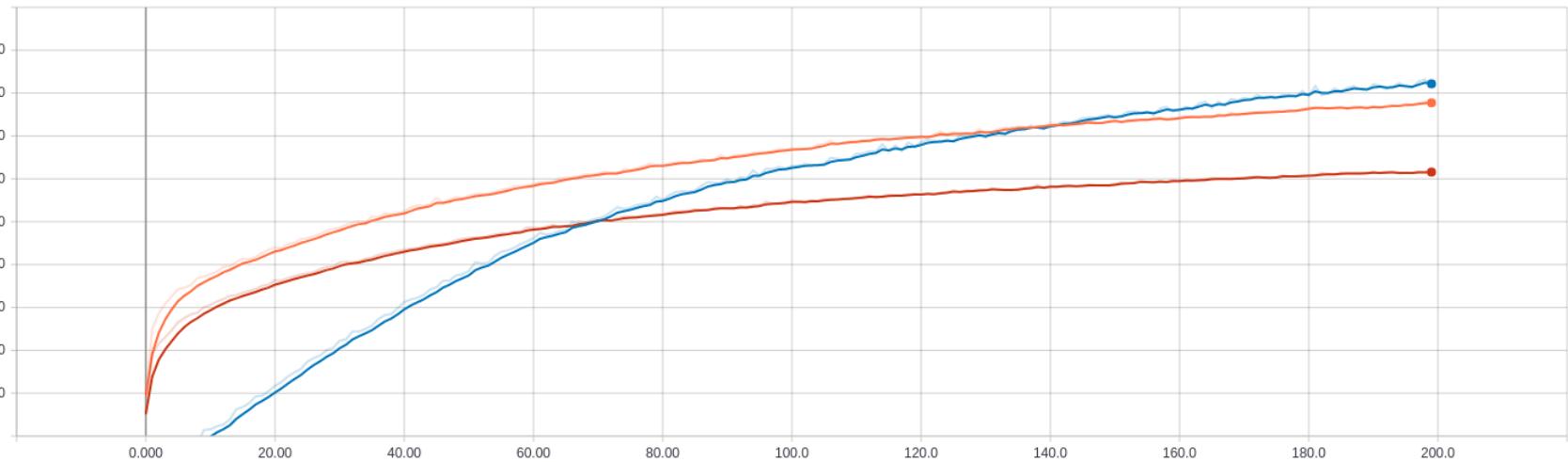


ANN1

- ▶ Accuracy maggiore di 0.80 nel trainset di tutti i dataset

Dataset 1 - CelebA
Dataset 2 – WheDataset
Dataset 3 – CelebA + WheDataset

epoch_accuracy



ANN1

epoch 200 e batch-size 64

- ▶ Dataset1 - CelebA

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	19161	6117
	negatives	6620	18763

Accuracy = 0.75

- ▶ Dataset2 - WheDataset

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	6435	4474
	negatives	3249	7713

Accuracy = 0.65

ANN1

epoch 200 e batch-size 64

- ▶ Dataset3 - CelebA + WheDataset

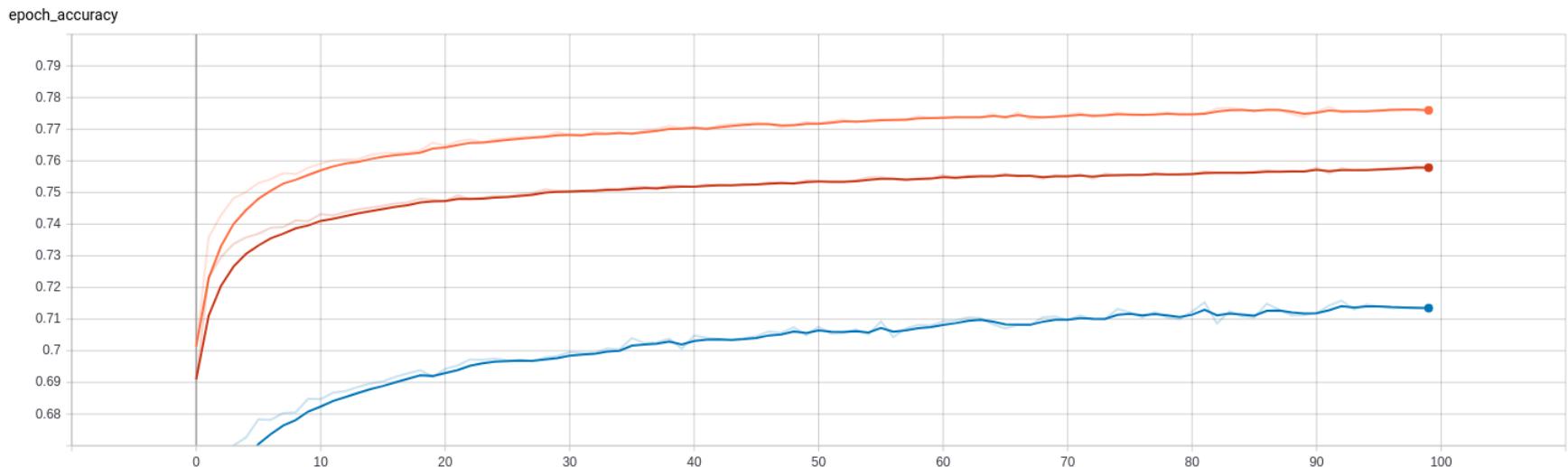
CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	33544	10985
	negatives	12196	32177

Accuracy = 0.74

ANN2

- ▶ Accuracy nel trainset su tutti i dataset con 100 epocha

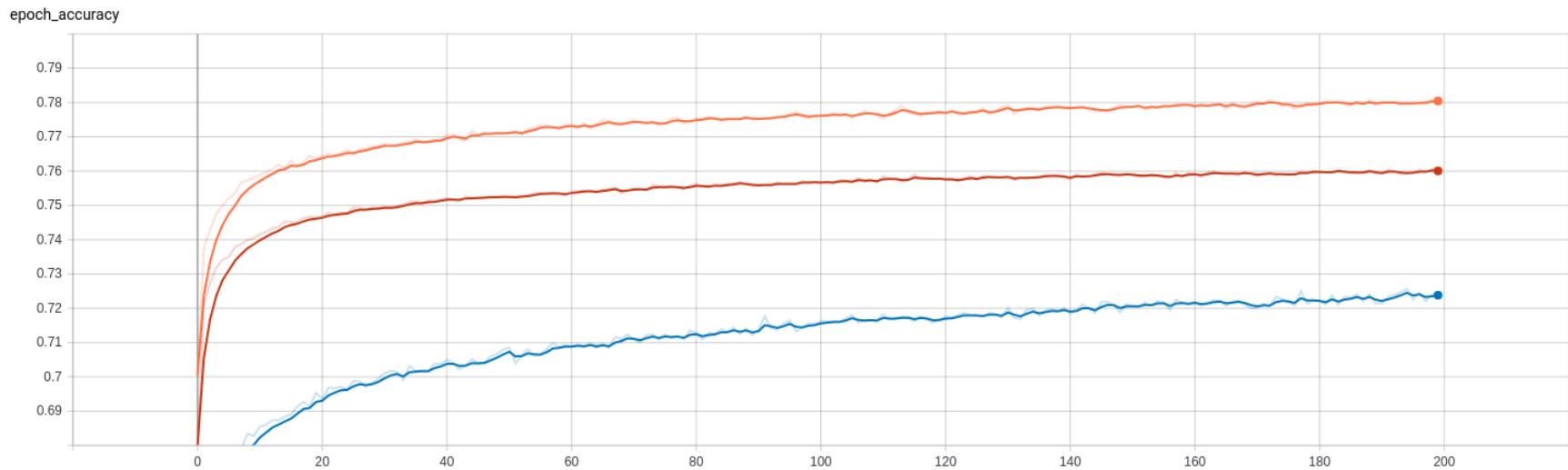
Dataset 1 - CelebA
Dataset 2 – WheDataset
Dataset 3 – CelebA + WheDataset



ANN2

- ▶ Accuracy nel trainset su tutti i dataset con 200 epocha

Dataset 1 - CelebA
Dataset 2 – WheDataset
Dataset 3 – CelebA + WheDataset



ANN2

epoch 100 e batch-size 32

- ▶ Dataset1 - CelebA

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	19575	5703
	negatives	5785	19598

Accuracy = 0.77

- ▶ Dataset2 - WheDataset

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	6839	4070
	negatives	2812	8150

Accuracy = 0.69

ANN2

epoch 100 e batch-size 64

- ▶ Dataset3 - CelebA + WheDataset

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	31957	12572
	negatives	9260	35113

Accuracy = 0.75

ANN2

epoch 200 e batch-size 64

- ▶ Dataset1 - CelebA

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	19943	5335
	negatives	6323	19060

Accuracy = 0.77

- ▶ Dataset2 - WheDataset

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	6842	4067
	negatives	2857	8105

Accuracy = 0.68

ANN2

epoch 200 e batch-size 64

- ▶ Dataset3 - CelebA + WheDataset

CONFUSION MATRIX		Actual Value	
Predicted Value		positives	negatives
	positives	32413	12116
	negatives	9721	34652

Accuracy = 0.75

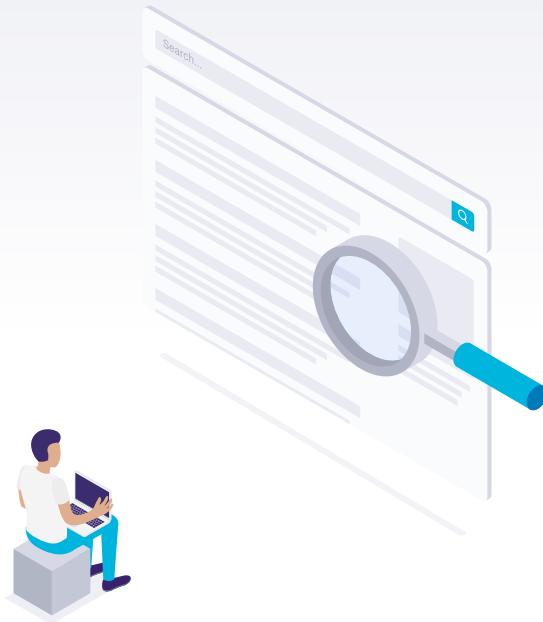
Test incrociati

- ▶ Utilizzo di ANN2 con un 100 epoche e batch size 32

TRAINING	TESTING	Accuracy
Dataset1	Dataset2	0.61
Dataset2	Dataset1	0.69

6

Conclusioni



Conclusioni

- ▶ La massima accuracy raggiunta dallo Spider Web Model è dello **0.77** sul dataset CelebA con una rete neurale di discreta qualità
- ▶ Andando ad aumentare le dimensioni della rete neurale in questione i risultati non si discostano da quelli ottenuti
- ▶ Lo Spider Web Model offrendo risultati soddisfacenti sulla classificazione del genere e della posa, potrà offrire in seguito migliori risultati anche sulla stima dell'età e delle espressioni facciali.





Grazie per l'attenzione

Grauso Fabio
matricola 0522500811