**Università degli Studi di Salerno**

**Dipartimento di Informatica**

Immagine che contiene regina

Descrizione generata automaticamente

**Corso di Laurea in Informatica**

Fondamenti di Visione Artificiale e Biometria

Progetto: Age Regression

|  |  |
| --- | --- |
| Relatori | Studenti |
| Prof. Michele Nappi  Dott.ssa Paola Barra | Domenico Trotta 0522500810  Michele Castellaneta 0522500824 |

**Anno Accademico 2019/2020**

1. Abstract
2. Scopo del progetto
3. ~~Dataset~~
   1. ~~Analisi dei dati~~
4. ~~Lavori precedenti~~
5. Lavoro svolto
   1. Caricamento dataset
   2. Pre-processing
   3. Creazione modello
   4. Training modello
   5. Testing modello
6. Valutazioni
7. Altri modelli
8. ~~Demo~~
9. Conclusioni

**Sommario**

[3. Dataset 4](#_Toc43720089)

[3.1 Analisi dei dati 5](#_Toc43720090)

[**3.1.1 Distribuzione età** 5](#_Toc43720091)

[**3.1.2 Distribuzione sesso** 6](#_Toc43720092)

[**3.1.3 Distribuzione etnia** 7](#_Toc43720093)

[4. Lavori precedenti 8](#_Toc43720094)

[4.1 UTKFace - age and sex prediction 9](#_Toc43720095)

[4.2 Age and Gender Estimation using CNN 10](#_Toc43720096)

[8. Descrizione Demo 12](#_Toc43720097)

[8.1 Age prediction con una sola persona 12](#_Toc43720098)

[8.2 Age prediction con due persone 13](#_Toc43720099)

[8.3 Schermata impostazioni 14](#_Toc43720100)

# 3. Dataset

UTKFace[[1]](#footnote-1) è un dataset di volti di persone con età che variano da 1 a 116 anni.

Il set di dati comprende oltre 20.000 immagini di volti (precisamente 23708) con annotazioni di età, sesso ed etnia. Le immagini presentano una grande varietà di pose, espressioni facciali, illuminazione, occlusioni, risoluzione, ecc. Inoltre, questo set di dati potrebbe essere utilizzato per un’ampia varietà di compiti, ad esempio face detection, age estimation, age progression/regression, lendmark localization, ecc.

Di seguito sono mostrate alcune immagini presenti nel dataset.

Immagine che contiene fotografia, largo, posando, mucchio

Descrizione generata automaticamente

Le etichette di ogni volto sono incorporate nel nome del file con il seguente formato:

[age] \_ [gender] \_ [race] \_ [data&time].jpg

* **[age]** è un numero intero compreso tra 0 e 116, che indica l'età;
* **[gender]** 0 (maschio) o 1 (femmina);
* **[race]** è un numero intero compreso tra 0 e 4, che indica occidentale, africana, asiatico, indiano e altri (come ispanico, latino, mediorientale);
* **[date&time]** è nel formato yyyymmddHHMMSSFFF, che mostra la data e l'ora in cui un'immagine è stata raccolta su UTKFace.

## 3.1 Analisi dei dati

Prima di iniziare ad effettuare una completa analisi del dataset abbiamo riscontrato che su alcune immagini l’operazione di crop e align non è stata effettuata correttamente e questo ha portato all’introduzione di immagini che non raffiguravano volti di persone.

Di seguito sono riportate alcune delle immagini sopra citate che abbiamo individuato e provveduto a rimuovere.

Immagine che contiene vicino, remoto

Descrizione generata automaticamenteImmagine che contiene tatuaggio, disegnando

Descrizione generata automaticamenteImmagine che contiene cibo

Descrizione generata automaticamenteImmagine che contiene cibo, latte

Descrizione generata automaticamenteImmagine che contiene persona, uomo, fotografia, guardando

Descrizione generata automaticamente

Chiaramente per il nostro scopo le etichette saranno l’età dei volti presenti nel dataset, mentre le caratteristiche che abbiamo considerato nella creazione dei diversi modelli (che verranno illustrati successivamente) oltre alle immagini stesse sono il sesso e l’etnia. Pertanto, di seguito gli attributi età, sesso ed etnia verranno analizzati singolarmente.

### **3.1.1 Distribuzione età**

Siamo partiti con l’analizzare la distribuzione dell’età dei volti presenti nel dataset.

Di seguito è mostrato il grafico che abbiamo realizzato.

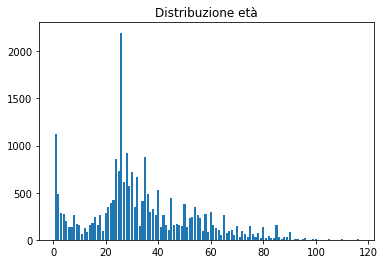


Figura 1 Grafico Distribuzione età.

Da una rapida analisi del grafico mostrato in *Figura 1* si può notare che:

* Nella fascia di età compresa tra i 25 e i 27 anni si ha un numero di volti molto maggiore rispetto alla media;
* Nella fascia d’età da 81 a 116 anni il numero di volti è molto basso rispetto alla media dei volti per ogni età.

### **3.1.2 Distribuzione sesso**

La seconda caratteristica sul quale ci siamo focalizzati è il sesso. Di seguito è mostrato il grafico che mostra il numero di volti per ciascun sesso.

Immagine che contiene screenshot

Descrizione generata automaticamente

Figura 2 Grafico Distribuzione sesso.

Come si può notare osservando il grafico mostrato in *Figura 2* la distribuzionedi immagini per i due sessi risulta essere piuttosto equilibrata. Abbiamo **12379** volti di uomini e **11316** volti di donne.

### **3.1.3 Distribuzione etnia**

Infine, l’ultima caratteristica che abbiamo analizzato è stata l’etnia. Come già accennato in precedenza, sono state individuate 5 differenti tipologie di etnia:

* Occidentale;
* Africana;
* Asiatica orientale;
* Asiatica centro – meridionale;
* Altro.

Di seguito è riportato il grafico che mostra il numero di volti per ognuna delle 5 categorie.

Immagine che contiene screenshot

Descrizione generata automaticamente

Figura 3 Distribuzione etnia.

Osservando il grafico mostrato in *Figura 3* si nota che per l’etnia “Occidentale” si ha il numero maggiore di volti, mentre tra le restanti etnie le differenze sono nettamente inferiori.

Precisamente, abbiamo **10070** volti per etnia Occidentale, **4526** per l’etnia Africana, **3431** e **3976** rispettivamente per l’etnia Asiatica orientale e Asiatica centro-meridionale e **1692** altro.

# 4. Lavori precedenti

In questa sezione, discuteremo di alcuni lavori[[2]](#footnote-2) che sono stati già svolti in questo ambito e che hanno utilizzato il dataset UTKFace per l’addestramento dei loro modelli.

Di seguito analizzeremo i modelli che sono stati realizzati e i risultati che si sono ottenuti.

Da notare che i lavori mostrati di seguito oltre ad effettuare l’operazione di age prediction hanno effettuato anche un operazione di gender prediction, pertanto focalizzeremo l’attenzione sui risultati ottenuti per quanto riguarda la sola operazione di age prediction.

I modelli che verranno mostrati sono stati costruiti con Keras utilizzando le Convolutional Neural Networks (CNN). Una rete neurale convoluzionale è un tipo speciale di deep neural network che funziona estremamente bene per la classificazione di immagini. Una CNN è essenzialmente costituita da un blocco di **input**, uno o più blocchi nascosti **(hidden layer)**, che effettuano calcoli tramite funzioni di attivazione (**ad esempio RELU**) e un blocco di **output** che effettua la classificazione/regressione vera e propria.

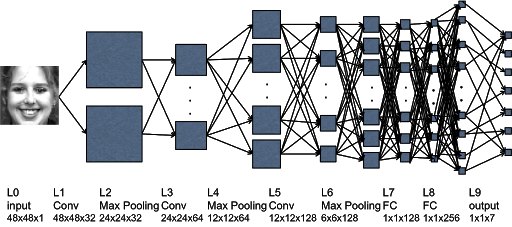


Figura 4 Struttura di una CNN.

Nella immagine in *Figura 4* è mostrata la struttura tipica di una CNN che risulta essere costituita da:

* Livelli di convoluzione;
* Livello ReLU (Rectified Linear Units);
* Livello di Pooling;
* Livello FC (o Fully connected, completamente connesso).

## 4.1 UTKFace - age and sex prediction

Nel primo dei due lavori esaminati abbiamo osservato che:

* Le immagini sono state lette a colori;
* Nell’estimation del sesso e dell’età la feature etnia non è stata utilizzata, pertanto, si sono utilizzate soltanto le immagini;
* La CNN costruita è strutturata nel seguente modo:
  + Convolutional layer; 32 nodi, kernel size 3
  + Convolutional layer; 32 nodi, kernel size 3
  + Pooling layer; Pool size 2
  + Convolutional layer; 64 nodi, kernel size 3
  + Convolutional layer; 64 nodi, kernel size 3
  + Pooling layer; Pool size 2
  + Convolutional layer; 84 nodi, kernel size 3
  + Fully connected layer; 64 nodi
  + Fully connected layer; 1 nodi
* É stato effettuato un training con batch size uguale a 240 e un numero di epoche pari a 2000.
* I risultati che si sono ottenuti effettuando l’addestramento su un training set composto da 15884 samples e un validation set di 7824 samples sono elencati di seguito[[3]](#footnote-3):
  + Loss: 4.8312
  + Val\_loss: 6.7540;
* L’evaluate del modello sul test set ha fornito una loss pari a 6.0371. Quindi, l’errore medio commesso dal modello è di circa 6 anni.



**Etichette:**

[‘0’]: sesso.

[‘26’]: età.

Come si può osservare, in questo esempio l’errore commesso dalla rete nello stimare l’età del soggetto risulta essere molto basso.

## 4.2 Age and Gender Estimation using CNN

Nel secondo dei due lavori esaminati abbiamo osservato che:

* Le immagini sono state lette in scala di grigi;
* Nell’estimation del sesso e dell’età la feature etnia non è stata utilizzata, pertanto, si sono utilizzate soltanto le immagini;
* La CNN costruita è strutturata nel seguente modo:
  + Convolutional layer; 32 nodi, kernel size 3
  + Convolutional layer; 64 nodi, kernel size 3
  + Pooling layer; Pool size 2
  + Convolutional layer; 128 nodi, kernel size 3
  + Pooling layer; Pool size 2
  + Fully connected layer; 128 nodi
  + Fully connected layer; 64 nodi
  + Fully connected layer; 32 nodi
  + Fully connected layer; 1 nodi
* É stato effettuato un training con batch size uguale a 128 e un numero di epoche pari a 25.

Visualizzando alcuni esempi pubblicati si può notare che per il sex prediction il modello funziona piuttosto bene, ma lo stesso non si può dire dell’age prediction.



Osservando l’immagine di cui sopra si può osservare come la rete non è riuscita a stimare l’età della persona mostrata in foto.

Poiché per tale modello non è stato effettuato l’evaluate sul test set non disponiamo di ulteriori dati.

# 

# 8. Descrizione Demo

In questa sezione verrà descritta la demo che abbiamo realizzato per testare il nostro modello. L’analisi che effettueremo di seguito verterà principalmente su quelle che sono le diverse impostazioni di test che vengono offerte nella demo. Quindi, tramite delle immagini verrà mostrata l’interfaccia della demo, mentre per la visione del codice si rimanda al file demoGUI.py.

## 8.1 Age prediction con una sola persona

Immagine che contiene persona, uomo, interni, fotografia

Descrizione generata automaticamente

**4**

**2**

**1**

**3**

**1**

Tasto per accedere alle impostazioni della demo descritte nel paragrafo 8.3.

**2**

Viene mostrato il nome del modello che si sta utilizzando, in questo caso il modello ***Img/Sex/Eth Colori***, ovvero un modello che fa uso non solo delle immagini, ma anche del sesso e dell’etnia per effettuare la prediction dell’età.

**3**

Tasto per chiudere la demo.

**4**

Per ogni frame catturato viene effettuato il prediction dell’età, ma prima di stampare il risultato della prediction viene effettuata una media con le acquisizioni precedenti. Si è scelto di optare per una media di tutte le acquisizioni poiché il numero stampato risultava essere piuttosto variabile a seconda dell’illuminazione e dell’orientamento del volto.

In questo modo, dopo qualche secondo il valore mostrato tende a stabilizzarsi rendendone più agevole la lettura.

## 8.2 Age prediction con due persone

Immagine che contiene persona, interni, fotografia, donna

Descrizione generata automaticamente

L’unica differenza rispetto al modello precedente è che in questo caso viene visualizzata una fascia d’età piuttosto che un numero.

Abbiamo optato per la scelta di una fascia d’età, poiché risulta essere più difficile associare la media ad ogni volto rilevato, in quanto, il primo volto identificato in un’acquisizione potrebbe essere rilevato come secondo volto in un’acquisizione successiva e viceversa.

Un’alternativa che abbiamo considerato è quella di avvalersi di un ulteriore modello, che effettui un’operazione di face recognition[[4]](#footnote-4) in modo tale che ogni età calcolata (calcolando la media delle varie acquisizioni che vengono fatte in real-time) sia associata univocamente ad uno degli *n* volti individuati. Pertanto, se vengono rilevati *n* volti avremo che per ognuno di essi vi sarà un media che rappresenterà l’età di quel determinato volto.

Con un approccio di questo tipo, è indifferente l’ordine con cui i volti verranno rilevati, poiché ogni soggetto (riconosciuto per mezzo di un’operazione di face recognition) avrà una propria media che verrà aggiornata man mano che i frame verranno acquisiti e valutati.

## 8.3 Schermata impostazioni

Il menu impostazioni della demo realizzata è mostrato nell’immagine seguente.

**Immagine che contiene screenshot

Descrizione generata automaticamente**

**1**

**2**

**5**

**3**

**4**

Figura 5 Impostazioni.

Con tale menu mostrato in Figura 5 si vuole dare la possibilità di poter effettuare uno switch tra i diversi modelli addestrati e valutarne i risultati in real-time. La schermata offre le seguenti opzioni:

**1**

Possibilità di selezionare il modello da utilizzare tra una serie di modelli proposti. È possibile selezionare modelli che effettuano la prediction basandosi solo ed esclusivamente sull’immagine acquisita oppure selezionare modelli che considerano anche altre caratteristiche come il sesso e l’etnia. Inoltre, è possibile anche scegliere se si vuole testare un modello che faccia uso di immagini a colori o in scala di grigi.

Possibilità di selezionare il sesso del soggetto nel caso in cui si sia selezionato un modello che faccia uso di tale caratteristica.

**2**

**3**

Possibilità di selezionare l’etnia del soggetto nel caso in cui si sia selezionato un modello che faccia uso di tale caratteristica.

**4**

Possibilità di modificare la soglia del riconoscimento facciale. Se il volto non viene riconosciuto vi è la possibilità di abbassare la soglia, mentre nel caso in cui si voglia essere più selettivi ed evitare il più possibile false alarms è possibile alzare tale soglia.

**5**

Per confermare le modifiche effettuate è necessario cliccare la spunta verde.

1. https://susanqq.github.io/UTKFace/ [↑](#footnote-ref-1)
2. I due lavori esaminati sono entrambi pubblicati sulla piattaforma Kaggle nella sezione kernel [1]. [↑](#footnote-ref-2)
3. Sono riportati soltanto i risultati per l’attività di age prediction. [↑](#footnote-ref-3)
4. https://www.pyimagesearch.com/2018/09/24/opencv-face-recognition/ [↑](#footnote-ref-4)