Project PRIME

Students: Fabrizio Mazzone, Mattia Proietto
Company: Ubble
Task: Face matching

Useful Datasets for face recognition

1. **YouTube Faces DB:** [Link](#)
   a database of face videos designed for studying the problem of unconstrained face recognition in videos.The data set contains 3,425 videos of 1,595 different people. All the videos were downloaded from YouTube. An average of 2.15 videos are available for each subject. The shortest clip duration is 48 frames, the longest clip is 6,070 frames, and the average length of a video clip is 181.3 frames.
   State: Free and easy to download
   Size: 24 GB
   Username: wolftau
   Password: wtal997

   Comments: after downloading the DB we saw that not all the existent pictures are exactly in the way we need. The character's face is not always depicted as a portrait and most of times the pictures, although including just a face, are much more wider.
   So we had to select some specific and useful group of subject to perform our algorithm (n.9)

2. **NRC-IIT Facial Video Database**: [link](#)
   This database contains pairs of short video clips each showing a face of a computer user sitting in front of the monitor exhibiting a wide range of facial expressions and orientations as captured by an Intel webcam mounted on the computer monitor.
   The video capture resolution is kept to 160 x 120. With the face occupying 1/4 to 1/8 of the image (measured by width), this translates into a commonly

observed on a TV screen situation when a face of a TV show occupies 1/8 to 1/16 of the screen. It could be good as an initial test for the algorithm
State: Free and easy to download

3. **Faces 1999**
   Frontal face dataset.450 face images. 896 x 592 pixels. Jpeg format.
   27 or so unique people under with different lighting/expressions/backgrounds.
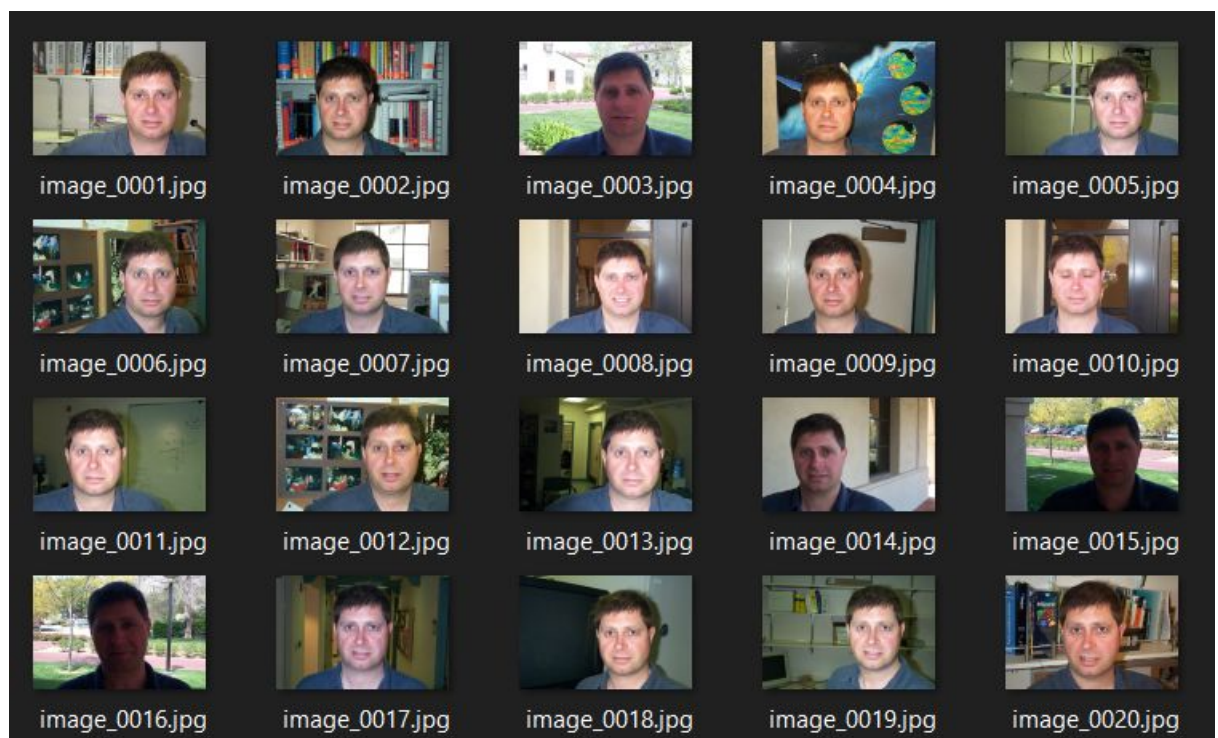   ImageData.mat is a Matlab file containing the variable
   SubDir_Data which is an 8 x 450 matrix. Each column of this matrix hold the coordinates of the bike within the image, in the form:
   [x_bot_left y_bot_left x_top_left y_top_left ...
    x_top_right y_top_right x_bot_right y_bot_right]
   State: Downloaded (Fabrizio's laptop)
   Size: < 100 MB
   Drawbacks: ok white and asian people, **no black ones**

   Comments: very useful. It corresponds exactly with what we are looking for. We trained algorithm 9 on it.

4. **Yale Face DataSet** : link
   The Yale Face Database (size 6.4MB) contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: centre-light, w/glasses, happy, left-light,

w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. Could be useful in the initial process.
State: free to download

5. **Face Recognition Data, University of Essex, UK**: link
   395 individuals (male and female), 20 images per individual. Contains images of people of various racial origins, mainly of first year undergraduate students, so the majority of indivuals are between 18-20 years old but some older individuals are also present. Some individuals are wearing glasses and beards.The data is held in four directories (faces94 , faces95 , faces96 , grimace), in order of increasing difficulty. Faces96 and grimace are the most difficult, though for two different reasons (variation of background and scale, versus extreme variation of expressions).
   State: free to download

- **Dataset 5 is not suitable for our purposes, the images are relatively small and all of them have been taken with a uniform green background.**

6. **AR Face Database**: link
   126 people (over 4,000 color images).
   Different facial expressions, illumination conditions and occlusions.
   Two sessions per person (2 different days).
   This face database was created in the Computer Vision Center (CVC) at the U.A.B. It contains over 4,000 color images corresponding to 126 people's faces (70 men and 56 women). Images feature frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. Each person participated in two sessions, separated by two weeks (14 days) time. The same pictures were taken in both sessions.
   This face database is publicly available and can be obtained from this web-site. It is free for academic use. Commercial distribution or any act related to commercial use of this database is strictly prohibited.
   State: Free but no easy to get (permission needed, with honour declaration)

7. **Disguised Faces in the Wild dataset**: link
   The DFW dataset consists of 1000 subjects and 11157 images. Pre-defined training and testing splits are provided, where 400 subjects form the training set, while the remaining 600 subjects form the test set. Each subject may contain normal, validation, **disguised**, and **impersonator** images.

So could be useful to test people that can look similar.
State: Free to donwload / downloaded but need passoword (Fabrizio's laptop)
Size: 4.5 GB

- **Dataset 7 still waiting for approval. Already downloaded but it needs passwords to be decrypted. We already asked and signed for permission but no answers were given.**

8. **VidTIMIT Audio-Video Dataset** : [link](#)
   The VidTIMIT dataset is comprised of video and corresponding audio recordings of 43 people, reciting short sentences.The dataset was recorded in 3 sessions, with a mean delay of 7 days between Session 1 and 2, and 6 days between Session 2 and 3. In addition to the sentences, each person performed a head rotation sequence in each session. The sequence consists of the person moving their head to the left, right, back to the center, up, then down and finally return to center.
   The video of each person is stored as a numbered sequence of JPEG images with a resolution of 512 x 384 pixels.
   State: Free to download

9. **UPNA Head Pose Database**: [link](#)
   Is a public database of videos for head tracking and pose estimation. The database consists of a set of 120 videos acquired with a standard webcam, corresponding to 10 different subjects (6 males and 4 females) and 12 videos each. Every set of 12 videos is composed of 6 guided-movement sequences and 6 free-movement sequences In the free sequences, the user moves the head at free will combining translations and rotations along the three spatial axes. Every video begins and ends with the head in a frontal position.
   State: Free download

10. **IST-EURECOM Light Field Face Database**: [link](#)
    Two separate sessions were performed for each subject with a temporal separation between 1 and 6 months. The database includes 20 image shots per person in each of the two sessions, with several facial variations including emotions, actions, poses, illuminations, and occlusions.
    The database can be used not only in the context of face recognition research but also for other research areas such as emotion recognition, gender classification, age estimation, ethnicity classification and face modeling.
    State: Free to download but need to ask for permission

Size: 25 part of 4 GB each

- **Dataset 10 still waiting for approval. All the images are encrypted and a password is needed, we already signed and sent several times the agreement but no answers were given.**

11. **I2Head**: link
    For each user, 8 sessions are recorded under controlled movements in static and free head movements' scenarios. Point grids containing 17 and 65 fixations are recorded. For each fixation point, the best 10 frames are selected providing an image and the head poses for each one of the samples. In addition, a simplified head model for each user is provided.
    Size: Around 24GB
    State: Free download

Useful Algorithms for face recognition

Eigenfaces (1991)
Local Binary Patterns Histograms (LBPH) (1996)
Fisherfaces (1997)
Scale Invariant Feature Transform (SIFT) (1999)
Speed Up Robust Features (SURF) (2006)

1. LBPH: link
   Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number.
   Advantages: easy, based on local features in the images, can achieve, great results,robust.
   Implementation: Go, OpenCV

2. PCA (from university studies)
   Language: Matlab

3. Self-organizing Map (from university studies)
   Language: Python

4. Face Recognition Algorithms: link
   This file contains a description of all the known method to detect and recognise face pictures.

5. Face recognition through Dlib: link

6. Face recognition through SVM: link
   Language: Python (Dlib)

7. Face recognition through KNN: link
   Language: Python (Dlib)

8. Face recognition live through Webcam: link
   Language: Python (Dlib)

9. Face recognition with OpenCV, Python, and deep learning: link
   Language: Python (Dlib)

10. Video-Based Face Recognition and Face-Tracking using Sparse Representation Based Categorization: paper link

11. Face Recognition in Video several approaches: paper link
    pages 5 to 15

Useful Algorithms from PapersWithCode

12. **DeepFace: Closing the Gap to Human-Level Performance in Face Verification**
    Paper: link
    Code: link

    They revisit face recognition standards, changing the used pattern: detect => align => represent => classify. They revisit both the alignment step and the representation step by employing explicit 3D face modeling in order to apply a piecewise affine transformation, and derive a face representation from a nine-layer deep neural network.

    Pro: > 90% accuracy over more than 40 000 subjects
    Cons: using VGG (convolutional neural network); it needs long trainings.

13. **Improving Face Anti-Spoofing by 3D Virtual Synthesis:**
    Paper: link
    Code: link

    in this paper, they present a method to synthesize virtual spoof data in 3D space to alleviate spoofing issues. Specifically, they considered a printed photo as a flat surface and mesh it into a 3D object,which is then randomly bent and rotated in 3D space.
    Cons: using CNN

14. **Stacked Dense U-Nets with Dual Transformers for Robust Face Alignment**
    Paper :link
    Code: link

    Novel scale aggregation network topology structure and a channel aggregation building block to improve the model's capacity without sacrificing the computational complexity and model size. With the assistance of deformable convolutions inside the stacked dense U-Nets and coherent loss for outside data transformation, our model obtains the ability to be spatially invariant to arbitrary input face images. Extensive experiments on many in-the-wild datasets, validate the robustness of the proposed

method under extreme poses, exaggerated expressions and heavy occlusions. Finally, we show that accurate 3D face alignment can assist pose-invariant face recognition where we achieve a new state-of-the-art accuracy on CFP-FP.

15. **FaceNet: A Unified Embedding for Face Recognition and Clustering**
    Paper:link
    Code:link

    This paper present a system, called FaceNet, that directly learns a mapping from face images to a compact Euclidean space where distances directly correspond to a measure of face similarity. Once this space has been produced, tasks such as face recognition, verification and clustering can be easily implemented using standard techniques with FaceNet embeddings as feature vectors. Our method uses a deep convolutional network trained to directly optimize the embedding itself, rather than an intermediate bottleneck layer as in previous deep learning approaches. To train, we use triplets of roughly aligned matching / non-matching face patches generated using a novel online triplet mining method. The benefit of our approach is much greater representational efficiency: we achieve state-of-the-art face recognition performance using only 128-bytes per face. On the widely used Labeled Faces in the Wild (LFW) dataset, our system achieves a new record accuracy of 99.63%. On YouTube Faces DB it achieves 95.12%. Our system cuts the error rate in comparison to the best published result by 30% on both datasets. We also introduce the concept of harmonic embeddings, and a harmonic triplet loss, which describe different versions of face embeddings (produced by different networks) that are compatible to each other and allow for direct comparison between each other.

16. **DocFace: Matching ID Document Photos to Selfies**
    Paper: link
    Code: link

    An automatic system for matching ID document photos to live face images in real time with high accuracy would speedup the verification process and remove the burden on human operators. In this paper, by employing the transfer learning technique, we propose a new method, DocFace, to train a domain-specific network for ID document photo matching without a large dataset. Compared with the baseline of applying existing methods for general face recognition to this problem, our method achieves considerable improvement.
    In the paper: parameter tuning informations for CNN