Computer Vision Final Project: Face Recognition

Abhinav Singh

COURSE ID : CPSC-6820-003

1 Introduction

The objective of the program will be to detect objects of interest(faces) in a video and recognize faces of celebrities and that were used in the training of the system using Deep Neural Networks or Histogram of oriented Gradients . As there is a plethora of data-set available to train celebrities' faces and it would be hard to train too many celeb faces because of limited resources. So a data-set is images of actors from a particular movie like Jurassic Park to show the network is being trained properly and can recognize faces of the actors and can name them.

2 Approach

Finding all the Faces:
The first step in the pipeline is face detection, to find all faces in an image frame. For this purpose a more reliable solution is (Histogram of Oriented Gradients)HOG rather than Paul Viola and Michael Jones's face detection method.
The option to use CNN instead of HOG is also provided, but the result are not as accurate as Neural Networks requires large data-set to come with accurate predication. So for large data-set CNN is more optimal than HOG. As our data-set is small HOG is more optimal here.

Posing and Projecting Faces
After isolating the faces in our image, using Face landmark estimation, which locates 68 specific points(called landmarks) on a face, and using affine transformations, which would help the system to recognize a face turned in different directions.
For extracting these feature I looked into dlib, found something better a package which is built on top of both dlib and deep learning model with accuracy of 99.83 percent i.e. Face Recognition.

Encoding Faces:

To avoid training the system again and again on new images of the same person, encoding is used, which takes out the basic measurements of the person's face. The best way of doing it is to let the system figure out which part is most important to measure, by using Deep Convolutional Neural Network and training the network to generate 128 measurements of each face embedding.

Finding the person's name from the encoding:
This can be done by training Support Vector Machine(SVM) or any other classification algorithm.

3 Implementation:

Face Recognition Libraries used:
-Dlib: The Dlib library, maintained by Davis King, contains our implementation of "deep metric learning" which is used to construct our face embeddings used for the actual recognition process.

-Face Recognition: The Face Recognition library, created by **Adam Geitgey**, wraps around dlib's facial recognition functionality, making it easier to work with.

Dataset:
Data-set less than 50 images of 6 actors in movie Jurassic Park, to create 128-d embedding for each face in the data-set. These embedding would be later used to detect actors in video streams.

Directories:
-dataset/: Contains face images for 6 actors of Movie Jurassic Park.
-examples/: 3 face images for testing that are not in the dataset.
-videos/: Input video clip from movie Jurassic park as input video stream.
-output/: .mp4 video output file.

Files and Usage:
-facial_encoding.py: Encoding 128-d vectors for faces are built with this script.
-encodings.pickle: Facial recognitions encodings are generated from your dataset via facial_encoding.py and then serialized to disk.
-Facial_recognition_image.py: Recognize faces in a single image (based on encodings from your dataset).
-Facial_recognition_video.py: Recognize faces in a video file residing on disk and output the processed video to /output/.

Flags and their Usage:

–**dataset: "-i"** for adding dataset location.
–**encoding: "-e"** output location of encodings.
–**detection-method: "-d"** for passing method either "cnn" or "hog".
–**display: "-y"** for adding displaying the video while it is processing.

4 References:

[1.]Kazemi, Vahid Sullivan, Josephine. (2014). One Millisecond Face Alignment with an Ensemble of Regression Trees. 10.13140/2.1.1212.2243.
[2.]https://cmusatyalab.github.io/openface/
[3.]Zhu, Qiang Yeh, Mei-Chen Cheng, K.-T. Tim Avidan, Shai. (2006). Fast Human Detection Using a Cascade of Histograms of Oriented Gradients. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2. 1491 - 1498. 10.1109/CVPR.2006.119.
[4.]Viola, Paul Jones, Michael. (2004). Robust Real-Time Face Detection. International Journal of Computer Vision. 57. 137-154. 10.1023/B:VISI.0000013087.49260.fb.
[5.]https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78