Face Detection, Recognition and Classification
Computer Vision and Image Processing (Fall 2017)
University at Buffalo

Prachi Jayakumar Patil
UB Person Number: 50247309
ppatil3@buffalo.edu

Smruti Venkatesh
UB Person Number: 50247212
smrutive@buffalo.edu

Abstract

The goal of this project is to study different face detection methods, utilize the better one in the implementation of face recognition and further classify it in different folders.

1. Introduction

A face recognition system is a computer application capable of identifying or verifying a person from a digital image. One of the ways to do this is by detecting the facial features from the image and comparing it to a face database available. Following process is followed in this paper: First step is face detection.Then comes data gathering and comparison of the unique facial characteristics of each face irrespective of variations in background.Next is face Recognition, i.e. identifying the person in image. Last one is classification of the images in different folders based on the obtained names.

2. Dataset

Large face datasets are important for advancing face recognition research. The FaceScrub dataset comprises a total of 106,863 face images* of male and female 530 celebrities, with about 200 images per person. As such, it is one of the largest public face databases.
The images were retrieved from the Internet and are taken under real-world situations (uncontrolled conditions). Name and gender annotations of the faces are included. We have take a chunk of these data and implemented for 8 celebrities. A sample of it can be seen in the figure below:-
We would be using the inbuilt opencv libraries of HAAR, LBP for the computing.

3. Face Detection

Face identification is a technology and an important process in recognition that identifies human faces in digital images. Some of the methods for face detection are Principal component analysis, detection using skin segmentation,using HOG, rejection based classifier etc. Factors to be considered while selection is to check if these methods are luminance invariant, connectivity present in the photo, etc.Skin segmentation, HAAR and LBP Classifiers are the methods studied and implemented.

3.1. Skin Color segmentation

The flow of this method can be seen in following figure

3.1.1 Skin segmentation

Skin color segmentation is the process of separation between skin and non-skin pixels. The image is converted from RGB to a different color model. In YCbCr color model, the distribution of skin areas is consistent across different races in the Cb and Cr color spaces. In HSV color model, Hue (H) is not reliable for the discrimination task when the saturation is low whereas the RGB color model is lighting sensitive. Thus, YCbCr is chosen. The range in which the skin region is chosen approximately and the ones that fall in this region are retained and others are rejected by assigning them black.

3.1.2 Morphological processing

Morphological Processing is an operation on a binary image that creates a new binary image in which the pixel has a non-zero value only if the test is successful at that location

Figure 1: Face scrub-vintage dataset

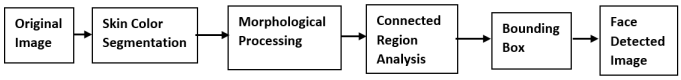

Figure 2: Flow for Skin Segmentation

Figure 3: Original Image of 2 women and its YCbCr converted image

in the input image. This helps us to remove the noise and clutter from the image obtained as an output of the previous step.
One of the most important things here is size of the disk that is used for performing the opening and filling operations. The disk size used for opening must be bigger than that of the filling process or else we might get a bigger connected region like an two arms combined that may be difficult to reject in other processes. The aim is to induce more number of holes in the image overall there by obtaining separate small regions and face would have holes for eyes nose mouth etc, which would help us in its detection.

3.1.3 Connected region analysis

In this phase we first reject all the small areas that are not potential phases. Area of any skin region is calculated by counting no of skin region pixels in the connected region.This basically helps to eliminate if there are small parts of legs or arms that have been retained after the previous processing. The output is as shown in figure 4. Human faces contain some holes like eyes, eyebrows, a mouth, mustaches etc. the two eyes of human face are major holes for human faces. If any skin region doesn't contain any holes then, we can safely discard these type all skin region.The rejection is

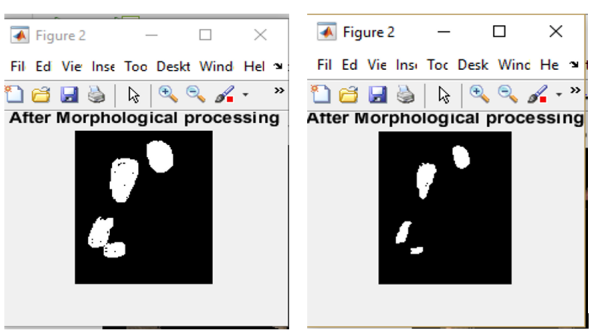

Figure 4: Output after morphological processing. Both figures have opening disk size 10.Left side figure has filling disk size 10 whereas the right one has filling disk size 4.

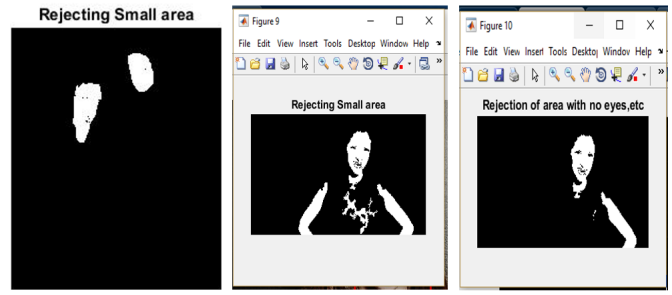

Figure 5: Fig 4a shows the two women after small area phase,Fig 4b is of anna kendrick image after the small areas are removed which is then given as input for euler number phase the output of which is fig 4c

done by Euler number method,which is the next step. The formula for Euler number is:

$$E = C - H$$

E is Euler no. of skin region is the total number of object in the image minus total number of holes in the object. C is the no. of connected component, and, H is no. of holes in a region. We can reject these type of skin regions which having Euler value greater than zero .
Here we can consider either 4 connectivity or 8 connectivity to consider the regions being analyzed for finding the number of holes. We prefer 4 connectivity as the smaller regions will be considered at a time unlike 8 connectivity where there is a possibility that face and neck together may considered there by not eliminating the neck.
Figure 4a does not exhibit any changes as the faces there are pretty much extracted.Although if we consider a different image the effect is evident as seen in figure 4b and 4c.

The next step in this process is rejection based on eccentricity. Generally, human face has an oval shape, so those regions which have shape similar to oval are not rejected by this method, and those regions whose shape is similar to line are rejected. For finding shape of skin regions, we used 'eccentricity based region properties' MATLAB function for each region. This function gives its eccentricity value. An ellipse (skin region) whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment. The oval shape of a face can be approximated by an ellipse so we calculate Eccentricity of all skin connected regions and to discard all skin regions whose eccentricity greater than by 0.89905.
The output of this phase can be observed in figure 5a. The final step in this chain is bounding box property. This stage rejects non human face skin region based on height to width ratio. Generally, height to width ratio of skin regions is measurable factor because it is also big factor for rejecting non human face like regions. If height to width ratio of skin region is less than by threshold value, then this skin region

will be discarding from class of probable human faces. Here we decided 1.902 threshold values for height to width ratio. For determining height to width ratio of each skin region, we used region properties based Bounding Box MATLAB function. The final output can be seen in fig 5b.

3.1.4 Disadvantages

- Hugely dependent on the choice of the color model.It retains non skin regions in case they fall near the spectrum of skin color. Cannot use HSV as the hue doesn't work properly if saturation is low. RGB is also not fully reliable because of it variance to the luminance factor. Even YCbCr cannot completely reject the non skin areas if the value of those colors fall in the skin value spectrum.An example of this is in visible in figure 6a. There we can see that the colors near the skin color spectrum are retained.

- While rejecting based on small area it may happen that the people who are standing far away, appear small and are rejected as seen in figure 6b.

- While rejecting the regions based on Euler number there is a possibility that if the people in the image are standing too close to each other they may be considered as a single region and be rejected. An example of this is shown in figure 7.

- If the output of the Euler phase is wrong, i.e. it gives long connected regions like a small part of hand that looks oval then it may be retained in this phase giving a false positive.

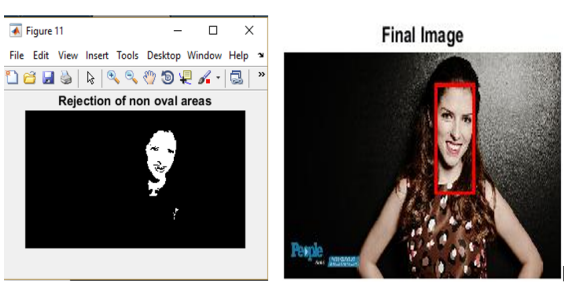

Figure 6: The fig 5a shows the image after rejecting the non oval shapes.Fig 5b shows the final image after bounding box

Figure 7: Fig 6a shows colors that fall near the skin color spectrum are retained and 6b shows that people standing far away are rejected during the small area phase

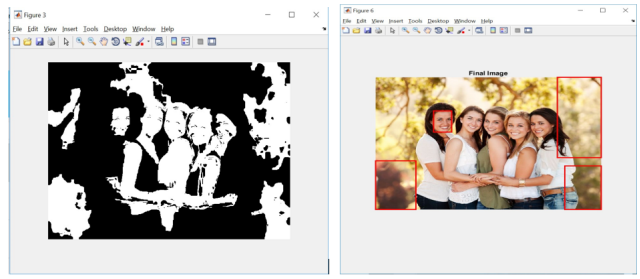

Figure 8: fig 7 shows how people standing too close are considered as a single region and rejected in euler phase

Thus skin segmentation is good in case of detecting single faces or fewer number of faces from sparsely connected images for which we get accurate results as in figures 8a and 8b.

3.2. HAAR and LBP classifier

This is another method using which is used for face detection. Face detection using skin segmentation is a thresholding method and whose disadvantages can be overcome using this method. We can reduce the number of false positives largely.
A computer program that decides whether an image is a positive image (face image) or negative image (non-face image) is called a classifier. A classifier is trained on hundreds of thousands of face and non-face images to learn how to classify a new image correctly.
Both of these classifiers process images in gray scales, basically because we don't need color information to decide if a picture has a face or not. As these are pre-trained in OpenCV, their learned knowledge files also come bundled with OpenCV opencv/data/.

3.2.1 HAAR classifier

The Haar Classifier is a machine learning based approach, an algorithm created by Paul Viola and Michael Jones. It starts by extracting Haar features from each image as shown by the windows in figure 9.
Each window is placed on the picture to calculate a single feature. This feature is a single value obtained by subtracting the sum of pixels under the white part of the window

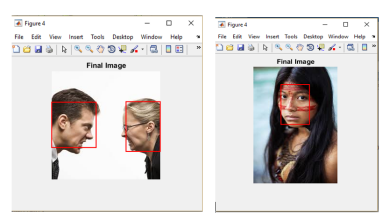

Figure 9: Output of face detection using skin segmentation

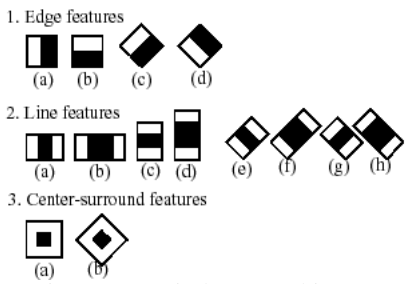

Figure 10: Windows used in HAAR

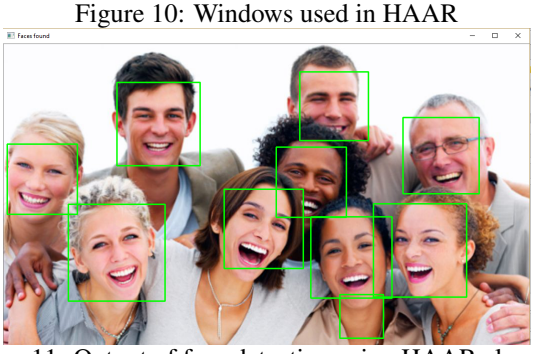

Figure 11: Output of face detection using HAAR classifier

from the sum of the pixels under the black part of the window.
Now, all possible sizes of each window are placed on all possible locations of each image to calculate plenty of features. But among all these features calculated, most of them are irrelevant. So we promptly discard irrelevant features and keep only those relevant with a fancy technique called Adaboost. AdaBoost is a training process for face detection, which selects only those features known to improve the classification (face/non-face) accuracy of our classifier.
After running the HAAR classifier we get the result as in figure 10.

3.2.2 LBP classifier

As any other classifier, the Local Binary Patterns (LBP in short) also needs to be trained on hundreds of images. LBP is a visual/texture descriptor, and thankfully, our faces are also composed of micro visual patterns. So, LBP features are extracted to form a feature vector that classifies a face from a non-face.
Each training image is divided into some blocks.
For each block, LBP looks at 9 pixels (33 window) at a time, and with a particular interest in the pixel located in the center of the window.
Then, it compares the central pixel value with every neighbor's pixel value under the 33 window. For each neighbor pixel that is greater than or equal to the center pixel, it sets its value to 1, and for the others, it sets them to 0.
After that, it reads the updated pixel values (which can be either 0 or 1) in a clockwise order and forms a binary number. Next, it converts the binary number into a decimal number, and that decimal number is the new value of the center pixel. We do this for every pixel in a block.
Then, it converts each block values into a histogram. Finally, it concatenates these block histograms to form a one feature vector for one image, which contains all the features we are interested. So, this is how we extract LBP features from a picture and we have observed that the output is same as that of HAAR. Yet there are differences which can be seen in the next section.

3.3. Comparisons

Each Open CV face detection classifier has its pros and cons, but the major differences are in accuracy and speed. So, in case more accurate detections are required, HAAR classifier is the way to go.But the LBP classifier is faster, therefore, should be used in mobile applications or embedded systems.

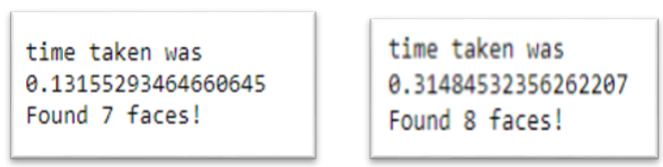

Figure 12: The time taken by HAAR (11a) and LBP (11b) classifiers

The few differences between HAAR and LBP classifier which are as follows:-

Pros-HAAR	Cons-HAAR
High accuracy	Computationally complex and slow
Pros-LBP	Cons-LBP
Computationally simple and fast	Low accuracy

From the comparison of three methods, we can see that skin segmentation could be used for accurate results in case of single face detection or fewer face detection. But Haar and LBP are better methods for face detection as they are more accurate.

4. Face Recognition

Open Cv has three built in recognizers which are as follows.

- **Eigen faces:**
  This algorithm considers the fact that not all parts of a face are equally important or useful for face recognition.In that sense, you are focusing on the areas of maximum change. When you look at multiple faces, you compare them by looking at areas like

eyes,nose,etc., because by catching the maximum variation among faces, they help you differentiate one face from the other.
This is how EigenFaces recognizer works. It looks at all the training images of all the people as a whole and tries to extract the components which are relevant and useful and discards the rest. These important features are called principal components.So, EigenFaces recognizer trains itself by extracting principal components, but it also keeps a record of which ones belong to which person.
However, one thing to note in above image is that EigenFaces algorithm also considers illumination as an important feature. In consequence, lights and shadows are picked up by EigenFaces, which classifies them as representing a face'.

- **Fischer Faces:**
  This algorithm is an improved version of the last one. Since EigenFaces also finds illumination as a useful component, it will find this variation very relevant for face recognition and may discard the features of the other people's faces, considering them less useful.
  FisherFaces face recognizer algorithm extracts principal components that differentiate one person from the others. In that sense, an individual's components do not dominate over the others.
  One thing to note here is that FisherFaces only prevents features of one person from becoming dominant, but it still considers illumination changes as a useful feature. We know that light variation is not a useful feature to extract as it is not part of the actual face.This is when we go to the next method

- **LBP:** This is the method that we will be using in this project for face recognition.
  This process works as follows:-
  Take a 33 window and move it across one image. At each move (each local part of the picture), compare the pixel at the center, with its surrounding pixels. Denote the neighbors with intensity value less than or equal to the center pixel by 1 and the rest by 0.
  After you read these 0/1 values under the 33 window in a clockwise order, you will have a binary pattern like 11100011 that is local to a particular area of the picture. When you finish doing this on the whole image, you will have a list of local binary patterns.
  Now, after obtaining a list of local binary patterns, convert each one into a decimal number using binary to decimal conversion and make a histogram of all of those decimal values. In the end, there will be one histogram for each face in the training data set .

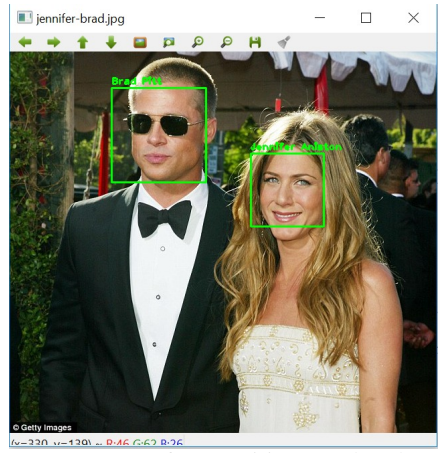

Figure 13: Output of recognition Brad and Jennifer

In order to recognize the faces using LBP we will implement the following steps:-

- Read the images from URL mentioned in the data file
- Histogram equalization before face detection is applied to improve detection accuracy for images
- Detect face in each image using the Haar-cascade classifier
- For each image, store the faces and corresponding labels
- Train LBP face recognizer by using the above data
- When a new image comes for testing, first detect the face using haar-cascader
- Next predict the faces detected using the trained recognizer and get the label
- Draw a rectangle around the face and display the label above it.

After applying the LBP face recognition algorithm we obtain the output as follows:-

From the figures 16 and 17we can see that if multiple faces are detected then that photo is copied across the folders of the people whose face are detected.

In case the image we send is not among the trained images then the classification is done based on another persons label whose features are similar to the image we have given.An example of this is as follows:-

Also training is done on images with single face as if we take multiple faces for training then there would be confusion in assignment of the label. The thing to be noted is that the larger dataset we take for training the better would be the accuracy.

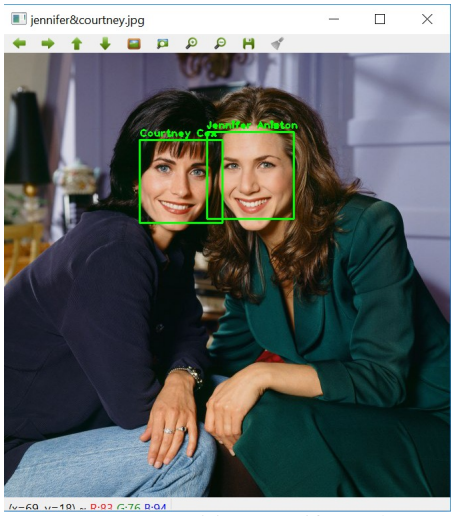

Figure 14: Face recognition Jennifer and Courteney

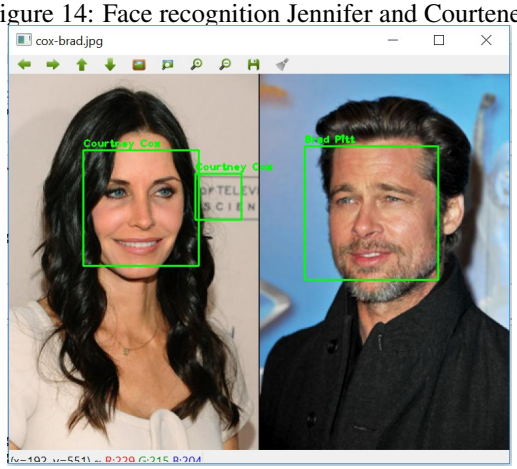

Figure 15: Face recognition Courteney and Brad

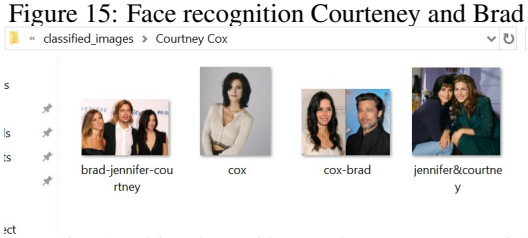

Figure 16: Classification of image in Courteney's folder

Figure 17: Classification of image in Jennifer's folder

5. Classification

Here we take the labels obtained from face recognition and based on them the photos are placed in the respective folders. In case multiple faces are detected in an image then the photo will be copied across all the folders whole labels

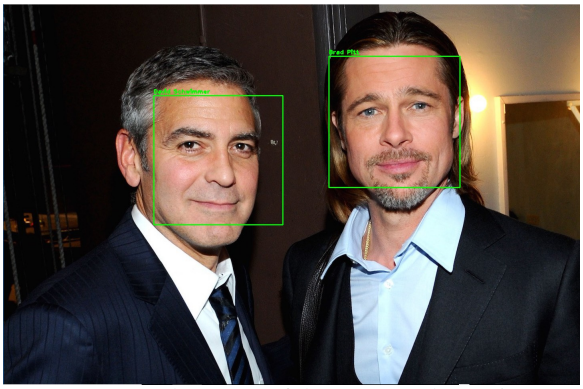

Figure 18: Case where one of the person whose picture was not in training but in testing

have been identified.For example if we have identified an image with Courtney Cox and Jennifer Aniston in it then that photo will be copied in two folders one contatining the images of Jennifer Aniston and the other containing the images of Courtney Cox

6. Conclusion

Through this project we have studied the different techniques of face detection and observed their pros and cons, along with the various face recognition techniques of which we have used the LBP classifier for face recognition and then classified the recognized images.This project can be further analysed in future with the help of convolutional neural network.

References

[1] Waqar Mohsin,Noman Ahmed,Chung-Tse Mar *Face Detection Project* May 26,2003

[2] Human Face Detection by using Skin Color Segmentation, Face Features and Regions Properties `https://pdfs.semanticscholar.org/4d2b/d65a4b24679d4c440afc6fe88f13229e644d.pdf`

[3] Face recognition using color local binary pattern from mutually independent color channels `https://arxiv.org/ftp/arxiv/papers/1501/1501.00105.pdf`

[4] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen *Face Recognition with Local Binary Patterns*

[5] Face Scrub-vintage http://vintage.winklerbros.net/facescrub.html

[6] H.-W. Ng, S. Winkler *"A Data-driven approach to cleaning large face datasets."* In Proc. IEEE International Conference on Image Processing (ICIP), Paris, France, Oct. 27-30, 2014.