

Face Detection and Recognition for Digital Forensics

Lidiya Sebastian

Dept. of Computer Science and Engineering
Amal Jyothi College of Engineering Kottayam,
Kerala, India

Merin M Jose

Dept. of Computer Science and Engineering
Amal Jyothi College of Engineering Kottayam,
Kerala, India

Rose Maria Jose

Dept. of Computer Science and Engineering
Amal Jyothi College of Engineering Kottayam,
Kerala, India

Sheba Ann Roy

Dept. of Computer Science and Engineering
Amal Jyothi College of Engineering Kottayam,
Kerala, India

Neenu R

Dept. of Computer Science and Engineering
Amal Jyothi College of Engineering Kottayam,
Kerala, India

Abstract—Surveillance methods are becoming widely popular in many organizations, including law enforcement, traffic control and residential applications. A reliable mechanism to record and track human activities is through analyzing images. In particular, investigations are performed by searching for specific people in pictures. As the numbers of such images are increasing, manual examination of all frames becomes cumbersome. Some degree of computerization is strongly recommended. This paper discusses the possibility of developing a new software to achieve this. On analyzing the potential of various programming languages for image processing, Python has been opted for development as it provides a range of image manipulation libraries. This way, software is created where photos can be imported from different locations in the system and clustered according to the faces in these images. The application designed is helpful during investigation processes and helps in validating claims and evidence.

Keywords—Face recognition; HOG; Feature extraction

I. INTRODUCTION

With the growth of computer vision into an advanced and promising branch of computer science, face detection and recognition has upstaged from an obscure to a well explored area. Face recognition is now one of the best and successful applications of image analysis and related applications. Because of growing interest in the area, researches are going on in this field and any advancement can hopefully aid other related applications. These face detection and recognition techniques can be used in image analysis and are taken advantage of to make the crime investigation process easier. A face recognition problem (in computer vision) can be generalized as follows: given photos of a scene, check whether the persons of interest were in it with the help of some database of faces. Facial recognition generally involves two stages: Face Detection in which a photo is searched for a face, then the image is cropped to extract the person's face for making recognition easier. Face Recognition where that detected and processed

face is compared to a database of faces that are known, to decide who that person is. The application that is developed uses OpenCV, an open source computer vision and machine learning software library for the recognition and related tasks. It does allows the user to choose the search image, an unknown folder and another folder which can be used to store the result. Once the known image is compared with the unknown images the matching images are placed into a result folder selected by the user.

II. METHODOLOGY

The common approach to implement clustering is using a visual pattern recognition pipeline that involves extracting the images followed by recognizing and clustering all images of a particular person.

First, the image is captured by the system, and then its various features are extracted. It is then compared with other unknown images and matching images are identified.

Once the required fields (known image, unknown folder and the result folder) are selected by the user, the application compares this selected search image with each image in the unknown folder, to find the matching images. The images are compared based on the features extracted by Histogram of Oriented Gradients. A copy of the matching images is then placed in the result folder specified by the user. This algorithm is used to detect faces in other images. It also has the ability to detect other body parts and can be trained to identify almost any object in an image as the method is based on identifying the edges.

The face detection stage is about identifying key parts of a face that will be similar for any photo of the same person and different for an image of anyone else. Histogram of oriented gradients, a feature descriptor used to extract the features of images and is used here.

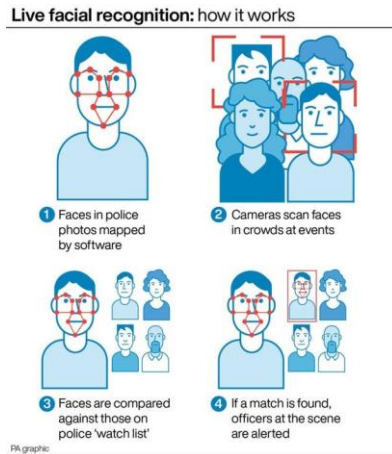


Figure: 1

HOG works through the steps of preprocessing, gradient calculation, histogram calculation, block normalization and calculation of HOG feature vector. In the preprocessing step, the faces are cropped out from the images. After this each of these cropped out faces are evaluated to calculate the gradient. The magnitude and angle direction-orientation- of gradients are calculated using the standard way of calculating gradients. These gradient values are then divided into 9 bins of bins size 20. For each of these 8 x 8 cells are created. After binning we can build a histogram based on it. The image and the faces might be containing certain imperfections due to improper lighting. Thus we need a step of normalization to make the images independent of lighting values. We use the L2 block normalization for this. In L2 normalization divides each element with RGB color vector length to perform the normalization. The 7 horizontal and 15 vertical positions will make up 105 positions of 16 x 16. Then we can calculate the HOG feature vector from the 36 x 1 block representation of each of the 16 x 16 blocks which gives out a giant feature vector.

The next step after this is taking some features that are most relevant. We use the 5-point predictor classifier in which feature descriptions of 5 points from the face are taken as the relevant ones. These points are two points each for the two eyes and one point for the nose. It is based on these features that the faces are compared for similarity.

An image can be considered as a function of spatial coordinates and intensity values of the form $f(x, y)$. The gradient of such an image is a vector of the form (f_x, f_y) and it represents a transition in the intensity values in the image. Such a transition occurs at edges in an image.

Histogram of Oriented Gradients will calculate the gradient image which means it is based on the edges.

As it is based on gradients and edges it can extract the faces from images precisely based on the features of interest. The search image and all other images are encoded based on this feature vector and these encodings are placed in an array.

A variable can be initiated which can contain the path that need to be checked, say `ipath`. These along with the known encoded images are passed as arguments to a

function, say `img_proc`. This function scans all the sub folders within the unknown folder and checks all the files. If the file is an image file, then it is compared with the encoding of the known image. A tolerance factor is added to get more accurate results. The smaller the tolerance value, the more accurate the results are. As soon as a similar face is detected, a copy of it is placed in the result folder.

III. PROPOSED SYSTEM

The application takes advantage of the image processing capability of python. user friendly GUI is provided where the user can import photos from different locations. Both the image to be searched as well as the folder on which we need to find the image should be manually given by the user. The result folder also can be specified by the user. The general flow of the system according to the approach as shown in figure 2.

As soon as the search image, the search folder and the result folder are submitted by the user the analysis process is formed. The search image undergoes detection phase and encoding of the face thus detected is generated. Meanwhile the application traverses through all the subdirectories of the search folder selected by the user, generating encodings of each of the image file. These encodings are compared with that of the search image using the Euclidean distance. A suitable tolerance parameter is used to improve the accuracy.

IV. RELATED WORKS

This section considers the various works that has previously been done in face clustering area. Over the last few years, face detection, recognition and clustering applications have gained immense acclamation due to its several applications. Many research work has been done in this area in order to build an efficient face clustering system. The aim is to select an algorithm that will increase the accuracy and speed up the Face recognition and clustering process.

G. M. Zafaruddin and H. S. Fadewar have examined various comprehensive approaches in face recognition. Their study summed up approaches like spatial matching detector method, neural networks method, eigenface based method and fuzzy theory based method for face recognition.[1]

R. Ranjan, S. Sankaranarayanan, C. D. Castillo and R. Chellappa presented a multi-task CNN-based system for concurrent detection, alignment, pose estimation, smile and gender categorization, age estimation and recognition. The research used MTL framework for training their algorithm. [2]

R. Feraud, O. Bernier, J. E. Viallet and M. Collobert proposed a search algorithm; the Constrained Generative Model to reduce processing time. The approach begins with simple processes, using the standard image processing ways and then goes to more sophisticated processes that involve statistical analysis. [3]

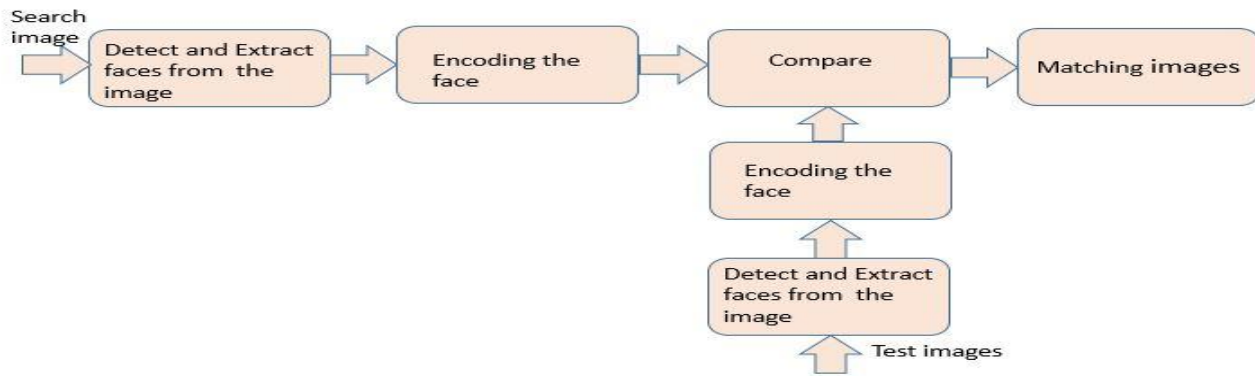


Figure: 2

Mrs. Madhuram.M, B. Prithvi Kumar, Lakshman Sridhar, Nishanth Prem, Venkatesh Prasad designed a real time face recognition method based on IP camera and employing image set algorithm by applying Python programming development and OpenCV. The paper is about to detect and recognise human faces. The dataset which contains the images of the person has been given and previously trained .Haar cascade algorithm was used for the purpose of detection. [4]

Bijl, Erik have described several clustering algorithms and applied them to various datasets. They mainly concluded that threshold clustering showed the best performance in terms of the f-measure and number of false positives. DBSCAN had also shown good performance during their experiments and is also considered a good algorithm for face clustering. They discouraged to use k-means and for big datasets Approximate Rank-Order. [5]

V. RESULTS AND DISCUSSIONS

On deploying the proposed model with various datasets, it seems to work well. Our classifier HOG will calculate the gradient of an image. The gradient is a vector of the form (fx,fy) and it represents a transition in the intensity values in the image- which means it is based on the edges. The advantage of using edges for the image processing is that edges are tolerant to variations that might be due to shakes or other disturbances. Evaluating on the basis of accuracy, it shows an acceptable value which is greater than 75 percent most of the time. Taking it to the digital forensic toolkit can be beneficial for the crime investigators. It shows good results even with a large amount of data. Also, when it comes to folders having a number of sub folders, the traversal time is a increased just a little. This can be reduced by using certain optimization techniques in the search.

VI. CONCLUSION

In this paper we propose an image processing system that can support the crime analysis process greatly. The face

recognition and clustering process performed using Histogram of oriented gradients and 5 point predictor classifier performs

well except for those images in which the faces are poorly illuminated or partially visible. These issues can be addressed in the future by employing various alignment correctors. Even though intended primarily to aid the inquiry process through slight changes it can be extended to meet other similar applications.

For improving the performance small features can be improved and used. In future, the advancements in technology can add more advanced features to the model and to broaden its application range.

ACKNOWLEDGMENT

First and foremost we thank God Almighty for showering his blessings upon us during this endeavor. We take this opportunity to express our sincere gratitude to the Dept. of Computer Science and Engineering, Amal Jyothi College of Engineering, Kanjirappally for providing us the opportunity to take up this venture. We owe a great depth of gratitude towards our Head of the Department Prof. Manoj T Joy for his support during the course of this project work. Next we like to offer our genuine thanks to all our teachers for their guidance and useful suggestions throughout the project. We also express our gratitude towards the supporting staff for their technical assistance. This project would not have been a great success without the assistance given by our loved ones in all the research work accomplished for the project and the enormous encouragement given to us.

REFERENCES

- [1] Ghazi Mohammed Zafaruddin, H. S. Fadewar: "Face Recognition: A holistic approach review", IEEE, International Conference on Contemporary Computing and Informatics, pp. 175-178, November 2014.
- [2] Ranjan R, Sankaranarayanan S, Castillo CD, Chellappa R: "An all-in-one convolutional neural network for face analysis." Proceedings of the 12th IEEE International Conference on Automatic Face Gesture Recognition Gesture Recognition 2017.
- [3] J. Feraud, O. Bernier, and M. Collobert: "A fast and accurate face detector for indexation of face images," In Proc. Fourth IEEE Int. Conf on Automatic Face and Gesture Recognition, Grenoble, 2000
- [4] Mrs. Madhuram.M, B. Prithvi Kumar, Lakshman Sridhar, Nishanth Prem, Venkatesh Prasad "Face Detection and Recognition Using OpenCV" Volume: 05 Issue: 10 , Oct 2018

- [5] Bijl and Erik, `` A comparison of clustering algorithms for face clustering."Bachelor's Thesis, Computing Science, 2018.
- [6] Isra's Abdul-Ameer Abdul-Jabbar,Jieqing Tan and Zhengfeng Hou, `` Face Recognition Enhancement Based on Image File Formats and Wavelet De-noising."
- [7] Varun Garg,Kritika Garg Information Technology Department Ch. Brahm Prakash Government Engineering College , `` Face Recognition Using Haar Cascade Classifier
- [8] M.H Ghaseminezhad and A. Karami , `` A novel self-organizing map(SOM) neural network for discrete groups of data clustering.",Applied Soft Computing Journal,no. 4,2011.
- [9] Dinalankara, Lahiru (2017), `` Face Detection \& Face Recognition Using Open Computer Vision Classifies."
- [10] Teddy Mantoro,Media A.Ayu,Suhendi, ``Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods"
- [11] Maliha Khan,Sudeshna Chakraborty,Rani Astya,Shaveta Khepra, ``Face Detection and Recognition Using OpenCV"