TEST Engineering & Management

May-June 2020
ISSN: 0193-4120 Page No. 5256-5260

Automated Event Monitoring System using Facial Recognition

Mayank Bajaj [1] , Madhav A[2] , Mohammad KC[3] , Ashwinth Singh P[4] , A Ajil[5]

[1,2,3,4,5] School of Computing and Information Technology, Reva University

Article Info
Volume 83
Page Number: 5256-5260
Publication Issue:
May-June 2020

Article History
Article Received:*19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *16 May 2020*

Abstract
In the current day and age, while organizing any event, one of the biggest concerns is that of security. Being able to monitor a large event requires a huge amount of security personnel and resources and due to human involvement, there will be cases of ignorance in security. In any event, there will be human many security personnel whose only job is to keep looking at the live security feed to avoid the entry of unauthorized individual. Our model demonstrates a way to minimize human involvement. By equipping all the cameras with facial recognition algorithm, we can solve this problem easily. As soon as system finds an unauthorized person inside the premise, it will inform the administrator who then will be able to take necessary actions in premise. This way, we can minimize the security mishappens due to human negligence and since a computer doesn't require to take break, this entire process can keep running without interruption for as long as administrator wants.

Keywords: *Facial Recognition, Embedding, Encodings, Live video feed, Imaging Hardware, Cameras, Face detection.*

1. Introduction

What is the easiest and most efficient way to find a person without the need of any advanced hardware? We can solve this problem by employing facial recognition. You need to involve an individual in order to scan any other biometric like fingerprint but facial recognition can be done with just security cameras and if given a big enough dataset then it can be as accurate as any other biometric.

Automated event monitoring is a system that employs facial recognition to find people with their faces in all the available cameras. It can be used to search for just one person or for multiple people in all the cameras. This project can be used as a concept to implement a similar system but on a much larger scale. China currently has more than 200 million cameras all over the country and all of the camera are designed to recognize faces.

All the current implementations of these idea are based on the Local Binary patterns [2] which works by labelling each pixel in an image and thresholding it each pixel and around it and finally considers result in the form of a binary number. This technique was developed in 1996 and is heavily hardware dependent i.e. the chances of the results being accurate depends on the imaging capabilities of the camera which makes it unsuitable for large scale implementations.

he system proposed in this paper uses a far more accurate method to recognize faces which is based on the FaceNet [1]. FaceNet proposes a system of recognizing faces by extracting 128 measurements from each face or encoding the face. These measurements are called an embedding. Our Facial recognition algorithm works by combing multiple embeddings of a single individual in various conditions. By merging all the embeddings we create an encoding that would extremely resembles the actual measurements and give us a significantly more accurate result than using the Local Binary Patterns [5].

One of the major advantages of this approach is the dependency on the hardware is significantly reduced as we are not using each individual pixel, rather we are extracting measurements from face of an individual which is possible even on a subpar imaging hardware.

2. Objective

Facial recognition has the ability to help detect threats more efficiently than relying on standard video surveillance and security personnel. Facial recognition will be able to determine the person attending event ,attendees as well as cross-referencing faces with criminal databases, looking for signs of potential dangers to your event.

Figure 1: Expected Working of our sytem.

Figure 1 shows an example of the system recognizing a person. This person was recognized by capturing his live feed and running his profile with existing database which had been inputted with his profile.

3. Literature Survey

One of the most famous paper that changed the landscape of face recognition is FaceNet: A Unified Embedding for Face Recognition and Clustering by Florian Schroff, Dmitry Kalenichenko, James Philbin published in 2015. In this paper they 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 FaceNetembeddings as feature vectors. [1]

The very first paper published on using local binary patterns as a facial descriptor was published in 2006 by Ahonen, T., Hadid, A. and Pietikäinen, M. Using Local Binary Pattern for a facial descriptor is a very simple and easy way of creating a description of a face. It works by labelling each pixel and thresholding each pixel around it to give output in the form of a binary number. This method relies heavily on the capability of the imaging hardware as the descriptor is completely based on the value of each pixel. [2]

Histograms of oriented gradients for human detection by N. Dalal and B. Triggs published in 2005 uses one of the most widely known method for face detection. It works by creating a gradient map of the image and then cross referencing the obtained image to the database that has thousands of gradient maps of faces. The reason for using this is because the accuracy of this method outclasses any other method for Face detection. [3]

4. Methodology

The proposed Event monitoring using facial recognition system has different process:

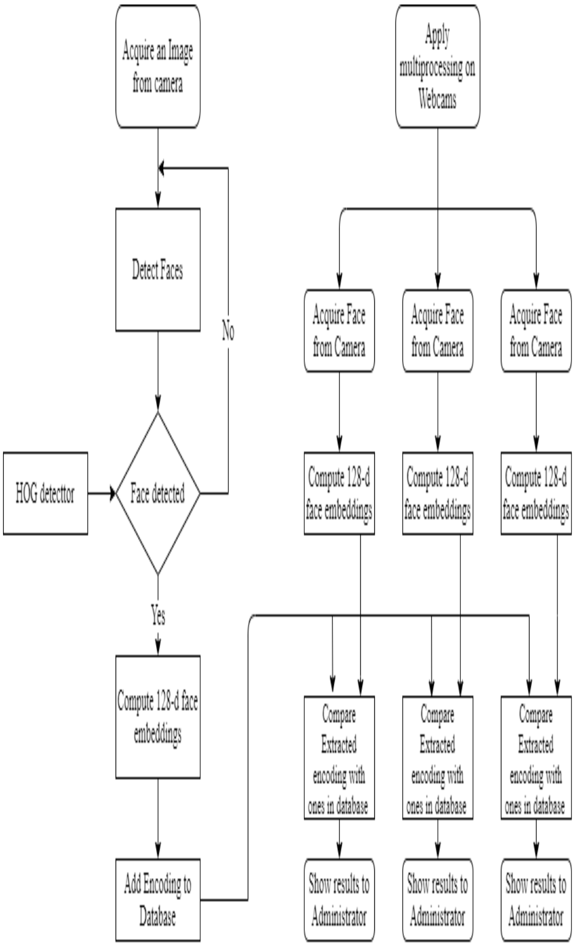
Acquire an Image from camera
Apply multiprocessing on Webcams
Detect Faces
No
Acquire Face from Camera
Acquire Face from Camera
Acquire Face from Camera
HOG detecttor
Face detected
Compute 128-d face embeddings
Compute 128-d face embeddings
Compute 128-d face embeddings
Yes
Compute 128-d face embeddings
Compare Extracted encoding with ones in database
Compare Extracted encoding with ones in database
Compare Extracted encoding with ones in database
Add Encoding to Database
Show results to Administrator
Show results to Administrator
Show results to Administrator

Figure 2: A flowchart which depicts the entire process of our System.

The first part of the chart i.e. the left side depicts the **Training** part where we create a database of each person whose faces we'll try to find in the live video feed.

It shows the creation of the database which has face embeddings of each person who is registered in the system.

The second part of the flowchart i.e. right side shows how we'll use multiprocessing on multiple video capturing devices to get frames from each device and processing each frame for face detection and face recognition which is done on the frames of all capturing devices at the same time.

The output is then displayed on the video feed of each device by editing each frame and adding a bounding box around each face with separate labels for faces that are recognized and those that are not recognized.

Training

Training is the process where the given dataset is used to create the embedded face vector of each face in the given

Multiple images of and individual is provided to the model. The faces are detected from each image and then the extracted face is converted to Face encoding.

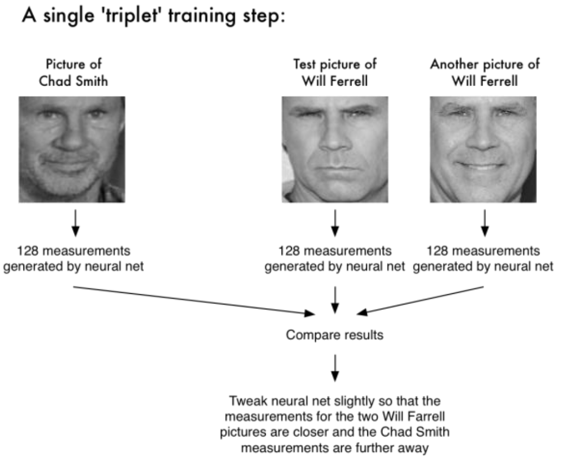
A single 'triplet' training step:
Picture of Chad Smith
Test picture of Will Ferrell
Another picture of Will Ferrell
128 measurements generated by neural net
128 measurements generated by neural net
128 measurements generated by neural net
Compare results
Tweak neural net slightly so that the measurements for the two Will Farrell pictures are closer and the Chad Smith measurements are further away

Figure 3: Creation of face encodings from an extracted face.

Conversion of face into its encodings is completed by training it to get 128 measurements for every face.

After repeating this step many times for many images of thousands of various people, the neural network learns to reliably generate 128 measurements for every person. Any ten different pictures of same individual should give roughly an equivalent measurement.

Machine learning people call the 128 measurements of every face an embedding. the thought of reducing complicated data sort of a picture into an inventory of computer-generated numbers comes up tons in machine learning (especially in language translation).

Capturing Live Feed

We use the function present in the OpenCV library to get the live video feed from all the available cameras. Each camera will have a unique camera ID and location will be stored in respective line in the text file.

We use the dlib and face_recognition python library to implement the next two phases of the implementation. dlib library contains all the resources for face detection and facerecognition. face_recognition library acts as a nice wrapper to make the coding part easier.

Real Time Face Detection

Once the live feed starts being captured, it will check each frame for a face and will try to detect using the HOG - face detector (Histogram Oriented Gradient + linear SVM classifiers).

To detect faces in an image, we'll first change image to grayscale. HOG detector works by comparing how dark a pixel is compared to pixels surrounding it and replacing the pixel with an arrow if pixel gets darker, this process is repeated for every pixel, replacing each pixel with a gradient. HOG detector only requires an essential flow of light/darkness which is achieved by resizing the image into small squares of 16x16 pixels each. In each square, count up what percentage gradients point in each major direction (how many point up, point up-right, point right, etc…) and replaces that square within the image with the arrow directions that were the strongest. The end result is we turn the original image into a very simple representation that captures the basic structure of a face. Detecting faces in it is done by finding the part of our image that looks the most similar to a known HOG pattern that already come in the library.

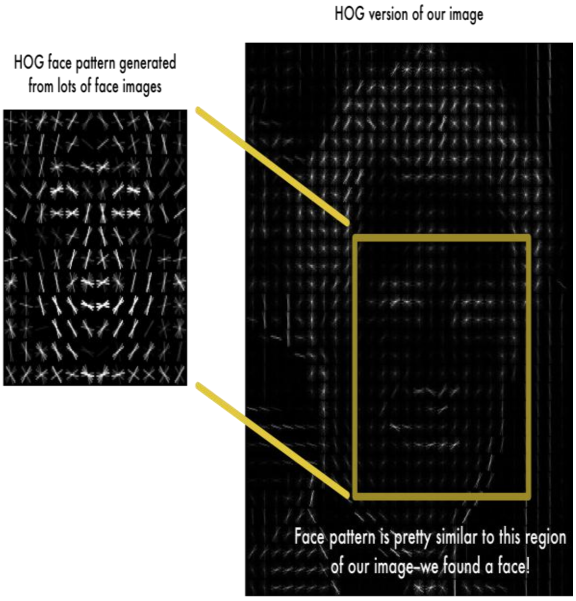
HOG version of our image
HOG face pattern generated from lots of face images
Face pattern is pretty similar to this region of our image-we found a face!

Figure 4: Gradient map created by HOG detector

The above figure shows the gradient map created by HOG detector and how a face is extracted from it.

A. Real Time Face Recognition

Just like in the training step, the face extracted from the previous step is converted to its embedding. This embedding is then compared to all the embeddings that are already stored in the database and the output is shown in the live video feed to the administrator.

5. Results

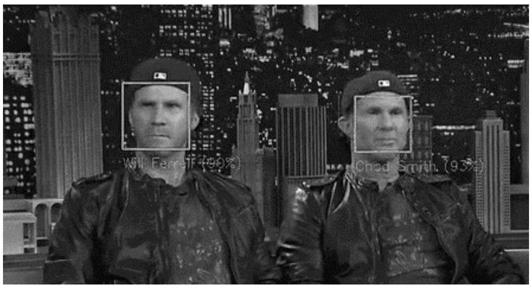

Figure 5: Working of our Face Recognizer

The above image depicts the working of our facial recognition algorithm where it successfully recognizes the faces of Will Ferrel and Chad Smith with an accuracy of over 90%.

This is really remarkable because most of the recognizer (like LBHP face recognizer that comes with OpenCV package and uses Local Binary Pattern [2]) fail to differentiate between the faces of Will Ferrel and Chad Smith. Compared to them, our algorithm was successful in recognizing and differentiating them.

6. Applications

Monitoring

Our system is based on the highly advanced method for facial recognition, entirely different from the one that come in OpenCV. The monitoring capabilities of such a algorithm is really high. Monitoring tasks become really secure using this,which may include:

- The application can be implemented to train the dataset everyone that needs to be present in a meeting or seminar and that dataset can be used to tell how many people are present.
- Monitoring students during exams in large examination centers.
- Monitoring employees in office.
- Unlocking a device using facial verification.

Multiple Searches simultaneously

If you have a database of multiple people then all of those people can be searched in all of the cameras greatly reducing the time by preventing the need to run it for each person one by one.

Advanced Surveillance System

China currently has more than 200 million cameras all over the country and all of those are equipped to perform facial recognition. This is monitoring on national scale. The entire country being monitored at all times greatly reduces any chances of crimes going unnoticed. Even more so, it makes the job of authorities easier as it will be really easy to locate a criminal offender or even an entire gang.

7. Future Enhancements

As previously mentioned, the purpose of this project was to develop a facial recognition system that can used to monitor a crowd in the event, but for future developments this can be implemented in a large scale and be used in traffic and catching criminals.

8. Conclusion

The paper aims to solve one of the most challenging situations in the modern era. Surveillance is a challenging task for any event and even more so when it’s almost completely based on the human involvement. The proposed system can identify participants more accurately than even a human can in some situations. It aims to minimize the danger and any mishappens during an event by informing the administrator upon the entry of an unauthorized individual inside the premise so necessary actions can be taken immediately.

The system is only limited by the hardware being provided to it. If given the right amount of hardware, it can be implemented on a nation wide scale that many countries hope to achieve as a very effective and robust way of surveillance.

Acknowledgement

Any given task achieved is never the result of efforts of a single person. There are always a bunch of people who play an instrumental role in leading a task to its completion. Our joy at having successfully finished our mini project would be incomplete without thanking everyone who helped us along the way. we would like to express our sense of gratitude to our REVA university for providing us the means of attaining our cherished goal.

We would like to thank our Hon’ble Chancellor, Dr. P. Shyama Raju, Hon’ble Vice-Chancellor, Dr. S. Y. Kulkarni for their immense support towards students to showcase innovative ideas.

We cannot express enough thanks to our respected Director, Dr. Sunilkumar S. Manvi for providingus with a highly conductive environment and encouraging the growth and creativity of each and every student. we would like to offer our sincere gratitude to our Mini project Coordinators, Prof. Kiran m, Prof. Sailaja Thota,

Prof. Ashok k. Patil for the numerous learning opportunities that have been provided. We would like totake this opportunity to express our gratitude to our Mini Project Guide, Prof. AjilAylwin for continuously supporting and guiding us in our every endeavor as well for taking a keen and active interest in the progress of every phase of our mini project. Thank you for providing us with the necessary inputs and suggestions for advancing with our mini project work. We deeply appreciate the wise guidance that mam has provided.

References

[1] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 815-823.

[2] T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, pp. 886-893 vol. 1.

[4] Adam Geitgey. *"Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning".* 2016. https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78