International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-7, Issue-6, March 2019

CNN Integrated With HOG For Efficient Face Recognition

R.Angeline, Kavithvajen.K, Toshita Balaji, Malavika Saji, Sushmitha.S.R

Abstract**:** ***Human faces in the video are subject to illumination variation, out-of-focus blur and pose variations during face recognition process in various applications. The proposed system aims to eradicate the problems mentioned above. This is done by utilizing Histogram of Oriented Gradients algorithm as a feature descriptor to detect faces. The training data is composed of still images and blurred images. For the system to learn pose variations, an additional dataset of artificially aligned images is fed by using Face landmark estimations algorithm. Convolutional Neural network is trained, and effective face recognition is obtained. Thus, can make surveillance applications work efficiently.***

Index Terms**:** ***Convolutional neural network, Face recognition, Histogram of oriented gradients, Support Vector machine.***

I. INTRODUCTION

A face recognition technology basically involves face detection, face alignment, feature extraction, and identification. It can be used in a variety of applications such as security systems, video surveillance, human-computer interaction and many more. A wide range of algorithms has been developed in the last few years for detection and recognition purpose. Most of the face recognition algorithms recognize or identify by extracting features of a human face from a digital image or video frame which is then used to find other pictures with identical characteristics. For example, eyes, lips, nose, etc. In different algorithms, images of faces to be trained are stored in a database and compared to the input facial image. Although face recognition technology has gained immense popularity due to its numerous applications and advantages in various fields, it remains one of the most challenging tasks to perform.

In video-based face recognition, a number of issues take place such as illumination variation, out- of-focus-blur, pose variations. There is a need to make recognition reliable under an uncontrolled environment where the images captured are likely to reduce quality and other useful details due to extreme lighting conditions such as underexposure or overexposure which are essential for the face recognition algorithm to function properly. This results in the appearance of the same image differently.

Therefore, the system misclassifies the input image while comparing it with similar models in the dataset. Pose variation occurs due to a change in viewing conditions. Just like illumination variation, pose variation also causes the same image to appear differently. Hence, making face recognition difficult.

Detection is an integral part of a face recognition system. There is a number of algorithms that have been developed for face detection. Some of them are histogram of oriented gradients, Local binary pattern histogram, eigenfaces, Fisher faces. In the histogram of oriented gradients (HOG), every pixel of the captured image is compared with other pixels in darkness. Based on that, an accurate image is obtained. It is unaffected by any kind of lighting problems. Hence, it's a less complicated and more efficient technique for face detection.

Revised Manuscript Received on March 25, 2019.

R.Angeline, Department of Computer Science, SRM Institute of Science and Technology, Ramapuram, Chennai India.

Kavithvajen.K, Department of Computer Science, SRM Institute of Science and Technology, Ramapuram, Chennai India.

Toshita Balaji, Department of Computer Science, SRM Institute of Science and Technology, Ramapuram, Chennai India.

Malavika Saji, Department of Computer Science, SRM Institute of Science and Technology, Ramapuram, Chennai India

Sushmitha.S.R, Department of Computer Science, SRM Institute of Science and Technology, Ramapuram, Chennai India

II. RELATED WORK

A. HOG

Histogram of Oriented Gradients (HOG) is widely used as a feature descriptor for detecting objects or human face in computer vision and image processing. The object search is based on the detection technique applied for the small images defined by sliding detector window that probes region by region of the original input image and its scaled versions. HOG detection is to divide the source image into blocks (for example 16×16 pixels). Each block is divided by small regions, called cells (for example 8×8 pixels). Blocks are usually overlapped which may result in several blocks consisting of the same cell. Vertical and horizontal gradients are obtained for every pixel present in the cell. Fig 1 represents the simple process of the algorithm. The simplest method to do that is to use vertical and horizontal operators [17]:

$$G_x(y, x) = Y(y,x+1) - Y(y,x-1); G_y(y,x) = Y(y+1,x) - Y(y-1,x) \quad (1)$$

$Y(y,x)$ - pixel intensity at coordinates x and y
$G_x(y,x)$ - horizontal gradient
$G_y(y,x)$ - vertical gradient.

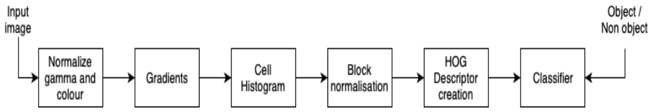

Fig 1: The sequence of Object detection using HOG

Magnitude and phase of the gradient are determined as:

$$G(y,x) = \sqrt{Gx(y,x)2 + Gy(y,x)2},$$

$$\theta(y,x) = \arctan(Gx(y,x)) \quad (2)$$

Histogram of Oriented Gradients is generally used along with Support Vector Machine (SVM) classifiers. In order to determine if the object was found or not, every descriptor of HOG is computed and then fed to an SVM classifier.

A. CNN

Convolutional Neural Network (CNN) basically refers to an artificial neural network used to perform tasks like image recognition or video recognition. The name convolutional neural network comes from a linear mathematical operation that takes place between matrixes known as convolution. In traditional algorithms like viola jones, specific filters called Haar cascades are required for every part of the face such as eyes, nose, etc. The significant advantage of CNN over other algorithms is that there is no need for a feature extraction algorithm. The network is trained to learn the filters, and many neurons use the same filter. CNN consist of multiple layers; like convolutional layer, non-linearity layer, pooling layer, and fully-connected layer. The CNN first identifies primary characteristics, then learn to recognize these features, and then to enhance its learning capabilities, combines these features with learning complicated patterns. The several levels of elements are obtained from multiple layers of the network. Every level has a particular count of neurons and constructed in three dimensions: height, width, depth.

As described in [17], Convolution layer takes a volume of size [W1×H1×D1], the output of the neurons is computed by applying the product of weights and the localized region. The resultant volume [W2×H2×D2] is known as Convolution maps. Convolution maps generate an output [W2×H2×D2] which is the volume and the equations are as follows (3), (4), (5):

$$W2 = (W1 – F + 2 + P)/S + 1 \quad (3)$$
$$H2 = (H1 – F + 2 + P)/S + 1 \quad (4)$$
$$D2 = D1 \quad (5)$$

Where, F is the dimensional extent of the filter, K equals number of filters, P represents zero padding and S equals stride. Zero Padding is a hyper-parameter that regulates the resultant volume. Stride is a hyper-parameter that controls shifting of filters.

Rectified linear unit layer is where the activation function max (0, x) is utilized.

POOL layer is applied in between successive Convolution layers, the spatial dimension's width and height is calibrated by coupling down sampling operation. MAX operation is employed to optimize the spatial size of the

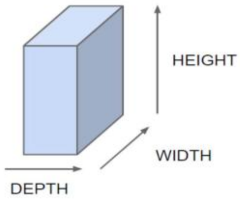

Fig 2: Three-Dimensional input representation of CNN.

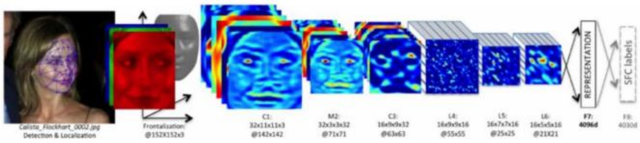
Fig 3: Visualizing Convolutional Deep Neural Network Layers

representation as well as reducing the number of parameters. Pool Layer generates a volume [W2×H2×D2] where W2, H2, D2 are given by applying equations (6), (7) and (8):

$$W2 = (W1 – F) /S + 1 \quad (6)$$
$$H2 = (H1 – F)/S + 1 \quad (7)$$
$$D2 = D1 \quad (8)$$

Finally, the result is a unique vector which is produced by a feature extractor vector or convolutional neural network code.

III. PROPOSED METHOD

The proposed method is classified into four modules and in the same order.

A. *Face Detection*

The first step of a face recognition system is to detect a human face. The algorithm chosen to do so is Histogram of Oriented Gradients (HOG). The following describes the process:

1) The colored image is first converted into black and white.
2) Every pixel of the image is then compared with its neighboring pixels and observed how dark the current pixel is with regard to the rest. This is depicted by replacing every pixel with an arrow that points to the darker pixel. These arrows are called gradients.
3) This process is repeated on every pixel of the image. The result becomes a highly detailed representation of pixels.
4) Hence to reduce the complexity, the images are broken into 16 x 16 squares.
5) The directions of the gradients are observed, and the dominant gradient is applied. Thus, converting the highly complex structure to a simple frame of a face.

This is called This is called a HOG image, which is compared with pre-trained HOG images in order to complete the detection process.

B. *Training Data*

The training data consist of three different data streams – still images of holistic faces, artificially blurred images and

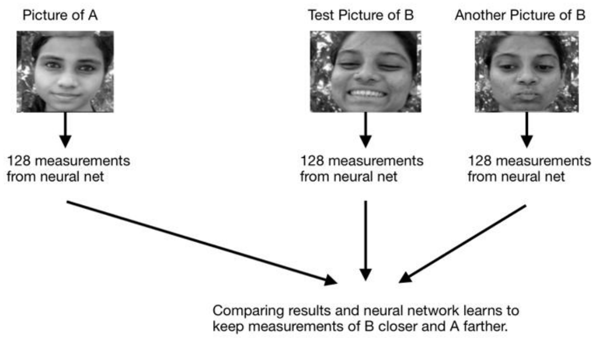

Fig 4: Triplet Training step

aligned images of pose variated faces.

The first stream of data is obtained from the previous step by feeding still images of the entire face. The second stream of data is received by applying an artificial blur to the still images. There are a number of kernels that can be used, but the Gaussian kernel is employed in this setup. [6]

$$f(x,y) = C.\ \exp\left(-\frac{x^2+y^2}{2k^2}\right)$$

The Gaussian kernel is of the above form where k denotes magnitude and C is a constant. The still image is blurred with the above blur kernel by changing the values of k. Hence, applying different blur types randomly to obtain a dataset of blurred images equal to the size of still images. The third stream of data is to identify faces that are turned or are not whole images. To ensure efficient detection, an algorithm called Face Landmark estimations is utilized. Here, specific points or landmarks on the face region is identified like the eyes and mouth. There exist 68 particular locations, but in images where the whole face is not revealed, the algorithm identifies eyes and mouth. These specific regions are then centered by scaling and rotating. This ensures uniformity among the dataset.

C. *Face Recognition*

Training a Convolutional neural network is the most efficient model of recognizing objects and faces too. Here, it is taught to generate 128 measurements for each face. Measurements to describe a few are, the distance between the eyes or nose or cheeks that are obtained to match with pre-trained measurements. The training process involves three inputs:

1) A known image of the face.
2) Another known image of the face.
3) Unknown new image.

The system is trained on thousands of such combinations. Fig 4 diagrammatically explains the comparison process. By doing so, the system learns to differentiate measurements of similar and dissimilar faces and accordingly learns to recognize accurately.

D. *Identification*

The final step is identifying the names of the recognized faces. For this, a simple linear Support Vector Machine (SVM) classifier. The network is trained to take the

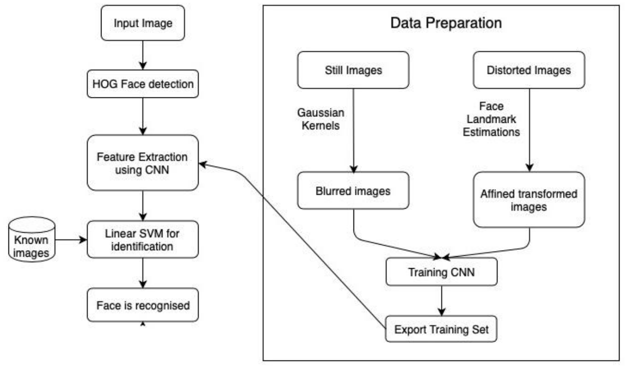

Fig 5: Block Diagram of the proposed model

measurements of the new image as input and display the closest match as output. This involves two steps:

1) Train SVM with known images of faces.
2) Run SVM classifier over new faces or a video for identification.

The total time for training this classifier takes milliseconds which return the name of the recognized face.

IV. SYSTEM DESIGN

Fig 5 represents the architecture of the proposed method. The data involved in training the Neural network is assembled at the initial stage. This is depicted in a separate block in the diagram where still, blurred and distorted images are constructed and rectified to train the network. Then the recognition process starts by detecting faces using the HOG algorithm. This is run on the pre-trained convolutional neural network which extracts features to recognize the detected faces. Lastly, SVM is used to identify the names of the known faces. This face recognition system is built in python. Histogram of Oriented Gradients and Face landmark estimations is implemented by importing functions from the dlib library. This program utilizes trained networks from OpenFace to train the CNN. This generates 128 measurements called embeddings of the face, and this program returns a CSV file containing the same. In the data preparation module, OpenFace scripts work on a folder containing faces images, that creates a folder with cropped and aligned images.

V. EXPERIMENTAL RESULTS

The experimental results show how the proposed system works under different conditions. Fig 6 is under usually lit conditions without any pose variations. Fig 7 shows accurate results even when the subject is not straight and subject to pose variations. Fig 8 includes a presence of an unknown face. Fig 9 was captured when the subject was under motion. Fig 10 is taken under dimly lit conditions with occlusions. These results show that the proposed model works well with pose variations, occlusions, and blurred input.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

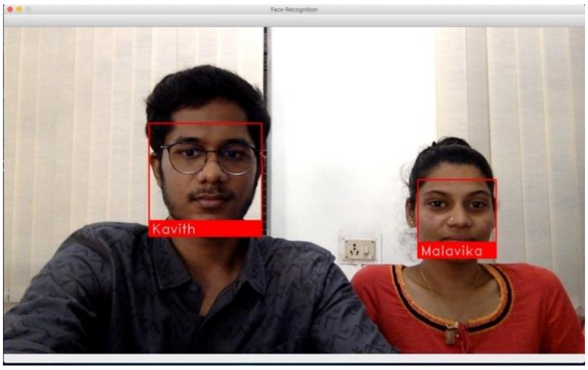

Fig 6: Under normal conditions

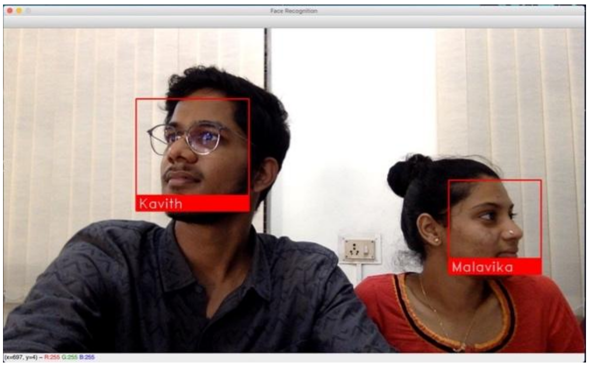

Fig 7: Pose variation

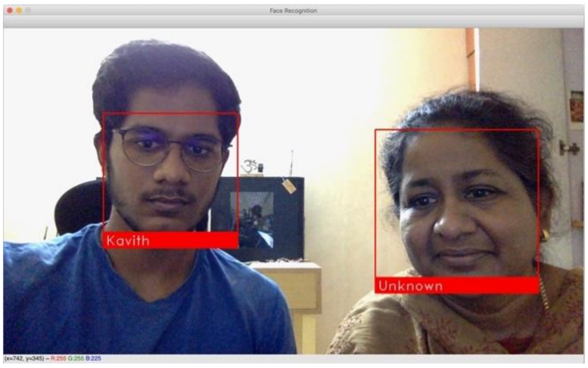

Fig 8: In presence of unknown face

VI. CONCLUSION AND FUTURE SCOPE

The proposed method in this paper combines Histogram of Oriented gradients detection algorithm and train it on CNN with a linear SVM classifier. This is implemented to eradicate motion-blur, occlusions and pose variations that decides the accuracy of recognition systems in today's applications. This was constructed on a self-made database, and the following results show the accuracy.

Moreover, the project is implemented on a simple machine without incurring high costs. Further enhancements can be made like:

- Increasing the training dataset.

-Including low resolution images that will help obtain better results in surveillance applications.

-For the above, different combinations of classifiers can be added to increase performance.

Such complex systems can pave way to a new form of recognition systems that can achieve high levels of precision at lower costs.

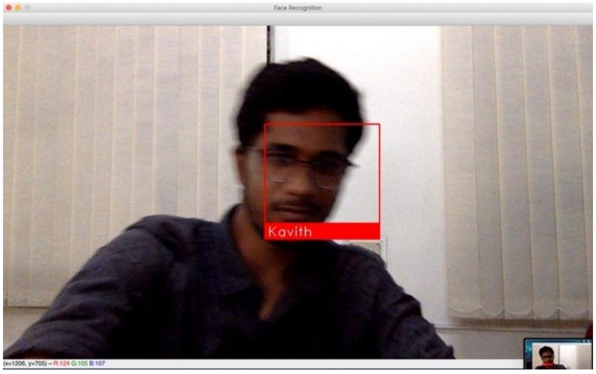

Fig 9: Severe motion blur

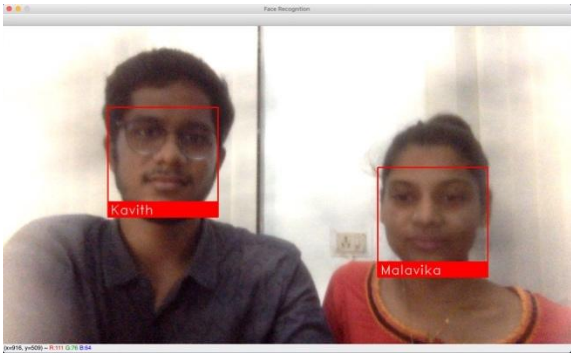

Fig 10: Occlusions

ACKNOWLEDGMENT

The authors would like to thank Project guide Mrs. R.Angeline for her support and timely suggestions.

REFERENCES

1. Vishakha Mehta; Sarika Khandelwal; Ashish Kumar Kumawat, *A Survey on Face Recognition Algorithm*, 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 11-12 May 2018.
2. Zuzana Képešiová; Štefan Kozák, *An Effective Face Detection Algorithm*, 2018 Cybernetics & Informatics (K&I), 31 Jan.-3 Feb 2018.
3. Nicolas Delbiaggio, *A Comparison of Facial Recognition's Algorithms.*
4. Yuqian Zhou ; Ding Liu; Thomas Huang, *Survey of Face Detection on Low-Quality Images*, 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018),15-19 May 2018.
5. Ali Almuhamadi, *Face Recognition System.*
6. Changxing Ding ; Dacheng Tao , *Trunk-Branch Ensemble Convolutional Neural networks for Video-based Face recogniton*, April 1, 2018.
7. V.K.N Kamlesh Pai; Manoi Balrai; Sachinkumar Mogaveera; Deepak Aeloor, *Face Recognition Using Convolutional Neural Networks*, 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 11-12 May 2018.
8. Liu Hui; Song Yu-jie, *Research Based Face Recognition Algorithm Based on Improved Convolutional Neural Network*, 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 31 May-2 June 2018.
9. Anush Ananthakumar, *Efficient Face And Gesture Recognition For Time Sensitive Application*, 2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 8-10 April 2018.
10. Anush Ananthakumar, *Efficient Face And Gesture Recognition For Time Sensitive Application*, 2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 8-10 April 2018.
11. Ekberjan Derman; Albert Ali Salah, *Continuous Real-Time Vehicle Driver Authentication Using Convolutional Neural Network based face recignition*, 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 15-19 May 2018.

12. Subham Mukherjee; Sumalya Saha; Sounak Lahiri; Ayan Das; Ayan Kumar Bhunia; Aishik Konwer; Arin, *Convolutional Neural Network based face detection*, 2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), 28-29 April 2017.
13. Xiao Han; Qingdong Du, *Research on Face Recognition Based on Deep Learning*,, 2018 Sixth International Conference on Digital Information, Networking, and Wireless Communications (DINWC), 25-27 April 2018.
14. Aftab Ahmed Jiandong Guo; Fayaz Ali; Farha Deeba; Awais Ahmed, *LBPH Based Improved Face Recognition at Low Resolution*, 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), 26-28 May 2018.
15. Katarina Knežević; Emilija Mandić; Ranko Petrović; Branka Stojanović, *Blur and Motion Blur Influence on Face Recognition Performance*,2018 14th Symposium on Neural Networks and Applications (NEUREL), 20-21 Nov. 2018.
16. Amritha Purushothaman; Suja Palaniswamy, *Pose and Illumination Invariant Face Recognition for Automation of Door Lock System*, 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 20-21 April 2018.
17. Nadia Jmour; Sehla Zayen; Afef Abdelkrim, "*Convolutional Neural networks for image classification*", 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 22-25 March 2018.

AUTHORS PROFILE

R.Angeline, Assistant Professor with keen interest in Computer Networks and Cloud Computing. Some of the papers published are "A Supervised web-scale forum crawler using URL Type Recognition", "Harmonizing presentation, exactness and correctness for protected Cloud transaction".

K.Kavithvajen, "Automated Attendance System using Facial recognition", "Hardware based customer feedback system with Modelling and Analysis of Data" are the papers published.

Toshita Balaji "Automated Attendance System using Facial recognition", "Hardware based customer feedback system with Modelling and Analysis of Data" are the papers published.

Malavika Saji, "Automated Attendance System using Facial recognition", "Hardware based customer feedback system with Modelling and Analysis of Data" are the papers published.

Sushmitha.S.R, "Hardware based customer feedback system with Modelling and Analysis of Data" , "Public information and welfare system" are the papers published.