Feature Based Face Recognition using Machine Learning Techniques'

Bhanushree K. J., Meenavathi M. B.

Abstract: ***Human Face has Numerous unique Features to Distinguish between each other. Face can Identified by distinguishing between face and non-face followed by Identification. Traditionally face recognition uses distinct features Comparison to Identify the Faces which is Complex for larger databases and ambiguous in many scenarios. To improve the accuracy and Scalability Proposed method uses machine learning based Haar Cascade technique for face detection and convolutional neural network is used for feature extraction followed by classification using Euclidean distance and cosine transformation to recognize the face. The results demonstrate the work is performed well in recognizing the face efficiently with different variations.***

Keywords: ***Feature Extraction. Discrete Cosine Transformation; Face Recognition Euclidean Distance; Face Detection;***

I. INTRODUCTION

Face recognition is identification of patterns and its unique structures to differentiate one face with others faces that humans perform routinely and effortlessly in their daily lives. In current digital transformation era face recognition is widely used in many applications to match the human face for biometric solutions, identify people in the humans from crowd and track the peoples in the large group movements for security surveillance and face match for people identity. the research work in the face recognition topic has been progressing towards using machine learning based deep learning approaches to recognize with the facial expression variants, light effects and partial visible areas. Face recognition is an umbrella term that includes both face detection and face verification.

Face detection refers to identify and mark the face and non-face regions and discard the non-face data from the large images and reduces the computation load during the face recognition processing. face detection can be implemented by training the large dataset of human faces using machine learning techniques and build the interested human face region feature matrix to differentiate the face and non-face regions.

Face verification refers to validating the one human face with large dataset of human faces. In reality during the live image capture and usually human face images are distorted due the variations in the lighting effects or partial visible positions. In addition to this facial expression and facial aesthetics will distract the image matching score which will increase the probability of ambiguity in face verification.

In the proposed method face recognition is implemented using machine learning based algorithms to build dynamic features set by exhaustive training data set and verify the face images.

II. FACE RECOGNITION STEPS

Face recognition has the following steps pre-processing is done on the database then face detection is performed followed by features extraction and classification finally, the ace is recognized.

A. Pre-Processing

In proposed method live captured color images will be converted into to grey scale images and resize the images to fit into 624x544 to match the Aberdeen dataset.

B. Face Detection

Face must be captured in order to recognize it, when compared with a new face captured on future. Paul Viola and Michael Jones [1] introduces Haar feature and define the cascade classifiers to detect objects. The haar features for a face are shown in Fig 1.

Haar Cascade is implemented by using Open Source Computer Vision library (OpenCV). Haar features will be used to find the existences of the features in the given input image. In the feature result set a one unique value is computed by subtracting total pixel under white surface area from total pixel under black surface area as shown in Fig 1. Haar like features are the rectangle features for rapid face detection.

Face is scanned from left to right corner of the image using Haar features for face detection. To optimize the features computation and increase the performance of algorithm integral image method is used. In this method total no of pixels computed by referring only four corner edge pixels. integral image the value at pixel (x,y) is the total number of pixels above and to the left of (x,y). The calculations are demonstrated in Fig. 2.

Revised Manuscript Received on February 18, 2020.
***** Correspondence Author
Bhanushree K. J.*, Department of computer science, Bangalore institute of technology, Bangalore, India.
E-mail: Bhanushreekj.bit@gmail.com
Meenavathi M. B.*, Department of electronics and instrumentation engineering, Bangalore institute of technology, Bangalore, India.
E-mail: Meenavathi@gmail.com

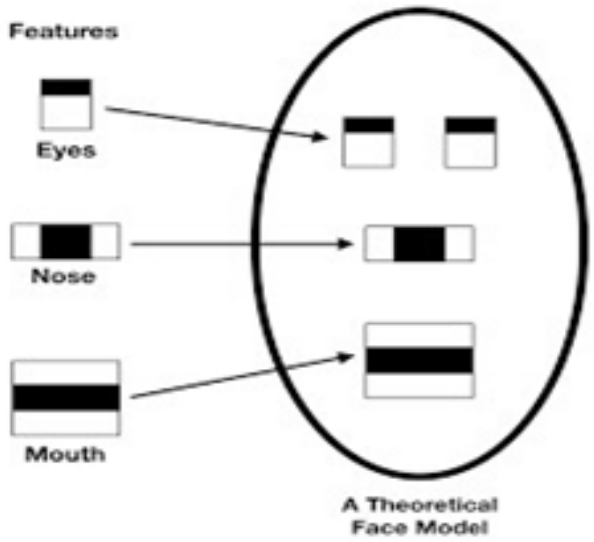


Fig 1: Haar Features

$$G_1 = A, G_2 = A + B, G_3 = A + C, G_4 = A + B + BC \quad (1)$$

$$G_1 + G_4 - G_2 - G_3 = A + A + B + C + D - A - B - A - C = D \quad (2)$$

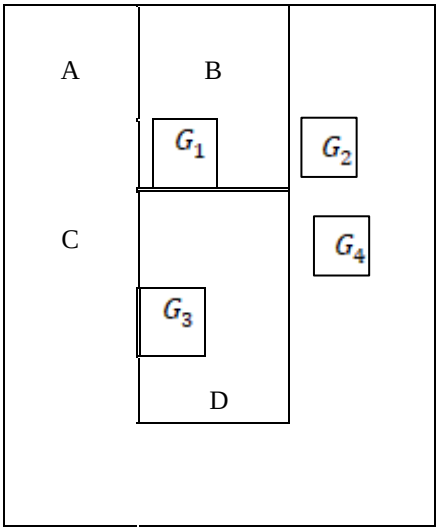


Fig 2: Integral Calculations

Viola, Paul, and Michael Jones proposed an algorithm to define and use base window size as 24×24[1] and verify these features in the input image. This method increases features set to 160000 with combinations of haar features like position, type and scale. Computing these many combinations nearly impossible with defined window size.

C. Feature Extraction

Face recognition success is dependent on useful unique feature data extracted from the input image. The set of distinct features with higher distance between feature will increase probability of success rate of image recognitions.

A convolution neural network (CNN) is like one of the Artificial neural network (ANN). In CNN convolution block is built instead of 1-dimension neurons layer in MLP. convolutional block is a sequence of convolutional layers. In convolution function 3X3 square matrix trainable filter is used find the inner product and find the sum of all entries by sliding on the overlapping regions of the input image [2].

Produce the activation maps by applying the ReLU (Rectified linear unit) and sigmoid on the convolution operation output. Spatial dimensionality and complexity are reduced using pooling layers between convolutional layers and generate the

global features and local features together in one go. Generate the vector of required dimension to define the output by appending the convolutional layers with fully deep networked dense layers. Fully connected neural network (FCN) is derived and modified from Convolution neural network. To resolve the scale invariance and input image size invariance at the end fully connected layers are replaced with convolution layers. In fully connected output layer Pre learning of the background scene will introduce the spatial invariance inherent to the convolution layers.

The detected face will be broken down into facial features using convolutional neural network. In which features data is extracted from input images in the form of numerical digits and it is difficult understand and corelate by individuals.

D. Classification

k-class classification which results in two values: Recognized or Not Recognized i.e.0/1 for each class. Euclidean distance and cosine transformation [3] are used for classification with the threshold values to compare with the trained data.

In 3-D space or 2-D plane Euclidian Distance between two points finds the length of segment connecting the two points. This is not the most common way to represent the distance between the points.

Length of the line segment is defined by The Euclidean distance between points p and q. For cartesian coordinates, two points p = (p1, p2,..., pn) and q = (q1, q2,..., qn) defines the Euclidian's n-spaces. Pythagorean formula [3] defines the distance (d) from p to q to q to p as shown in below equation.

$$d(p,q) = d(q,p) = \sqrt{(q_1 + p_1)^2 + (q_2 + p_2)^2 + \cdots + (q_n + p_n)^2} \quad (3)$$

$$= \sqrt{\sum_{i=1}^{n} (q_i + p_i)^2} \quad (4)$$

VGG16[4] is enhanced to deep neural network using DCT (Discrete Cosine Transform). DCT-Net ignores distortion type and its level in the testing and training input images. During the DCT train process input information which contributes to the high frequency is discarded. In general, DCT-Net is trained only once and applied on to the algorithm without repeat the training.

For image classification Convolutional Neural Network is one of the good methods but it is vulnerable to image quality degradation. Performance of CNN architecture is affected by small amount of Distortion such as noise or blur [12]. Improve the pre-trained CNN by training the mutually exclusive distorted image data. Usually deep learning network is trained with single input image and classify the output /label the image.

III. VALIDATION AND VERIFICATION

For all the values of a test data set, a machine learning model only predicts the outcomes as either positive or negative.

The evaluation parameters for any classification are based on only four outcomes. Significant meaning of four outcomes is shown in the Table. I
Correct positive prediction: True positive (TP)
Incorrect positive prediction: False positive (FP)
Correct negative prediction: True negative (TN)
Incorrect negative prediction: False negative (FN)

Table I: Significant Meaning of TP, FP, FN, TN

System's Predication	*Correct Classification*	*TP*	*FP*	*FN*	*TN*
No	No	0	0	0	1
No	Yes	0	0	1	0
Yes	No	0	1	0	0
Yes	Yes	1	0	0	0

The terms positive and negative refer to the classifier's prediction, and the terms true and false refer to whether that prediction matches to the external knowledge.

A. Confusion Matrix

Confusion Matrix calculates the number of correct & incorrect predictions which is further summarized with the number of count values and breakdown into each classis. It can be used to get precision, accuracy, recall etc. Confusion matrix is provided by the metrics module of the sklearn. Table II shows the confusion matrix with TP, FP, FN, TN.

Table II: Confusion Matrix

		Actual Values	
		Positive (1)	**Negative (0)**
Predicted values	**Positive (1)**	**TP**	**FP**
	Negative (0)	**FN**	**TN**

B. Accuracy Score

Accuracy is the prediction of the model. Accuracy score is provided by the metrics module of sklearn or we can calculate it using the confusion matrix.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

C. Sensitivity

Sensitivity (SN) is calculated as the number of correct positive predictions divided by the total number of positives. It is also called recall (REC) or true positive rate (TPR). The best sensitivity is 1.0, whereas the worst is 0.0.

$$SN = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (6)$$

D. Specificity

specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. It is called as true Negative

$$SP = \frac{TN}{TN + FP} = \frac{TN}{N} \quad (7)$$

Confusion Matrix and Accuracy matrix for the Face Recognition model is shown in Table III

Table III: Confusion Matrix for the Face recognition model

	Predicted Positive	**Predicted Negative**
Actual Positives	124	8
Actual Negatives	0	2

IV. RESULTS AND DISCUSSIONS

The prediction of the Face Recognition with different variations of facial images of Aberdeen dataset is considered.

A. Effect of face recognition with illumination and different orientations

Here the database is containing images with illumination effect. so, the accuracy has been reduced.
Measures:M
Sensitivity:ST
Specificity: SF
Predicted Positive Values: PPV
Predicted Negative Values: PNV
Accuracy: A
Proposed Application with varied illumination: PAVI

Table IV: Performance measure of application

M	**ST**	**SF**	**PPV**	**PNV**	**A**
PAVI	0.837	1	1	.0769	84%

B. Effect of face recognition with illumination and different orientations

Here the database is containing face images can be recognized with illumination effect, but darker illuminated images are not considered.so, the accuracy has been increased.

Table V: Performance measure of application

M	**ST**	**SF**	**PPV**	**PNV**	**A**
PA	0.939	1	1	0.2	94%

C. Effect of face recognition without darker illumination

Here the database is containing face images can be recognized with illumination effect, but darker illuminated images are not considered .so, the accuracy has been increased.

Table VI: Performance measure of application

M	ST	SF	PPV	PNV	A
PA	0.979	1	1	0.1	98%

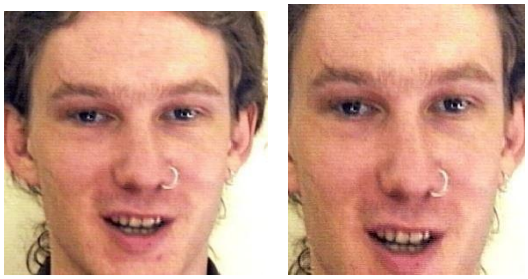
The following example shows how the dataset is working with different orientations of face images. Below representation of face with frontal face and different face angles are considered.

(a) (b)

(c) (d)

(e)

Fig 4. (a) Original Frontal Face (b) Detected face (c) Face with 20-degree deviation (d) Recognized face with 20-degree deviation. (e) Face with greater than 50 degree is not recognized



(a) (b)

Fig. 5: (a) Input face (b) Identified with smile expression

The following examples in Fig. 6 to Fig. 13 shows the faces with different illumination and different degree of orientation.

(a) (b)

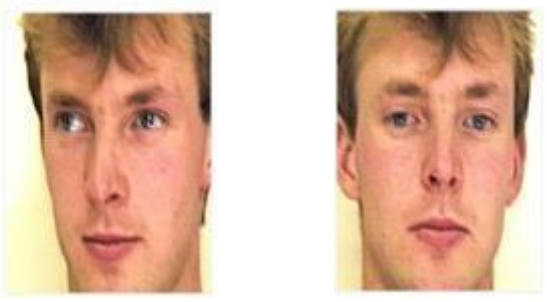
Fig. 6: (a) Input face (b) Identified face with illumination

(a) (b)

Fig7: (a) Input face (b) Identified face with Illumination

(a) (b)

Fig 8: (a) Input face (b) Identified Face with Normal Illumination



(a) (b)

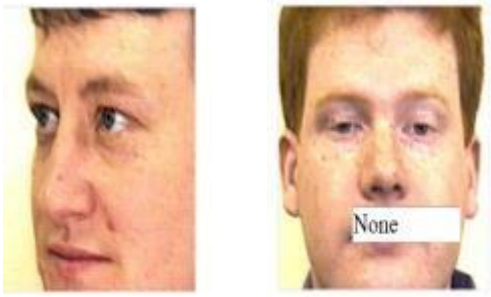
Fig. 9: (a) Input face (b) Identified Face with Brighter Illumination

(a) (b)

Fig. 10: (a) Input face (b) Identified Face with Illumination

(a) (b)

Fig11: (a) Input face (b) Identified face with Expression



(a) (b)

Fig 12: (a) Input face (b) Identified Face with Illumination

(a) (b)

Fig 13: (a) Input face (b) True Negative case (there is no individual in dataset whose image is shown on left)

V. CONCLUSIONS

Face Recognition application is a standalone application which allows the user to take an input image and do face detection and if a face has been detected then the detected face is then Recognized. If the face matches with the faces in the stored database, then it gives recognized face as the output or else displays face not found. In this paper discussed about different variations such as deviation in pose, illumination and different face contours. This is an effective application in the sense it does not require pretraining again and again of the Convolutional Neural Network. Its effective for wider range of scenarios.

REFERENCES

1. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." CVPR (1) 1 (2001): 511-518.
2. Zhang, Tong, et al. "A deep neural network-driven feature learning method for multi-view facial expression recognition." IEEE Transactions on Multimedia 18.12 (2016): 2528-2536.
3. Aggarwal, Swati, Nitika Agarwal, and Monal Jain. "Performance Analysis of Uncertain K-means Clustering Algorithm Using Different Distance Metrics." Computational Intelligence: Theories, Applications and Future Directions-Volume I. Springer, Singapore, 2019. 237-245.
4. Hossain, Md Tahmid, et al. "Distortion robust image classification with deep convolutional neural network based on discrete cosine transform." arXiv preprint arXiv:1811.05819(2018).
5. Z. Wang, R. Huang, W. Yang, and C. Sun "Face recognition rate using different classifiers methods", in Proceedings of the 33rd Chinese Control Conference, , July 2017, pp. 4636–4640.
6. Balcoh, Naveed Khan "Algorithm for efficient attendance management: Face recognition based approach"., et al ,IJCSI International Journal of Computer Science Issues 9.4 (2012): 146-150.
7. "Labeled faces in the wild": A survey, E. Learned-miller, G. Huang, A. Roychowdhury, H. Li, and G. Hua, pp. 1– 55, 2016.
8. Kamencay, Patrik, et al. "A new method for face recognition using convolutional neural network." (2017). Volume : 15 Number: 4 | 2017
9. Ranjan, Rajeev, et al. "Deep learning for understanding faces: Machines may be just as good, or better, than humans." IEEE Signal Processing Magazine 35.1 (2018): 66-83.
10. Wang, Hao, et al. "Cosface: Large margin cosine loss for deep face recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
11. Sun, Xudong, Pengcheng Wu, and Steven CH Hoi. "Face detection using deep learning: An improved faster RCNN approach." Neurocomputing 299 (2018): 42-50.
12. Farfade, Sachin Sudhakar, Mohammad J. Saberian, and Li-Jia Li. "Multi-view face detection using deep convolutional neural networks." Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. ACM, 2015.

AUTHORS PROFILE

Bhanushree K. J., working as Assistant Professor in department of Computer science and Engineering at Bangalore Institute of Technology. she is pursuing Ph.D. at Visvesvaraya Technological University, Belgaum, Karnataka, India. in the area of Face recognition

Meenavathi M. B., working as Professor and Head, Department of Electronics and Instrumentation Engineering, Bangalore Institute of Technology, Bangalore India. Her research interest includes Image filter designs, Image restoration and segmentation