Facial Recognition Using Aggregation and Random Forest Classification Method

K. Aishwarya¹, B. Suresh Kumar², Dr. S. Viswanadha Raju³

¹Computer Science and Engineering, Vardhaman College of Engineering, Shamshabad, India
²Assistant Professor in CSE, Vardhaman College of Engineering, Hyderabad, India,
³Professor in CSE, JNTU Hyderabad, India,
aishwaryakoukuntla@gmail.com¹, sureshkumargoud2006@gmail.com², Svraju.jntu@gmail.com³

Abstract—Face detection and recognition performs an essential role in computer. There was tremendous increase in face recognition during the last years. There are numerous applications that require the face detection. As this face detection is the first step. There are many growing applications such as bank authentication, security access in system, enforcement of law, verification of credit cards, biometric authentication which works based on face detection. The goal of this paper is to presents a facial recognition system with deep learning methodologies. In this paper machine learning aggregation method is used to store the feature of detecting image and random forest algorithm also used to classify a detected image which is applied on face database and compared with previous algorithm which shows accurate ratio.

Keywords— Local Binary Pattern Histogram (LBPH), Principal Component Analysis (PCA), Histogram Oriented Gradient (HOG), Linear Discriminant Analysis (LDA), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN).

1. INTRODUCTION

Many researchers around the world, face recognition is a well-known and challenging filed. Commonly many people often use faces to recognize people and advances in computer skills in recent decades to automatically enable similar recognitions. The process of identifying people through facial images, face recognition, has many real-time applications in the fields of bio-metric, security of information, access control, law enforcement, smart cards and video surveillance.

In relatively restricted scenes, current face recognition systems will in general experience the ill effects of varieties in posture, brightening or outward appearance in videos. Major advances and initiatives over the last 10 to 15 years have brought the technology of face recognition to the fore. For both verification and identification, face recognition can be used. Face recognition system if the faces in the images and videos are automatically identified. It is divided into two different modes:

1. Verification of the face
2. Identification of the face.

In verification of the face, a balanced match contrasts the substance of an inquiry and the essence of a layout whose personality is asserted. There are many images compared by one image in face recognizable proof or acknowledgement that contrasts a question face image and the whole format face images in database to decide the identity of the inquiry. The third case is likewise considered when a question could possibly be in the database accessible. For this situation, we can compute the similitude score and match based on the most noteworthy likeness score.

From video surveillance, video is converted into frames. Edge detection is applied to each frame as shown in Fig1.
Fig. 1: A Facial Recognition System

Face detection for images captured in favorable situations can be done in real time. Face detection and matching is essential for extracting facial capabilities and calculating exactness. Face recognition remains a challenge to understand faces in movement images, variations in posture, twin. There are different add-ons which include bread, color of hair and special facial expressions in lightning conditions, mild intensity and speckle.

A facial recognition [1] is a generation which could perceive or affirm a person from a video frame. There are numerous methods in which facial popularity structures work; however, they typically work by evaluating selected facial traits of a given image with faces in a database. Face recognition analyzes the characteristics of the input of face images of a person via a digital video camera. It measures the overall facial structure, including eye, nose, mouth and jaw distances. These measurements are kept in a database and used as a comparison when a user is standing in front of the camera. This biometric has been widely praised, and perhaps wildly, as a fantastic system for recognizing potential threats, but has not been widely accepted in high-level use. Biometric facial recognition technology [2] is anticipated to before long assume control over the biometrics of fingerprints as the most type of client validation. Each face has numerous particular tourist spots, the different pinnacles and valleys that make up the face. Each human face had around 80 nodal focuses and separation between the eyes, width of the nostril, and profundity of the eye attachments, the state of the cheekbones, and the length of the stunning. These nodal points are measured to create a numerical code, called a face print, representing the face in the database.

The integral image conjointly referred to as a summed space table. An associate degree formula for quickly and with efficiency computing the total of values in a very parallelogram set of a grid. It absolutely was introduced to the graphics field. For fast computation, the integral image is used. Critical component in the Viola-Jones detector [3] is the attentional cascade. Most sub-windows are therefore rejected in the detector’s early stages, making the detection process extremely efficient. The degenerate tree is formed by classifying a window which is named as “cascade”. To learn and perform a face model, classification is used which is adopted from the machine learning algorithms. In this process, there are two keys: one is what features to extract, and which algorithm to apply.

2. RELATED WORK

The strategy proposed [4] is Eigen faces. The PCA is the premise of the Eigen faces technique. The principal component analysis (PCA) is the basis of the Eigen faces method. To represent the face images efficiently the Eigen faces and PCA is used by sirovich and Kirby. Initially they started with a group of face images which are original and calculated the compression vector system. Generally, the PCA is a method which is used for pattern recognition. The aim of the PCA is to replace large scale correlated vectors with the smaller, uncorrelated vectors. Another goal is to calculate the dataset basis. The main advantages of the PCA are its low noise sensitivity, reduced memory and capacity requirements, and increased efficiency due to operation in smaller spaces. The Eigen faces approach appeared to be an appropriate method for face recognition because of its simplicity, speed and ability to learn.

The method proposed in [5] is the histogram oriented gradient. The HOG’s basic concept is to replace the original pixel value to adjacent pixel gradient direction to improve illumination, variance and robustness. The oriented analysis is robust since the lightning. Histogram gives invariance to translations. The HOG function summarizes the measurement distribution in the image regions and is especially important for the recognition of deformable shaped textured objects. The method is also
simple and fast. The histogram can be calculated fast. It is used to extract information on the edge and works well during variations in lighting conditions and poses. HOG works well in such challenging situations. The descriptor for HOG functions can be used together with other descriptors. The facial features of an input image containing face are marked by a HOG descriptor forming a facial cell feature. Each cell is grouped into blocks and if the gradient of an image contributes to the facial features, blocks are assigned with positive weights. All these blocks are clustered together and the input image matches the face visualization model. HOG helps us visualize face from +90 degrees to -90 degrees in any direction.

The strategy proposed in [6] is the Linear Discriminant Analysis (LDA), which is a notable framework for removing highlights and diminishing measurements. It has been broadly utilized in high-dimensional information applications; for example, face acknowledgment and recuperation of the picture. The purported peculiarity issue is a characteristic impediment of traditional LDA, that is, it comes up short when all disperse lattices are special. LDA finds the vectors that best segregate between classes in the basic space.

The strategy proposed [7] is Convolutional neural network. In the field of image processing, deep learning has taken place. CNN is a typically deep learning method that has a unique superiority in image processing due to its special local weight sharing structure. In particular, the multi-dimensional input vector image can be entered directly into the network, preventing the extraction of features and reducing the complexity of data reconstruction in the classification process. Convolution neural network has been successfully applied to the recognition of character, face recognition, estimation of human pose and target detection. It differs from the traditional extraction of features which is carried out layer by layer, and after multi-layer nonlinear mapping, the network can automatically learn to form the feature extractor and classifier suitable for the recognition task from the training samples without special pre-treatment. This method reduces the training sample requirement and the more network layers are set, the more global characteristics can be learned.

The technique proposed in [8] is the recurrent neural network which contrasts to feed ahead neural networks. This may be successfully operated on facts sequences with variable input length. Because of these RNNs uses expertise of its previous nation as an input for its cutting-edge prediction, and this process can be repeated for an arbitrary variety of steps that permit the network to propagate information over the years via its concealed state. This is basically giving a short-term memory to a neural network. This feature makes RNNs very effective in working with data sequences that occur over time, such as time-series data, such as changes in stock prices, a sequence of characters, such as typing a stream of characters into a mobile phone. Gated RNNs and long-term short-term memory RNNs (LSTMs) are the two variants of the basic RNN architecture that help to solve a common problem with training RNNs. Both variants use a memory form to help predict sequences over time. The main difference between a Gated RNN and an LSTM is that Gated RNN has two gates to control its memory: an update gate and a reset gate, while an LSTM has three entryways: an info gate, a yield entryway and an overlook gate. RNNs work well for applications involving a time-changing sequence of data. These applications include the processing of natural languages, speech recognition, language translation, image titration, conversation modeling.

The approach proposed in [9] is the support vector classifier that maps information factors in vicinity. This statistics factor’s rectangular degree is accessed via analyzing levels of similarities. While points have comparable attributes, the distance among them in area is tiny. Classes are assigned to these records points by using maximizing the gap among points of various instructions. Even though this technique is inherently a linear classifier, using special functions, the classifier may be made non-linear in nature. The SVM classifier changed into used alongside the function extractor that changed into utilized in the proposed methodology. It had been found that the accuracies had been inferior to the accuracies acquired as soon as the ML Perceptron classifier became used.
3. PROPOSED METHOD

A. Edge Detection

In object-based image processing there are many applications like robotic and industrial vision, medical imaging, remote sensing etc. Edge detection [10][11] is one of the fundamental but most important tasks.

There are specific part detectors which vary in phrases of location precision, intricacy, ongoing applications, noise reduction and so on. There are detectors like sobel, prewitt and many others, which can be based totally on gradient magnitude. Those are easy to put in force however there is a bad in noise immunity. Detectors which are based on Gaussian mask like LoG, DoG reduces noises. The discovery of lines that mark the limit and the image appearance is divided from the other places or things.

The Laplacian [12][13] is an isotropic which is a 2-D proportion of spatial subsidiary of the picture. The Laplacian is utilized for edge identification; mainly it highlights regions of an image with rapid intensity change. To remove noise in the image, Gaussian smoothing filter [14] is used. More often than not, the operator takes single grey degree picture as input and this operator produces an output as another grey degree photograph [15].

The Laplacian of a picture with pixel power esteems is given by

\[ Lp(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \]

This can be determined utilizing a convolution channel [16]. Convolution is a central and straightforward scientific operator to other normal picture preparing operators. This provides two arrays with multiplication having different sizes, but the dimensionality should be the same; so that a third array is produced with the same dimensionality. The convolution is utilized in picture preparing for actualizing the operators. The yield pixel esteems are of straightforward direct mixes with the estimations of information pixels.

The info picture is demonstrated as a gathering of discrete pixels. The beneath are the ordinarily utilized parts.

The LoG of 2-D fixated on zero and with Gaussian standard deviation \( \sigma \) has the structure:

\[ log(x, y) = \frac{1}{\pi \sigma^4} \left[ 1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \]

As Laplace administrator may recognize edges just as commotion (disengaged, out-of-go) it might be attractive to smooth the picture first by a convolution with a Gaussian portion of width \( \sigma \).

\[ G_\sigma(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2 + y^2}{2\sigma^2}} \]

To stifle the commotion before utilizing Laplace for edge recognition:

\[ \Delta[G_\sigma(x, y) * f(x, y)] = [\Delta G_\sigma(x, y)] * f(x, y) = log * f(x, y) \]

The primary equivalent sign is because of the way that

\[ \frac{d}{dt} [h(t) * f(t)] = \frac{d}{dt} \int f(\tau)h(t - \tau)d\tau = \int f(\tau) \frac{d}{dt} h(t - \tau)d\tau = f(t) \frac{d}{dt} h(t) \]
So we can get the Laplacian of Gaussian first and after that convolve it with the info picture. To do as such, first consider
\[
\frac{\partial}{\partial x} G(\sigma, x, y) = \frac{\partial}{\partial x} e^{-\frac{(x^2+y^2)}{2\sigma^2}} = \frac{-x}{\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]
And
\[
\frac{\partial^2}{\partial^2 x} G(\sigma, x, y) = \frac{x^2}{\sigma^4} e^{-\frac{(x^2+y^2)}{2\sigma^2}} - \frac{1}{\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} = \frac{x^2 - \sigma^2}{\sigma^4} e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]
Note that for effortlessness we precluded the normalizing coefficient \(1/\sqrt{(2\pi\sigma^2)}\).correspondingly we can get.
\[
\frac{\partial^2}{\partial^2 x} G(\sigma, x, y) = \frac{y^2 - \sigma^2}{\sigma^4} e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]
Presently we have LoG as a convolution portion characterized as
\[
\log \triangleq \Delta G(\sigma, x, y) + \frac{\partial^2}{\partial x^2} G(\sigma, x, y) + \frac{\partial^2}{\partial y^2} G(\sigma, x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]
The rims inside the picture can be received by those steps:

- making use of log to the picture
- Detection of zero-crossings in the photograph
- Threshold the 0-crossings to maintain simplest those strong ones (large distinction between the effective most and the bad minimum).

The ultimate step is needed to suppress the susceptible 0-crossings most probable caused by noise.

B. Vector of Locally Aggregated Descriptors

After detecting edges using LoG, the VLAD [17] is applied to these features. In the wake of extricating the descriptors from the picture, we tend to combine to exploit the neighborhood criterion within the feature area in a compact vector illustration. For each visual word \(c_i\), VLAD someone aggregates the distinction between the native descriptor \(x\) and furthermore the nearest visual word \(c_i\) as \(x - c_i\). This might be in profound aggravation describing the center related appropriation of the vector. The measurement D of the model is \(k \ast d\) where \(d\) and \(k\) the component of the local descriptor as the assortment of visual words. \(v\) is imagined by:
\[
v_{i,j} = \sum_{X\text{suchthatNN}(X)\in c_i} x_i - c_{i,j}
\]
where \(v_{i,j}\) gives the descriptor, \(i\) mean the visual word and the \(j\) mean the local descriptor component. Likewise, \(x_j\) represents the descriptor’s \(f^{th}\) element \(x\) and\(c_i, j\)'s corresponding visual word\(c_i\).
The descriptors acquired from the VLAD square measure are organized and distributed comparatively, ultimately positioned within the same cluster by the upper values of the square descriptors. The quantum vector is supplemented by supervised algorithms for the classification exploitation model. The output of this method provides text.

C. Random Forest Algorithm

Inside the next level, random forest algorithm is applies on this newsletter, to classify the features. Random wooded area can be flexible, easy to apply gadget learning algorithmic rule that produces, even while no longer hyper-parameter standardization, an notable result maximum of the time. It’s conjointly one of the foremost used algorithms, because of its simplicity and also the plain truth that it could be used for every category and regression duties. Random wooded area is an algorithm of supervised mastering. It creates woodland, as you could see from its name, and makes it comes what may random. The "wooded area" it builds is an ensemble of bushes of decision, primarily skilled with the method of "bagging." the overall concept of the bagging method is that the overall result is improved through a combination of mastering fashions. Virtually placed, random forest builds and merges multiple decision trees to obtain a greater correct and strong prediction

\[
P_{X,Y}(P_\theta(h(X,\theta) = Y) - \max_{j\neq Y} P_\theta(h(X,\theta) = j) < 0
\]

Where, X, Y are demonstrative of the likelihood over the space X, Y. This shows why irregular timberlands don't overfit when the quantity of trees is expanded. Arbitrary woods additionally give a compelling technique to gauge missing qualities. Maybe a couple of the real focal points are that they don't overfit, they needn't bother with cross-approval to get a fair-minded gauge of the test set blunder and they are quick.
4. EXPERIMENT ANALYSIS

All the experiments are executed in an Intel i5 processor and used with Ubuntu as platform. The programming languages used while doing these experiments are Python with its necessary packages like numpy, os, and openCV etc., are used. To do this we have selected a Grimace dataset. In this dataset totally consist of 360 images. In that 20 different pictures of 18 individual members are used. The size of each image is 180*200. In this paper, the proposed method compared with previous popular methods which are used for face recognition those are LBP, PCA, LDA etc.

In the initial step LoG method is applied on dataset. The main purpose of LoG method is to detect the edge of the faces. The Fig2 shows the Edge detection technique using LoG.

![Edge detection technique](image)

**Fig.3: the Edge detection technique**

VLAD method is used, the purpose of this method is to convert feature of the image into text. Random Forest is applied on the output of the Aggregation method. The images of Grimace dataset are taken in complex background. The following results shows the accuracy rate when compare with other proposed methods.

| Method/Dataset | Grimace Dataset |
|----------------|-----------------|
| PCA            | 73.82%          |
| LDA            | 78.24%          |
| LBP            | 81.56%          |
| Proposed Method| 84.76%          |

The results above show that the proposed results are good and accurate compared to other methods proposed.

5. CONCLUSION

Face recognition system plays vital role in the present research area. The proposed method uses edge detection, aggregation method, and classification. In aggregation method, it will convert all features into text, on that we are applying classification techniques. This method provides very good accurate rate when compared to the previous work.

6. REFERENCES

[1] B C. Zhang and Z. Zhang, A survey of recent advances in face detection. Technical report, Microsoft Research, 2010.
[2] M.-H. Yang, D. Kriegman, and N. Ahuja, “Detecting faces in images: A survey,” IEEE Trans. On PAMI, vol.24, pp. 34–58, January 2002.

[3] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features", IEEE Conference on Computer Vision and Pattern Recognition, 2001.

[4] T. M. Mitchell. Machine Learning. McGraw-Hill International Editions, 1997. D. Pissarenko. Neural networks for financial time series prediction: Overview over recent research. BSc thesis, 2002.

[5] Park, K.-Y., & Hwang, S.-Y. (2014). An improved Haar-like feature for efficient object detection. Pattern Recognition Letters.

[6] Intelligent Face Recognition techniques: A Comparative Study” by S.T Gandhe, K T Talele, A G Keskar.

[7] Y. Sun, D. Liang, , and X Wang, “Deepid3: Face recognition with very deep neural networks,” http://arxiv.org/abs/1502.00873, 2015.

[8] Recurrent Neural Networks for prediction: Learning Algorithms, Architectures and Stability John Wiley & Sons, Inc. New York, NY, USA 2001.

[9] M. O. a. M. A. M. H. Faruqe, ""Face recognition using PCA and SVM," Anti-counterfeiting, Security, and Identification in Communication," in 3rd International Conference on. IEEE., 2009.

[10] A. Nomura, M. Ichikawa, R. H. Sianipar, and H. Miike, “Edge detection with reaction-diffusion equations having a local average threshold,” Pattern Recognition and Image Analysis, Vol. 18, pp. 289–299, 2008

[11] P. Perona and J. Malik. “Scale-space and edge detection using anisotropic diffusion”, IEEE TPAMI, vol 12, pp 629-639, July 1990

[12] J.S.Chen etc. ( 1987) , Fast Convolution with Laplacian of Gaussian Masks. IEEE Trans. on PAMI, No. 4: 584- 590.

[13] https://homepages.inf.ed.ac.uk/rbf/HIPR2/log.htm

[14] Raman Maini, Dr. Himanshu Aggarwal, “Study and Comparison of Various Image Edge Detection Techniques”, International Journal of Image Processing,Volume(3):Issue(1)

[15] Pinaki Pratim Acharjya, Ritaban Das & Dibyendu Ghoshal, “Study and Comparison of Different Edge Detectors for Image Segmentation”, Global Journal of Computer Science and TechnologyGraphics & Vision Volume 12 Issue 13 Version 1.0 Year 2012.

[16] https://homepages.inf.ed.ac.uk/rbf/HIPR2/convolve.htm

[17] Jgou H, Douze M, Schmid C, Prez P. Aggregating local descriptors into a compact image representation in Computer Vision and Pattern Recognition (CVPR), 2010 Jun 13 (pp.3304-3311), IEEE.

[18] Radenovi F, Tolias G, Chun O. CNN image retrieval learns from BoW: Unsupervised fine-tuning with hard examples. InEuropean Conference on Computer Vision 2016 Oct 8 (pp. 3-20), Springer, Cham.

[19] Z ZAIN, “High Speed And Lowpower Gdi Based Full Adder”,Journal of VLSI Circuits And Systems, 1 (01), 5-9,2019

Pb Agus Ristono*,"Design Of Reliable And Efficient Manchester Carry Chain Adder Based 8-Bit Alu For High Speed Applications",Journal Of VLSI Circuits And Systems, 1 (01), 1-4,2019