10-30-2020

Face Recognition: A Comparative Approach from Traditional to Recent Trends

Jayanthi Raghavan
University of Windsor

Face Recognition: A Comparative Approach from Traditional to Recent Trends

By

Jayanthi Raghavan

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Electrical and Computer Engineering
in Partial Fulfillment of the Requirements for
the Degree of Master of Applied Science
at the University of Windsor

Windsor, Ontario, Canada

2020

Face Recognition: A Comparative Approach from Traditional to Recent Trends

By

Jayanthi Raghavan

APPROVED BY:

A. Fartaj
Department of Mechanical, Automotive and Materials Engineering

E. Abdel-Raheem
Department of Electrical and Computer Engineering

M. Ahmadi, Advisor
Department of Electrical and Computer Engineering

September 18^{th}, 2020

DECLARATION OF CO-AUTHORSHIP AND PREVIOUS PUBLICATIONS

I. Co-Authorship

I hereby declare that this thesis incorporates material that is the result of research conducted under the supervision of my supervisor, Dr. M. Ahmadi. Results related to this research are reported in Chapters 2 through 4. I am aware of the University of Windsor's Senate Policy on Authorship and I certify that I have properly acknowledged the contributions of the other researchers to my thesis, and I have obtained written permission from my co-author to include the aforementioned materials in my thesis. I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

Chapter 2 to 4 of the thesis was co-authored and supervised by Dr.M. Ahmadi. In all cases, the key ideas, primary contributions, experimental designs, data analysis, interpretation, and writing were performed by Jayanthi Raghavan and the contribution of co-author was primarily providing feedback on refinement of ideas and editing of the manuscript.

II. Previous Publication

I declare that this thesis includes three original papers that have been previously published/ submitted for publication in peer reviewed conferences and journals as follows:

Thesis Chapter	Publication title and full citation	Publication Status
Chapter 2	Performance Evaluation of Entropy based LBP for Face Recognition.	Submitted to 2020 High Performance Computing and Simulation Conference
Chapter 3	Performance Evaluation of Weighted Entropy Based Fusion Technique for Face Recognition with Different Pre-processing Techniques	Accepted for publication to FICC 2021
Chapter 4	Performance Evaluation of CNN for Face Recognition	Submitted to International Journal of Pattern Recognition and Artificial Intelligence

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

III. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices.

Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Face recognition, an important biometric method used extensively by researchers, has become more popular recently due to development of mobile applications and frequent usages of facial images in social media. A major development is attained in facial recognition methods due to the emergence of deep learning methods. As a result, the performance of face recognition systems reached a matured state.

The objectives of this research are to improve the accuracy rate of both traditional and modern methods of face recognition system under illumination variation by applying various preprocessing techniques. In the proposed face recognition approach, various preprocessing methods like SQI, HE, LTISN, GIC and DoG are applied to the Local Binary Pattern (LBP) feature extraction method and by using the Weighted Entropy based method to fuse the output of classifiers on FERET database, we have shown improvement in recognition accuracy of as high as 88.2 % can be obtained after applying DoG .

In a recently used approach, deep CNN model is suggested. The Experiments are conducted in Extended Yale B and FERET Database. The suggested model provides good accuracy rates. To improve the accuracy rates further, preprocessing methods like SQI, HE, LTISN, GIC and DoG are applied to both the models. As a result, higher accuracy rates are achieved in deep CNN model both in Extended Yale B Database and FERET Database. Extended Yale B Database provides the highest accuracy rate of 99.8% after the application of SQI and an accuracy rate of 99.7% is achieved by applying HE.

DEDICATION

Dedicated to my beloved sister Dr. Anandhi Ray and my family

ACKNOWLEDGEMENTS

First and foremost, I would like to give my sincere thanks to my advisor Dr. Majid Ahmadi for his marvellous help, wonderful guidance, and great support throughout this thesis and for his faith in my abilities. His constant interest motivated me at every stage to enable this thesis to be completed. Also, I would like to thank my committee, Dr. Esam Abdel-Raheem and Dr. Amir Fartaj for the important feedback that helped to improve the quality of this work. I would like to thank Dr. Mitra Mirhassani for her wonderful guidance and support. I am always thankful to my sister, kids, friends and my husband and for their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/SYMBOLS

AHE	Adaptive Histogram Equalization
ANN	Artificial Neural Networks
AWPPZMA	Adaptively Weighted Patch PZM Array
BN	Batch Normalization
BoW	Bag of Words
BP	Back Propagation
CLAHE	Contrast-Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
DCS	Decorrelation Stretching
DoG	Difference of Gaussian
FERET	Facial Recognition Technology
FG-NET	Face and Gesture recognition Network
FLD	Fishers Linear Discriminant
GIC	Gamma Intensity Correction
GPU	Graphical Processing Unit
HE	Histogram Equalization
HMM	Hidden Markov Model
HOG	Histogram of Gradients
JAFFE	Japanese Female Facial Expression database.
KLDA	Kernel Linear Discriminative Analysis
K-NN	K-Nearest Neighbor
KPCA	Kernel Principle Component Analysis

LBP	Local Binary Pattern
LDA	Linear Discriminative Analysis
LFW	Labeled Faces in the Wild
LoG	Laplacian-of-Gaussian
LPQ	Local Phase Quantization
LRC	Logistic regression classifier
LTP	Local Ternary Pattern
LTISN	Locally Tuned Inverse Sine Nonlinear
MFM	Max-Feature-Map
MMD	Manifold-Manifold Distance
NB	Naïve Bayes
1D	One Dimensional
2D	Two Dimensional
PP	PreProcessing
PCA	Principle Component Analysis
PZM	Pseudo-Zernike Moments
ReLU	Rectified linear unit
RF	Random Forest
ROI	Region of Interest
SIFT	Scale Invariant Feature Transform
SQI	Self-Quotient image
SURF	Speeded-Up Robust Features
SVM	Support Vector Machine
VGG	Visual Geometric Group
VLD	VanderLugt Correlator
WOAPP	Without Application of Preprocessing

ZM	Zernike Moments
OSH	Optimal Separating Hyperplane.
MAP	Maximum A Posteriori
GM	Gaussian mixture

Chapter 1

Introduction

1.1. Introduction

Biometric methods apply physiological and/or behavioral characteristics of people to confirm their identities. There are various biometric modalities like fingerprints, face, voice, hand geometry and iris recognition which can be employed to identify people.

Each biometric has its own advantages and disadvantages. Speech recognition applies the acoustic nature of voice that has been discovered to be different for different individuals. It is a simple biometric method and it needs less training in comparison to other methods. The drawback of the speech recognition method is, it is affected by changes in quality of voice due to aging process [1]. This issue needs to be taken care of by the speech recognition systems. In finger print recognition, the unique ridges and patterns present in fingers help to identify the individuals. It is extensively used in the field of forensic science. The accuracy of fingerprint recognition is affected by problems like dry skin and injuries present in the fingers. Even though the iris recognition is stable with the age, it is an intrusive biometric method [52]. The signature recognition is relatively simple but it is not stable and changes during the course of time.

In comparison to other biometric techniques, face recognition is pretty simple, cost effective and a non-intrusive approach. Face recognition has wide applications in the areas of advertising, health care, virtual reality, border security, biometrics, access control, and criminal investigation.

1.2 Face Recognition Steps

For human brain, recognizing face is a very simple task and can be performed very fast. In computer vision, face recognition is a very challenging task. Even though the face recognition research is in advanced state, till now it is not possible to obtain results on par with humans.

Recent days, deep learning methods are becoming popular because they can be trained with huge databases to learn important features to represent the data [54].

Face Recognition consists of various steps as shown in the (Fig.1.1). Preprocessing methods are applied to the captured input image to improve recognition accuracy. The important features that represent the facial image are extracted from the processed image. The extracted features are assigned to particular class using classification process and then feature matching is done. Finally, outcome is deduced based on the feature matching.

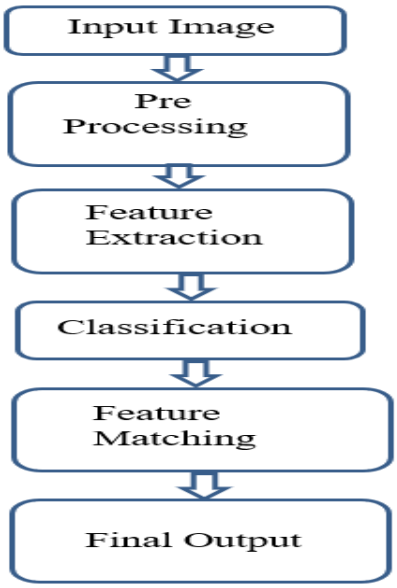


Fig.1.1. Various steps involved in Face Recognition

1.2.1 Preprocessing Phase:

For better face recognition under illumination variation conditions, the vital features responsible for differentiating two different faces should be retained. The shadows produced in facial images due to variation in lighting directions may cause loss of important facial features, which are helpful for recognition. A preprocessing method normally applied to the images to remove the background noise [15]. In addition, the applied methods must enhance the intensity in the regions of inadequately illuminated and decreased intensity in the densely illuminated regions while retaining the intensity in the fairly illuminated portions [11]. Moreover, preprocessing methods help to classify the images accurately [3]. In the following paragraphs, few important of preprocessing methods are discussed.

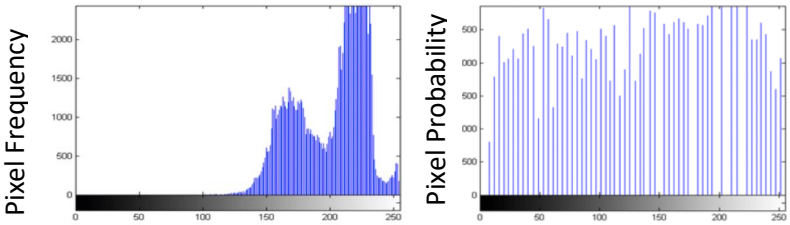
- **Histogram equalization (HE)**

It is very effective and simple global image processing technique applied in spatial domain [9]. The image histogram yields details about the intensity distribution of the pixels present in the image.

For image I(x,y) with discrete k gray values, histogram is defined by the probability of occurrence of a gray level I, is given by equation (1.1) as follows.

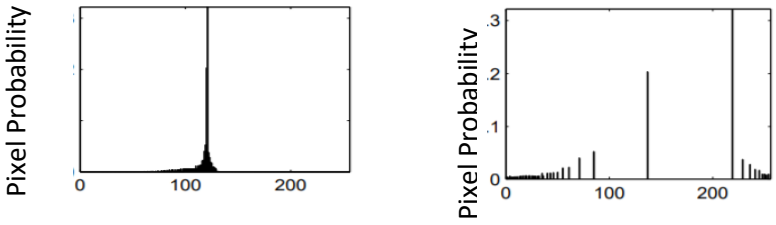
$$p(i) = n_i / N \quad (1.1)$$

Where i ∈ 0, 1…k -1 grey level and N is total number of pixels.



(a) (b)

Fig.1.2. Histogram plot by 1st method (a) Before applying HE (b) After applying HE



(a) (b)

Fig.1.3. Histogram plot by 2nd method (a) Before applying HE (b) After applying HE

In Histogram equalization approach, the image intensities are adjusted to improve the image contrast. Histogram of an image can be constructed by two ways. (i) It can be constructed by plotting pixel intensities against the pixel frequencies as shown in (Fig.1.2). (ii) It can also be constructed by plotting pixel intensities against pixel probabilities as shown in (Fig.1.3). HE improves the global contrast of the image

- **Contrast-Limited Adaptive Histogram Equalization (CLAHE)**

It is suggested to handle the problem of HE [57]. HE is based on the underlying assumption that the overall quality of the image is not varying and single grayscale mapping provides uniform improvement for all regions of the image. But when there is a variation in allocation of grayscale values from one area to another, the assumption fails. In this case, an adaptive histogram equalization technique can perform better than the standard HE. In CLAHE, the image is divided into a limited number of regions and the same histogram equalization technique is applied to pixels in each region [55].

The images treated with CLAHE will have a natural look. However, the low contrast enhancement of CLAHE may obstruct the capability of an observer to find the existence of some significant gray-scale contrast [60].

Even though histogram equalization (HE) is a simple approach, it is not effective in handling the non-uniform variations [9]. In addition to that, it may also eliminate minute details present in the image and can produce hidden noise in the processed image. Moreover, HE is not suitable for the images affected by gamma distortion. Due to technical restrictions, many imaging equipment may not display the real image of objects [10]. This technical constraint is known as gamma distortion.

- **Gamma Intensity Correction (GIC)**

It is an image enhancement technique, which can handle the problem of gamma distortion [58]. It enhances the local dynamic span of the image in darker or under exposed areas and at the same time, squeezes the brighter portions. Hence it changes the overall tone of an image either lightly or darkly based on the value of gamma and at the same time, it maintains the dynamic range of the image. It is a nonlinear gray-level transformation that substitutes gray-level I with the gray-level $I^{\frac{1}{\Upsilon}}$, given by the equation (1.2). There are some limitations in gamma correction technique. it does not eliminate the effects of shadows.

$$I = I^{\frac{1}{\Upsilon}} \tag{1.2}$$

- **Difference of Gaussian (DoG)**

It is a popular and effective in handling the shadowing effect [8]. DoG is essentially acting as band-pass filter, which eliminates both high frequency components that comprises of noise and also some low frequency components that constitutes the homogeneous areas in the image. The frequency components present in the pass band are essentially edges present in the images.

In DoG filter, image is convolved with two different Gaussian kernels. The output image is acquired by computing the difference of the two convolved images [12].

Gaussian kernels having differing standard deviations, which is given in the equation (1.3).

$$DOG(x,y) = \frac{1}{2\pi{\sigma_1}^2} e^{-\frac{x^2+y^2}{2{\sigma_1}^2}} - \frac{1}{2\pi{\sigma_2}^2} e^{-\frac{x^2+y^2}{2{\sigma_2}^2}} \tag{1.3}$$

$where,\ \sigma_1, \sigma_2$ are Gaussian kernel widths.

There are some drawbacks in DoG operation. DoG filtering diminishes the contrast of the image as a whole and hence the contrast of the image needs to be improved after the application of DoG filtering. Moreover, gamma normalization is a pre-requisite for DoG. If DoG is applied without gamma normalization, it may affect the local contrast in shadowed regions.

Most of the preprocessing approaches essentially concentrate on either global or local improvements which may not be suitable for all kinds of images. Especially these methods do not consider the nature of the image, whereas images corrupted with different kinds of degradation might require different types of treatments.

One of the suggested solutions for presenting high contrast image in a bounded contrast range equipment is to squeeze the image contrast without loss of data. Global preprocessing methods are suitable to execute this idea. But there are some performance related problems associated with global methods. While applying these methods, minute details are lost or some regions remain un-enhanced [5]. Local contrast enhancement techniques retain or improve fine image details.

- **Locally Tuned Inverse Sine Nonlinear (LTISN)**

It is suggested to handle the problem of limiting the dynamic range as well as improving the local contrast [11]. It can also handle severe illumination conditions using a series of operations and a nonlinear intensity transformation. It is capable of enhancing the intensity of pixels in the dark regions and at the same time reducing the intensity of pixels in the bright regions. In addition to that, it also improves the local contrast with a single image dependent parameter.

The LTISN is a nonlinear pixel wise image enhancement technique, where the improved intensity values are calculated by applying the inverse sine function with a tunable parameter based on the nearby pixel values given in the equations (1.4), (1.5), (1.6), (1.7) and (1.8) below. The intensity range of the image is rescaled to [0 1], followed by a nonlinear transfer function.

$$I_{enh}(x,y) = \frac{2}{\pi} sin^{-1}(I_n(x,y)^{q/2}) \tag{1.4}$$

$$kernel_i(n_1,n_2) = \frac{h_g(n_1,n_2)}{\sum_{n_1}\sum_{n_2} h_g(n_1,n_2)} \tag{1.5}$$

$$h_g(n_1,n_2) = e^{\frac{-\left(n1^2+n2^2\right)}{2\sigma^2}} \tag{1.6}$$

$$I_{M,i}(x,y) = \sum_{m=-\frac{M_i}{2}}^{\frac{M_i}{2}} \sum_{n=-\frac{N_i}{2}}^{\frac{N_i}{2}} I(m,n) kernel_i(m+x, n+y) \tag{1.7}$$

$$q = \begin{cases} tan(\frac{\pi}{C_1} I_{M_n}(x,y)) + C_2 & I_{M_n}(x,y) \geq 0.3 \\ \frac{1}{C_3} ln(\frac{1}{0.3} I_{M_n}(x,y)) + C_4 & I_{M_n}(x,y) < 0.3 \end{cases} \tag{1.8}$$

The main pitfall of the LTISN technique is, over enhancement in under exposed regions creates black artifacts near the edges and does not recover back the color in over exposed regions.

- **Self-Quotient image (SQI)**

It is one of the popular illumination invariant algorithms suggested for handling both shadow and lighting changes [59]. SQI algorithm is based on the reflectance-illumination

models shown in Equation (1.9) and (1.10) bellow. Haitao Wang and others investigated about the application of self- quotient approach in handling illumination variation in face recognition [6] [7].

$$Q(\mathrm{x},\mathrm{y}) = \frac{I(x,y)}{S(x,y)} \tag{1.9}$$

$$= \frac{I(x,y)}{F(x,y)} * I(x,y) \tag{1.10}$$

The main disadvantages of SQI are, it is not having capacity to differentiate between shadow and low albedo area. As a result, critical information representing identity is ignored. Another problem in SQI is requirement of partitioning the area, where pixel values change gradually is a difficult challenge [14].

1.2.2 Feature Extraction:

In this step, important features are extracted from the face image. Feature extraction algorithms convert the given input data into the set of features, consists of useful information from the original data. Feature extraction methods retain the good classification accuracy level by removing unnecessary features [53]. Various feature extractors are discussed below. Different face recognition methods are as shown in (Fig. 1.4).

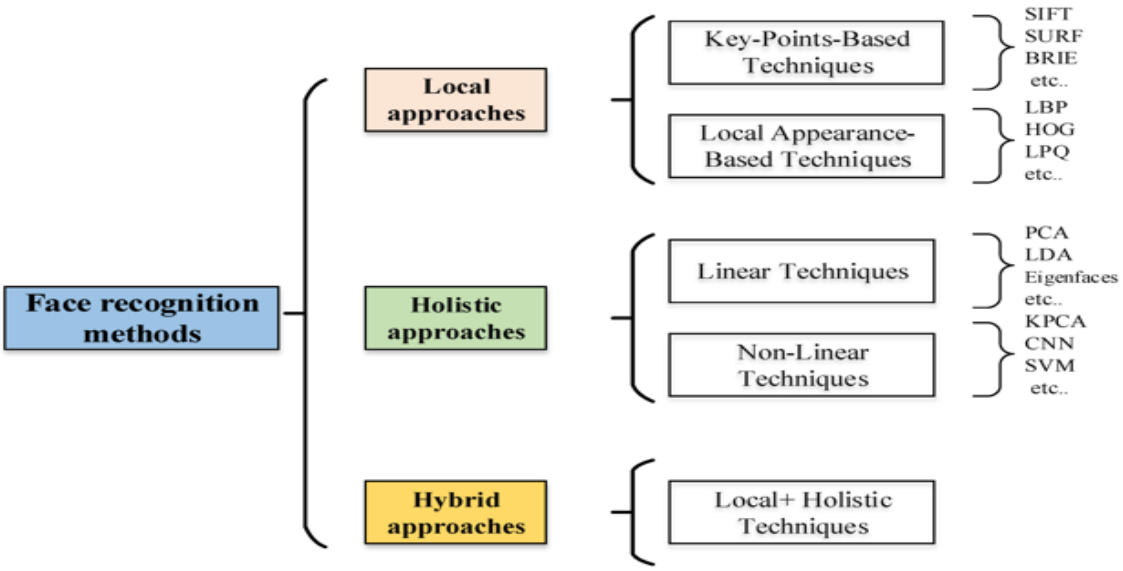


Fig.1.4. Face Recognition Approaches

Key-Points-Based Feature Extraction

The key-points-based techniques are used to find particular geometric features based on the geometric data of the facial image [19] such as gap between the eyes, nostril's center, and mouth's extremities. There are two steps in key-points-based techniques, key-point detection and feature extraction. The key points are primarily populated in the regions of eyes, eyebrows, nose and mouth. Some popular key point feature extractors are discussed below.

- **Scale Invariant Feature Transform (SIFT)**

It is based on the Difference-of-Gaussians (DoG) approach [16], which is an approximation of Laplacian-of-Gaussian (LoG) as shown in (Fig.1.5). Feature points are discovered by finding local maxima using DoG operations at different scales of the images. This approach extracts a 16x16 surroundings around each detected feature and further

divides the area into sub-regions, yielding a total of 128 bin values. SIFT is invariant to rotation, scaling and limited affine changes but its main problem is it is slow due to is high computation.

- **Speeded-Up Robust Features (SURF)**

It is based on determinant of Hessian Matrix and it uses integral images to increase the feature-detection operation speed [16]. Subsequently, Harr-wavelet responses are computed for the pixels within a circular neighborhood. Then weighted Gaussian function is carried out for the feature points and higher weights are assigned to the feature points, which lies close to the circle's center and contributes a lot in orientation. The prominent orientation is judged by computing the sum of overall responses within a sliding orientation window encompassing the angle of 60°. The x and y axes responses within the window are added up to create a new vector. Therefore, the longest vector's orientation is chosen as the feature point with prominent orientation through searching the entire circular neighborhood. SURF features are invariant to rotation and scale but they have little affine invariance. The main advantage of SURF over SIFT is, it is not computationally intensive [17].

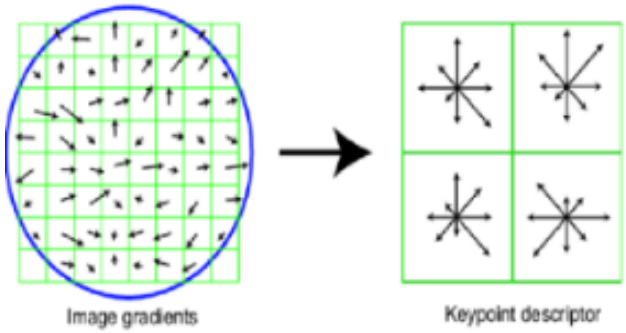


Fig.1.5. SIFT Feature Extraction

Holistic Feature Extraction

Holistic face recognition employs global information from facial image to recognize the face. The global information from faces is fundamentally represented by a limited feature which are derived from the pixels present in the face images [24].

Holistic methods apply the whole face as an input. These methods do not extract feature points, instead represent the face image by pixel matrix. Finally, pixel matrix is transformed into feature vectors and these feature vectors are represented in low dimensional space. The main problem with holistic method is, it is sensitive to changes in facial expressions, illumination, and poses. Holistic methods can be divided into categories, linear and non-linear techniques based on the method used to represent the subspace.

Linear Methods

Linear methods are recommended for handling linearly separable data. Few examples for linear method are discussed in the paragraph below.

- **Principal Component Analysis (PCA)**

It is non-parametric approach to handle the problem of high-dimensional data [20] while retaining vital features as shown in (Fig.1.6). It is a statistical technique that uses orthogonal transformation to transform a set of observations of correlated variables into a linearly uncorrelated variable. It is a popular technique for finding useful patterns in data of high dimension. It converts huge number of correlated variables into a limited number of uncorrelated variables called principal components [21], which act as summaries of features. Along the direction of principal components, the data variation is maximum. High-dimensional data is a common problem in data mining. The dimensionality reduction

helps to reduce the computational expense. PCA aims at capturing most possible of the variance in data while reducing the dimension given in equations (1.11) and (1.12)

$$S_T = \sum_{t=1}^{N}(x_i - \mu)(x_i - \mu)^T \quad (1.11)$$

$$S_T V_i = u_i V_i \quad (1.12)$$

Where S_T=Variance μ= Average and x_i =Face Image

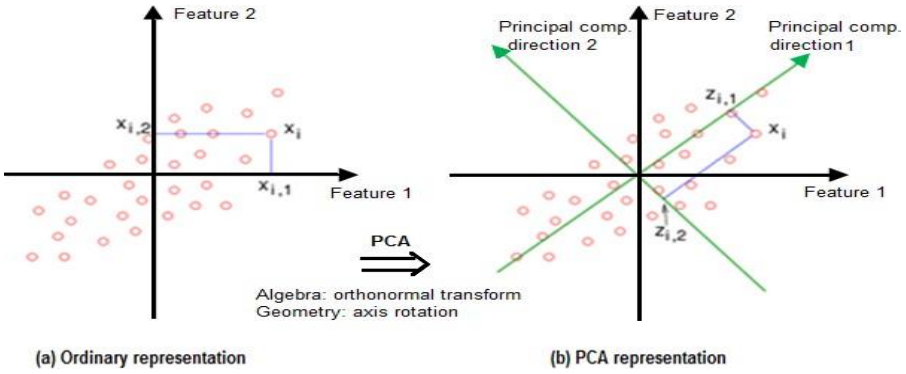


Fig.1.6. Principal component analysis (PCA)

- **Linear Discriminant Analysis (LDA)**

  It is commonly used for dimensionality reduction [68] as shown in (Fig.1.7). LDA is a supervised learning technique. The goal of LDA is to find direction, which maximizes the between class variance of individual faces while minimizing the within class variance of individual faces [26] as shown in equation (1.13). The Images are projected on the calculated hyperplanes called fisher faces, which leads to better classification and higher recognition rates in most of the experiments.

$$\left.\begin{aligned} S_B &= \sum_{i=1}^{c} |\chi_i|(\mu_i - \mu)(\mu_i - \mu)^T \\ S_W &= \sum_{i=1}^{c} \sum_{x_k = \chi_i} (x_k - \mu_i)\,(x_k - \mu_i)^T \end{aligned}\right\} J(W) = \frac{W^T S_B W}{W^T S_w W} \qquad (1.13)$$

S_B is called between class variance and S_w is called within class variance

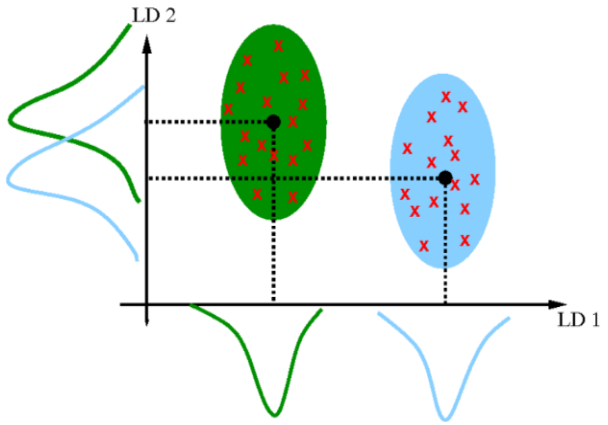


Fig.1.7. LDA Maximizing the Component Axes for Class Separation

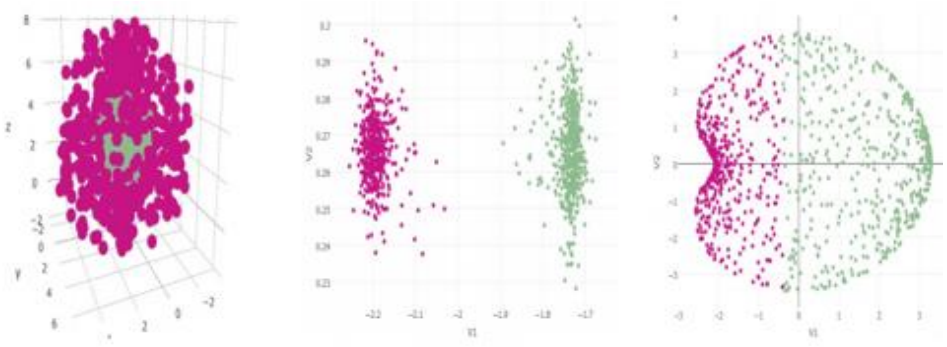
Non-Linear Methods

Not all the data in real time is linearly separable. Especially facial images show non-linearity. To handle non-linear data, non-linear methods are recommended. Few Examples for non-linear methods are discussed in the subsequent paragraphs.

- **Kernel PCA and Kernel LDA (KPCA and KLDA)**

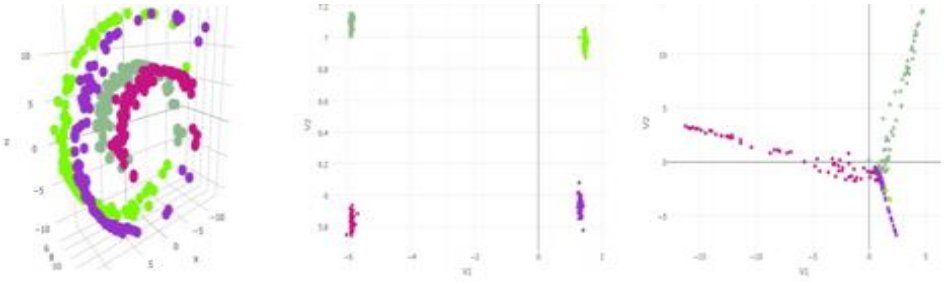
Both are extensions of conventional PCA and LDA that use kernel methods to project the data onto a high-dimensional feature space by using the "kernel trick" approach [18]. Both KPCA and KLDA are designed to handle the non-linearity in binary-class as well as multi-class problems as shown in (Fig.1.8) and (Fig.1.9). In KPCA and KLDA, the points

are projected to non-linear feature space instead of linear space using polar coordinates to deal with circle. After projecting onto nonlinear feature space, in KPCA principle components are extracted and KLDA projections are linearly separable with very good variation between the classes and also very less variation within the classes.



(a) (b) (c)

Fig.1.8. Binary class problem (a) Data (b) KLDA in 2D (c) KPCA in 2D



(a) (b) (c)

Fig.1.9. Multi class Problem (a) Data (b) KLDA in 2D (c) KPCA in 2D

Local Appearance-Based Feature Extraction

In local appearance-based technique, the face image is partitioned into smaller sub-image blocks and features are extracted from each block. The accuracy of face recognition rate increases when the image is divided into smaller blocks and this method is highly recommended especially for images under occlusion, variation in illumination and different pose conditions [23]. The holistic method is sensitive to local variation. Hence the local appearance-based method performs better than holistic method. Few important feature extractors are discussed below.

- **Histogram of Gradients (HoG)**

Descriptors are widely applied for human detection [67]. HoG is a local descriptor that uses a gradient vector orientation histogram as shown in (Fig.1.10). The feature extraction process can be classified into four steps: 1) Applying the processing technique like GIC to the data. 2) Partitioning the image into small areas called cells and for each cell, compute a histogram of gradient directions for the pixels within the cell. 3) Grouping cells to form blocks and 4) Computing a histogram of gradient orientations over the blocks [29]. HoG is mainly employed in object appearance and shape of the object is characterized by pixel gradient distribution.

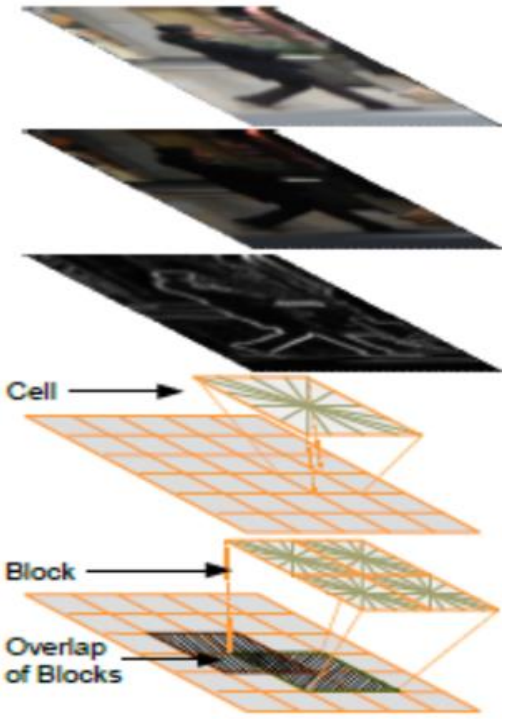


Fig.1.10. HoG Feature Extraction

Local Binary Pattern Feature Extraction (LBP)

It is a powerful texture descriptor. In LBP [27], each pixel of the input image is compared with its eight surrounding pixels in 3x3 neighborhoods by subtracting the middle pixel's value as shown in (Fig.1.11). The outcome of negative values is encoded with 0 and non-negative values are coded with 1. A binary number is obtained by combining all these binary codes in a clockwise direction starting from the top-left one and its corresponding decimal value is used for labeling [22]. The derived binary numbers are called LBP patterns or codes.

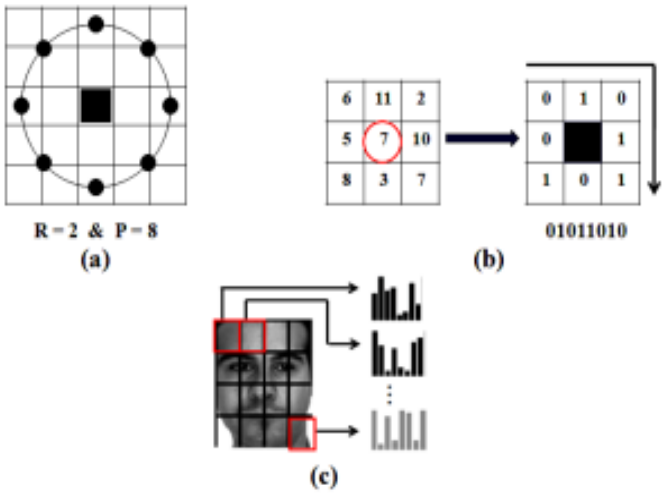


Fig.1.11. Different steps of LBP. (a) Radius and numbering of neighbors. (b) Comparison between each pixel and its neighbors. (c) Using binary codes to make histograms for image sub blocks.

- **Local Ternary Pattern (LTP)**

It is designed by Tan and Triggs [49] to handle the problem of LBP. LBP is affected by noise, since the threshold value of the operator is exactly present at the middle value [22]. To overcome this problem Tan and Triggs modified the LBP by updating with 3-value codes, which is called LTPs as shown in equation (1.14).

$$S(i_{n,}\, i_{n,}, t) = \begin{cases} 1, & i_n \geq i_c + t \\ 0, & |i_n - i_c| < t \\ -1, & i_n \leq i_c - t \end{cases} \qquad (1.14)$$

where t is a user-selected threshold parameter. The LTP codes are immune to noise, but not invariant to gray-level transformations. The problem with LTP is, selection of correct threshold value is difficult.

- **Local Phase Quantization (LPQ)**

It was devised by Ojansivu and Heikkila as a texture descriptor [51]. The design of LPQ is based on the blur invariance characteristics of the Fourier phase spectrum [50]. It employs the local phase data extracted by applying the 2-D short-term Fourier transform (STFT) calculated over rectangular surroundings at each pixel position of the image. In LPQ, only four complex coefficients are considered, relating to 2-D frequencies.

Hybrid methods

Hybrid methods apply combination of the local and holistic features to recognize a face. These methods have the capacity to provide better performance in comparison to independent holistic or local methods, since in-depth details could be employed for the recognition process [25].

Moments Based Feature Extraction

Moments are very good descriptors of the image's content. Moments are pure statistical measure of pixel distribution around the center of gravity of the image and they permit capturing global shape information [30]. They describe numerical quantities at some distance from a reference point or axis. Moments are often applied in statistics to represent the distribution of random variables. Invariant moment was originally introduced by Hu [61]. Hu's set of seven moment invariants have been extensively used in the field of pattern recognition but it exhibits redundant feature. Hence rebuilding the image is a challenging task [31]. Teague [32] suggested Zernike and Legendre orthogonal moments in image processing applications as a resolution to the redundancy problem involved in the non-orthogonal moments such as geometric moments and Hu's invariants. The orthogonality

characteristics allow us to apply the corresponding moments the property of minimum redundancy, meaning that various orders of moments represent various part of the image [28]. Although higher order moments represent minute details of an image, they are affected by noise badly. Therefore, the decision of optimal order of moments plays a vital role in image analysis. Both Zernike moments (ZM) and Pseudo-Zernike Moments (PZM) work better than geometric moments and complex moments. Zernike moments-based system is highly immune to noise in comparison with Legendre moments-based system. PZMs perform better than Zernike moments in the presence of noise. However, PZMs are computationally intensive in comparison to Zernike moments [33].

1.2.3 Classification

Extracted features of an image obtained in the feature extraction phase are fed to the classifier phase to compute the similarity measure between gallery and probe image. Based on the similarity, extracted features are assigned to a particular class. Different classifiers applied for the classification process are discussed below.

- **K-Nearest Neighbour (K-NN)**

It falls under the category of instance-based learning algorithms [34]. It is among the simplest of all machine learning algorithms It is a supervised approach for classifying objects based on the closest training examples in the feature space as shown in (Fig.1.12). It classifies the object based on majority vote of its surrounding neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor. The Euclidean distance metric [35] is often selected to make a decision

about the proximity of the data points in KNN. Euclidean distance is defined as the straight-line distance between two pixels.

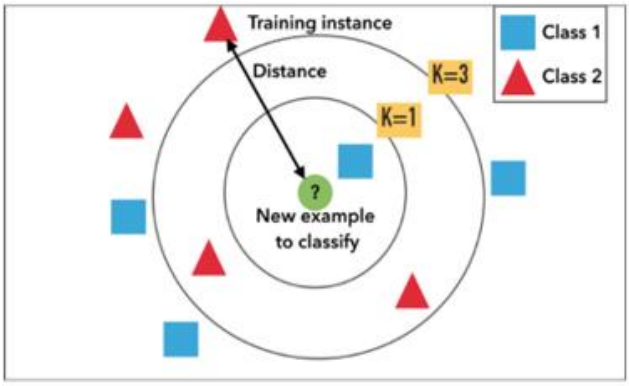


Fig.1.12. K-NN Classification

Advantages of K-NN Algorithm are described as follows.

1. No Training Period required. KNN is called Lazy Learner.

2. New data can be added seamlessly.

3. Easy to implement.

- **Support Vector Machine (SVM)**

It is designed to handle both binary and multi class pattern recognition problem [36]. SVM can minimize the classification error and maximize the geometric margin simultaneously. Hence SVM is called Maximum Margin Classifiers.

SVM model represents the examples as points in space mapped so that the examples of the separate categories are partitioned by a clear space that is as broad as possible.

For example, given a set of points belonging to either one of the two classes, the SVM discover a hyperplane having the greatest possible fraction of points of the same class on the same plane. This dividing hyperplane is called optimal separating hyperplane (OSH).

SVMs perform pattern recognition between two classes by choosing a decision line that maximizes the distance to the closest points in the training set which are known as support vectors. The separation of data using SVM classifier is illustrated in (Fig.1.13)

Even though there are many linear classifiers available for classifying the data, SVM maximizes the margin [37] and it is termed as optimal separating hyperplane (OSH).

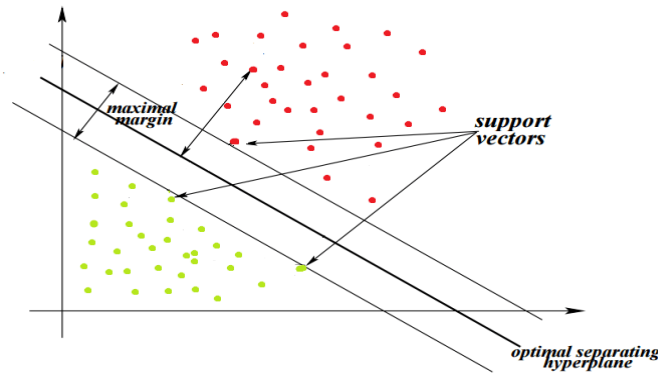


Fig.1.13. Separation of data using SVM.

Naïve Bayes Classifier (NB)

It is trained using the data of the training images in a supervised learning setting [38]. It is very simple but efficient probabilistic classifier. NB classifies the object based on the Bayes theorem. It predicts membership probabilities for each class such as the probability that given record or data point belongs to a particular class. The class with the highest probability is considered as the most likely class. This is also known as Maximum A Posteriori (MAP).

- **Ensemble Classifier**

The ensemble classifiers [63] are a set of classifiers whose individual decisions are fused by voting or averaging, to classify new examples as shown in (Fig.1.14).

There are two categories of ensembles, one is homogeneous and the other one is heterogeneous. In a homogeneous ensemble type, the same learning algorithm is implemented by each member of the ensemble and they are compelled to produce non-correlated results using different training sets. However, in heterogeneous ensemble, it merges more than one learning algorithms.

Various approaches have been proposed for constructing ensembles, such as bagging, boosting, error-correcting. Homogeneous ensembles with manipulation of input features is widely used [35].

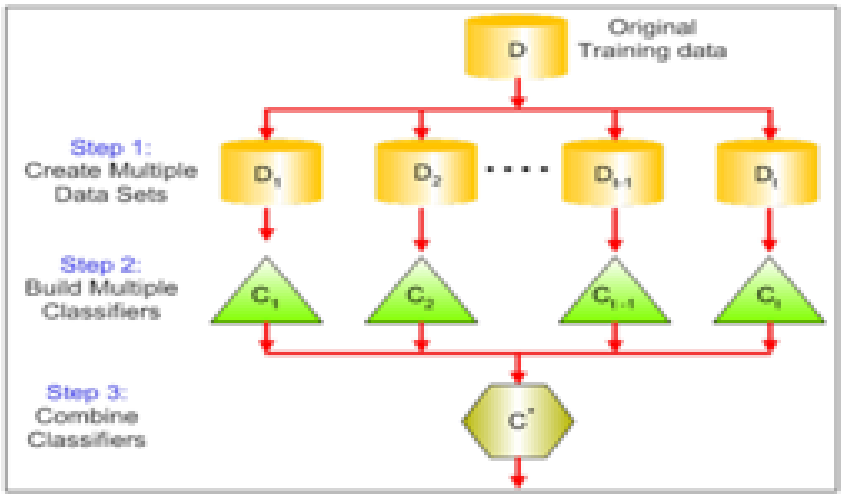


Fig.1.14. Ensemble Classifier

- **Chi Square Distance**

It is a nonlinear metric [64] and is extensively applied to compare histograms. There are multiple ways to classify the histogram. One of the best approaches is to apply the histogram similarity measure for the comparison of gallery and probe image blocks like intersection, log-like hood and Chi square distance [39]. Chi square distance is defined in equation (1.15) as follows.

$$\chi^2(p_b, g_b) = \sum_{j=1}^{p(p-1)+3} \frac{\left(p_{b_j} - g_{b_j}\right)^2}{p_{b_j} + g_{b_j}} \quad (1.15)$$

Where, $p_{b_j}$and $g_{b_j}$are j[th] histogram bins in the probe sub-block p_b and gallery sub-block g_b respectively.

- **Artificial Neural Network based Classifier (ANN)**

It is widely applied in the areas of pattern recognition and image processing [66]. ANN is a set of inter linked neurons that forms a network with input and the output in which weight is linked with each connection as shown in (Fig.1.15). It contains one input layer, one or more hidden layer and one output layer. It starts learning with initial value and learning of neural network is performed by modifying the weight of the connection. Weights gets modified on each iteration. By adjusting the weight repeatedly, the performance of network is improved. Based on the connection type, ANN can be classified into two types, feed-forward network and recurrent network [40].

The performance of the neural network is affected by various factors such as learning rule, architecture, initial parameter setting, weight, bias and transfer function. Neurons are activated by the weighted sum of input. The activation signal is then passed

through transfer function to create a single output of the neuron. Network's non linearity is achieved by this transfer function.

Generally, log-sigmoid and tang-sigmoid functions are commonly used as a transfer function in both hidden and output layers. Log-sigmoid function is used widely in face recognition applications.

There are two types of learning algorithms, supervised and unsupervised. Unsupervised learning model identifies the pattern class information heuristically and in reinforcement method of learning, the network learns by trial and error method [42]. There is no label available with unsupervised learning [44] and the machine tries to find interesting patterns in the data.

In the supervised type, the weights of the network are randomly initialized at the beginning and then gradually updated to obtain desired outputs for training data inputs. The difference between the target output and the actual output is called the error and it is computed for every input. The weights of the network are modified in proportion to the error factor. The entire process is repeated until the network error is decreased to minimum and the accuracy of the recognition reaches maximum [43].

Classification is a data mining method mainly applied for predicting object class. Classification is the process of applying a model to find unknown output variable by making use of known input values [40].

ANN with Back Propagation (BP) is proposed [41] for the classification of data. In BP, the error is propagated back to the input layer through the intermediate layer in the network to achieve the final target output. The gradient descent method is employed to compute the weight of the network and it modifies the weight of interconnections to

minimize the error. The error function is defined in equation (1.16) as follows [45]. The back-propagation algorithm is applied to train the neural network to classify the image.

$$E = \frac{1}{2}\sum_k (T_k - A_k)^2 \qquad (1.16)$$

T_k and A_k are known as target and actual output values and k is the output neuron.

In ANN, the number of neurons should be properly selected to get higher accuracy rates. When the number of neurons is low, the error will be high and as a result, low accuracy rate is obtained. When the number of neurons is high, it results in overfitting. The overfitted system performs very well with the training data and does not perform well on testing data [56]. Hence selection of correct number of neurons plays a vital role in network accuracy.

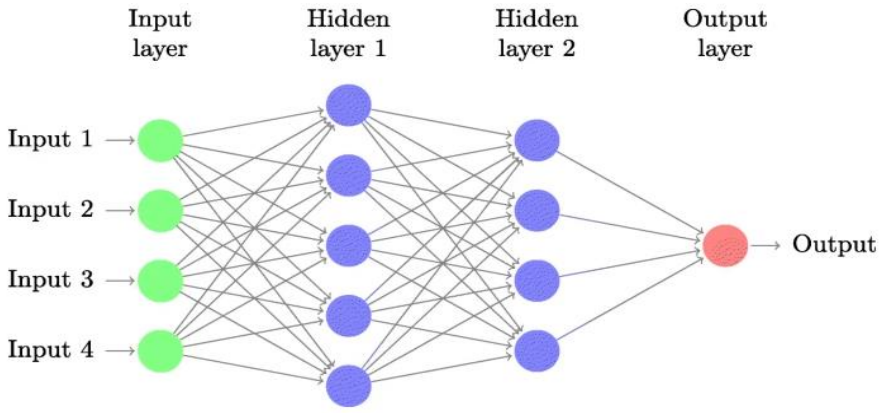


Fig.1.15. Artificial Neural Network

- **Random Forest (RF)**

It is an ensemble learning, in which the multiple week classification methods are combined to produce strong classifier with better accuracy [62]. RF can be applied for both classifications and regression problems. Random forest is a collection of huge number of separate decision trees that perform as an ensemble. Each individual tree in

the random forest spits out a class prediction and the class with the majority voting will be selected as model's prediction. The classification prediction is determined the majority votes of the predicted values using the Decision Trees [46]. There are many advantages with Random Forest Classifier.

(i) Prediction rate is high (ii) Training period is short (iii) Random Forest can handle the imbalanced datasets very well.

- **Hidden Markov Model (HMM)**

It is based on statistics [65]. It has two processes and it is applied in modeling sequences of events in order. HMM has wide application in speech recognition. The HMM consists of states, where the probability of transition from one state to another is based only on two states [48] and does not depend on any other state [47]. Each state in the HMM is represented as a part of the sequence. Thus, each state can be treated as probability density function model with its own parameters. The Gaussian mixture (GM) modeling is often used for each state.

1.2.4 Feature Matching:

In this step, features extracted from the input image during the feature extraction step are compared with features of known faces stored in a specific database for the matching to arrive at final decision.

1.3 Face Recognition Applications

Face Recognition has extensive applications in the areas of surveillance, border security, identification of individuals, police investigation and advertising [4].

Law Enforcement: Criminal Investigation, Video Surveillance, Missing Children Identification.

Access Control: Information Security, Data Privacy.

Biometrics: Travel document, Border Control, Driver's Licenses, Forensic Applications.

Security: ATM machines, Airport Checkpoints.

1.4. Face Recognition Challenges

The face recognition task becomes challenging due to illumination conditions, occlusion, pose, similar faces, aging, low-resolution facial expression variations and availability of quality database.

1.4.1 Pose Variation

Pose variations in an image is an important factor, which affects the face recognition. The pose of a face changes with head movements, angle of rotation and viewing angle [2]. The large variation in these parameters poses a serious challenge for recognition of a facial image and results in failure in facial recognition process. The (Fig.1.16) below shows the images with different poses from Indian Face Database

Fig.1.16. Images with different poses from Indian Face Database (Females)

1.4.2 Illumination Changes

The variation in illumination is one of the main challenging problems which affect the performance of the face recognition system. If the lighting conditions present in the gallery image is different from the probe image, then the process of face recognition may completely fail. (Fig.1.17) shows the sample images from Extended Yale B database with different illumination conditions.

Fig.1.17. Images from Extended Yale B database with different illumination conditions.

1.4.3 Presence of Occlusion

When the facial image is not fully visible in face recognition system, then it is called occlusion. This is due to presence of various structuring elements such as sunglasses, spices, beard and scarf. In real time, it is not possible to gather a large-scale training dataset with every possible occlusion scenario to apply facial recognition techniques. Therefore, the problem of face recognition under occlusions poses a great challenge in face recognition. The images under occlusion from AR database are shown in (Fig.1.18).

Fig.1.18. Images from AR Database with occlusion

1.4.4 Aging

Another important parameter that affects face recognition is ageing process. Since the appearance changes with time, the pictures captured in different phases of time have impact in face recognition. The (Fig.1.19) shows the appearance changes in images from FG-NET ageing database.

Fig.1.19. Images with appearance changes taken in different phases of time from FG-NET ageing database.

1.4.5 Similar Faces

Different individuals may have an identical look. It is self-evident that the identical twins have the similar face shape, size, and features. Sometimes it is difficult to differentiate people with similar appearance. The (Fig.1.20) shows the sample images from the ND-TWINS-2009-2010 Dataset. The images of people with identical appearances add complexity in facial recognition process.

Fig.1.20 Identical Images from ND-TWINS-2009-2010 Dataset

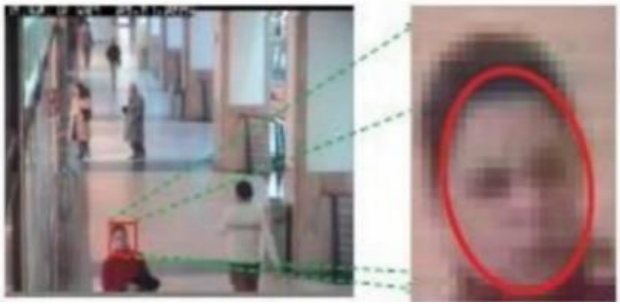
1.4.6 Facial Expression Variation

Appearance of the face changes due to the facial expressions like anger, disgust, fear and happiness affects the accuracy of face recognition process. Fig. (1.21) shows the sample images with different facial expression from The Japanese Female Facial Expression (JAFFE) database.

Fig.1.21 Images form JAFFE database with different facial expression

1.4.7 Low Resolution

The problem of low-resolution arises, when the image is captured in the uncontrolled manner. The quality of image acquisition device also plays a major role in the resolution of the image. The low-resolution is very common in surveillance applications. In such applications, the user is not aware of being captured by the system. In addition to that, the camera may not be kept in an appropriate position, which results in poor quality of the image as shown in (Fig.1.22). The images may suffer due to low-resolution problem, especially when the resolution of the image is below 10*10.



(a) (b)

Fig.1.22. (a) Surveillance Video (b) Face Region

1.4.8 Availability of the Quality Database

For successful face recognition process, a complete reliable, accurate and high-quality database is necessary. Complete database should contain images with illumination and expression variation and an accurate database should have a collection of facial images of same persons with different ages [13].

1.5 Research Objective and Contributions

The objective of this research is to apply different type of preprocessing techniques to both traditional and modern face recognition methods to increase their accuracy rate.

In traditional method, block-based classification approach has been applied for facial recognition. To improve the accuracy rate of face recognition, a local entropy-based adaptive-weight Local Binary Pattern (LBP) is used. A comparative analysis of elliptical and rectangular cropping has been integrated into the methodology for analyzing the recognition rate in different cropping scenarios. Local Binary Pattern-based feature extraction is used to extract unique features from multiple sub-image blocks from the face Region of Interest (RoI) image blocks. K-Nearest Neighbor (K-NN) classifier is applied to classify individual blocks. A Local entropy principle is used to assign weights to classifier outputs from individual sub-blocks to give more weightage to relevant areas in the face image. The local entropies of individual blocks are calculated to assess the contribution of each region towards the final face recognition accuracy. Finally, a weighted majority voting scheme is used for decision fusion. Highest and lowest accuracy rates are computed in FERET database for both rectangular and elliptical cropping. Rectangular cropping performs better than elliptical cropping. The point at which, elliptical cropping completely fails is also analyzed.

To increase the accuracy rates of face recognition further, preprocessing methods are applied to both whole image and block sized images to improve the performance of face recognition. In holistic method, preprocessing methods like Locally Tuned Inverse Sine Nonlinear (LTISN), Self-Quotient Illumination (SQI), Gamma Intensity Correction (GIC),

Histogram Equalization (HE), Contrast-Limited Adaptive Histogram Equalization (CLAHE) and Difference of Gaussian (DoG) are applied to the complete image (holistic) and then the images are block sized. Subsequently, K- Nearest Neighbor (K-NN) classifier and weighted entropy-based fusion method are applied. In the block-based method, input image is divided into non-overlapping blocks. Then above-mentioned preprocessing techniques are applied to each block. Subsequently K-NN classifier and weighted entropy-based fusion method are applied to the output of each classifier. The recognition accuracy is computed for various block sizes and threshold values. This thesis presents comparison of various pre-processing methods applied to both the holistic and block-based methods to mitigate the effects of illumination variation and preserve the essential details that are needed for recognition. The applied preprocessing methods improves the face recognition accuracy on FERET dataset under lighting variations. The block-based weighted entropy-based fusion method performs better than holistic method in FERET database.

In the modern approach, deep CNN model has been suggested for face recognition. CNN architecture is employed to extract unique facial features and softmax classifier is applied to classify images in CNN. The experiments conducted in Extended YALE B and FERET databases showed that the deep CNN model in Extended Yale B has improved the face recognition accuracy. To further improve the accuracy rates of CNN model, preprocessing techniques like Locally Tuned Inverse Sine Nonlinear (LTISN), Self-Quotient Illumination (SQI), Gamma Intensity Correction (GIC), Histogram Equalization (HE), Contrast-Limited Adaptive Histogram Equalization (CLAHE) and Difference of Gaussian (DoG) are applied to images. After the application of preprocessing techniques, the accuracy rates are improved for both the models.

The chapters are organized as follows. In chapter 2, block-based weighted entropy-based LBP approach with both elliptical and rectangular cropping for human face recognition is discussed. In chapter 3, the effect of applying various preprocessing techniques in block-based weighted entropy-based fusion method on face recognition is evaluated. In chapter 4, the effect of applying various preprocessing techniques in deep CNN model for face recognition is analyzed. Finally, the work is concluded in chapter 5.

References

1. M.S Kumar and D.E Walia “Various Biometric Authentication Techniques: A Review” International Journal of Computer Science and Information Technologies, Vol. 2 (4), pp 1595-97, 2011.

2. S. Ohlyan, Ms. S. Sangwan and T. Ahuja “A Survey on Various Problems & Challenges in Face Recognition” International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 6, pp 2533-2538, June 2013.

3. Calvo G., Baruque and B. Corchado E. Study of the Pre-processing Impact in a Facial Recognition System. In: Pan JS., Polycarpou M.M., Woźniak M., de Carvalho A.C.P.L.F., Quintián H., Corchado E. (eds) Hybrid Artificial Intelligent Systems. HAIS 2013. Lecture Notes in Computer Science, vol 8073. Springer, Berlin, Heidelberg 2013.

4. R. Jafri and H. Arabnia “A Survey of Face Recognition Techniques” Journal of information processing society Vol 5, no. 2, Korea information processing society pp 41-58, June 2009.

5. S. Arigela and V. K. Asari “self-tunable transformation function for enhancement of high contrast color images” Journal of Electronic Imaging Volume 22, Issue 2, 2013.

6. H. Wang, S. Z. Li, Y. Wang and J. Zhang, "Self quotient image for face recognition," International *Conference on Image Processing, ICIP '04.*, Singapore, pp. 1397-1400 Vol.2, doi: 10.1109/ICIP.2004.1419763, 2004.

7. H. Wang, S.Z. Li and Y. Wang "Face Recognition Under Varying Lighting Conditions Using Self Quotient Image" Proc. of IEEE International conference on Automatic FGR (AFGR), pp 819- 824,2004.

8. S. Anila, K. Arulsukanya, K. Divya, C. Kanimozhi, and D. Kavitha “An efficient Preprocessing Technique for Face Recognition under difficult Lighting Conditions,” in Proc. National

Conference on Emerging Trends in Computer Communication and Informatics (ETCCI-2011), March 10-11, 2011.

9. V. Karthikeyan, K. Vijayalakshmi and P. Jeyakumar "An Efficient Method for Face Recognition in Various Assorted Conditions" Biometrics and Bioinformatics, 4(12) ,543-547,2012.
10. S. A. Amiri and H. Hassanpour "A preprocessing approach for image analysis using gamma correction" International Journal of Computer Applications, Volume 38, Issue 12, Pages 38-46 January 2012.
11. P.A Koringa, S.K Mitra and V.K Asari "Handling Illumination Variation: A Challenge for Face Recognition" In: Raman B., Kumar S., Roy P., Sen D. (eds) Proceedings of International Conference on Computer Vision and Image Processing. Advances in Intelligent Systems and Computing, vol 460. Springer, Singapore, 2017.
12. S.Anila and N. Devarajan. "Preprocessing technique for face recognition applications under varying illumination conditions." Global Journal of Computer Science and Technology 12, no. 11-F (2012).
13. J. I. Olszewska Automated Face Recognition: Challenges and Solutions, Pattern Recognition - Analysis and Applications, S. Ramakrishnan, Intech Open, December 14th 2016.
14. M. Nishiyama T. Kozakaya, and O. Yamaguchi "Illumination Normalization using Quotient Image-based Techniques," Recent Advances in Face Recognition, I-Tech, Vienna, Austria, pp. 236, December 2008.
15. S. Muthu Selvi and P.Prabhu "Digital Image Processing Techniques – A survey" Golden Research Thoughts Journal, Vol 5, Issue 11, pp 1-7, May 2016.
16. S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Sukkur, pp. 1-10, 2018.

17. S. J. Chen, S. Z. Zheng, Z. G. Xu, C. C. Guo, X. L. Ma, “An Improved Image Matching Method Based on Surf Algorithm‟ International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. Vol. 42 Issue 3, pp 179-184, Beijing, China, 2018.
18. K. Kempfert, Y. Wang, C. Chen and S. W. Wong “A Comparison Study on Nonlinear Dimension Reduction Methods with Kernel Variations: Visualization, Optimization and Classification”. Intelligent Data Analysis, pp 1 – 24, October 2019.
19. Y. Kortli, M. Jridi, A.A. Falou and M. Atri “Face Recognition Systems: A Survey.” *Sensors (Basel, Switzerland)* 20 (2020).
20. R. Shobha Patil and S. Pandey “Principal Component Analysis” Global Journal of Mechanical Engineering and Computational Science, Vol.1 (3) pp.25-29, 2011.
21. Mr. S. Omprakash and S. Gokila “Principal Component Analysis - A Survey” International Journal of Advanced Research in Computer and Communication Engineering Vol. 7, Issue 8, pp 63-66, August 2018.
22. D. Huang, C. Shan M. Ardebilian, Y. Wang, and L. Chen “Local Binary Patterns and Its Application to Facial Image Analysis: A Survey” IEEE Transactions on Systems, Man and Cybernetics, part C, volume 41, pp 765-781, 2011.
23. H. K. Ekenel and R. Stiefelhagen, "Block Selection in the Local Appearance-based Face Recognition Scheme,” Conference *on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, New York, NY, USA, pp. 43-43, 2006.
24. S. Karamizadeh, S. Abdullah and S.M.Zamani “An Overview of Holistic Face Recognition” International Journal of Research in Computer and Communication Technology, Vol 2, Issue 9, pp 738-741, September 2013.
25. X. Tan, S. Chen, Z. Zhou and F. Zhang “Face Recognition from a Single Image per Person: A Survey” Pattern Recognition, Volume, pp 1725-1745, September 2006.

26. A. Martnez and A. Kak “PCA versus LDA” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No.2, pp. 228-233, 2001.
27. T. Ahonen, A. Hadid, and M. Pietikainen, “Face recognition with local ¨ binary patterns,” in Proc. Euro. Conf. Computer Vis, pp. 469–481, 2004.
28. G. A. Papa Kostas, D.E. Koulouri Otis, E.G. Karakasis and V. Tourassis “Moment Based Local Binary Pattern: A Novel Descriptor for Invariant Pattern Recognition Applications” Neurocomputing, Vol. 99 No. 1, pp. 358-371 Jan 2013.
29. J. Cruz, E. Shiguemori, L. Guimarães “ A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery” Journal of Computational Interdisciplinary Sciences, Volume: 6, pp 121-136, 2016.
30. R. C. Gonzalez and R. E. Woods, “Digital Image Processing”, Vol. 1, Parson, 2008.
31. P. Rao and P. Kumar “Feature Extraction Using Zernike Moments”, International Journal of Latest Trend in Engineering and Technology, vol. 2, no. 2, pp 228-234, Mar 2013.
32. M. R. Teague. “Image Analysis via the General Theory of Moments*.” *JOSA*, Optical Society of America, Vol. 70, No.8, 1, pp. 922-930, Aug 1980.
33. C.W. Chong, P. Raveendran and R. Mukundan “The Scale Invariants of Pseudo-Zernike Moments.” Pattern Analysis & Applications, vol. 6, no. 3, pp. 176–184, Jan 2003.
34. M. Kaur Dhriti and “K-Nearest Neighbor Classification Approach for Face and Fingerprint at Feature Level Fusion”, International Journal of Computer Applications Volume 60, pp 13-17, 2017.
35. C. Martínez and O. Fuentes “Face Recognition Using Unlabeled Data” Computer Systems, Vol. 7 No. 2 pp. 123 – 129, 2003.
36. M. O. Faruqe, M. A. Hasan “Face Recognition Using PCA and SVM” 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, pp. 97-101, 2009.

37. V. Anuja Kumari and R. Chitra “Classification of Diabetes Disease Using Support Vector Machine” International Journal of Engineering Research and Applications, Vol. 3, Issue 2, pp.1797-1801, March -April 2013.

38. A. Devi, K.K. Gaurav and H. Gupta “Naive Bayes Classification Based Facial Expression Recognition with Kernel PCA Features”, International Journal of Engineering Development and Research, Volume 5, Issue 3. pp 326-330, 2017.

39. S. Nikan and M. Ahmadi, "Human face recognition under occlusion using LBP and entropy weighted voting," *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, pp. 1699-1702, Tsukuba 2012.

40. R. Bala and D.D Kumar **“**Classification Using ANN: A Review” International Journal of Computational Intelligence Research Volume 13, Number 7, pp. 1811-1820, (2017).

41. H. Bischof, W. Schneider, and A. J. Pinz, “Multispectral classification of Landsat- images using neural networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp 482–490, 1992.

42. R. Sathya and A. Abraham “Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification”, International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No. 2, pp 34-38, 2013.

43. N. Kannathal, U.R. Acharya, C. M. Lim, P. Sadasivan and S. Krishnan Classification of cardiac patient states using artificial neural networks. *Exp Clin Cardiol*, 8(4): pp 206-211, 2003.

44. I. Muhammad and Z. Yan “Supervised Machine Learning Approaches A survey”, ICTACT Journal on Soft Computing, Vol. 05, Issue. 03, pp. 946-952, 2015.

45. T. Lee “Back-propagation neural network for long-term tidal predictions” Journal of Ocean Engineering, Volume 31, Issue 2, Pages 225-238, February 2004.

46. P. A. Resende and A. C. Drummond. “A Survey of Random Forest Based Methods for Intrusion Detection Systems.” *ACM Computing Surveys (CSUR)* 51 pp 1 – 36, (2018).

47. L. Rabiner “A tutorial on Hidden Markov Models and selected applications in speech recognition”, Proceedings of the IEEE, Vol.77, No.2, pp.257-286, 1989.
48. Hidden Markov Model Toolkit (HTK), S. Young et http://htk.eng.cam.ac.uk/
49. X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” in Proc. Anal. Model. Faces Gestures, pp. 168–182, 2007.
50. V.Ojansivu and J. Heikkilä “Blur Insensitive Texture Classification Using Local Phase Quantization” 5099. pp 236-243. (2008).
51. S. Brahnam, L. Nanni, J. Shi and A. Lumini “Local Phase Quantization Texture Descriptor for Protein Classification” Proc. International Conference on Bioinformatics and Computational Biology. pp 159-165, (2010).
52. S Latifi and N Solayappan “A Survey of Unimodal Biometric Methods” International Conference on Security and Management, SAM, pp. 57-63,2006.
53. Zahraddeen Sufyanu, Fatma Susilawati Mohamad, Abdulganiyu Abdu Yusuf, Abdulbasit Nuhu Musa and Rabi’u Abdulkadir “Feature Extraction Methods for Face Recognition” International Review of Applied Engineering Research (IRAER) ISSN 2248-9967 Volume 5, Number 3 pp. 5658-5668, (2016).
54. D. Saez Trigueros, L. Meng and M. Hartnett “Face Recognition: From Traditional to Deep Learning Methods” J. Theory. Appl. Inf. Technol. 2018, 97, 3332–3342.
55. A. Reza. "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement." Journal of VLSI Signal Processing, vol. 38, pp 35-44, 2004.
56. I. Nusrat and S Jang. “A Comparison of Regularization Techniques in Deep Neural Networks.” Symmetry 10 (2018): 648.
57. B. S. Min, D. K. Lim, S. Kim and J. H. Lee “A Novel Method of Determining Parameters of CLAHE Based on Image Entropy” International Journal of Software Engineering and Its Applications Vol.7, No.5, pp.113-120, (2013).

58. B. Vinoth Kumar “Gamma Correction Technique Based Feature Extraction for Face Recognition System”, International Journal of Computational Intelligence and Informatics, Vol. 3: No. 1, pp 20-26, April - June 2013.
59. M.Biglari, F. Mirzaei, H and Ebrahimpour-Komeh “Illumination invariant face recognition using SQI and weighted LBP histogram”. Pattern Recognition and Image Analysis (PRIA), 2013 First Iranian Conference on. IEEE, pp 1-7, 2013.
60. John B. Zimmerman, Steve B. Cousins, Karin M. Hartzell, Mark E. Frisse, and Michael G. Kahn “A Psychophysical Comparison of Two Methods for Adaptive Histogram Equalization” Journal of Digital Imaging, Vol 2, No 2, pp 82-91, (May), 1989.
61. M. Hu, "Visual pattern recognition by moment invariants," in *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, February 1962.
62. X. Feng, Z. Xiao, B. Zhong, J. Qiu and Y. Dong “Dynamic ensemble classification for credit scoring using soft probability”. Applied Soft Computing 65, 139–151, 2018.
63. N.K. Oza and K. Tumer, “Classifier ensembles: Select real-world applications.” Journal of Information Fusion, 9, pp. 4-20, 2008.
64. W.Yang, L. Xu, X. Chen, F. Zheng and Y.Liu “Chi-Squared Distance Metric Learning for Histogram Data” Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2015, Article ID 352849, pp 1-12, 2015.
65. F. H. Alhadi, W. Fakhr and A. Farag Hidden Markov Models for Face Recognition ASTED International Conference on Computational Intelligence, Calgary, Alberta, Canada, July 4-6, 2005.
66. D. Patterson, 1996. “Artificial Neural Networks, Theory and Applications”, Singapore: Prentice Hall.
67. T. Kobayashi, A. Hidaka and T. Kurita “Selection of Histograms of Oriented Gradients Features for Pedestrian Detection” In: Ishikawa M., Doya K., Miyamoto H., Yamakawa T.

(eds) Neural Information Processing. ICONIP 2007. Lecture Notes in Computer Science, vol 4985. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69162-4_62.

68. O. Saini and S. Sharma "A Review on Dimension Reduction Techniques in Data Mining" International Journal of Pattern Recognition and Artificial Intelligence Vol. 33, No. 10, 1950017 (2019).

Chapter 2

Performance Evaluation of Entropy Based LBP for Face Recognition

2.1. Introduction

Face recognition is one of the most popular biometric techniques used extensively in the areas of border security, access control, personal security, criminal investigation, identity checks, and law enforcement [25]. The face recognition system has the advantage of being a passive, non-intrusive system for checking the individual's identity [8].

The precision of facial recognition is affected by multiple factors such as occlusion, degradation, illumination and pose variation. It would be challenging to recognize a face, if some portion of the face is under occlusion. For example, sunglasses and specs can partially or fully hide the eyes, earrings and hair can hide ears, scarfs can hide the portion of the face, mustaches and beard of boys can affect the appearance of the face. These factors can affect the overall performance of the face recognition system [14].

There are three approaches in face recognition technique based on their detection and recognition method [26]. The first one is holistic, which makes use of global information to extract feature vector. The feature vector extracted from other regions of face such as shoulders and background will create a huge data for processing and creates unnecessary burden for the classifier part [1]. The second approach is local based method in which the image is broken into smaller blocks and features are extracted from each block.

The accuracy of face recognition rate increases when the image is divided into smaller blocks and this method is highly recommended, specially, for images under occlusion, variation in illumination and different pose conditions [23]. Even though the performance is good, the image's partitions in a block-based scheme produces partitioning effect, where each block creates dark edges that affect the quality of the image during image rebuilding phase [2]. The third approach is model-based approach in which 3D facial model can be acquired. The common image acquisition method is applying infrared as an input, which projects the laser beam onto an object and records the reflection results in the best and accurate 3D model recognition [3]. The problem with the model-based approach is that the construction of face model is a difficult task.

To overcome the effect of occlusion, Y. Deng, Q. Dai, and Z. Zhang [10] presented the work by applying a Graphical model and Laplace equation to whole images. Wright [9] proposed sparse representation. Kanan, Faez, and Gao [16], suggested Adaptively Weighted Patch PZM Array (AWPPZMA) to handle the problems of both controlled and varying lighting conditions, single image per person, different facial expressions and partial occlusion towards recognizing frontal face images. This method produced better results in less computational time [16]. A local probabilistic subspace method was suggested by Martinez [17] for occluded face recognition with one sample per subject. Xiao yang Tan et al. presented work in single image per person [21].

In the block-based approach, the accuracy rate increases when the individual blocks are assigned with proper weights. In real time scenario, only certain portion of image is affected by occlusion but not the whole image. To overcome the problem of occlusion, an

effective approach is to assign the affected image blocks with lower weights compared to normal blocks [19].

The weights of individual blocks can be computed using different methods. It is possible to calculate the weights of image blocks using Fishers Linear Discriminant (FLD). This method is similar to PCA, but it is supervised machine learning algorithm, which computes the directions that will represent the axes that maximize the separation between various classes [21]. This approach is used to compute the discrimination power of each block. But the problem with this approach for the construction of the subspace is, it requires comparatively bigger facial database for better outcome [18].

Nabatchian, Abdel-Raheem and Ahmadi proposed Weighted Voting Scheme for recognition of faces with illumination variation using multi class SVM classifier. This method provides very high accuracy rate, close to 100%. [5]. Zhao Ruzhe, Fang Bin, Wen Jing suggested an adaptive-weight approach [20]. This method applies the principle of variance to compute the weights of image blocks. The accuracy rate of face recognition is considerably high in this approach. On the other hand, this variance based weighted method takes huge time for computing the weight of image sub blocks, particularly when the number of training facial images is high [19].

Nikan and Ahmadi proposed local entropy-based weight computation of image sub blocks in Human Face Recognition under Occlusion using LBP and Entropy Weighted Voting [18]. This method assigns sufficient weights to image blocks according to the information property of image blocks.

2.2 Local Binary Pattern Descriptor

The LBP operator assigns labels to the pixels of an input image with decimal numbers, known as LBP codes, which encode the local structure around each pixel. Each pixel of the input image is compared with its eight surrounding pixels in 3x3 proximity by subtracting the middle pixel's value. The outcome of negative values is encoded with 0 and non-negative values are coded with 1. A binary number is obtained by combining all these binary codes in a clockwise direction starting from the top-left one and its corresponding decimal value is used for labeling. The derived binary numbers are called LBP patterns or codes, which is illustrated in (Fig 2.1). The reason behind using the LBP features is that the face images can be viewed as constitution of micro-patterns which are not variants to monotonic grey scale transformations. Unifying these micro-patterns, results in a global description of the face image [22]. One constraint with basic LBP operator is that its tiny 3x3 proximity could not capture vital features with large scale structures. To deal with this issue, LBP operator with variation in neighborhood was introduced by Heikkila [11].

In the proposed modified method, initially, the face images are cropped to remove the background regions. To evaluate the performance variation, we used two types of cropping: A rectangular cropping and elliptical cropping. The elliptical cropping helps to crop the oval face region and may help to remove the irrelevant information from the face samples. After the rectangular or elliptical cropping, each face image Region of Interest (ROI) is divided into multiple non-overlapping sub-blocks or patches, and LBP features are extracted from each of that sub-block for classification. In this method, K-NN is used as classifier for assigning classification labels to the face patches. The classified image

blocks are then assigned with appropriate weights based on local image entropy according to the information property of each sub-block.

2.3. Proposed Modified Approach

The proposed modified algorithm in this chapter is given in (Fig. 2.2), which consists of six steps. (i) The face image (after rectangular or elliptical cropping) is divided into non-overlapping image blocks of size $m \times m$ pixels. (ii) LBP operator is applied on each image sub block to obtain local features of the input image. (iii) Image blocks are given as the input to K-NN classifier and they are classified separately by K-NN. (iv) In the weighting step, entropy of image sub blocks is computed to find out the information content as well as their contribution towards the final decision. (v) Decision fusion is achieved using a voting scheme to determine the majority vote. (vi) The final outcome is retrieved by defining a threshold**.** In case its value is smaller than the value of threshold, the classification process is performed again for the highest n ranks.

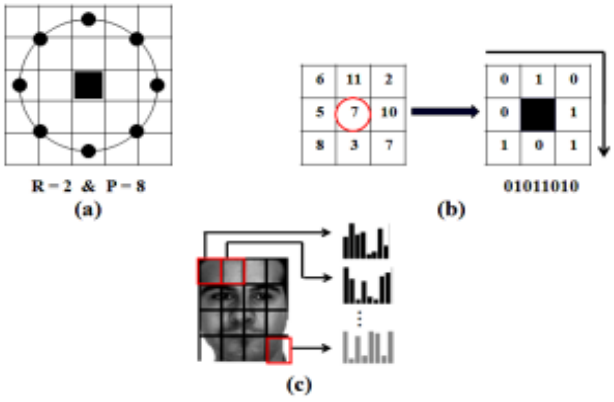


Figure 2.1. Different steps of LBP. (a) Radius and number of neighboring. (b) Comparison between each pixel and its neighbors. (c) Using binary codes to make histograms for image sub blocks.

The size of image blocks affects accuracy of the face recognition. Bigger block size does not yield better accuracy in block-based approach. On the other hand, smaller size adds the computational complexity. In this chapter we proposed $m \times m$ rectangular image blocks. If the image size is $R \times C$ pixels, then the number of image blocks can be computed as follows given in equation (2.1).

$$N = (R/m) * (C/m) \qquad (2.1)$$

Since the number of image sub blocks is greater than the contaminated blocks, the accuracy rate of recognition will be high in the democratic vote counting approach.

2.4 Classification

Features of each image block obtained in the last section are fed to the K-NN classifier to compute the similarity measure between gallery and probe sub-block. Based on the similarity, each block is assigned to a particular class.

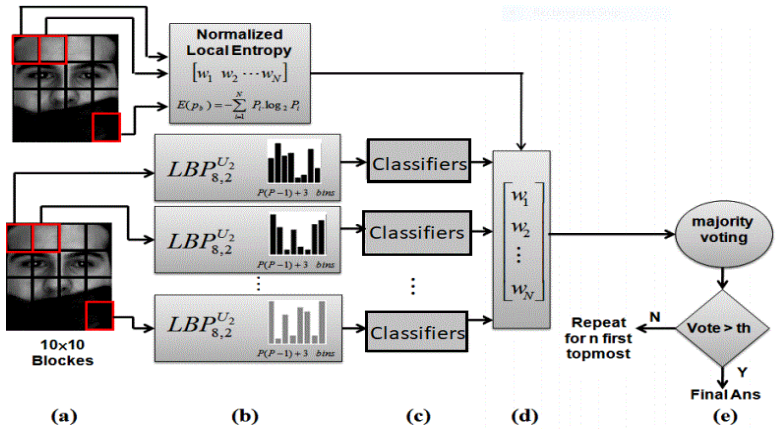


Figure 2.2. Suggested algorithm (a) Dividing the image. (b) Applying LBP operator on each block. (c) Classification of image blocks (d) Generating local entropy and summing the weighted votes. (e) Thresholding the majority votes to find final decision.

2.4.1 K Nearest Neighbour (K-NN)

K Nearest Neighbour is among the simplest of all machine learning algorithms. K-NN falls under the category of instance-based learning algorithms [24]. It is a supervised approach for classifying objects based on closest training examples in the feature space [4]. It classifies the object based on majority vote of its surrounding neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor [6].

The Euclidean distance metric is often selected to make a decision about the proximity of the data points in KNN.

Advantages of K-NN Algorithm are described as follows.

1.No Training Period required.

2. KNN is called Lazy Learner.

3. New data can be added seamlessly and it is easy to implement K-NN.

2.5 Local Entropy Weighting

The benefit of the block-based approach is its ability to reduce the effect of contaminated parts of image on the final decision by assigning lesser weights to the image blocks since the corrupted blocks are not providing effective contribution in final recognition. The proposed recognition system applies the local image entropy principle, which provides adequate weights to the image blocks according to the information property of each block and it is given as follows in equation (2.2). Where N_t is the number

$$E(p_b) = \sum_{i=1}^{N_t} P_i \, log_2(P_i) \qquad (2.2)$$

of levels in the probe image sub-block and P_i is the i^{th} gray value's probability in that block. The normalized entropy values are between 0 and 1 [7].

2.6 Majority Voting Scheme

In order to fuse the results of classifier and obtain the final class of the probe image, we need to choose a decision fusion method, which gives more accurate results in comparison with feature fusion method [12]. Weighted sum rule and majority voting scheme are two fixed rules integrated in the proposed approach, which are simple and do not demand training requirement [13]. In this chapter a weighted majority voting scheme is suggested, which is based on the local entropy, we sum the obtained weights of votes for each class. The final decision is the label of the majority vote. To obtain better accuracy rate, the maximum vote is compared with a threshold, which is achieved by trial and error basis. In case it is below the threshold value, the classification and majority voting process are repeated for the topmost number of votes. This method is not only simple, but also is not computationally expensive [7].

2.7 FERET Database

The FERET database was collected in 15 sessions between August 1993 and July 1996. The database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images as shown in (Fig.2.3). A duplicate set is a second set of images of a person already in the database and was usually taken on a different day [15].

Fig.2.3. images from FERET database.

2.7.1 Experiment

The actual FERET dataset having a resolution of 384x256 pixels and consist of images with both frontal and profile views. While conducting this experiment with elliptical cropping, images with frontal view are chosen. The images are initially resized to 128x192 for a better trade-off between accuracy and computation cost. Since the width of the image is up to 128 pixels, the values for the minor axis are selected within the range of 60 to 120. The ellipse with minor axis less than 60 are discarded because it couldn't represent most of the facial features. The major axis values are chosen as 1.2 times the minor axis length to keep a general aspect ratio of face. Elliptical binary mask was created corresponding to each set of minor and major axis with all values inside ellipse is '1' and all values outside elliptical curve is '0'. Then the value of binary mask is multiplied with the input images to make all pixels outside the elliptical region as '0'. Finally, the area inside elliptical regions are cropped and resized into a standard 128x128 pixel resolution and passed to the feature extraction part. The experiment is conducted for various block size and threshold values.

2.8. Results

The analysis carried out for a different number of sub-blocks per image. We considered the cases while dividing the face images into 4 blocks (2 blocks each across the row and column direction), 16 blocks (4 blocks each across the row and column direction), 64 blocks (8 blocks each across the row and column direction), and 100 blocks (10 blocks each across the row and column direction). The highest classification accuracy is observed when the image is divided into 100 sub-blocks. The accuracy obtained for elliptical cropping is 76.7% at a major and minor axis value of 90 and 108 pixels and with an entropy threshold of 2.5. For the same threshold value, the accuracy obtained for rectangular cropping is 84.1%.

Table 2.1 Recognition rate percentages (%) of the experiment with different block sizes and threshold values.

Block Size	**Threshold Values**				**Cropping Type**
	2.5	**5.0**	**7.0**	**9.5**	
10*10	84.1	80.0	78.5	74.3	Rect
	76.7	72.1	70.8	67.7	Elliptic
8*8	82.1	78.1	77.0	72.6	Rect
	67.9	64.8	60.7	57.0	Elliptic
4*4	77.7	75.1	73.6	72.4	Rect
	47.3	39.1	26.8	11.2	Elliptic
2*2	74.2	73.8	71.1	70.8	Rect
	41.0	29.8	13.9	9.4	Elliptic

The elliptical cropping completely fails, when the image is divided into 4 sub-blocks per face image. The lowest accuracy rate of 9.4% is observed, when the minor and major axis values are 60 and 72 at an entropy threshold of 9.5. For the same conditions, the accuracy of rectangular cropping is 70.8%. While dividing each face image into 64 blocks, it gives moderate performance as that of the previous two cases.

In general, the performance of the elliptical cropping and rectangular cropping increases with the number of sub-patches in the image.

In addition, the performance of elliptical cropping doesn't show any consistent relation with respect to the ellipse dimension.

The results are summarized in the Table 2.1 for various block sizes and threshold values.

2.9 Comparison with Other Approaches

The proposed method outperforms all the methods except Eigenfaces and DoG filter, which shows almost equal performance [27] compared to proposed method. The popular LeNet-5 architecture [31] in FERET database yields accuracy rate is 81.25% [29]. VanderLugt Correlator and Difference of Gaussian **(VLC+DoG)** approach yields 81.95% [27]. Histogram of Gradients and Manifold-manifold distance **(MMD)** method provides the accuracy rate of 68% [28]. Bag of Words **(BoW)** approach gives the accuracy rate of 82.3%. The comparison of the proposed method with various approaches is summarized in Table 2.2

Table 2.2 Comparison of Accuracy rates of different approaches

Method used	Database	Accuracy
Proposed method (Rectangular, Elliptical cropping)	FERET	84.1%, 76%
BoW [30]	FERET	82.3%,
HoG and MMD [28]	FERET	68%
VLC+DoG [27]	FERET	81.95%
Eigenfaces and DoG filter [27]	FERET	84.26%
CNN-LeNet [29]	FERET	81.25%

2.10 Conclusion

In this chapter we proposed entropy-based sub-image analysis face recognition with K-NN classifier. The classification performance is analyzed for both elliptical and rectangular cropping. In elliptical cropping, we tried different ellipses to crop the face ROI with minor axis ranges from 60 to 120 pixels. As results show, the rectangular cropping gives comparatively better results than elliptical cropping in all cases.

In general, the classification accuracy rate increases with the increase in the number of sub-blocks in case of rectangular cropping. It is because when there are large number of blocks, the percentage of good blocks which carry the relevant information (with high entropy weights) will be more. Since the blocks with irrelevant information are assigned with lesser entropy weights, the impact of such contaminated blocks will be less in the overall accuracy.

The selection of threshold value helps to choose the significant sub images in terms of entropy. The sub blocks will have high entropy values if it represents significant information. For corrupted or uniform background regions, entropy will be less. Hence the entropy threshold value can be selected to filter the insignificant sub blocks during the analysis. As the threshold increases, more sub blocks will be eliminated from voting scheme. The entropy threshold value is chosen empirically to retain the best possible accuracy in each case.

The proposed method performs relatively better compared to existing methods.

References

1. L. Chen, H. Liao, J. Lin and C. Han "Why Recognition in a Statistics-Based Face Recognition System Should Be Based on the Pure Face Portion: A Probabilistic Decision-Based Proof." *Pattern Recognition*, vol. 34, no. 7 pp. 1393–1403, 2001.

2. G. A. Papa Kostas, D.E. Koulouri Otis, E.G. Karakasis and V. Tourassis "Moment Based Local Binary Pattern: A Novel Descriptor for Invariant Pattern Recognition Applications" Neurocomputing, Vol. 99 No. 1, pp. 358-371 Jan 2013.

3. M. Naeem, I. Qureshi and F. Azam "Face Recognition Techniques and Approaches: A Survey", Sci. Int. (Lahore), 27(1), pp. 301-305, 2015.

4. M. Kaur and Dhriti "K-Nearest Neighbor Classification Approach for Face and Fingerprint at Feature Level Fusion", International Journal of Computer Applications Volume 60, pp 13-17, 2017.

5. A. Nabatchian, E. Abdel-Raheem, M. Ahmadi "A Weighted Voting Scheme for Recognition of Faces with Illumination Variation" ICARCV pp 896-899 2010.

6. V.Vaidehi, S. Vasuhi, R. Kayalvizhi, K. Mariammal, Raghuraman.M. B, Sundara Raman.V, Meenakshi.L, Anupriyadharshini.V, Thangamani.T "Person Authentication Using Face Detection" WCECS San Francisco, USA, pp. 222-224, 2008.

7. A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi. "Illumination invariant feature extraction and mutual information-based local matching for face recognition under illumination variation and occlusion. Pattern Recognition" 41(10-11) pp. 2576–2587, October 2011.

8. S. A. Nazeer, N. Omar and M. Khalid, "Face Recognition System using Artificial Neural Networks Approach," International *Conference on Signal Processing, Communications and Networking*, pp. 420-425, doi: 10.1109/ICSCN.2007.350774. *2007.*

9. C. A. Perez, L. A. Cament, L. E. Castillo “Local Matching Gabor Entropy Weighted Face Recognition”. *ICIG*, pp.179–184, August 2011.

10. Y. Deng, Q. Dai, and Z. Zhang. “Graph Laplace for Occluded Face Completion and Recognition”. *IEEE Trans. Image Process.*, 20(): pp.2329–2338, August 2011.

11. D. Huang, C. Shan, M. Ardebilian, Y. Wang, and L. Chen “Local Binary Patterns and Its Application to Facial Image Analysis: A Survey” IEEE Transactions on Systems, Man and Cybernetics, part C, volume 41, pp 765-781, 2011.

12. B. Topc¸ and H. Erdogan. “Decision fusion for patch-based face recognition”. ICPR, pp.1348–1351, August 2010.

13. F. Roli1, J. Kittler, G. Fumera, and D. Muntoni “An experimental comparison of classifier fusion rules for multimodal personal identity verification systems”. MCS, pp. 325–335, 2002.

14. S.Anwarul and S.Dahiya “A Comprehensive Review on Face Recognition Methods and Factors Affecting Facial Recognition Accuracy” In: Singh P., Kar A., Singh Y., Kolekar M., Tanwar S. (eds) Proceedings of ICRIC 2019. Lecture Notes in Electrical Engineering, vol 597. Springer, Cham (2020).

15. P. J. Phillips, S. A. Rizvi and H. Moon “The FERET Evaluation Methodology for Face-Recognition Algorithms” IEEE *Trans. Pattern Analysis. Machine Intelligence*, 22(10): pp. 1090–1104, Sep 2000.

16. H. R. Kanan, K. Faez, and Y. Gao. “Face recognition using adaptively weighted patch PZM array from a single exemplar image per person” *Pattern Recogni*tion 41(12): pp.3799–3812, 2008.

17. A. Martinez. “Recognizing imprecisely localized, partially occluded and expression variant faces from single sample per class”. *IEEE Trans. Pattern Anal. Mach. Intell*, 24(6): pp. 748–763, June 2002.

18. S. Nikan and M. Ahmadi. “Human Face Recognition under Occlusion using LBP and Entropy Weighted Voting” 21st International Conference on Pattern Recognition ICPR pp. 1699-1702, 2012.

19. J. Xiaoping, R. Guangxing and Z. Hua “Local entropy-based adaptive-weight LBP for face recognition” Volume 10 Issue 17 BTAIJ, 10(17), pp. 9532-9537, 2014.

20. Ruzhe Zhao, Bin Fang, Jing Wen; “Face recognition with single training sample per person based on adaptive weighted LBP” Computer Engineering and Applications, pp. 146-149, 2012.

21. X. Tan, S. Chen, Z. Zhou and F. Zhang “Face Recognition from a Single Image per Person: A Survey” Pattern Recognition 39(9) pp. 1725-1745, March 2006.

22. T. Ahonen, A. Hadid and M. Pietikäinen “Face Recognition with Local Binary Patterns” ECCV 2004, 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, Proceedings, Part I, 2004.

23. S. Nikan, M. Ahmadi “Modified technique for face recognition under degraded conditions” Journal of Visual communication and image Representation Vol. 55, pp. 742-755 2018.

24. C.Martínez and O.Fuentes “Face Recognition Using Unlabeled Data” Computer Systems, Vol. 7 No. 2 pp. 123 – 129, 2003.

25. N. S Devi and K. Hema Chandran, “Automatic Face Recognition System using Pattern Recognition Techniques: A Survey”, International Journal of Computer Applications (0975 – 8887), Volume 83 – No 5, pp 10-13, December 2013.

26. Y. Kortli, M. Jridi, A. A. Falou and M. Atri Face Recognition Systems: A Survey. *Sensors (Basel)*. 2020;20(2):342. Published on Jan 7. doi:10.3390/s20020342, 2020.

27. A. Ghorbel, I. Tajouri, W. Aydi and N. Masmoudi “A comparative study of Gabor Ordinal Measures, Uniform Local Binary Pattern, Vander Lugt Correlator and fractional Eigenfaces for face recognition” In Proceedings of the 2016 International Image Processing,

Applications and Systems (IPAS), Hammamet, Tunisia, 5–7 IEEE: Piscataway, NJ, USA, pp. 1–5. November 2016.

28. M. Karaaba, O. Surinta, L. Schomaker and M.A. Wiering. “Robust face recognition by computing distances from multiple histograms of oriented gradients”. In Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence, Cape Town, South Africa, 7–10 December 2015; IEEE: Piscataway, NJ, USA, pp. 203–209, 2015.

29. S. Ahmad Radzi, K.H Mohamad, S.S. Liew and R. Bakhteri “Convolutional neural network for face recognition with pose and illumination variation”. International Journal of Engineering and Technology (IJET), 6(1), pp.44-57, 2014.

30. Z. Cui, W. Li , D. Xu S.Shan and X. Chen “Fusing robust face region descriptors via multiple metric learning for face recognition in the wild”. In Proceedings of the IEEE conference on computer vision and pattern recognition, Portland, OR, USA, 23–28 June 2013; pp. 3554–3561, 2013.

31. R.Haridas and R.L. Jyothi “Convolutional Neural Networks: A Comprehensive Survey” International Journal of Applied Engineering Research Volume 14, Number 3, pp. 780-789, (2019).

Chapter 3

Performance Evaluation of Weighted Entropy Based Fusion Technique for Face Recognition with Different Pre-processing Techniques

---

3.1 Introduction

In recent years, face recognition has become one of the most attractive biometric techniques compared to other biometric applications. Face recognition is used widely in the areas of advertising, health care, virtual reality, border security, biometrics, access control, and criminal investigation [1]. In addition to that, it is also used as personal identification in driver's license and travel document.

Face recognition task becomes challenging due to illumination conditions, occlusion, pose and facial expression variations [4]. The variation in illumination is one of the main challenging problems which affect the performance of the face recognition system. Among all, shadowing effect, underexposure, and overexposure conditions are challenging problems that need to be addressed in the face recognition process [23]. If the lighting conditions present in the gallery image is different from the probe image, then the process of face recognition may completely fail [8]. A good face recognition system should be able to give accurate recognition rate under the different illumination conditions between images of the same face [2]. Image enhancement algorithms play a great role in handling the illumination variation.

The main aim of preprocessing is to remove features that obstruct the process of classifying the images of the same person (within-class differences), thereby boosting the difference of them with others (between-class differences) [14].

For better face recognition under uncontrolled and illumination variation conditions, the vital features responsible for differentiating two different faces require to be retained. The shadows produced in facial images due to variation in lighting directions may cause loss of important facial features which are helpful for recognition. A preprocessing method must enhance the intensity in the regions of inadequately illuminated and decrease the intensity in the densely illuminated regions while retaining the intensity in the illuminated portions [11].

Preprocessing techniques have wide applications in the areas of image analysis, object detection, background foreground subtraction, and medical image processing, measuring tissue volumes, and locating objects in satellite imaging, traffic control, biometrics and agriculture [3].

- **Histogram equalization (HE**)

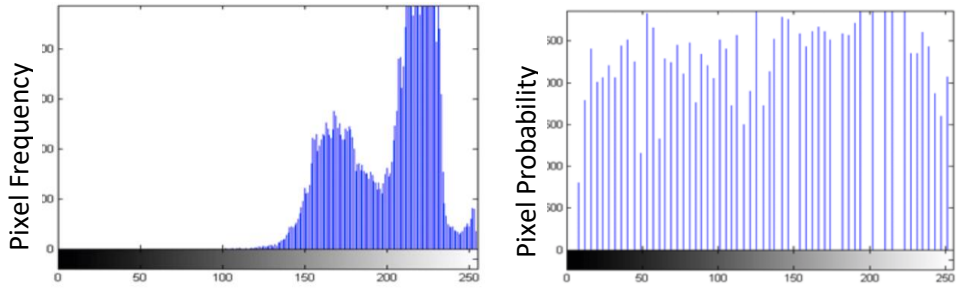
It is very effective and simple image processing technique [9] applied in spatial domain. The image histogram yields details about the intensity distribution of the pixels present in the image.

For image I(x, y) with discrete k. gray values histogram is defined by the probability of occurrence of the gray level I [39] is given by equation (3.1) as follows.

$$p(i) = n_i / N \qquad (3.1)$$

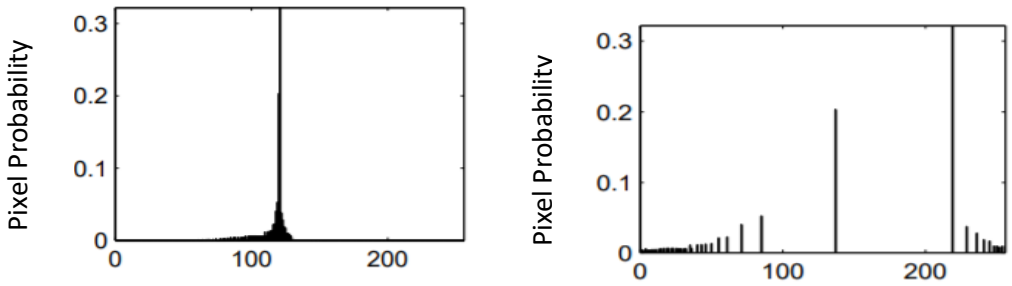
Where $i \in 0, 1 \ldots k-1$ grey level and N is total number of pixels in the image.

HE is the most popular and simplest global image processing technique applied in spatial domain. The image histogram yields detail about the intensity distribution of the pixels present in the image. It modifies image intensities to enhance the image contrast using histogram of image. HE improves the global contrast of the image.



(b) (b)

Fig.3.1. Histogram plot by 1st method (a) Before applying HE (b) After applying HE



(a) (b)

Fig.3.2. Histogram plot by 2nd method (a) Before applying HE (b) After applying HE

In Histogram equalization approach, the image intensities are adjusted to improve the image contrast. Histogram of an image can be constructed by two ways. (i) It can be constructed by plotting pixel intensities against the pixel frequencies as shown in (Fig.3.1).

(ii) It can also be constructed by plotting pixel intensities against pixel probabilities as shown in (Fig.3.2).

- **Contrast-Limited Adaptive Histogram Equalization (CLAHE)**

  It is recommended to handle non-uniform variation [27]. CLAHE can perform better than the standard HE. HE is based on the underlying assumption that the overall quality of the image is not varying, and single grayscale mapping provides uniform improvement for all regions of the image. But when there is a variation in allocation of grayscale values from one area to another, the assumption fails. In CLAHE, both block division of image and histogram clipping techniques are suggested to reduce the effect of noise amplification.

  The images treated with CLAHE will have a natural look. However, the low contrast enhancement of CLAHE may obstruct the capability of an observer to find the existence of some significant gray-scale contrast [26].

  HE is not suitable for the images affected by gamma distortion. Due to technical restrictions, many imaging equipment may not display the real image of objects [10]. This technical constraint is known as gamma distortion.

- **Gamma Intensity Correction (GIC)**

  It is an image enhancement technique, which can handle the problem of gamma distortion [25]. It enhances the local dynamic span of the image in darker or under exposed areas and at the same time, squeezes the brighter portions. Hence it changes the overall tone. of an image either lightly, or darkly, based on the value of gamma and at the same time, it maintains the dynamic range of the image. There are some

limitations in GIC technique as it does not eliminate the effects of shadows. In addition to that, choosing the correct gamma value is a little bit tedious process.

- **Difference of Gaussian (DoG)**

It is an effective preprocessing technique in mitigating the shadowing effect [8]. DoG is essentially acting as band-pass filter, which eliminates both high frequency components that comprises of noise and some low frequency components that constitutes the homogeneous areas in the image. The frequency components present in the pass band are essentially edges present in the images.

In DoG filter, image is convolved with two different Gaussian kernels. The output image is acquired by computing the difference of the two convolved images [12].

There are some drawbacks in DoG operation. DoG filtering diminishes the contrast of the image as a whole and hence the contrast of the image needs to be improved after the application of DoG filtering. Moreover, gamma normalization is a pre-requisite for DoG. If DoG is applied without gamma normalization, it may affect the local contrast in shadowed regions.

Most of the preprocessing approaches essentially concentrate on either global or local improvements which may not be suitable for all kinds of images. Especially these methods do not consider the nature of the image, whereas images corrupted with different kinds of degradation might require different types of treatments.

One of the suggested solutions for presenting high contrast image in a bounded contrast range equipment is to squeeze the image contrast without loss of data. Global preprocessing methods have been suggested to execute this idea. But there are some performance related problems associated with these methods. While applying the

methods, minute details are lost, or some regions remain un-enhanced [5]. Local contrast enhancement techniques retain or improve fine image details.

- **Locally Tuned Inverse Sine Nonlinear (LTISN)**

  It is suggested to handle the problem of limiting the dynamic range [11] as well as improving the local contrast. It can also handle severe illumination conditions using a series of operations and a nonlinear intensity transformation. It can enhance the intensity of pixels in the dark regions and at the same time reducing the intensity of pixels in the bright regions. In addition to that, it also improves the local contrast with a single image dependent parameter.

  The main pitfall of the LTISN technique is, over enhancement in under exposed regions creates black artifacts near the edges and does not recover back the color in over exposed regions.

- **Self-Quotient Illumination (SQI)**

  It is one of illumination invariant algorithm suggested for handling both shadow and lighting changes [13]. SQI algorithm is based on the reflectance-illumination model. Haitao Wang and others investigated about the application of self- quotient approach in handling illumination variation in face recognition [6] [7].

  The main disadvantages of SQI are, (i) is not having capacity to differentiate between shadow and low albedo area. As a result, critical information representing identity is ignored. (ii) Another problem in SQI is that partitioning the area, where pixel values change gradually is a difficult challenge [24].

3.2 Preprocessing Methods

The very first step in most face recognition systems as well as many other computer vision applications is preprocessing phase.

3.2.1 Histogram Equalization

It is necessary to improve the image quality an image that has high brightness or darkness, when images are taken in unconstraint environment. HE flattens the histogram of the image [38] and expands the dynamic range of the pixel intensity values by employing cumulative density function.

Histogram equalization normally applied to enhance the visual appearance of an image. It is most widely employed in contrast enhancement technique due to its simplicity and ease of implementation [32].

3.2.2 Self Quotient Image (SQI)

It is defined as the ratio of the input image and its smooth versions as given in equations (3.2) and (3.3).

$$Q(\mathrm{x},\mathrm{y}) = \frac{I(x,y)}{S(x,y)} \tag{3.2}$$

$$= \frac{I(x,y)}{F(x,y)} * I(x,y) \tag{3.3}$$

Where, *I (x, y)* is the face image and *S (x, y)* is a smoothed version of the image and $*$ is the convolution operation. F is the smoothing kernel which in this case is a weighted Gaussian filter and Q is the Self Quotient Image since it is derived from one image and has the same quotient form as that in the quotient image method [13].

3.2.3 Locally Tuned Inverse Sine Nonlinear (LTISN)

It is a nonlinear and pixel by pixel approach, where the improved intensity values are calculated by applying the inverse sine function with a tunable parameter based on the nearby pixel values [21] given in the equations (3.4), (3.5), (3.6), (3.7) and (3.8). The intensity range of the image is rescaled to [0 1], followed by a nonlinear transfer function.

$$I_{enh}(x,y) = \frac{2}{\pi} sin^{-1}(I_n(x,y)^{q/2}) \tag{3.4}$$

$$kernel_i(n_1,n_2) = \frac{h_g(n_1,n_2)}{\sum_{n_1}\sum_{n_2} h_g(n_1,n_2)} \tag{3.5}$$

$$h_g(n_1,n_2) = e^{\frac{-(n1^2+n2^2)}{2\sigma^2}} \tag{3.6}$$

$$I_{M,i}(x,y) = \sum_{m=-\frac{M_i}{2}}^{\frac{M_i}{2}} \sum_{n=-\frac{N_i}{2}}^{\frac{N_i}{2}} I(m,n) kernel_i(m+x, n+y) \tag{3.7}$$

$$q = \begin{cases} tan(\frac{\pi}{C_1} I_{M_n}(x,y)) + C_2 & I_{M_n}(x,y) \geq 0.3 \\ \frac{1}{C_3} ln(\frac{1}{0.3} I_{M_n}(x,y)) + C_4 & I_{M_n}(x,y) < 0.3 \end{cases} \tag{3.8}$$

3.2.4 Gamma Intensity Correction (GIC)

It is a nonlinear gray-level transformation that substitutes gray-level I [12] with the gray level $I^{\frac{1}{\Upsilon}}$, given by the equation (3.9).

$$I = I^{\frac{1}{\Upsilon}} \tag{3.9}$$

As shown in the (Fig.3.3), for the values of gamma less than 1.0 darkens the image and for the values of gamma greater than 1.0 lightens the image. When gamma value is 1.0, it does not produce any effect.

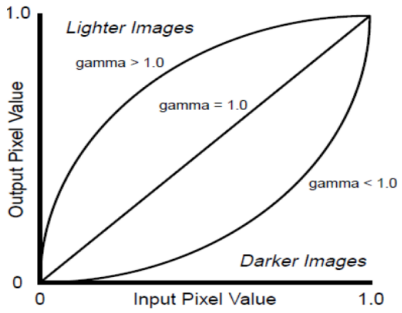


Fig. 3.3. Gamma Intensity Correction

3.2.5 Difference of Gaussian (DoG)

It is a grayscale image enhancement algorithm that involves the subtraction of one blurred version of an original grayscale image from another less blurred version of the original. The blurred images are obtained by convolving the original grayscale image with Gaussian kernels having differing standard deviations [8], which is given in the equation (3.10).

$$DOG(x,y) = \frac{1}{2\pi{\sigma_1}^2} e^{-\frac{x^2+y^2}{2{\sigma_1}^2}} - \frac{1}{2\pi{\sigma_2}^2} e^{-\frac{x^2+y^2}{2{\sigma_2}^2}} \qquad (3.10)$$

σ_1, σ_2 are Gaussian kernel widths respectively.

3.2.6 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

It works on small areas in the image, known as tiles, rather than the whole image [30]. Individual tile's contrast is improved. Hence histogram of the output area is roughly matching with the histogram specified by the distribution parameter. The tiles present in

the neighborhood regions are then joined by applying bilinear interpolation to minimize the effect of artificially induced border line [29]. In CLAHE, the image is divided into a limited number of regions and the same histogram equalization technique is applied to pixels in each region [20].

3.3. Proposed Approach

The proposed algorithm in this chapter is given in (Fig.3.4), which consists of six steps. (i) The face image is divided into non-overlapping image blocks of size $m \times m$ pixels. (ii) Various preprocessing methods like HE, CLAHE, SQI, DoG, GIC and LTISN are applied before and after block sizing the images to improve the recognition accuracy (iii) LBP operator is applied on each image sub block to obtain local features of the input image. (iv) Image blocks are given as the input to the classifier and they are classified separately by K-NN. (v) In the weighting step, entropy of image sub blocks is computed to find out the information content as well as their contribution towards the final decision. (vi) Decision fusion is achieved using a voting scheme to determine the majority vote. (vii) The outcome is retrieved by defining a threshold**.** In case its value is smaller than the value of threshold, the classification process is performed again for the highest n ranks.

The size of image blocks affects accuracy of the face recognition. Bigger block size does not yield good accuracy in block-based approach. On the other hand, smaller size adds to the computational complexity. In this paper we proposed $m \times m$ rectangular image blocks. If the image size is $R \times C$ pixels, then the number of image blocks can be computed as follows given in equation (3.11).

$$N = (R/m) * (C/m) \quad (3.11)$$

Since number of image sub blocks are greater than the contaminated blocks, the accuracy rate of recognition will be high in the democratic vote counting approach.

3.4 Classification

Features of each image block obtained in the last section are fed to the K-NN classifier to compute the similarity measure between gallery and probe sub-block. Based on the similarity, each block is assigned to a class.

3.4.1 K-Nearest Neighbour (K-NN)

K-NN is the supervised machine learning algorithm for classifying objects based on closest training examples in the feature space [17]. It classifies the object based on majority vote of its surrounding neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbor. [17].

The Euclidean distance metric is often selected to decide about the proximity of the data points in KNN.

Advantages of K-NN algorithm are described as follows.

1. No Training Period is required. KNN is called Lazy Learner.

2. New data can be added seamlessly.

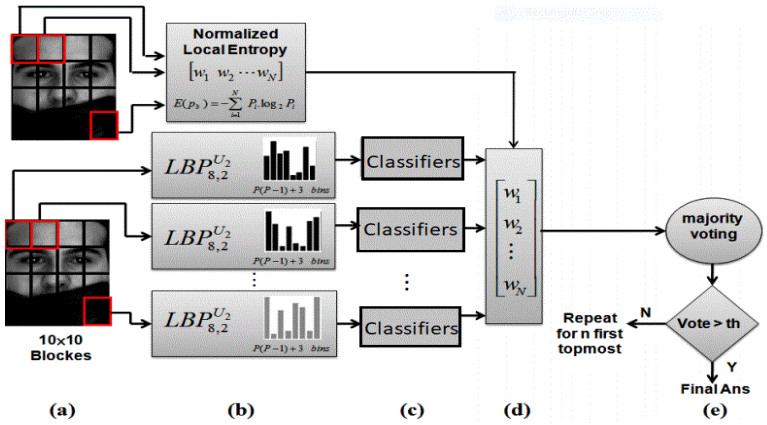


Fig. 3.4. Suggested algorithm. (a) Dividing the image (b) Applying LBP operator on each block. (c) Classification of image blocks using classifiers. (d) Generating local entropy and summing the weighted votes. (e) Thresholding the majority votes to find final decision.

3.5 Local Entropy Weighting

The benefit of the block-based approach is its ability to reduce the effect of contaminated parts of image on the final decision by assigning lesser weights to the image blocks since the corrupted blocks are not providing effective contribution in final recognition. The proposed method applies the local image entropy principle, which provides adequate weights to the image blocks according to the information property of each block and it is given as follows in equation (3.12).

$$E(p_b) = \sum_{i=1}^{N_t} P_i \, log_2(P_i) \qquad (3.12)$$

Where N_t is the overall number of gray levels in the probe image sub-block and P_i is the i^{th} gray value's probability in that block. The normalized entropy values are between 0 and 1 [16].

3.6 Majority Voting Scheme

In order to fuse the results of classifier and to obtain the final class of the probe image, we need to choose a decision fusion method which gives more accurate results in comparison with feature fusion method [18]. Weighted sum rule and majority voting scheme are two fixed rules integrated in the proposed approach, which are simple and do not demand training requirement [19]. In this chapter a weighted majority voting scheme is suggested, which is based on the local entropy. We sum the obtained weights of votes for each class to arrive at the decision. The final decision is the label of the majority vote. To obtain better accuracy rate, the maximum vote is compared with a threshold, which is achieved by trial and error basis. In case it is below the threshold value, the classification and majority voting process are repeated for the topmost number of votes. This method is not only simple, but also not computationally expensive [16].

3.7 Database

The FERET database was collected in 15 sessions between August 1993 and July 1996. The database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images as shown in (Fig.3.5). A duplicate set is a second set of images of a person already in the database and was usually taken on a different day [15].

3.7.1 Experiment

In the current study, the impact of different pre-processing algorithms is evaluated to recognize face images. The input images are taken from FERET dataset and many of the images are having illumination variations. Five different pre-processing techniques are

applied to enhance the accuracy of recognition process. The overall experiments can be divided into two types. In block-based method, the image is block sized and then pre-processing methods are applied to each block. After that KNN classifier and weighted entropy-based fusion technique are applied. The recognition accuracy is computed for various block sizes and threshold values. In holistic method, after the application of preprocessing methods to the entire image, it is block sized. Subsequently classifier and weighted entropy-based fusion technique method are applied. Finally, recognition accuracy is computed for various block sizes and threshold values.

The enhancement methods used in this study are LTISN, DoG, GIC, SQI, HE and CLAHE.

In LTISN, the intensity of the input image pixel is enhanced by a quotient parameter q, which is generated non-linearly based on the input image pixel characteristics. The values of $of\ C_1, C_2, C_3$ and C_4 mentioned in equation (9) are decided on experimental basis by testing with typical images. It is found experimentally that $I_{M_n}(x, y)$ is 1 and q value is adjusted between 5 and 6. C_1, C_2, C_3 and C_4 values are chosen in the experiment as 2.255, 0.162 60.6 and .6061 for better enhancement [21].

The second method uses HE, which uses the contrast normalization and illuminance correction. In some scenarios of the experiment, especially in holistic method, the HE does not provide consistent results due to non-uniform illumination variation. To avoid this problem, CLAHE is applied. In CLAHE, the idea of clip limit besides the block size is applied to minimize the noise problems. The method minimizes the amplification of noise by clipping the histogram at a predetermined value before calculating the Cumulative

Distribution Function. The clipping is based on the normalization of the histogram in turn depends on the surrounding area's size [28].

The third one is DoG filtering-based enhancement which uses a couple of processing stages. The first stage uses a gamma correction with a gamma value greater than 1. The value 1.5 gives better accuracy rates. A DoG filtering is used then with two Gaussian kernels of size 5 and 10 pixels, and standard deviation of 1 and 2 respectively. Finally, the contrast is enhanced based on the DoG image. The fourth enhancement is gamma correction. In this experiment, a gamma value chosen is 1.2

The fifth method uses SQI based intrinsic filtering with series of filtering stages.

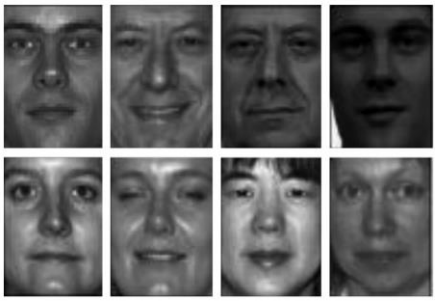


Fig. 3.5. Sample Images from FERET Database

3.8 Results

The experiment is classified into two types. In the first type, the preprocessing techniques are applied before block division of image (holistic). In the second type, the preprocessing techniques are applied after block sizing the image.

3.8.1 Holistic Method

In this method, the preprocessing methods are applied to the complete image. Then the image is block sized. Subsequently, K-NN classifier and weighted entropy-based fusion

technique is applied. Accuracy rates are computed for the various block sizes and threshold values.

Sample images from FERET database shown (Fig.3.6) which show the effect of preprocessing in holistic method.

Among all the preprocessing techniques applied, SQI gives maximum accuracy of 74.6% for the block size of 2*2, when the threshold value is 2.5. DoG performs equally good. DoG gives maximum accuracy of 74.4%. Both DoG and SQI are found to be very effective for illumination- invariant face recognition under severe lighting conditions.

LTISN and HE (CLAHE) perform moderately well and achieve highest performance of 74.2 % and 72.9% in 2*2 block size for the threshold value of 2.5. Then GIC, achieves the least accuracy rate of 72.9%. The maximum accuracy obtained without applying preprocessing techniques (denoted as WOAPP in table) is 72.8% for the block size of 2*2 and for the threshold value of 2.5. Table 1 shows the accuracy rates of block size of 2*2 computed after applying the preprocessing techniques in holistic method.

Similarly, for the block size of 4*4, SQI and LTISN and perform well and achieve the accuracy rate of 82.3% and 81.1% at the threshold value of 2.5. DoG achieves 81.2%. GIC and CLAHE perform almost equally and achieve accuracy rate of 80.5% and 79.9%. The maximum accuracy obtained without applying preprocessing techniques is 77.9% for the block size of 4*4. Table 2 shows the accuracy rates of block size of 4*4 computed after applying the preprocessing techniques in holistic method

For the block size of 8*8, both DoG and SQI perform very well for the threshold value of 2.5. By applying SQI, the accuracy rate increases to 85% and DoG increases the accuracy up to 86.9%.

LTISN and HE (CLAHE) perform moderately good and achieves highest performance of 82.7% and 84.1% in 8*8 block size for the threshold value of 2.5. Then GIC, achieves the least accuracy rate of 82.2%. The maximum accuracy obtained without applying preprocessing techniques is 77.9% for the block size of 8*8. Table 3 shows the accuracy rates of 8*8 computed after applying the preprocessing methods in holistic method.

Among all the applied preprocessing methods, SQI and DOG perform consistently better in comparison to other techniques for all the block sizes and for all the threshold values. DoG and SQI yields the accuracy rate of 86.9% and 85% for the threshold value of 2.5. After that, LTISN, HE and GIC perform equally good.

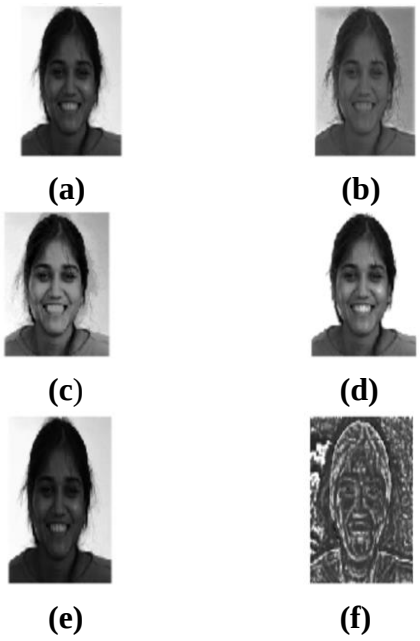


Fig. 3.6. Sample image from FERET database showing the effect of preprocessing in holistic approach (a) Input Image (b) LTISN (c) HE (d) DoG (e) GIC (f) SQI

Table 3.1. The accuracy rates of block size of 2*2 computed after applying the preprocessing techniques in holistic method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
2*2	LTISN	74.2	74.7	71.8	71.1	PP applied Before Block sizing
	HE	72.9	72.5	70.1	70.1	
	DoG	74.4	76.4	70.3	69.1	
	GIC	72.9	72.2	74.7	67.9	
	SQI	74.6	72.5	71.2	68.9	
	None	72.8	71.9	69.1	67.8	WOAPP

Table 3.2 shows the accuracy rates of block size of 4*4 computed after applying the preprocessing techniques in holistic method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
4*4	LTISN	81.1	77.9	75.1	72..1	PP applied Before Block sizing
	HE	79.9	75.5	74.0	73.3	
	DoG	81.2	80.8	74.1	72.7	
	GIC	80.5	75.7	75.5	72.1	
	SQI	82.3	82.2	74.9	71.9	
	None	77.9	74.9	731	71.7	WOAPP

Table 3.3 The accuracy rates of block size of 8*8 computed after applying the preprocessing techniques in holistic method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
8*8	LTISN	82.7	76.7	75.9	70.1	PP applied Before Block sizing
	HE	84.1	75.7	74.6	71.7	
	DoG	86.9	73.9	73.7	72.9	
	GIC	82.2	72.1	72.3	71.9	
	SQI	85.0	81	74.2	72	
	None	77.9	74.9	72.1	71.8	WOAPP

3.8.2 Block-based Method

In the block-based approach, the image is block sized and then preprocessing techniques are applied to each block to improve the accuracy rate. Subsequently K-NN classifier and weighted entropy-based LBP method is applied. Accuracy rates are then computed for various block sizes and threshold values. Sample image from FERET database (Fig. 3.7) showing effect of preprocessing in block-based approach.

For the block size of 2*2, DoG and LTISN perform well and achieve the accuracy rate of 80.1% and 79.1% for the threshold of 2.5. HE (CLAHE) and GIC perform equally and achieve accuracy rate of 74.8% and 76.0% Finally, SQI achieves 76.8%. The maximum accuracy obtained without applying preprocessing techniques is 74.2% for the block size of 2*2. Table 4 shows the accuracy rates of block size of 2*2 computed after applying the preprocessing techniques in block-based method.

For the block size of 4*4 DoG, SQI and GIC perform equally well and achieve the accuracy rate of 81.4%, 81.3% and 81.2% for the threshold of 2.5. HE (CLAHE) and LTISN performed equally well and achieved the accuracy rate of 79.1% and 79.9% respectively. The maximum accuracy obtained without applying preprocessing techniques is 77.7% for the block size of 4*4. Table 5 shows the accuracy rates of block size of 4*4 computed after applying the preprocessing techniques in block-based method.

For 8*8 DoG, SQI and LTISN perform equally well and achieve the accuracy rate of 88.2%, 88.0% and 84.2% for the threshold of 2.5. HE and GIC does not have much effect in increasing the accuracy rate. HE and GIC achieve 82.1% and 82.2%. Table 6 shows the accuracy rates of block size of 8*8 computed after applying the preprocessing techniques in block-based method

DoG and SQI perform extremely good for all the block sizes and for all the threshold values in both the approaches. DoG and SQI gives maximum accuracy of 88.2% and 88% for the block size of 8*8, when the threshold value is 2.5 in the block-based method.

Table 3.4 The accuracy rates of block size of 2*2 computed after applying the preprocessing techniques in block-based method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
2*2	LTISN	79.1	76.8	73.1	71.9	PP applied After Block sizing
	HE	74.8	73.9	72.7	71.8	
	DoG	80.1	78.1	76.3	71.3	
	GIC	76.0	74.2	74.4	71.1	
	SQI	76.8	74	75.9	70.9	
	None	74.2	73.8	71.1	70.8	WOAPP

Table 3.5 The accuracy rates of block size of 4*4 computed after applying the preprocessing techniques in block-based method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
4*4	LTISN	79.9	80	76.1	73.1	PP applied After Block sizing
	HE	79.1	76.3	74.6	73.5	
	DoG	81.4	79.0	76.4	73.2	
	GIC	81.2	79.1	78.7	72.8	
	SQI	81.3	78.7	77.2	72.6	
	None	77.7	75.1	73.6	72.4	WOAPP

Table 3.6 The accuracy rates of block size of 8*8 computed after applying the preprocessing techniques in block-based method.

Block Size	PP Type	Threshold Values				Types of Application of PP
		2.5	5.0	7.0	9.5	
8*8	LTISN	84.2	80.5	77.5	73.4	PP applied After Block sizing
	HE	82.1	78.2	77.1	72.9	
	DoG	88.2	78.4	80.0	73.4	
	GIC	82.2	78.3	77.2	72.7	
	SQI	88	83	84	73	
	None	82.1	78.1	77	72.6	WOAPP

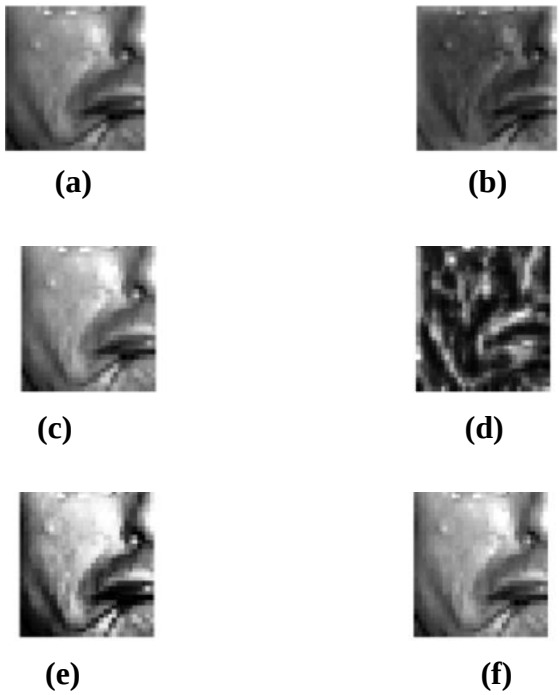


Fig. 3.7. Sample image from FERET database showing the effect of preprocessing in block-based approach (a) Input Image (b) LTISN (c) GIC (d) SQI (e) HE (f) DoG

As the threshold increases, the accuracy rate started decreasing in both block-based and holistic methods. Especially for the threshold values above 7, the applied preprocessing methods do not have much effect in increasing the accuracy. Compared to holistic method, block-based method performs slightly better.

3.9 Comparison with Existing Methods

The proposed method performs very well in comparison to existing methods**.** The existing method applies Gamma Intensity Correction **(GIC),** Difference of Gaussian **(DoG),** Adaptive Histogram Equalization **(AHE),** Decorrelation Stretching **(DCS)** and their combination in FERET database [22]. The proposed method achieves better accuracy in comparison to all methods except the combination of GIC, Gaussian Blur, AHE and DCS, which achieves the accuracy rate of 88.3% almost equal to the proposed method. The accuracy rates of different methods are given in Table 7.

Table 3.7 Comparison of Accuracy Rates different approaches

Method	Accuracy
Gaussian Blur	85.42%
AHE	84.59%
GIC+ Gaussian Blur	86.00%
GIC+Gaussian Blur+DCS	86.33%
GIC+Gaussian Blur+AHE	86.00%
GIC+Gaussian Blur+AHE+DCS	88.30%
Proposed Method	88.20%

3.10 Conclusion

The applied preprocessing techniques to holistic and block sized images in weighted entropy-based LBP method help to enhance the face recognition accuracy. Especially in comparison to all other techniques, SQI and DoG perform very well for all the block sizes and threshold values for both the approaches. Even though both approaches enhance the recognition accuracy, block-based method gives consistently good performance in comparison to holistic approach.

SQI, DoG and LTISN perform well and maximum accuracy of 88.0%, 88.2% and 84.2% are achieved for 8*8 block size and for the threshold value of 2.5 in block-based method. With the increase in block size, the face images preprocessed by HE and GIC became smooth. At the same time, minute facial details are lost. Hence there is not much improvement in accuracy with increase in block size especially with HE and GIC in block-based method.

In general, the classification accuracy rate increases with the increase in the number of sub-blocks. It is because when there are many blocks, the percentage of good blocks which contain the relevant information (with high entropy weights) will be more. Since the contaminated blocks are assigned with lesser entropy weights, the impact of such contaminated blocks will be less in the overall accuracy. The application of preprocessing techniques further helps to increase the accuracy rates further for majority of cases.

The preprocessing techniques improve the recognition accuracy till the threshold value is 7.0. When the threshold value reaches beyond 7.0, the applied techniques do not have much effect in the recognition accuracy. This is because the selection of threshold value helps to select the significant sub images in terms of entropy. The sub blocks will have high entropy values if it represents significant information. For corrupted or uniform

background regions, entropy will be less. Hence the entropy threshold value can be selected to filter the insignificant sub blocks during the analysis. As the threshold increases, more sub blocks will be eliminated from voting scheme. The entropy threshold value is chosen empirically to retain the best possible accuracy in each case. Hence the application of preprocessing techniques do not help in improving the accuracy.

The proposed method performs very well as compared to existing techniques in FERET database.

References

1. R. Jafri and H. Arabnia “A Survey of Face Recognition Techniques” Journal of information processing systems, Vol 5, no. 2, Korea information processing society, pp 41-68, June 2009.

2. Y. Admi, Y. Moses and S. Ulman “Face Recognition: The problem of compensating for changes in illumination direction”, IEEE Transaction. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 721–732, Jul. 1997.

3. S. Muthu Selvi and P. Prabhu “Digital Image Processing Techniques – A survey” Golden Research Thoughts Journal, Vol 5, Issue 11, pp 1-7, May 2016.

4. Soodeh Nikan and M. Ahmadi. “Human Face Recognition under Occlusion using LBP and Entropy Weighted Voting” 21st International Conference on Pattern Recognition ICPR pp. 1699-1702, 2012.

5. S. Arigela, V. K. Asari “self-tunable transformation function for enhancement of high contrast color images” Journal of Electronic Imaging Volume 22, Issue 2,2013.

6. H. Wang, S. Z. Li, Y. Wang and J. Zhang "Self-quotient image for face recognition," *2004 International Conference on Image Processing, 2004. ICIP '04.*, Singapore, pp. 1397-1400 Vol.2, 2004.

7. H. Wang, S.Z. Li, Y. Wang "Face Recognition Under Varying Lighting Conditions Using Self Quotient Image" Proc. of IEEE International conference on Automatic FGR (AFGR), pp 819-824, 2004.

8. S. Anila, K. Arulsukanya, K. Divya, C. Kanimozhi, and D. Kavitha “An efficient Preprocessing Technique for Face Recognition under difficult Lighting Conditions,” in Proc. National Conference on Emerging Trends in Computer Communication and Informatics (ETCCI-2011), March 10-11, 2011.

9. V. Karthikeyan, K. Vijayalakshmi, P. Jeyakumar “An Efficient Method for Face Recognition in Various Assorted Conditions” Biometrics and Bioinformatics, 4(12) ,pp 543-547,2012.

10. S.A. Amiri and H. Hassanpour “A preprocessing approach for image analysis using gamma correction” International Journal of Computer Applications, Volume 38, Issue 12, Pages 38-46 January 2012.

11. P.A. Koringa, S.K. Mitra and V.K. Asari “Handling Illumination Variation: A Challenge for Face Recognition”. In: Raman B., Kumar S., Roy P., Sen D. (eds) Proceedings of International Conference on Computer Vision and Image Processing. Advances in Intelligent Systems and Computing, vol 460. Springer, Singapore, 2017.

12. S.Anila, and N. Devarajan. "Preprocessing technique for face recognition applications under varying illumination conditions." Global Journal of Computer Science and Technology 12, no. 11-F (2012).

13. M. Biglari, F Mirzaei and H. Ebrahimpour-Komeh “Illumination invariant face recognition using SQI and weighted LBP histogram”. Pattern Recognition and Image Analysis (PRIA), 2013 First Iranian Conference on. IEEE, pp 1-7, 2013.

14. G. Calvo G, B.Baruque and E.Corchado “Study of the Pre-processing Impact in a Facial Recognition System”. In: Pan JS., Polycarpou M.M., Woźniak M., de Carvalho A.C.P.L.F., Quintián H., Corchado E. (eds) Hybrid Artificial Intelligent Systems. HAIS, Lecture Notes in Computer Science, vol 8073. Springer, Berlin, Heidelberg, 2013.

15. P. J. Phillips, S. A. Rizvi and H. Moon. “The FERET Evaluation Methodology for Face-Recognition Algorithms” IEEE *Trans. Pattern Analysis. Machine Intelligence*, 22(10): pp. 1090–1104, Sep 2000.

16. A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi. Illumination invariant feature extraction and mutual information-based local matching for face recognition under illumination variation and occlusion. Pattern Recognition., 41(10-11) pp. 2576–2587, October 2011.

17. M.Kaur and Dhriti “K-Nearest Neighbor Classification Approach for Face and Fingerprint at Feature Level Fusion”, International Journal of Computer Applications Volume 60, pp 13-17, 2017.

18. B. Topc¸ and H. Erdogan. “Decision fusion for patch-based face recognition”. ICPR, pp.1348–1351, August 2010.

19. F. Roli1, J. Kittler, G. Fumera, and D. Muntoni. An experimental comparison of classifier fusion rules for multimodal personal identity verification systems. MCS, pp. 325–335, 2002.

20. A. Reza. "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement." Journal of VLSI Signal Processing, vol. 38, pp 35-44, 2004.

21. E Krieger, V. K Asari, and S Arigela. Color image enhancement of low-resolution images captured in extreme lighting conditions. In *SPIE Sensing Technology and Applications*, pages 91200Q–91200Q. International Society for Optics and Photonics, 2014.

22. K. Varadarajan, P. Suhasini, K. Manikantan, and S. Ramachandran “Face Recognition Using Block Based Feature Extraction with CZT and Goertzel-Algorithm as a Preprocessing Technique” Procedia Computer Science, 46, 1458-1467, 2015.

23. H. Han, S. Shan, X. Chen and W.Gao “A comparative study on illumination preprocessing in face recognition” Pattern Recognition., Volume 46, Issue 6, Pages 1691-1699, June 2013.

24. M. Nishiyama, T. Kozakaya, and O. Yamaguchi, “Illumination Normalization using Quotient Image-based Techniques,” Recent Advances in Face Recognition, I-Tech, Vienna, Austria, pp-98-108, December 2008.

25. B. Vinoth Kumar, “Gamma Correction Technique Based Feature Extraction for Face Recognition System”, International Journal of Computational Intelligence and Informatics, Vol. 3: No. 1, pp 20-26, April - June 2013.

26. J. B. Zimmerman, S. Cousins, K. M. Hartzell, M. Frisse and M. Kahn “A Psychophysical Comparison of Two Methods for Adaptive Histogram Equalization” Journal of Digital Imaging, Vol 2, No 2, pp 82-91, (May), 1989.

27. S. Saurav, S. Singh, R. Saini, A. K Saini and N. Sharma “A Comparative Analysis of Various Image Enhancement Techniques for Facial Images”, International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp 2279-2284, India, Aug. 2015.

28. B.S. Min, D.K. Lim, S.J. Kim and J.H Lee, “A Novel Method of Determining Parameters of CLAHE Based on Image Entropy,” International Journal of Software Engineering and Its Applications vol.7, no.5, pp.113-120, 2013.

29. P. Suganya, S. Gayathri, N. Mohanapriya “Survey on Image Enhancement Techniques” International Journal of Computer Applications Technology and Research, Volume 2– Issue 5, 623 - 627, 2013.

30. N. Longkumer, M. Kumar and R. Saxena “Contrast Enhancement Techniques using Histogram Equalization: A Survey” International Journal of Current Engineering and Technology, Vol.4, No.3, E-ISSN 2277 – 4106, P-ISSN 2347 – 5161 (June 2014).

31. A. Sri Krishna, G. Srinivasa Rao and M. Sravya “Contrast Enhancement Techniques using Histogram Equalization Methods on Color Images with Poor Lightning”, International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.3, No.4, August 2013.

32. T. Celik and T, Tjahjadi. Automatic image equalization and contrast enhancement using Gaussian mixture modeling. IEEE Transactions on Image Processing. 2012; 21(1):145-56.

33. R. C. Gonzalez and R. E. Woods, “Digital Image Processing”, Vol. 1, Parson, 2008.

Chapter 4

Performance Evaluation of CNN for Face Recognition

4.1 Introduction

For human brain, recognizing face is a very simple task and can be performed fast. In computer vision, face recognition is a very challenging task. Even though the face recognition research is in an advanced state [7], till now it is not possible to obtain results on par with humans. Due to the emergence of modern GPU's, deep learning-based methods provide improved speed rates in pattern recognition related problems.

To date**,** many approaches have been suggested for facial recognition. One of the oldest approaches is holistic method, which considers the whole image for processing. Holistic method works by projecting facial images onto a low-dimensional space, which neglects surplus details and variations that are not needed for the facial recognition [11]. One of the famous methods under this category is PCA [2]. The holistic methods are sensitive to local distortions like facial expression or illumination variation.

Subsequently, progress in the field of computer vision led to the growth of feature-based method in which features are extracted from various parts of a face image. Feature-based methods are robust to local variations such as intensity variation and face expression

changes [22]. With the development of the local feature descriptors, feature based methods gained popularity. Local Binary Pattern (LBP) [6] is an extensively applied local feature descriptor in face recognition.

The recent trend is towards neural network-based approach [30]. Deep learning-based methods achieve excellent results in many fields like robotics and autonomous driving cars [3]. Deep learning methods are based on convolutional neural networks (CNNs). CNNs are slightly different from normal neural network. In CNN, neurons in convolutional layer are thinly connected to the neurons in the next layer based on their relative location. CNNs are multilayer network trained from end to end with raw image pixel values assigned to classifier outputs.

The main advantage of deep learning methods is that they can be trained with very big datasets to learn the vital features to represent the input data. The main issue with deep learning method is models trained with small datasets are having the problem of poor generalization, which results in over-fitting.

Generalization term indicates the performance difference of a network model when assessed on earlier viewed training data against the testing data, the network has never viewed before [8]. Models with poor generalizability have overfitted the training data.

Overfitting is a term used when the network model functions extremely good with the training data, but could not work well with the test data. In the over fitted network validation error goes up while the training error comes down [21].

To reduce the overfitting, regularization process is employed on the network model. *Regularization* is a method of making minute changes to the actual network model and the

learning algorithm, so that the model functions better in both training and testing data set. Regularization is defined as "Allowing to generalize well to unseen data even when training on a finite training set or with an imperfect optimization procedure" [4]. There are various regularization techniques available in machine learning like Dropout, Data Augmentation, Early stopping, batch normalization etc. [10]. Some of the regularization techniques are explained below.

One of the well-known techniques is called dropout. *Dropout* means removing units temporarily in a neural network, together with all the incoming and outgoing links [23]. Dropout can be explained as the regularization technique by including noise to the hidden units of the network.

Another popular technique is batch normalization. *Batch Normalization* operates by deducting the batch mean from each activation and dividing by the standard deviation of the batch [8]. The normalization technique together with standardization is used as a typical combination in the preprocessing of pixel values. Batch normalization technique can be employed to any individual layer within the network. Hence it is powerful [5].

Data augmentation is an extensively applied technique used to artificially inflate the size of the training dataset by the methods called data warping or oversampling. The augmented data is a representation of detailed set of feasible data points, thus minimizing the distance between the training and the validation set.

The main aim of augmentation technique is to diminish the effect of overfitting on models using traditional transformations to manipulate the training data. As an example, a labeled image in the dataset can be increased by various processes like flipping, rotation,

morphing and zooming. After some time, trained network gets exposure to such modifications and then it can identify the same object with different variations.

Optimizers update the weight parameters to minimize the cost function. *Cost function* is defined as the difference between predicted and actual output. One of the popular optimizers is Adam optimizer [1].

Adam is an adaptive learning rate optimization algorithm, which calculates learning rates for different parameters individually. Adam uses computations of first and second moments of gradient to adjust the learning rate for each weight of the network.

Application of regularization methods help to improve the accuracy rates. To increase the accuracy rates further, preprocessing techniques are applied to deep CNN architecture.

4.2 Preprocessing Methods

Face recognition task becomes challenging due to illumination conditions, occlusion, pose and facial expression variations [33]. The variation in illumination is one of the main challenging problems which affect the performance of the face recognition system. Among all, shadowing effect, underexposure, and overexposure conditions are challenging problems that need to be addressed in the face recognition process [35]. If the lighting conditions present in the gallery image is different from the probe image, then the process of face recognition may completely fail [34]. A good face recognition system should be able to give accurate recognition rate under the different illumination conditions between images of the same face [32]. Image enhancement algorithms play a great role in handling the illumination variation.

The main aim of preprocessing is to remove features that obstruct the process of classifying the images of the same person (within-class differences), thereby boosting the difference of them with others (between-class differences) [36].

For better face recognition under uncontrolled and illumination variation conditions, the vital features responsible for differentiating two different faces require to be retained. The shadows produced in facial images due to variation in lighting directions may cause loss of important facial features which are helpful for recognition. A preprocessing method must enhance the intensity in the regions of inadequately illuminated and decrease the intensity in the densely illuminated regions while retaining the intensity in the fairly illuminated portions [37]. Few important preprocessing techniques are discussed below.

4.2.1 Histogram Equalization (HE)

HE [38] flattens the histogram of the image and expands the dynamic range of the pixel intensity values by employing cumulative density function.

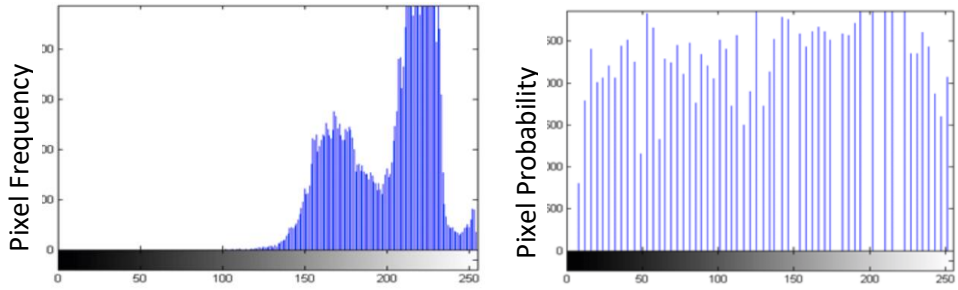
For image I(x, y) with discrete k. gray values histogram is defined by the probability of occurrence of the gray level I [39] is given by equation (4.1) as follows.

$$p(i) = n_i /N \quad (4.1)$$

Where $i \in 0, 1 \ldots k-1$ grey level and N is total number of pixels in the image.

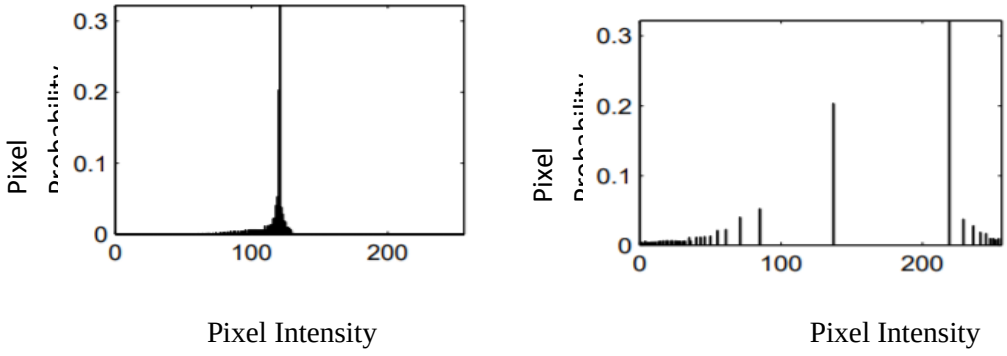
HE is the most popular and simplest global image processing technique applied in spatial domain. The image histogram yields detail about the intensity distribution of the

pixels present in the image. It modifies image intensities to enhance the image contrast using histogram of image. HE improves the global contrast of the image.



(a) (b)

Fig.4.1. Histogram plot by 1st method (a) Before applying HE (b) After applying HE



(a) (b)

Fig.4.2. Histogram plot by 2nd method (a) Before applying HE (b) After applying HE

In Histogram equalization approach, the image intensities are adjusted to improve the image contrast. Histogram of an image can be constructed by two ways [45]. (i) It can be constructed by plotting pixel intensities against the pixel frequencies as shown in (Fig.4.1). (ii) It can also be constructed by plotting pixel intensities against pixel probabilities as shown in (Fig.4.2).

4.2.2 Self Quotient Image (SQI)

SQI [40] is one of the illumination invariant algorithm suggested for handling both shadow and lighting changes. It is defined as the ratio of the intensity of the input image to its smooth version as given in equations (4.2) and (4.3).

$$Q(\mathrm{x},\mathrm{y}) = \frac{I(x,y)}{S(x,y)} \tag{4.2}$$

$$= \frac{I(x,y)}{F(x,y)} * I(x,y) \tag{4.3}$$

where *I (x, y)* is the face image and *S (x, y)* is a smoothed version of the image and $*$ is the convolution operation. F is the smoothing kernel which in this case is a weighted Gaussian filter and Q is the Self Quotient Image since it is derived from one image and has the same quotient form as that in the quotient image method.

4.2.3 Locally Tuned Inverse Sine Nonlinear (LTISN)

It is a nonlinear and pixel by pixel approach [41] , where the improved intensity values are calculated by applying the inverse sine function with a tunable parameter based on the nearby pixel values given in the equations (4.4), (4.5), (4.6), (4.7) and (4.8). The intensity range of the image is rescaled to [0 1] followed by a nonlinear transfer function.

$$I_{enh}(x,y) \quad = \frac{2}{\pi} sin^{-1}(I_n(x,y)^{q/2}) \tag{4.4}$$

$$kernel_i(n_1,n_2) = \frac{h_g(n_1,n_2)}{\sum_{n_1}\sum_{n_2} h_g(n_1,n_2)} \tag{4.5}$$

$$h_g(n_1,n_2) = e^{\frac{-\left(n1^2+n2^2\right)}{2\sigma^2}} \tag{4.6}$$

$$I_{M,i}(x,y) = \sum_{m=-\frac{M_i}{2}}^{\frac{M_i}{2}} \sum_{n=-\frac{N_i}{2}}^{\frac{N_i}{2}} I(m,n) kernel_i(m+x, n+y) \tag{4.7}$$

$$q = \begin{cases} tan(\frac{\pi}{C_1} I_{M_n}(x,y)) + C_2 & I_{M_n}(x,y) \geq 0.3 \\ \frac{1}{C_3} ln(\frac{1}{0.3} I_{M_n}(x,y)) + C_4 & I_{M_n}(x,y) < 0.3 \end{cases} \tag{4.8}$$

4.2.4 Gamma Intensity Correction (GIC)

It is a nonlinear gray-level transformation [42] that substitutes gray-level I with the gray level $I^{\frac{1}{\Upsilon}}$, given by the equation (4.9).

$$I = I^{\frac{1}{\Upsilon}} \tag{4.9}$$

As shown in the (Fig.4.3), for the values of gamma less than 1.0 darkens the image, values greater than 1.0 lightens the image and when the value is 1.0 there is no effect.

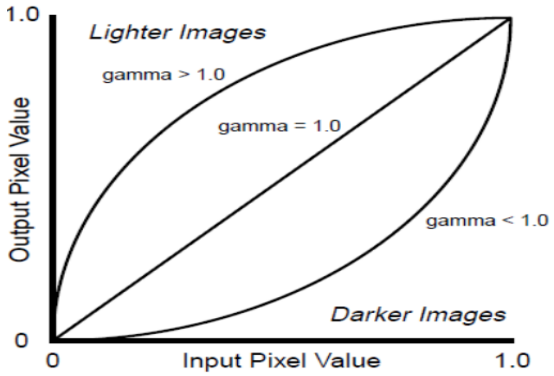


Fig.4.3. Gamma Intensity Correction

4.2.5 Difference of Gaussian (DoG)

It is a grayscale image enhancement algorithm [42] that involves the subtraction of one blurred version of an original grayscale image from another less blurred version of the original. The blurred images are obtained by convolving the original grayscale image with

Gaussian kernels having differing standard deviations [34], which is given in the equation (4.10).

$$DOG(x,y) = \frac{1}{2\pi{\sigma_1}^2} e^{-\frac{x^2+y^2}{2{\sigma_1}^2}} - \frac{1}{2\pi{\sigma_2}^2} e^{-\frac{x^2+y^2}{2{\sigma_2}^2}} \quad (4.10)$$

σ_1, σ_2 are Gaussian kernel widths.

4.2.6 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

CLAHE works on small areas in the image, known as tiles, rather than the whole image [43]. Individual tile's contrast is improved. Hence histogram of the output area is roughly matching with the histogram specified by the distribution parameter. The tiles present in the neighborhood regions are then joined by applying bilinear interpolation to minimize the effect of artificially induced border line [39]. In CLAHE, the image is divided into a limited number of regions and the same histogram equalization technique is applied to pixels in each region [44]

4.3. Convolutional Neural Network

Deep learning-based methods have shown better performances in terms of accuracy and speed of processing in image recognition

4.3.1 CNN Basics

CNN is biologically inspired by visual cortex in brain [20]. Different phases of learning procedure in CNN are similar to the visual cortex. The *visual cortex* has small group of cells that are reactive to particular regions of the visual field. Hubel and Wiesel [17] experimentally showed that particular group of neurons in the cat's brain reacted to the appearance of edges of a certain orientation as shown in the (Fig.4.4). Further they illustrated that one specific set of neurons were activated, when exhibited to vertical edges

and also another set of neurons fired when exposed to horizontal or diagonal edges. All of these categories of neurons are arranged in a columnar configuration and collectively these neurons can produce visual perception.

Different portions of the visual cortex are classified as V1, V2, V3, and V4. In general, V1 and V2 regions of visual cortex are having close resemblance with convolutional and subsampling layers, whereas inferior temporal region resembles the higher layers of CNN [16]. During the process of training, CNN learns with the help of backpropagation algorithm by making adjustments in weights with respect to the target.

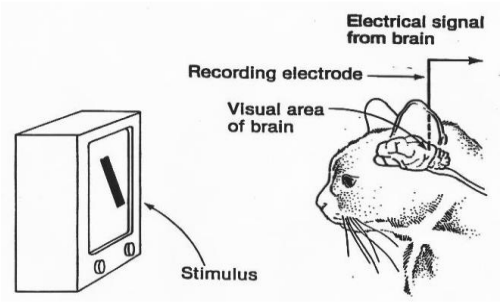


Fig.4.4. Hubel and Wiesel experiment

4.3.2 CNN Architecture

In CNN architecture, network layers are divided into three types: the convolutional, pooling and fully connected layers [20].

4.3.2.1 Convolutional Layer

The architecture of CNN is as shown in the (Fig.4.5). In CNN, every neuron in the convolutional layer is linked only to a small portion of the neurons in the preceding layer, which is in square shape area across the height and width dimensions.

The size of this square is a hyperparameter (controllable parameter) and called as *Receptive Field.* For the depth dimension, there is no hyperparameter available since the convolution operations are normally carried out for the whole depth. Generally, the depth dimension of the input describes the various colors of the image. Hence it is usually required to link them in order to bring out necessary information.

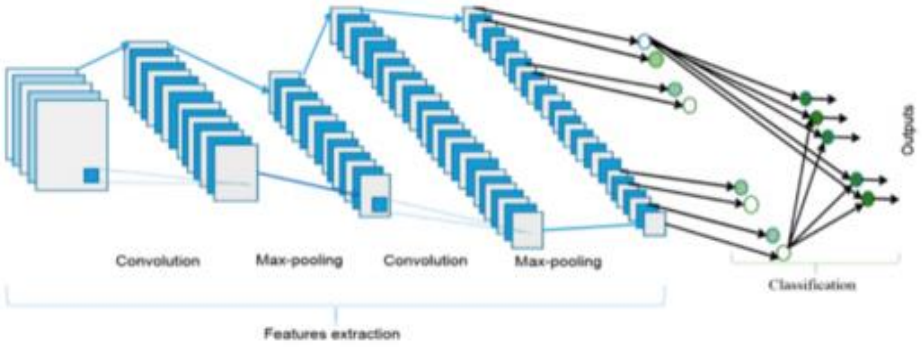


Fig.4.5. The CNN architecture

Neurons present in the convolution operator can recognize certain local patterns of the previous layer's output. Once features are obtained, its actual position is not important and all the neurons are expected to identify the same pattern. This is achieved by forcing all the neurons to have a common single set of parameters known as *Parameter Sharing* [9].

In order to identify various unique features within one layer, it is necessary to have multiple *Filters*, where each filter is a group of neurons that recognize a certain pattern at different locations in the image.

4.3.2.2 Pooling Layer

The main aim of pooling layer is to reduce the complexity of CNNs. The neurons present in the pooling layer form a square shaped area across the width and height

dimensions of the preceding layer. Even though it is very similar to the convolutional layer, it is different from convolutional layer because the pooling layer is *Non-Parametrized Layer*. It means that neurons present in this layer neither have weights nor biases.

The function carried out by this layer is known as subsampling or down sampling. During this process, contraction in size results in concurrent loss of data. On the other hand, such a loss is helpful to the network because the reduction in size not only reduces the computational burden for the succeeding layers of the network and also it reduces the effects of overfitting [12].

Max pooling and average pooling are the generally used techniques shown in (Fig.4.6). *Max Pooling* chooses the largest element within each receptive field [14] whereas *Average Pooling* computes the average among the output neurons within the pooling window. Max-pooling chooses the most prominent feature in a pooling window. On the other hand, average-pooling method selects whole features into consideration. Thus, max-pooling method keeps texture related information, while average pooling method retains the background related data [24]. Pooling operation does not combine neurons with different depth values. Instead, the resulting pooling layer will have the uniform depth as the previous layer and it will only combine local areas within a filter.

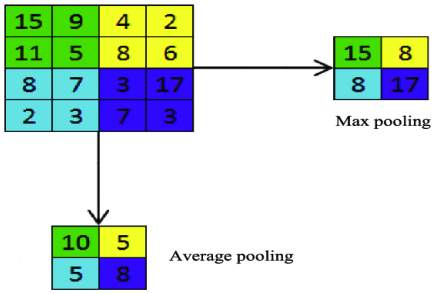


Fig.4.6. Max Pooling and Average Pooling Operations Illustration.

4.3.2.3 Fully Connected Layer

The filters and neurons present in this layer are connected to all the activation in the preceding layers resulting in a completely connected structure. The output feature maps of the final convolution or pooling layer is converted into a one-dimensional (1D) array of numbers [18]. High-level reasoning is carried out via fully connected layers [19].

The final fully connected layer has the same number of output nodes as the number of classes. Each fully connected layer is followed by a nonlinear function, such as (Rectified Linear Units). *ReLU* is an activation function operates by thresholding values at 0, i.e. f (x) = max (0, x). In other words, it outputs 0 when x < 0, and contrarily, it outputs a linear function with a slope of 1 when x ≥ 0 [15] as shown in the (Fig.4.7).

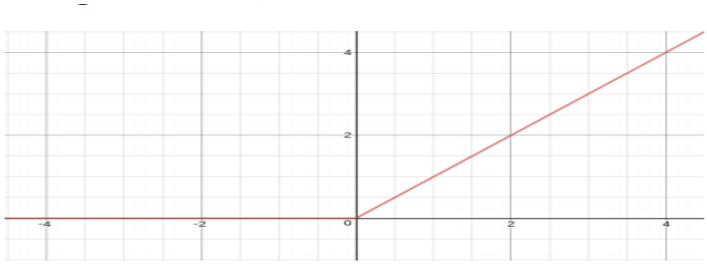


Fig.4.7. The Rectified Linear Unit (ReLU)

4.3.3 CNN Operation

Based on local receptive field, each component in a convolutional layer accepts inputs from a set of adjacent units belonging to the preceding layer layer. This way neurons are proficient in extracting rudimentary features like edges or corners. These features are then linked by the succeeding convolutional layers in order to further extract high level features.

The components of a convolutional layer are arranged in planes. All units of a plane share the same set of weights. Thus, each plane is in charge for building a particular feature. The results obtained from each plane are termed as feature maps. Each convolutional layer consists of several planes, so that multiple feature maps can be constructed at each location. The most significant features derived are passed from initial layers to higher layers. As the features are passed to the higher layer, there is a dimensionality reduction in features determined by kernel size of the convolutional and max-pooling layers. On the other hand, there is an increase in number of feature maps for representing better features of the input images for ensuring classification accuracy [13]. The derived feature vector either could be an input for classification task or could be treated as a feature vector for next level processing.

The network designed to do the classification task consists of several convolution pooling layers followed by some fully-connected layers. The first two layers mentioned above perform convolution and pooling operation in order to extract high-level features. The final layer's output of CNN is applied as the input to a fully connected network, which does the task of classification.

The output layer for classification task consists of one neuron for each class and the values of these neurons describe the score of each class. If score distribution is chosen, the score range will be between zero and one. The summation of class scores is one. The values of each neuron can be presumed as the probability of occurrence of the class.

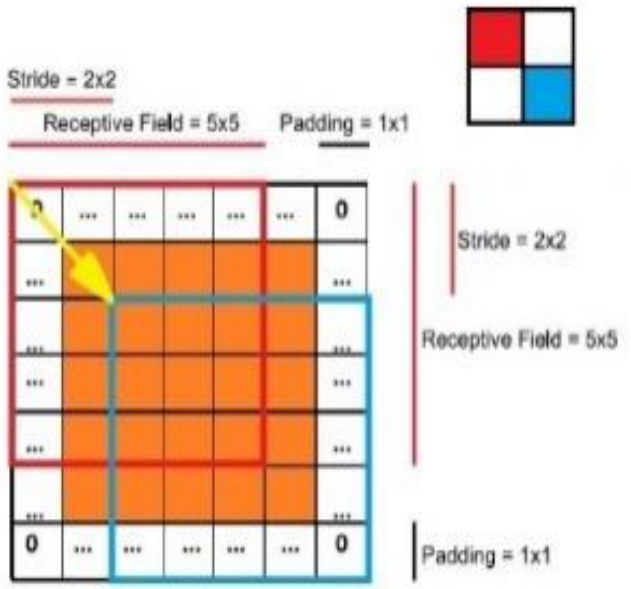


Fig.4.8. A visual representation of the various hyperparameters of convolutional layers: receptive field, stride and padding

4.4 Experimental Setup

The implementations are carried out in MATLAB 2019b on a workstation with windows 10 OS, AMD processor with 2.0 GHz, hard drive with 8 GB RAM. The experiments are conducted in Extended Yale B and FERET Database with deep CNN model**.**

4.4.1 Extended Yale Database

The extended Yale Face Database B contains 16128 images of 28 human subjects under 9 poses and 64 illumination conditions as shown in (Fig.4.9). The data format of this database is the same as that of the Yale Face Database B.

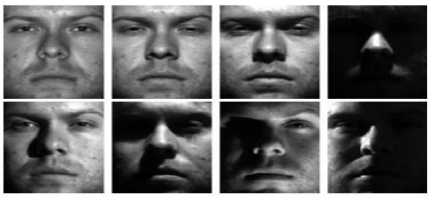


Fig.4.9. Images from Extended Yale B database.

4.4.2 FERET Database

The FERET database was collected in 15 sessions between August 1993 and July 1996. The database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images as shown in (Fig.4.10). A duplicate set is a second set of images of a person already in the database and was usually taken on a different day.

Fig.4.10. Images from FERET database

4.4.3 Experiment

The model uses a deep CNN network with six convolution layers (with size 3x3), two maxpooling and two fully convolutional layers. The first two convolutions use 8 filters,

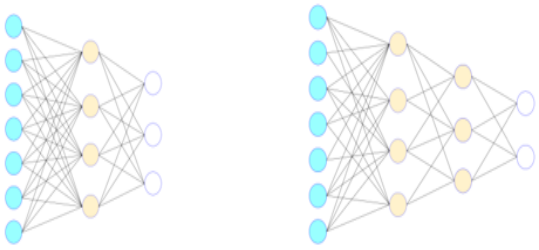
next two uses 16 and the last convolution layers uses 32 each in order to extract complex features.

The model uses batch normalization and dropouts in all convolutional layers along with ReLU activation function. In order to reduce the overfitting, the dropout rate is kept higher on the final convolutional layers. Two maxpooling layers are introduced after second and fourth convolutional layers to reduce the feature space dimension. The two fully connected layers use 128 and 28 neurons respectively. The final classification layer uses softmax activation with categorical cross entropy loss function. The *softmax* layer is used to produce the classification scores, in which each score is the probability of a particular class for a given instance [31].

The dropout principle is employed on the convolutional neural networks model and the value of drop out is chosen by trial and error method. The (Fig,4.11) shows the network model with drop out and the values of the drop out is different for different layers. The model also uses Adam optimizer with a learning rate of 0.001 which was found empirically after trying different combinations.

The feature maps are subsampled with maxpooling layers with a stride of 2x2. *Stride* is the number of pixels which shifts over the input matrix. When the stride is 2, it means the filter is moved by 2 pixels at a time. The (Fig.4.8) illustrates the stride of 2x2. The number of neurons in this output layer is limited to 28 as we have images of 28 different classes in Extended Yale B database. In FERET database, the number of neurons in the output layer is 994, as we have 994 different classes present in the FERET database.

The training is designed to use different batch sizes to do a fair trade off between accuracy and training time.



(a) (b)

Fig.4.11. (a) The neural networks model (b) The model after applying drop out

4.5. RESULTS

The CNN trained with back-propagation algorithm in batch mode with a various batch sizes of 4, 8,16 and 32 for different epochs are computed for Extended YALE B with an image size of 128×128 for 50 epochs. The same experiment is conducted on FERET database for the frontal face image size of 128×128 for 300 epochs.

For the Extended YALE B Database, the maximum accuracy rate of 97.2% is achieved for the batch size of 4 without applying preprocessing technique as shown in (Fig.4.12). After applying the preprocessing techniques, the maximum accuracy is improved to 99.8% as shown in (Fig.4.13).

SQI, HE and GIC perform very equally well and yield the maximum accuracy of 99.8%, 99.7% and 99.6%. After that, LTISN and DOG perform well and provide the accuracy rate of 99.2% and 99.1% respectively.

For the batch size of 8, the maximum accuracy rate of 97.0% is achieved without applying any preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 99.4%.

For the batch size of 8, SQI, LTISN and HE perform very equally well and yield the maximum accuracy of 99.4%, 99.3% and 99.1%. After that, DOG and GIC perform well and provide the accuracy rate of 99.0% and 98.9% respectively.

For the batch size of 16, the maximum accuracy rate of 96.8% is achieved without applying preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 99.1%.

For the batch size of 16 HE, SQI and GIC perform equally well and yield the maximum accuracy of 99.1%. 98.8% and 98.8% and After that, DOG and LTISN perform well and provide the accuracy rate of 98.7% and 98.4% respectively.

For the batch size of 32, the maximum accuracy rate of 96.2% is achieved without applying preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 98.7% for the same batch size.

For the batch size of 32, GIC, SQI, and HE perform equally well and yield the maximum accuracy of 98.7%, 98.3% and 98.3%. After that LTISN and DoG perform well and provide the accuracy rate of 98.1% and 97.9 % and respectively. Table 1 shows the accuracy rates of Extended Yale B database for various batch sizes before and after application of preprocessing techniques.

For the FERET Database, the maximum accuracy rate of 71.4% is achieved without applying preprocessing technique, for the batch size of 4, as shown in (Fig.4.14). After applying the preprocessing techniques, the maximum accuracy improved to 76.6% as shown in (Fig.4.15).

DoG, HE and LTISN perform very equally well and yield the maximum accuracy of 76.6%, 76.3% and 75.6%. After that SQI and GIC perform well and provide the accuracy rate of 74.9% and 74.2% respectively.

For the batch size of 8, the maximum accuracy rate of 71.2% is achieved without applying preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 76.4%.

For the batch size of 8, HE, SQI and LTISN perform very equally well and yield the maximum accuracy of 76.4%, 74.6% and 72.5%. After that, DOG and GIC perform well and provide the accuracy rate of 72.6% and 71.3% respectively.

For the batch size of 16, the maximum accuracy rate of 71.0% is achieved without applying preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 76.1%.

For the batch size of 16 HE, LTISN and SQI perform very equally well and yield the maximum accuracy of 76.1%, 72.0% and 71.5%. After that, DOG and GIC perform well and provide the accuracy rate of 71.3% and 71.0% respectively.

For the batch size of 32, the maximum accuracy rate of 68.7% is achieved without applying preprocessing technique. After applying the preprocessing techniques, the maximum accuracy is improved to 71.0%.

For the batch size of 32, SQI, GIC, LTISN and DoG perform equally well and yield the maximum accuracy of 71%, 70.9%, 70.8% and 69.9%. Finally, HE performs well and provides the accuracy rate of 69.2%. Table 2 shows the accuracy rates of FERET database for various batch sizes before and after application of preprocessing techniques.

SQI and HE perform equally well and provide best results for all the batch sizes in both Extended Yale B and FERET database.

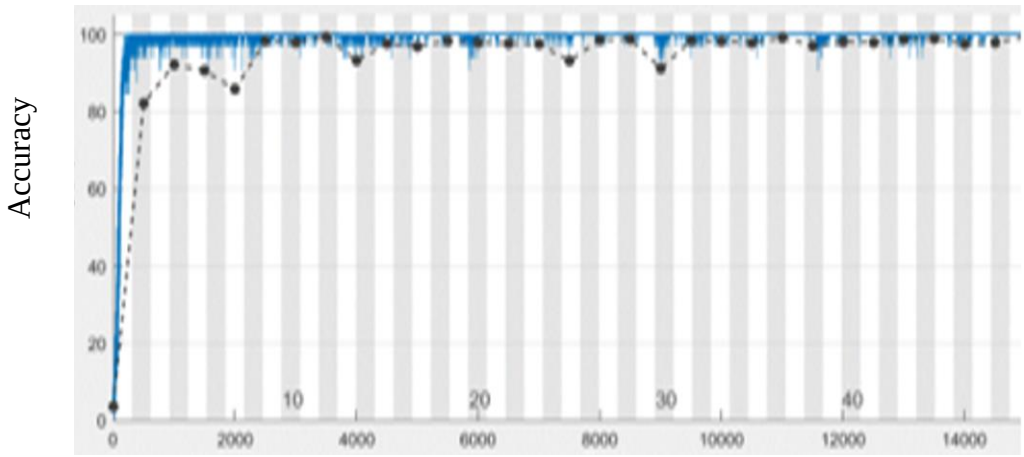


Fig.4.12. using Extended YALE B Database before application of preprocessing techniques

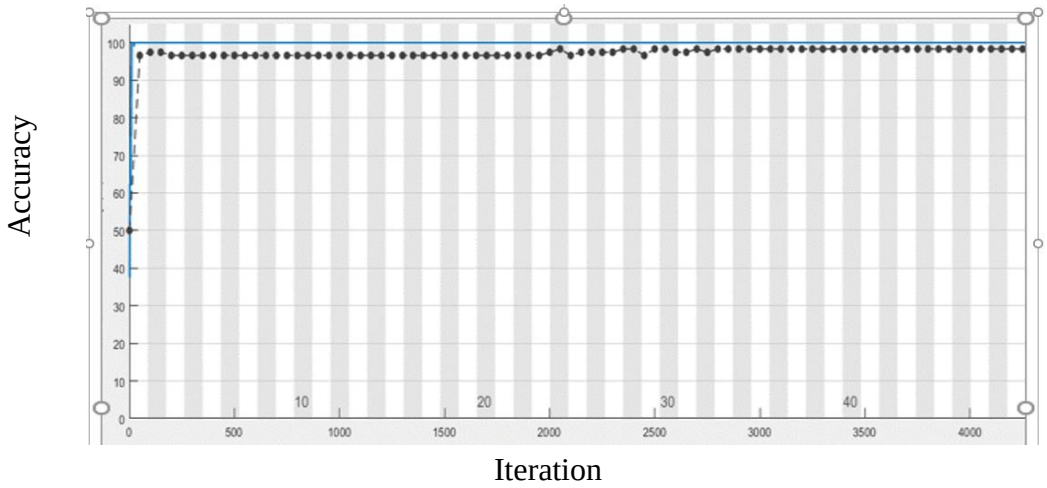


Fig.4.13. Accuracy using Extended YALE B Database after the application of preprocessing techniques

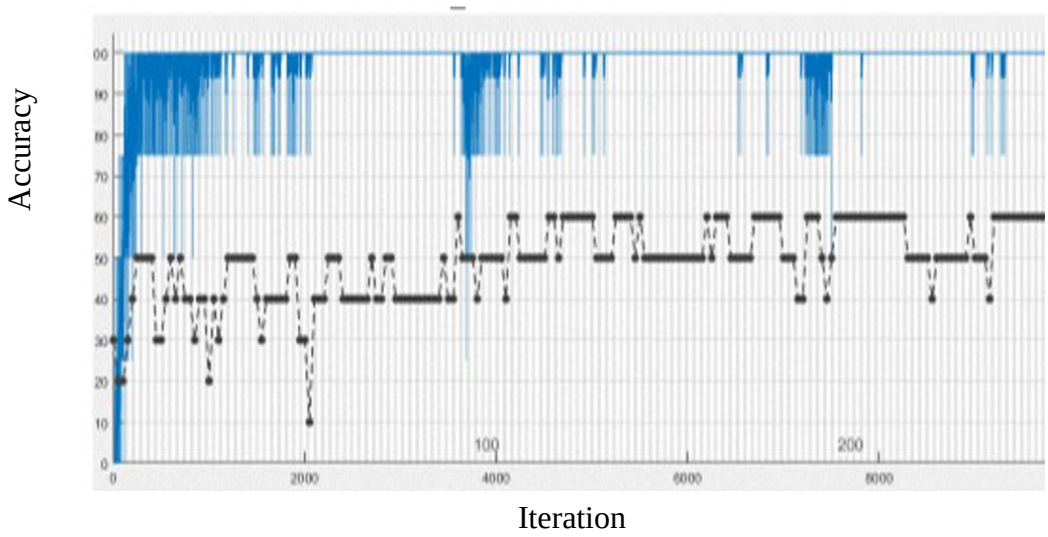


Fig.4.14. Accuracy using FERET Database before the application of preprocessing techniques.

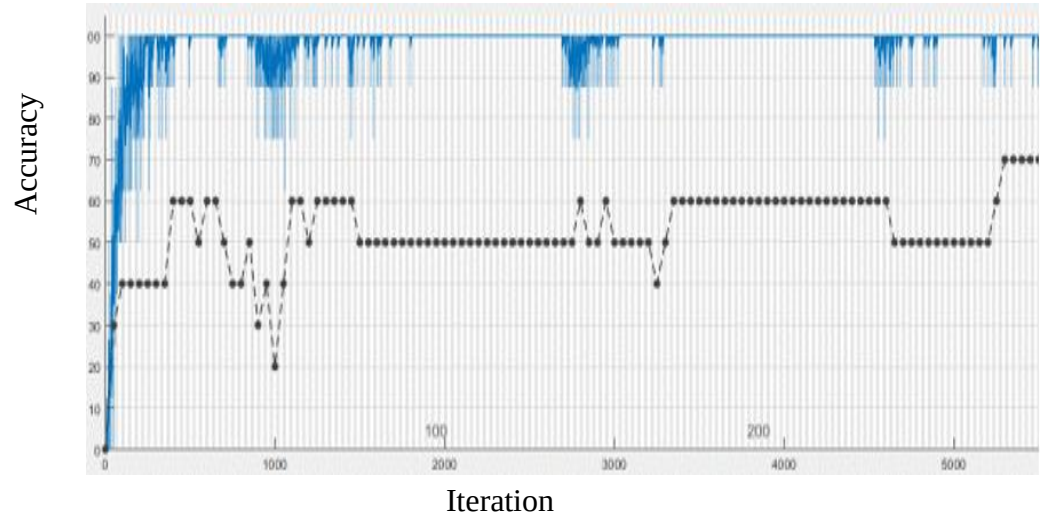


Fig.4.15. Accuracy using FERET Database after application of preprocessing techniques.

Table 4.1 Accuracy rates of Extended Yale B database for various batch sizes before and after application of preprocessing techniques.

Database	**Accuracy Rates**			
	Batch Size			
	4	**8**	**16**	**32**
Extended Yale B (WOAPP)	97.2	97.0	96.8	96.2
DoG	99.1	99.0	98.7	97.9
SQI	99.8	99.4	98.8	98.3
LTISN	99.2	99.3	98.4	98.1
HE	99.7	99.1	99.1	98.3
GIC	99.6	98.9	98.8	98.7

Table 4.2 Accuracy rates of FERET database for various batch sizes before and after the application of preprocessing techniques

Database	**Accuracy Rates**			
	Batch Size			
	4	**8**	**16**	**32**
FERET (WOAPP)	71.4	71.2	71	68.7
DoG	76.6	72.6	71.3	69.9
SQI	74.9	74.6	71.5	71.0
LTISN	75.6	72.5	72.0	70.8
HE	76.3	76.4	76.1	69.2
GIC	74.2	71.3	71.0	70,9

4.5.1 Number of trainable parameters calculation

The number of trainable parameters is computed for each layer as below. The deep CNN model architecture is as shown in (Fig.4.16).

C_1 Layer

Input image 128×128×1 and having one channel.

Number of weights per filter=3×3×1

There are totally eight filters in C1 layer and there is one bias parameter for each filter.

Total number of trainable parameters in C1 Layer is = (3×3×1+1) × 8= 80.

C_2 Layer

Input feature maps to C_2 layer having 8 channels.

Number of weights per filter is 3×3×8.

There are 8 such filters in the layer. Hence total number of weights = ((3×3×8) +1) ×8 and there is one bias for each filter.

Total weights in C_2 = ((3×3×8) +1) ×12=584.

S_1 Layer

S_1 layer is using max pooling operation. Hence no trainable parameter available for this layer

C_3 Layer

Input to C_3 has eight feature channels.

Number of weights per filter is 3×3×8.

There are 16 such filters in the layer. Hence total number of weights = ((3×3) ×8+1) ×16 and there is one bias for each filter.

Total weights in C_3 = (3×3×8+1) ×16=1168.

C_4 Layer

Input to C_4 has sixteen feature channels.

Number of weights per filter is 3×3×16.

There are 16 such filters in the layer. Hence total number of weights = ((3×3) ×16+1) ×16 and there is one bias for each filter.

Total weights in C4 = (3×3×16+1) ×16=2320.

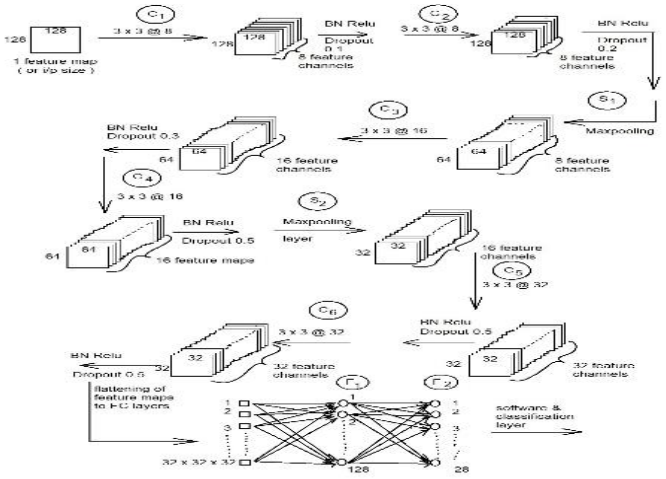


Fig.4.16. Deep CNN Model Architecture (BN-Batch Normalization)

S_2 Layer

S_2 layer is using max pooling operation. Hence no trainable parameter available for this layer

C_5 Layer

Input to C_5 has sixteen feature channels.

Number of weights per filter is 3×3×16.

There are 32 such filters in the layer. Hence total number of weights = ((3×3) ×16+1) ×32 and there is one bias for each filter.

Total weights in C_4 = (3×3×16+1) ×32=4640.

C_6 Layer

Input to C_6 has 32 feature channels.

Number of weights per filter is 3×3×32

There are 32 such filters in the layer. Hence total number of weights = (3×3 ×32+1) ×32 and there is one bias for each filter.

Total weights in C_4 = (3×3×32+1) ×32=9248.

F_1 Layer

In the fully connected layer, the input is having total feature space of size 32×32 and 32.

Input samples to F_1 layer=32×32×32

There are 128 neurons in F_1 layer and each have a bias.

Total Number of trainable parameters for F_1 layer is = (32×32×2+1) ×128=4194432.

F_2 Layer

The input to F_2 layer is the output from F_1 layer. It is connected to 28 neurons (since there are 28 classes). So, total number of trainable parameters = (128 +1) ×28= 16512.

Feature maps extracted for different batch size and epochs are as shown in the (Fig.4.17,4.18,4.19).

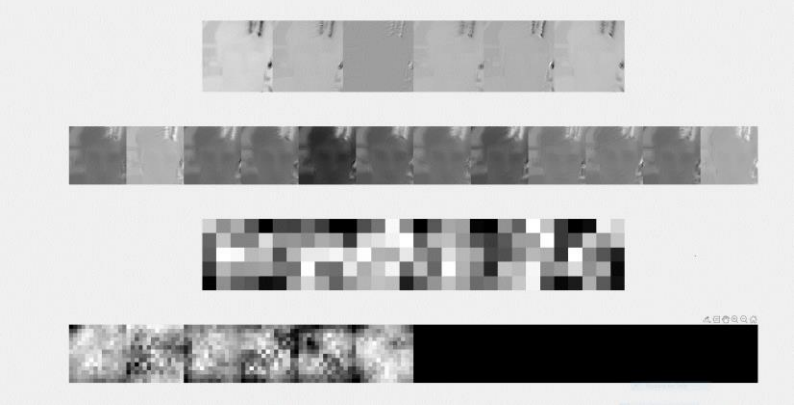


Fig.4.17. Feature map obtained for the batch size 32 and epoch 2 in Extended Yale B database.

Fig.4.18. Feature map obtained for the batch size 32 and epoch 5 in Extended Yale B database.

Fig.4.19. Feature map obtained for the batch size 32 and epoch 25 in Extended Yale B database.

4.6. Comparison with Other Approaches

In this section, the proposed approach is compared with the various approaches. Hybrid system, which combines Convolutional Neural Network (**CNN**) and a Logistic Regression Classifier **(LRC)** yields the accuracy of 80%. A CNN is trained to recognize facial images and LRC is used to classify the features learned by the convolutional network [26]. The neural network with the nonlinear theories, such as wavelet theory and fuzzy set is a new research idea, when applied in face recognition, gives accuracy around

93% in Wavelet-BP [25]. CNN architecture along with Max-Feature-Map **(MFM)** could extract compact vital information yields the accuracy rate of 98% [28]. LeCun proposed popular LeNet-5 architecture [3]. In FERET database, it yields the accuracy rate is 81.25% [27]. The pre-trained CNN model along with the VGG-Face **(Visual Geometric Group)** provides face verification rate accuracy of 86.85% and 83.32% on FRGC and LFW databases respectively [29].

The proposed method is compared with existing approaches are given in Table 3. The proposed method performs better in comparison to all other approaches in Extended Yale B. In FERET database, **CNN-LeNet** method performs better than the proposed method.

Table 4.3. Comparison of Accuracy rates of different approaches.

Architecture used	Dataset	Accuracy
Proposed method (CNN)	Extended Yale B FERET	99.80% 76.30%
VGG+CNN [29]	FRGC, LFW	86.85%, 83.32%
CNN with MFM [28]	LFW	98.00%
Wavelet-BP [25]	AT&T	93.00%
CNN-LRC [26]	Yale	80.00%
CNN-LeNet [27]	FERET	81.25%

4.7. Conclusion

In this chapter, deep convolution neural network (CNN) is applied for extracting features and classification. The performance of the network is assessed before and after

applying preprocessing techniques for various batch sizes and epochs on YALE B and FERET dataset.

Adam optimizer with a learning rate of 0.001 is applied in this experiment. The batch sizes used in this experiment are [4,8,16,32];

The optimum batch sizes are chosen using trial and error method. Initially, multiple batch sizes of [4,8,16,32,64,128] are selected for this experiment. Batch size of 4 with learning rate of 0.001 and number of epochs 50 give the best results. Batch sizes of [4,8,16,32] provide consistent and best results. The batch sizes are normally selected in power of 2. Choosing smaller batch size with low value of learning rate yields best result.

The results showed the maximum accuracy rate of 97.2% of achieved without using preprocessing techniques. When the preprocessing techniques are applied, the improved accuracy rates are achieved up to 99.8%.

The same experiments are conducted in FERET. dataset also. For the frontal face, the maximum 71.4% is achieved without optimization. After applying the preprocessing techniques, the accuracy rate increased to 76.3%.

SQI and HE perform equally well and provide best results for all the batch sizes in both Extended Yale B and FERET database.

The proposed approach performs very well as compared to existing methods.

References

1. S. Setiowati, Zulfanahri, E. L. Franita and I. Ardiyanto and "A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods" 9*th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Phuket, pp. 1-6, 2017.

2. A. Marttinez and A. Kak, "PCA versus LDA". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No.2, pp. 228-233, 2001.

3. R. Haridas and R.L. Jyothi "Convolutional Neural Networks: A Comprehensive Survey" International Journal of Applied Engineering Research Volume 14, Number 3, pp. 780-789, (2019).

4. I.J. Goodfellow, Y. Bengio and A. Courville Deep Learning MIT Press, (2016).

5. J. Wang and L. Perez The Effectiveness of Data Augmentation in Image Classification using Deep Learning. In: Stanford University research report, 2017.

6. D. Huang, C. Shan, M. Ardebilian, Y. Wang, and L. Chen "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey" IEEE Transactions on Systems, Man and Cybernetics, part C, volume 41, pp 765-781, 2011.

7. A. Shailaja Patil and P. J. Deore "Face Recognition: A Survey" Informatics Engineering, an International Journal (IEIJ), Vol.1, No.1, December 2013.

8. C. Shorten and T.M. Khoshgoftaar "A survey on Image Data Augmentation for Deep Learning" Journal of Big Data, - Springer, 2019.

9. F. Allenberger and C. Lenz "A Non-Technical Survey on Deep Convolutional Neural Network Architectures" Computer Vision and Pattern Recognition, 2018.

10. S.V.G. Reddy, K. Thammi Reddy, V. ValliKumari "Optimization of Deep Learning using various Optimizers, Loss functions and Dropout" International Journal of Recent Technology and Engineering (IJRTE) Volume-7 Issue-4S2, pp 448-455, December 2018.

11. D. Saez Trigueros, L. Meng and M. Hartnett “Face Recognition: From Traditional to Deep Learning Methods” J. Theory. Appl. Inf. Technol. 2018, 97, 3332–3342.

12. A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis “Deep Learning for Computer Vision: A Brief Review” Computational Intelligence and Neuroscience, pp 1-13, 2018.

13. Z. Alom, T. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, M. Hasan, B.V. Essen, A. Awwal and V. Asari “A State-of-the-Art Survey on Deep Learning Theory and Architectures”. *Electronics* 2019, *8,* 292.

14. W. Rawat and Z. Wang “Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review” Volume 29, pp 2352-2449 September 2017.

15. A. F. Agarap. Deep learning using rectified linear units (relu). CoRR, abs/1803.08375, URL http://arxiv.org/abs/1803.08375 , 2018.

16. A. Khan, A. Sohail, U. Zahoora and A.S. Qureshi, “A Survey of the Recent Architectures of Deep Convolutional Neural Networks”, published in Artificial Intelligence Review, pp 1-62, (2019).

17. D. H. Hubel and T. N. Wiesel. “Receptive fields of single neurons in the cat’s striate cortex” The Journal of Physiology, 148(3):574–91, 1959.

18. R. Yamashita, M. Nishio, R.K.G. Do, K. Togashi “Convolutional neural networks: an overview and application in radiology” Insights Imaging 9, pp 611–629, 2018.

19. A. Masys, M. Alazab and M. Tang “Deep Learning Applications for Cyber Security”, Springer, ISBN 978-3- 030-13057-2, May 2019.

20. N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," *2017 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, pp. 0588-0592, doi: 10.1109/ICCSP.2017.8286426, 2017.

21. I. Nusrat and S. Jang. “A Comparison of Regularization Techniques in Deep Neural Networks.” *Symmetry* 10 (2018): 648.

22. B. S. Manjunath, R. Chellappa and C. von der Malsburg, "A feature-based approach to face recognition," *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Champaign, IL, USA, pp. 373-378, 1992.

23. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov "Dropout: a simple way to prevent neural networks from overfitting." *J. Mach. Learn. Res.* 15, pp 1929-1958, (2014).

24. Q. Zhao, Q. Lyu and S. Zhang "Multi activation Pooling Method in Convolutional Neural Networks for Image Recognition." *Wireless Communications and Mobile Computing* (2018): n. page. *Wireless Communications and Mobile Computing*. Web, 2018.

25. J. Kittler, P. Koppen, et. al, "Conformal Mapping of a 3D Face Representation onto a 2D Image for CNN Based Face Recognition," 2018 International Conference on Biometrics (ICB), Gold Coast, QLD, pp. 124-131, 2018.

26. H. Khalajzadeh, M. Manthouri and M. Teshnehlab, 'Face recognition using convolutional neural networks and simple logistic classifier', Advances in intelligent systems and computing, 223, pp 197-207, (2014).

27. A.R. syafeeza, M.K. Hani, S.S. Uiw and R. Bakhteri,' Convolutional neural network for face recognition with pose and illumination variation', International journal of engineering and Technology, 6(1), pp 44-57, (2014).

28. Y. Xu, F. Liang, G. Zhang and H. Xu, 'Image intelligent detection based on the Gabor wavelet and the neural network', Symmetry, 8(130), (2016).

29. X. Wu, R. He and Z. Sun "A Lightened CNN for Deep Face Representation." *ArXiv* abs/1511.02683 (2015): n. pag.

30. Z. Lu, X. Jiang and A. Kot, "Enhance deep learning performance in face recognition," 2*nd International Conference on Image, Vision and Computing (ICIVC)*, Chengdu, pp. 244-248, 2017.

31. S. Pouyanfar, S. Sadiq, Y. Yan, H.Tian, Y. Tao, M.P. Reyes and S.S. Iyengar A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys*, *51*(5), (2018).

32. Y. Admi, Y. Moses and S. Ulman "Face Recognition: The problem of compensating for changes in illumination direction", IEEE Transaction. Pattern Anal. Mach. Intell., vol. 19, pp 721–732, Jul. 1997.

33. Soodeh Nikan and M. Ahmadi. "Human Face Recognition under Occlusion using LBP and Entropy Weighted Voting" 21st International Conference on Pattern Recognition ICPR pp. 1699-1702, 2012.

34. S. Anila, K. Arulsukanya, K. Divya, C. Kanimozhi, and D. Kavitha "An efficient Preprocessing Technique for Face Recognition under difficult Lighting Conditions," in Proc. National Conference on Emerging Trends in Computer Communication and Informatics (ETCCI-2011), March 10-11, 2011.

35. H. Han, S. Shan, X. Chen and W.Gao "A comparative study on illumination preprocessing in face recognition" Pattern Recognition., Volume 46, Issue 6, Pages 1691-1699, June 2013.

36. G. Calvo, B. Baruque and E. Corchado (2013) Study of the Pre-processing Impact in a Facial Recognition System. In: Pan JS., Polycarpou M.M., Woźniak M., de Carvalho A.C.P.L.F., Quintián H., Corchado E. (eds) Hybrid Artificial Intelligent Systems. HAIS, Lecture Notes in Computer Science, vol 8073. Springer, Berlin, Heidelberg, 2013.

37. P.A. Koringa, S.K. Mitra and V.K Asari "Handling Illumination Variation: A Challenge for Face Recognition" In: Raman B., Kumar S., Roy P., Sen D. (eds) Proceedings of International Conference on Computer Vision and Image Processing. Advances in Intelligent Systems and Computing, vol 460. Springer, Singapore, 2017.

38. A. Sri Krishna, G. Srinivasa Rao and M. Sravya "Contrast Enhancement Techniques using Histogram Equalization Methods on Color Images with Poor Lightning", International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.3, No.4, August 2013.

39. P. Suganya, S. Gayathri, N. Mohanapriya “Survey on Image Enhancement Techniques” International Journal of Computer Applications Technology and Research, Volume 2– Issue 5, 623 - 627, 2013.

40. M. Biglari, F Mirzaei and H. Ebrahimpour-Komeh “Illumination invariant face recognition using SQI and weighted LBP histogram”. Pattern Recognition and Image Analysis (PRIA), 2013 First Iranian Conference on. IEEE, pp 1-7, 2013.

41. Krieger, E, V. Asari, and S. Arigela “Color image enhancement of low-resolution images captured in extreme lighting conditions” *Sensing Technologies and Applications*, International Society for Optics and Photonics, 2014.

42. S. Anila and N. Devarajan. "Preprocessing technique for face recognition applications under varying illumination conditions." Global Journal of Computer Science and Technology Volume 12, issue 11, Version 1, 2012.

43. N. Longkumer, M. Kumar and R. Saxena “Contrast Enhancement Techniques using Histogram Equalization: A Survey” International Journal of Current Engineering and Technology, Vol.4, No.3, E-ISSN 2277 – 4106, P-ISSN 2347 – 5161 (June 2014).

44. A. Reza "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement." Journal of VLSI Signal Processing, vol. 38, pp 35-44, 2004.

45. R. C. Gonzalez and R. E. Woods, “Digital Image Processing”, Vol. 1, Parson, 2008.

Chapter 5

Conclusion

5.1 Conclusion

The research work conducted in this thesis addresses the challenge of improving accuracy rates of recognizing faces. Preprocessing methods like **DoG, HE, CLAHE, LTISN, SQI and GIC** help to improve the accuracy rate of face recognition process. This research covers two steps of computing accuracy rates of face recognition process as follows.

First, the performance of face recognition algorithms, namely weighted entropy-based fusion method and deep CNN model, are evaluated in FERET and Extended Yale B databases. The results show that weighted entropy-based fusion method provides the better accuracy rates without applying preprocessing techniques in FERET database in comparison to deep CNN model.

In the second step, above mentioned preprocessing techniques are applied to both weighted entropy-based fusion method and deep CNN model to further improve the accuracy rates.

After the application of preprocessing techniques, the accuracy rates improved in both weighted entropy-based fusion method and deep CNN model. Especially, weighted entropy-based fusion method provides better accuracy rates in comparison to deep CNN model in FERET database.

The accuracy increases with increase in block size and with the threshold value up to 7 in weighted entropy-based fusion method before and after application of preprocessing techniques. The highest accuracy rates are achieved

with the block size of 8*8 and for the threshold value of 2.5. The accuracy decreases with decrease in block size and increase in threshold beyond the value 7. The lowest accuracy rates are achieved with the block size of 2*2 and for the threshold value of 9.5.

In general, the classification accuracy rate increases with the increase in the number of sub-blocks. It is because when there is a large number of blocks, the percentage of good blocks which carry the relevant information (with high entropy weights) will be more. Since the blocks with irrelevant information are assigned with lesser entropy weights, the impact of such contaminated blocks will be less in the overall accuracy.

The selection of threshold value helps to choose the significant sub images in terms of entropy. The sub blocks will have high entropy values if it represents significant information. For corrupted or uniform background regions, entropy will be less. Hence the entropy threshold value can be selected to filter the insignificant sub blocks during the analysis. As the threshold increases more sub-blocks will be eliminated from voting scheme. The entropy threshold value is chosen empirically to retain the best possible accuracy in each case.

SQI, DoG and LTISN perform very well for all the block sizes and threshold values up to 7. Next to that HE and GIC perform well. Beyond the threshold value 7, the applied preprocessing technique do not have any effect in increasing the accuracy rate of weighted entropy-based fusion method.

In deep CNN model, the accuracy rate increases with low value of learning rate and small batch sizes, before and after application of preprocessing techniques. The highest accuracy rates are achieved with batch size of 4 and learning rate of 0.001 for 50 epochs. The accuracy rates are consistent and stable with smaller batch sizes and it decreases as the batch size increases. The highest accuracy rates are achieved with batch size of 4 and learning rate of 0.001 for 50 epochs.

The optimum batch sizes are chosen using trial and error method. Initially, multiple batch sizes of [4,8,16,32,64,128] are selected for this experiment. Batch

size of 4 with learning rate of 0.001 and number of epochs 50 give the best results. Batch sizes of [4,8,16,32] provide consistent and best results. The batch sizes are normally selected in power of 2. Choosing smaller batch size with low value of learning rate yields best result.

Both HE and SQI perform equally good for all the batch sizes in deep CNN model. After the application of preprocessing techniques, very high accuracy rate is achieved in deep CNN model in Extended Yale B database.

5.2 Future Work

This research addresses the challenge of improving the accuracy rate of human face recognition with the help of preprocessing techniques. Therefore, an appropriate and more challenging databases like FERET and Extended Yale B databases are chosen for this work. Future work can address working with bigger database with powerful GPU's to further increase the performance of face recognition process. Performance of deep learning method in face recognition is higher for large dataset with high dimensionality.

VITA AUCTORIS

NAME:	Jayanthi Raghavan.
PLACE OF BIRTH:	Vellore, Tamil Nadu, India.
YEAR OF BIRTH:	1972.
EDUCATION:	Madurai Kamaraj University, ME., India, 2002.