

# An Empirical Evaluation of the Local Texture Description Framework-Based Modified Local Directional Number Pattern with Various Classifiers for Face Recognition

R. Reena Rose<sup>1\*</sup>, K. Meena<sup>2</sup>, A. Suruliandi<sup>3</sup>.

<sup>1</sup>Department of Computer Applications, St. Xavier's Catholic College of Engineering, Nagercoil, India;

<sup>2</sup>Department of Computer Science and Engineering, VelTech Dr. R.R. & Dr. S.R. Technical University, Chennai, India; <sup>3</sup>Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India.

## ABSTRACT

*Texture is one of the chief characteristics of an image. In recent years, local texture descriptors have garnered attention among researchers in describing effective texture patterns to demarcate facial images. A feature descriptor titled Local Texture Description Framework-based Modified Local Directional Number pattern (LTDF\_MLDN), capable of encoding texture patterns with pixels that lie at dissimilar regions, has been proposed recently to describe effective features for face images. However, the role of the descriptor can differ with different classifiers and distance metrics for diverse issues in face recognition. Hence, in this paper, an extensive evaluation of the LTDF\_MLDN is carried out with an Extreme Learning Machine (ELM), a Support Vector Machine (SVM) and a Nearest Neighborhood Classifier (NNC) which uses Euclidian, Manhattan, Minkowski, G-statistics and chi-square dissimilarity metrics to illustrate differences in performance with respect to assorted issues in face recognition using six benchmark databases. Experimental results depict that the proposed descriptor is best suited with NNC for general case and expression variation, whereas, for the other facial variations ELM is found to produce better results.*

**Key words:** Face Recognition; Texture Descriptors; Local Texture Description Framework; Local Texture Description Framework-Based Modified Local Directional Number Pattern; Classifiers; Distance Metrics

---

\* Author for correspondence: mailto:reenarose@gmail.com

## INTRODUCTION

Face recognition systems include two key steps: feature extraction and classification. Features are effective representations of face images that can decrease the computational burden of classifiers, while classifiers distinguish face images based on their features. These features, as well as the classifiers, greatly impact the performance of human face recognition systems. It is essential, therefore, to choose a good feature extractor-and-classifier combination, in turn maximizing the effectiveness of face recognition systems.

In face recognition researches, the Nearest Neighborhood Classifier (NNC) [1], Support Vector Machine (SVM) [2] and Extreme Learning Machine (ELM) [3] are the most frequently used classifiers. The NNC is a simple yet powerful classifier extensively used in sundry fields, proving to be robust on many datasets [4, 5, 6] in face recognition. The SVM is a sophisticated classifier with a wide range of applications, particularly in computer vision, where it has been used for an array of applications such as face recognition [7], personal identification [8], background subtraction [9], and hand gesture detection [10]. However, in situations where the exact points are known, basic SVM models can be used [11]. Though the SVM dominates existing computational intelligence techniques in the field, it faces certain challenges in terms of low learning speed, less computational scalability and trivial human intervention. The ELM, a piece of emerging technology, has overcome several of these issues to perform better than SVM in solving applications in the fields of regression and classification [12]. A key characteristic of this classifier is that the hidden layer needs no tuning [13].

In the NNC classification algorithm, the kernel part is the matching process which defines the resemblance or disparity between the features of a probe image and that of a gallery image. There exist numerous distance measures for the matching process. An empirical evaluation for texture classification was carried out earlier by Rubiner et al. [14]. The success of the NNC depends heavily on the distance measures used, given that the performance of these different distance metrics depends on factors such as the datasets, experimental setup and applications. Hence, in this paper, the NNC's performance is reviewed with different distance metrics such as chi-square, Manhattan, G-statistics, Euclidean and Minkowski.

### Motivation and Justification

Features are essential to characterize face images. Facial features include facial skin color, skin texture, the geometrical shape of facial components and so on. Of the existing facial features, skin texture is found to work well in face recognition with different facial variations. Texture description can be done in two ways. One describes the entire face, a gray level co-occurrence matrix [15] being an instance of such a method. The other describes a local region in the face and collects all the local information available into a histogram to represent the whole face. The second type of method, known as a local texture descriptor, is effective with pose variation and scaling, the local binary pattern (LBP) [16] being an instance of such. This descriptor is computationally simple and encodes rotational invariant local texture patterns. The LBP encodes the first-order circular derivative of local patterns by concatenating binary gradient directions [17]. Several local texture descriptors have been applied for face recognition. Depending on whether the descriptors encode intensity transition or the directional information of the local neighborhood's intensity transition, they can be categorized into three. The first category, such as LBP, encodes the local neighborhood's intensity transition while the second, such as the local directional number pattern (LDN) [4] encodes directional information and the third, such as local tetra patterns (LTrPs) [18],

encode both. Although several categories of descriptors have evolved, many real-time challenges continue to remain unaddressed. Further, it is found that almost all start the description with eight neighbors, but pixels close to each other can have similar intensity values because of the likeness of skin tone in regions nearby. Hence, a framework named local texture description framework (LTDF) [19] was proposed earlier by us to direct the descriptors available, in the selection of sampling points from unlike regions so as to encode effective local patterns in face images. It is also noticed that certain descriptors like the local directional number patterns (LDN) [4] utilize the Kirsch mask, which complicates LTDF use in descriptors. Therefore, in view of replacing the Kirsch mask in face recognition, a vector equation is introduced by us while proposing a modified version of the LDN, the LTDF-based modified local directional number pattern (LTDF\_MLDN) [20]. This descriptor has performed better under different circumstances and has been evaluated only with the NNC, which uses the chi-square dissimilarity metric. However, a descriptor's performance can vary with different classifiers and distance metrics. Motivated by this, an attempt is made in this paper to review the best suited classifier for the LTDF\_MLDN descriptor for varied real-time challenges in face recognition.

So far, no research has been carried out to evaluate the performance of different classifiers under sundry real-time challenges in face recognition. However, a classifier's performance depends entirely on the training and testing datasets used. Owing to the use of the NNC, SVM and ELM classifiers in numerous, recent face recognition researches, the performance of the proposed descriptor LTDF\_MLDN is analyzed, in this work, with the aforesaid classifiers to review their effectiveness in face recognition under different circumstances.

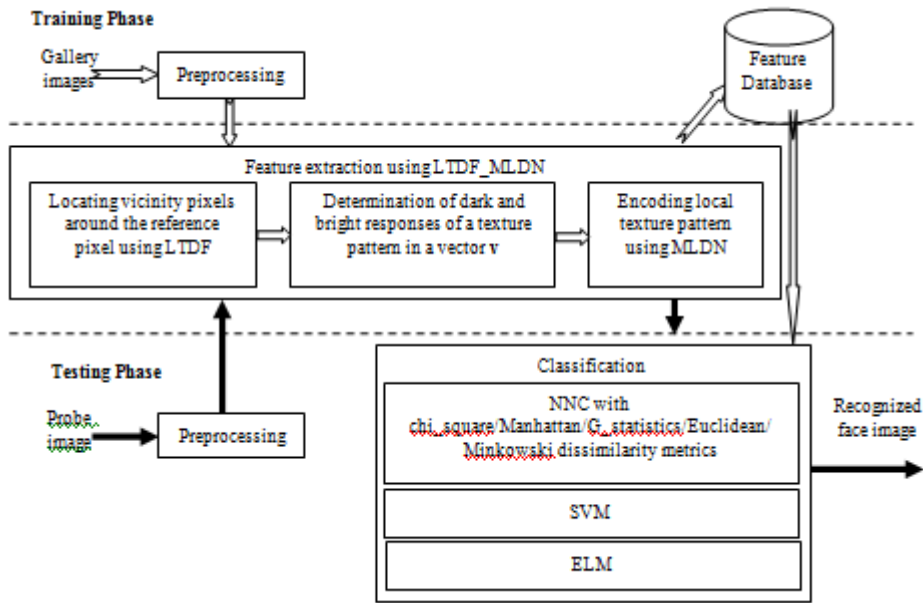
### **Outline of the Work**

The overall processing of face recognition is illustrated in Figure 1. As part of preprocessing, all images are eye aligned into the same canonical pose following which certain regions covering the image's essential features are cropped. Thereafter, the system is trained with the LTDF\_MLDN texture patterns of a gallery set of images by storing the features in the form of a histogram, in which the occurrence frequency of every texture pattern is collected for every image separately in a database.

While a probe image is given to the system, the LTDF\_MLDN features are extracted from it and the classification carried out separately with the NNC, SVM and ELM classifiers. With respect to the NNC, five different dissimilarity metrics are used to match the probe image features against those of all images in the database.

### **Organization of the Paper**

The remaining part of the paper deals with the different sections and their contents. Section 2 reports the feature extraction process of the LTDF framework and LTDF\_MLDN descriptor. Section 3 presents a brief review of the NNC, SVM and ELM classifiers. The distance measures used along with NNC are explained in Section 4. Section 5 deals with experimental results and discussions. Section 6 concludes the work.



**Figure 1.** Graphical illustration of the face recognition process using the LTDF\_MLDN feature.

## FEATURE EXTRACTION

### Local Texture Description Framework

Local texture description by almost all current local texture descriptors is encoded with certain vicinity pixels closest to a reference pixel in a region. However, when facial skin tone is considered, it can be similar in neighboring regions, resulting in little or no variation in the intensity value of pixels adjacent to each other. Hence, the local texture description framework (LTDF) proposed earlier recommended choosing vicinity pixels a certain distance from each other.

This framework is in two flavors: single ring (LTDFs) and multiple rings (LTDFm). The LTDFs framework recommends the selection of  $np$  (normally eight) vicinity pixels  $v_i$  evenly spaced in a circular (LTDFsc) or elliptical (LTDFse) ring with radii  $hd$  and  $vd$  from the reference pixel. The radii of the ring, in terms of the number of pixels, are greater than 1 ( $hd > 1$  &  $vd > 1$ ), in turn aiding the selection of vicinity pixels from dissimilar regions. The optimum value of the horizontal and vertical distances ( $hd$  &  $vd$ ) depends purely upon the size and type of the images used. The coordinates of the pixels along each angle  $\theta_k$  can be computed using Equations (1) - (3).

$$x_{i_2} = x_i + R_k \cos \theta_k, y_{i_2} = y_j + R_k \sin \theta_k \quad (1)$$

and

$$R_k = \sqrt{\frac{hd^2 * vd^2}{hd^2 \sin^2 \theta_k + vd^2 \cos^2 \theta_k}} \quad (2)$$

where,

$$\theta_k = \left[ \frac{360}{np} * (k-1) \right]^\circ, k = 1 \text{ to } np, \quad (3)$$

In Equation (1),  $x_i$  and  $y_j$  refers to the coordinates of the reference pixel,  $i$  &  $j$  vary from 0 to  $x$  and 0 to  $y$  respectively, and  $y \times x$  is the size of the image.

The LTDFm framework can also be described either in circular (LTDFmc) or elliptical (LTDFme) shapes. Owing to its similarity with the DAISY descriptor [21], the LTDFmc is referred to as loosely coupled DAISY. The LTDFm comprises several rings at a certain distance ( $h_d$ ,  $v_d$ ) from each other. Texture description can be done for an individual ring separately using any local texture descriptor, and then concatenated into a single histogram to represent the whole pattern. Figure 2 depicts the structure of the LTDFs for an  $np$  of 8.

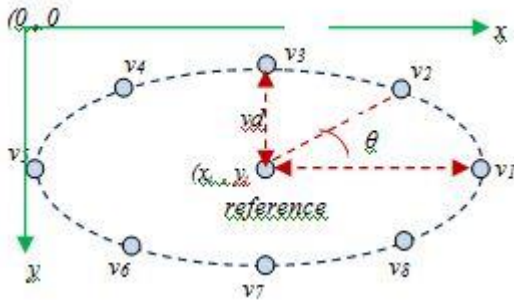


Figure 2. Selection of vicinity pixels by the LTDFs in a local region.

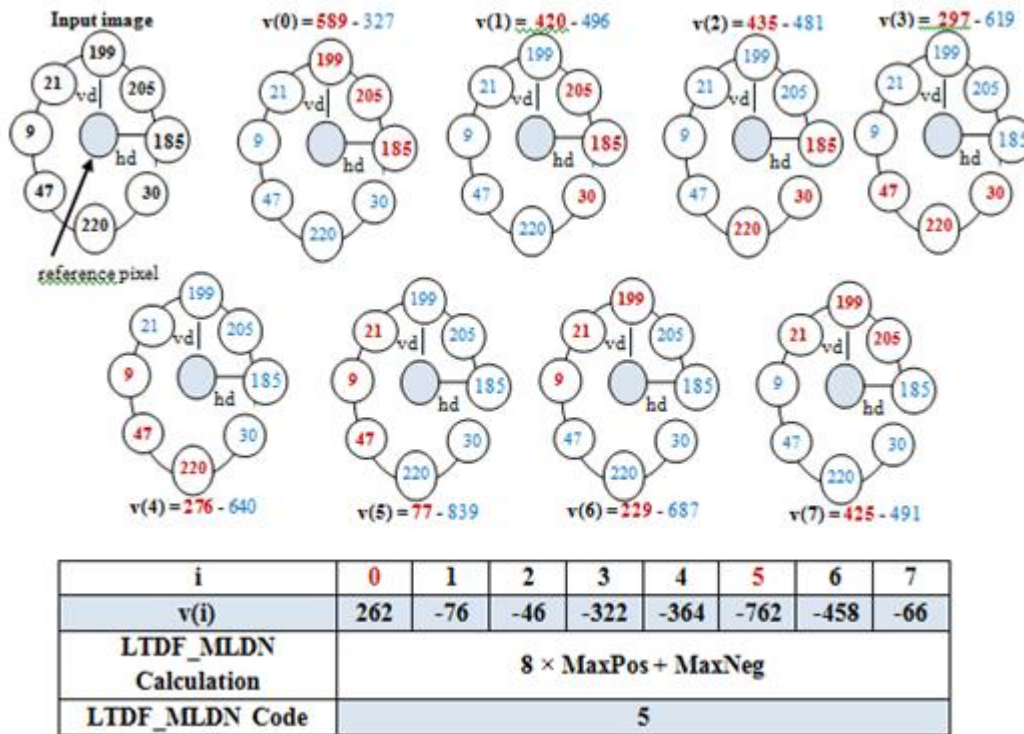
#### Local texture description framework-based modified local directional number pattern (LTDF\_MLDN)

The local texture description framework-based modified local directional number pattern [20] is proposed to describe effective local texture patterns using eight vicinity pixels located in dissimilar regions. The method of encoding this pattern around every pixel in an image encompasses four phases:

1. determining eight vicinity pixels using the LTDFes [19]
2. computing the elements of a vector  $v$  to find dark and bright edge responses of a pattern in each direction  $i$
3. finding the directional indices for which the vector value is minimum and maximum, and
4. encoding the pattern value.

The method of encoding the LTDF\_MLDN pattern is illustrated in Figure 3. The single ring elliptical local texture description framework (LTDFes) recommends that all local texture descriptors encode texture patterns using eight vicinity pixels evenly located on an elliptical ring with radii greater than one from the reference point. This framework effectively describes facial skin texture. Thanks to its ability to enhance the performance of any local texture description that uses closest neighbors, the LTDF\_MLDN chooses vicinity pixels suggested by the LTDFes.

After the vicinity pixels are selected, a signed vector  $v$  which characterizes a texture pattern with the darkness and brightness of edge responses is calculated using Equation (4). An element of the vector in direction  $i$  is determined as the difference in the sum of the intensities of pixels in two parts of the pattern: one with three projected vicinity pixels along direction  $i$ , and another with the remaining. This element is positive when the sum of the three projected vicinity pixels' intensity value is greater than the sum obtained for the other part; otherwise, it is negative.



**Figure 3.** Illustration of vector elements' calculation for a reference pixel, with its vicinity pixels obtained by the LTDFes and the LTDF\_MLDN code determination.

$$MaxNeg_{x,y} = \arg_i \min \{v(i) | 0 \leq i \leq 7\}, \quad (4)$$

$$G = \sum g_{Dir_i}, i = 0, 1, \dots, 7 \quad (5)$$

The vector equation has been derived from the Kirsch compass masks operation by replacing weights 5 & 3 with 2 & 1 respectively, resulting in a decrease in multiplication and addition operations. The proposed equation can replace the use of Kirsch compass masks when weights 5 and 3 in the mask are not significant.

The LTDF\_MLDN pattern is encoded with an index of the maximum positive and maximum negative vector elements, as in Equation (6). This descriptor produces a maximum of 62 texture patterns.

$$LTDFMLD(x,y) = 8 * MaxPos + MaxNeg \quad (6)$$

where

$$v(i) = 2(g_{Dir_i} + g_{Dir_{(i+1)\%8}} + g_{Dir_{(i+2)\%8}}) - G \quad (7)$$

and

$$MaxPos_{x,y} = \arg_i \max \{v(i) | 0 \leq i \leq 7\}, \quad (8)$$

## CLASSIFIERS

### Nearest Neighborhood Classifier

Face recognition is a multiclass problem involving a small sample size. Consequently, many systems utilize the nearest neighborhood classifier (NNC) to arrive at a decision. The goal of the NNC is to match a feature of a probe image with that of a gallery image.

A major issue in NNC design is measuring the similarity. There are two possible ways to measure the similarity: one with distance measures and the other with similarity measures. These two measures are the inverse of each other. There exist several similarity and distance measures. The distance measures used along with the NNC are presented in section 4, and the classification principle using the NNC is given below. Figure 4 explains the algorithm in detail.

Algorithm

Input : Probe image and gallery images

Output: Recognized image

Step 1: Extract global features of gallery images and store them in a database.

Step 2: Extract global features of the probe image.

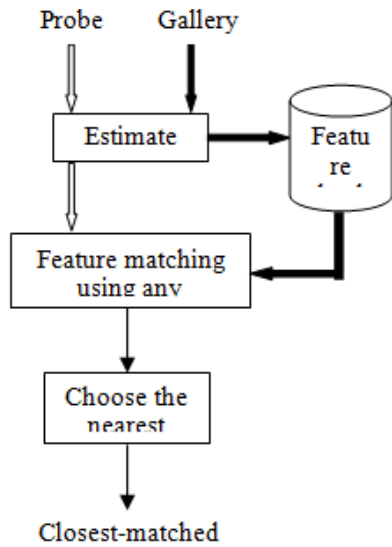
Step 3: Match features of the probe image with corresponding features of the gallery images using a similarity or dissimilarity measure.

Step 4: Choose the closest-matched gallery image as the recognized one.

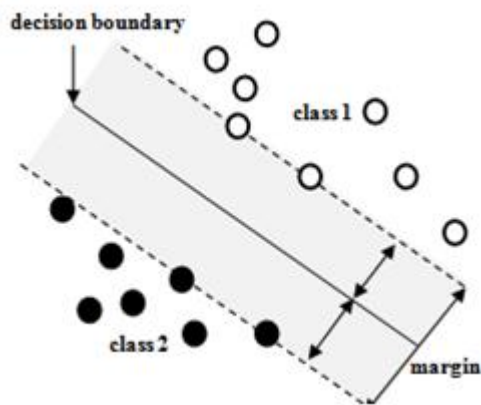
Support Vector Machine (SVM)

The SVM [2] is a binary classifier using the supervised learning paradigm that constructs a decision boundary from the training inputs of two classes. This classifier aims at maximizing the discrimination margin, which is the distance between the decision boundary and the training samples closest to the margin, as depicted in Figure 5. Training samples used to establish a decision boundary are referred to as support vectors. For linearly separable data, there exists an optimal boundary hyperplane that separates the two classes, class 1 and class 2.

The complete platform is generalized into a nonlinear case in which the samples are mapped into a suitable high-dimensional space called feature space. Building a separating hyper plane in the said space results in the construction of a nonlinear decision boundary in input space. Given that the dimensionality of feature space can be very high, the SVM adopts the kernel function. There are assorted kernel functions - polynomial, linear, sigmoid, and radial basis function (RBF), though the most commonly used one is the RBF. In face recognition systems using local texture descriptors, the input to the SVM is a global description of a face image in the form of a 2D histogram, and the output is the class label.



**Figure 4.** Illustration of nearest neighbor classification.



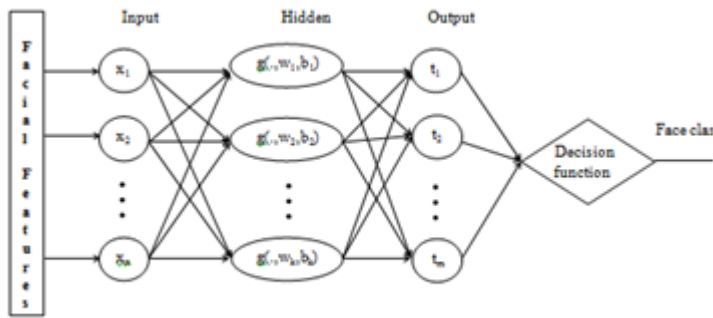
**Figure 5.** Support Vector Machine.

sigmoid, and radial basis function (RBF), though the most commonly used one is the RBF. In face recognition systems using local texture descriptors, the input to the SVM is a global description of a face image in the form of a 2D histogram, and the output is the class label.

#### Extreme Learning Machine (ELM)

The ELM is a single hidden layer feed forward neural network (SLFN) proposed by Huang et al. (2006) to tackle issues in traditional gradient descent approaches. Figure 6 shows the architecture of an ELM classifier, a very fast learning method which uses infinitely differentiable activation functions and selects input weights  $w_j$  and the bias of the hidden layer  $b_j$  randomly. The weights for the neurons in the output layer are determined analytically via the Moore-Penrose pseudo inverse of the hidden layer output matrix.





**Figure 6.** Structure of ELM classifier

Since the hidden neurons' weights and biases are assigned randomly, the ELM's recognition rate may change with different trials. Therefore, the larger the number of trials, the greater recognition rate produced by the ELM network.

#### DISTANCE METRICS

##### Chi-Square Distance Metric

The chi-square ( $\chi^2$ ) distance metric is a distance measure used to match two different feature vectors,  $\mathbf{H}_G$  and  $\mathbf{H}_P$ . In face recognition using local texture descriptors,  $\mathbf{H}_G$  and  $\mathbf{H}_P$  are global texture descriptions of facial local texture features of the gallery and the probe image respectively. This measure for the two feature vectors with feature length  $m$  can be defined as:

$$\chi^2(\mathbf{H}_G, \mathbf{H}_P) = \sum_{i=1}^m \frac{(\mathbf{H}_G(i) - \mathbf{H}_P(i))^2}{\mathbf{H}_G(i) + \mathbf{H}_P(i)} \quad (9)$$

##### Manhattan Distance Metric

The Manhattan distance between two feature vectors is the sum of the differences of their corresponding features. The formula for this distance measure is as follows:

$$D(\mathbf{H}_G, \mathbf{H}_P) = \sum_{i=1}^m |\mathbf{H}_G(i) - \mathbf{H}_P(i)| \quad (10)$$

In the equation above,  $m$  is the total number of bins in the histogram  $\mathbf{H}$ , which represents the total number of feature patterns encoded by the descriptors. The element in every bin  $i$  is the occurrence frequency of the corresponding feature pattern in the image.

##### G-Statistics Distance Metric

G-statistics is a nonparametric statistical measure, also known as the log likelihood ratio, used in the classification process.

$$G = 2 \left[ \begin{array}{c} \left[ \sum_{\mathbf{H}_G, \mathbf{H}_P} \sum_{i=1}^m f_i \log f_i \right] \\ - \left[ \sum_{\mathbf{H}_G, \mathbf{H}_P} \left( \sum_{i=1}^m f_i \right) \log \left( \sum_{i=1}^m f_i \right) \right] \\ - \left[ \sum_{i=1}^m \left( \sum_{\mathbf{H}_G, \mathbf{H}_P} f_i \right) \log \left( \sum_{\mathbf{H}_G, \mathbf{H}_P} f_i \right) \right] \\ + \left[ \left( \sum_{\mathbf{H}_G, \mathbf{H}_P} \sum_{i=1}^m f_i \right) \log \left( \sum_{\mathbf{H}_G, \mathbf{H}_P} \sum_{i=1}^m f_i \right) \right] \end{array} \right] \quad (11)$$

where HG is the histogram of the testing sample and HP the histogram of the trained sample, m the total number of bins in the histogram and  $f_i$  the frequency count of the feature pattern at every bin  $i$ .

### Euclidean Distance Metric

The distance between two points in the Euclidean space generally means the shortest distance between them. Since this measure is derived from the Pythagorean theorem, it is often called the Pythagorean metric. If HG and HP are the histograms representing the features of the gallery and probe image respectively, then the Euclidean distance measure  $D$  for  $m$  number of feature patterns can be described as below.

$$\begin{aligned} D(\mathbf{H}_G, \mathbf{H}_P) &= |\mathbf{H}_G - \mathbf{H}_P| \\ &= \sqrt{\sum_{i=1}^m |\mathbf{H}_G(i) - \mathbf{H}_P(i)|^2} \end{aligned} \quad (12)$$

### Minkowski Distance Metric

The Minkowski distance of order  $p$  ( $p$ -norm distance) between two feature vectors HG and HP of the gallery and probe images can be computed using Equation (13).

$$D(\mathbf{H}_G, \mathbf{H}_P) = \left( \sum_{i=1}^m |\mathbf{H}_G(i) - \mathbf{H}_P(i)|^p \right)^{1/p} \quad (13)$$

In the equation,  $m$  is the total number of feature patterns and the element at  $i$  the number of occurrences of the  $i$ th feature pattern in the image.

## EXPERIMENTS AND PERFORMANCE ANALYSIS

### Experimental setup and performance metrics

To investigate the strength of different classifiers along with the proposed descriptor LTDF\_MLDN, an extensive experimental analysis is carried out covering face recognition under different real-time challenges using AT & T [22], Georgia Tech [23], YALE B [24], JAFFE[25], Essex [26] and Indian Face databases [27]. The databases description is given in Table 1. As regards the LTDFes, the  $hd$  and  $vd$  values are varied from 1 to 5 to find out the optimum values. With respect to the Minkowski measure, the order  $p$  is assigned the value 3, and the SVM classifier with the RBF kernel function is used. With regard to the ELM classifier, the activation function used is sigmoid and the hidden nodes are additive hidden nodes. The performance metrics used for the evaluation are reported below.

## Empirical Evaluation for Face Recognition

$$\text{Recognition Rate (RR)} = \frac{\text{Number of probe images correctly recognized}}{\text{Total number of probe images}} \times 100 \quad (14)$$

$$\text{Average Recognition Rate (Avg. RR)} = \frac{\sum_{i=1}^n \text{RR}}{n} \quad (15)$$

$$\text{Standard Error (SE)} = \frac{SD}{\sqrt{n}} \quad (16)$$







In the equations above,  $n$  refers to the number of runs and  $SD$  to standard deviation.

Evaluation of NNC with different distance measures on various issues in face recognition

## Results on Facial Expressions

Face recognition with different facial expressions is a most difficult task, given that facial expressions result in temporarily distorted facial features, leading to false recognition. To evaluate the effectiveness of the LTDF\_MLDN under the NNC with different distance metrics, an experiment is conducted with expression-variant images using the JAFFE and Essex (grimace) databases. The JAFFE database contains images with seven varying facial expressions comprising neutral, sadness, surprise, anger, disgust, fear and joy. The grimace dataset is a subset of the Essex database comprising 360 facial images of 18 subjects with different expressions. A N-fold cross-validation principle is adopted for both databases and the results are depicted in Table 2.

Table 1 - Database Description

Database	Number of images in the database	Number of images used	Experiments Conducted	Sample images
AT & T	400	400	General face recognition and partial occlusion	
Georgia Tech	750	750	General face recognition	
YALE B	640 (frontal pose images with illumination variation)	640 (frontal pose images with illumination variation)	Illumination variation and pose variation	
	90 (10 subjects × 1 image per pose)	90 (10 subjects × 1 image per pose)		
JAFFE	213	213	Expression variation	
Essex	360 (grimace)	360 (grimace)	Expression variation, illumination variation and partial occlusion	
	1440 (faces95)	439 (faces95)		
Indian Face	242 (female)	242 (female)	Pose variation	

**Table 2.** Recognition Rate for expression-variant images by the LTDF\_MLDN with the NNC.

Distance Metrics	Avg. RR±SE(%)	
	Database	
	JAFFE	Essex_grimace
chi_square	<b>100</b>	<b>100</b>
Manhattan	<b>100</b>	<b>100</b>
G_statistics	99.45±0.54	96±1.04
Euclidean	99.45±0.54	99.72±0.27
Minkowski	99.45±0.54	99.44±0.55

From the experimental results, it is noted that both the chi-square and Manhattan distance metrics are capable of achieving 100% accuracy on the JAFFE and Essex (grimace) databases. Hence, these two distance metrics could be the best choice for the NNC to recognize face images with different facial expressions. Though the other distance metrics produce identical results for the JAFFE database, they produce different ones for the Essex (grimace) database, and the G-statistics measure is the one which gives the least results.

### Results on Illumination Variation

Recognition of face images under different lighting conditions is a challenge in computer vision, with changes in illumination greatly affecting classification. Frontal face images of 10 persons with 64 illumination variant image per person from the YALE B database, and frontal face images of 27 persons with controlled illumination variation from the Essex (faces95) database are considered, so as to conduct experiments on illumination-variant images. The subset of Essex is named Essex\_illu in this work. The results obtained for N-fold cross-validation are shown in Table 3.

The results show the effectiveness of the G-statistics metric in the recognition of face images with different lighting conditions. The NNC classifier achieves the maximum recognition accuracy of nearly 98.04% and 99.09% respectively for YALE B and Essex databases, proving the efficacy of the G-statistics metric over other tested distance metrics. The results indicate that the performance of distance metrics on the NNC can fluctuate with variations in face images.

### Results on Partial Occlusion

To analyze the capability of the LTDF\_MLDN with different distance metrics on the NNC for partially occluded face images, frontal face images of 15 and 13 individuals (wearing spectacles) are taken respectively from the AT & T and Essex (faces95) databases. The subset of Essex is named Essex\_po in this work. The experimental results for N-fold cross-validation are given in Table 4.

Among the distance metrics tested along with the NNC, G-statistics yields the best recognition accuracy of nearly 95.99% for AT & T database and 96% for Essex database, showing the effectiveness of the G-statistics distance metric over the others in the recognition of partially occluded face images.

**Table 3.** Recognition rate for face images under illumination variation by the LTDF\_MLDN with the NNC..

Distance Metrics	Avg. RR±SE(%)	
	Database	
	YALE B	Essex_illu
chi_square	97.39±0.45	97.73±0.35
Manhattan	97.57±0.42	97.11±0.6

Distance Metrics	Avg. RR±SE(%)	
	Database	
	YALE B	Essex_illu
G_statistics	<b>98.04±0.32</b>	<b>99.09±0.2</b>
Euclidean	95.83±0.95	96.53±0.62
Minkowski	97.3±0.29	95.79±0.6

**Table 4 . Recognition rate for face images under partially Occlusion by the LTDF\_MLDN using the NNC...**

Distance Metrics	Avg. RR±SE(%)	
	Database	
	AT & T	Essex_po
chi_square	94.66±2.49	94.47±0.96
Manhattan	94.66±2.49	93.43±0.71
G_statistics	<b>95.99±1.63</b>	<b>96±0.65</b>
Euclidean	93.33±2.1	92.79±0.89
Minkowski	90.66±4.98	91.66±1.05

### Results on Pose Variation

All the images from YALE B database (9 different poses of 10 individuals), and the images in female directory of the Indian face database are experimented with to evaluate the performance of the LTDF\_MLDN under the NNC with different distance metrics for face recognition with pose variant images. There are 7 different poses for each person in Indian face database. The results of the N-fold cross-validation scheme are shown in Table 5.

The Manhattan distance metric yields the highest output for the NNC in the recognition of pose variant images, evident from its effectiveness in recognizing face images with different poses. Experimental results clearly show that the performance of the distance metrics can vary with the type of input data given for classification. Due to pose variation of face images, certain information is lost. Even with the available information Manhattan distance is able to give better results.

### Results on General Face Recognition

Recognizing face images with all kinds of variations is, in fact, a challenge. Hence, the proposed descriptor is experimented for general face recognition using AT&T and Georgia Tech databases. N-fold cross-validation scheme was adopted for each database separately and the results given in Table 6.

It is worth noting that the NNC with the chi-square distance metric yields better results for both AT&T and Georgia Tech databases, proving that the combination of LTDF\_MLDN descriptor with NNC using chi-square distance metric is most suitable for general face recognition when all sorts of facial variations are considered.

### Performance evaluation of various classifiers for face recognition under real-time challenges

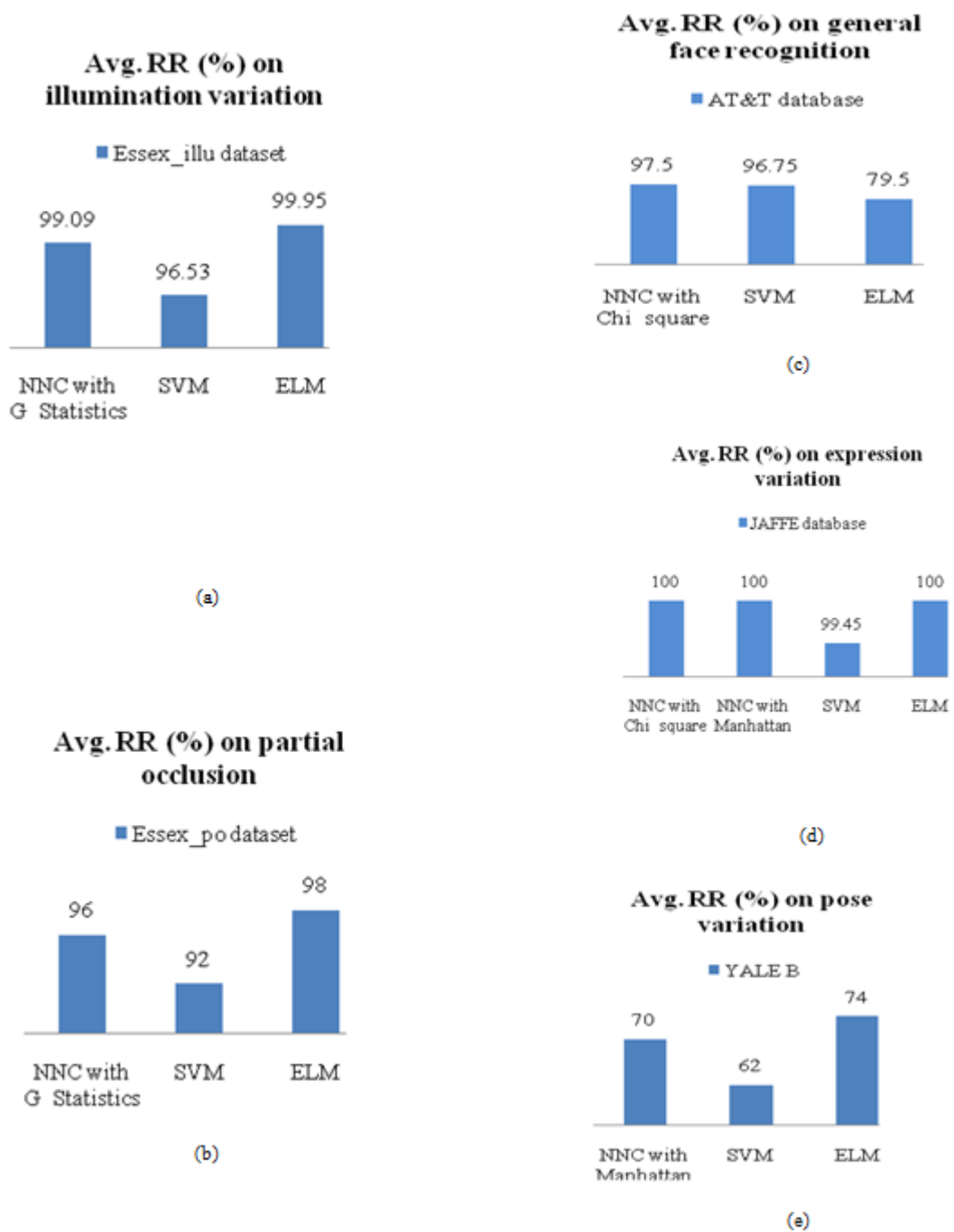
After conducting experiments on NNC with different distance metrics, the classifier which gives better results are compared against the one obtained with the SVM and ELM, and the results are shown in Figure 7.

**Table 5** . Recognition rate for pose variant images by the LTDF\_MLDN with the NNC..

Distance Metrics	Avg. RR±SE(%)		
	Database		
	YALE B	Indian database	Face
chi_square	58±3.57	50.64±5.91	
Manhattan	<b>70±2.1</b>	<b>61.68±6.8</b>	
G_statistics	67±3.2	59.73±7.53	
Euclidean	54±2.23	46.74±5.41	
Minkowski	65±3.8	38.95±5.05	

**Table 6** . Recognition rate for general face recognition using the LTDF\_MLDN with the NNC.

Distance Metrics	Avg. RR±SE(%)	
	Database	
	AT & T	Georgia Tech
chi_square	<b>97.5±1.04</b>	<b>83.73±2.03</b>
Manhattan	95.75±1.34	81.99±1.78
G_statistics	95.75±1.22	81.06±1.82
Euclidean	95±1.77	77.86±2.45
Minkowski	92±1.45	71.46±1.69



**Figure 7.** Performance comparison of the NNC, SVM and ELM.

As depicted in Figure 7, the NNC with chi\_square is found to perform better in terms of general face recognition and for face recognition with different variations in expression. The NNC with the Manhattan distance metric works well for issues relating to variations in expression. The SVM classifier works better in general cases and in expression variations. Though the ELM performs well for expression variation, illumination variation, partial occlusion and pose variation, it is found to perform rather worse than other classifiers in general face recognition with AT & T database. The results indicate that different classifiers work differently for different issues in face recognition and, consequently, it is essential to choose the

best feature descriptor-and-classifier combination for the type of face images used for face recognition.

## CONCLUSION

In this work, the performance of the proposed local texture descriptor LTDF\_MLDN is analyzed with various classifiers such as the NNC, SVM and ELM, to evaluate their effectiveness with respect to various real-time challenges in face recognition. The NNC is tested along with the chi-square, Manhattan, G-statistics, Euclidean and Minkowski distance metrics. Face recognition issues - such as facial expression variations, partial occlusion, pose variation, illumination variation and general cases are experimented with more than 3000 face images from six bench mark databases.

Experimental results show that the performance of a descriptor can vary with the classifier which is tested along with it. The results obtained from various databases reveal that the proposed descriptor LTDF\_MLDN is more suitable in general cases, with NNC using chi-square distance metric. It is noted that the SVM is able to achieve closer results to that of the NNC for general face recognition, while the ELM is the one found to perform well, separately, for all face recognition issues. Compared to other classifiers, however, the ELM with additive hidden nodes produces inferior results where general cases are concerned. However, a combination of the ELM with the RBF or polynomial nodes can produce better results.

## REFERENCES

- Kuncheva, L.I. and Jain, L.C. (1999) Nearest neighbor classifier: simultaneous editing and feature selection. *Pattern Recognition Letters*, 20, 1149-1156.
- Cortes, C. and Vapnik, V. (1995) Support-vector networks. *Machine learning*, 20, 273-297.
- Huang, G.B., Zhu, Q.Y. and Siew, C.K. (2006) Extreme learning machine: Theory and applications. *Neurocomputing*, 70, 489-501.
- Ramirez Rivera, A., Rojas Castillo, J. and Chae, O. (2013) Local directional number pattern for face analysis: Face and expression recognition. *IEEE Transactions on Image Processing*, 22, 1740-1752.
- Liu, F., Tang, Z. and Tang, J. (2013) Weber local binary pattern for local image description. *Neurocomputing*, 120, 325-335.
- Tang, H., Yin, B., Sun, Y. and Hu, Y. (2013) 3D face recognition using local binary patterns. *Signal Processing*, 93, 2190-2198.
- Wei, J., Jian-qi, Z., and Xiang, Z. (2011) Face recognition method based on support vector machine and particle swarm optimization. *Expert Syst Appl*, 38, 4390-4393.
- Prosser, B., Zheng, W.S, Gong, S., Xiang, T., and Mary, Q. (2010) Person re-identification by support vector ranking. *BMVC*, 2, 6.
- Han, B. and Davis, L.S. (2012) Density-based multifeature background subtraction with support vector machine. *IEEE Trans Pattern Anal Mach Intell*, 34, 1017-1023.
- Dardas, N.H. and Georganas, N.D. (2011) Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. *IEEE Transactions on Instrumentation and Measurement*, 60, 3592-3607.
- Wang, X. and Pardalos, P.M. (2014) A Survey of Support Vector Machines with Uncertainties. *Annals of Data Science*, 1.
- Huang, L.L. and Shimizu, A. (2006) A multi-expert approach for robust face detection. *Pattern Recognition*, 39, 1695-1703.
- Huang, G.B., Wang, D.H. and Lan, Y. (2011) Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, 2, 107-122.



- Rubner, Y., Puzicha, J., Tomasi, C. and Buhmann, J. M. (2001) Empirical evaluation of Dissimilarity Measures for Color and Texture. *Computer vision and image understanding*, 84, 25-43.
- Haralick, R.M., Shanmugam, K. and Dinstein, I.H. (1973) Texture features for image classification. *IEEE Transactions on Systems Man and Cybernetics*, 8, 610–621.
- Ojala, T., Pietikainen, M., Maenpaa, T. (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 971–987.
- Nanni, L., Lumini, A. and Brahnam, S. (2012) Survey on LBP based texture descriptors for image classification. *Expert Systems with Applications*, 39, 3634–3641.
- Murala, S., Maheshwari, R.P. and Balasubramanian, R. (2012) Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE Transactions on Image Processing*, 21, 2874–2886.
- Rose, R. R., Suruliandi, A. and Meena, K. (2014) Local texture description framework for texture based face recognition. *ICTACT Journal on image and video processing*, 4, 773-784.
- Rose, R. R., Suruliandi, A. and Meena, K. (2015) Local texture description framework-based modified local directional number pattern: a new descriptor for face recognition. *International Journal of Biometrics*, 7, 147–169.
- Tola, E., Lepetit, V. and Fua, P. (2010) Daisy: an efficient dense descriptor applied to wide\_baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, 815–830.
- Samaria, F.S. and Harter, A.C. (1994) Parameterisation of a stochastic model for human face identification. 2nd IEEE Workshop on Applications of Computer Vision, December 1994, 138–142.
- Nefian, AV. Georgia Tech Face Database. <ftp://ftp.ee.gatech.edu/pub/users/hayes/facedb>  
[http://vision.ucsd.edu/datasets/yale\\_face\\_dataset\\_origin\\_al/yalefaces.zip](http://vision.ucsd.edu/datasets/yale_face_dataset_origin_al/yalefaces.zip)
- Lyons, M., Akamatsu, S., Kamachi, M. and Gyoba, J. (1998) Coding facial expressions with gabor wavelets. *Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, April 1998, 200–205.
- Spacek, L. University of Essex Face Database  
<http://cswww.essex.ac.uk/mv/allfaces/index.html>,
- Asfaw, Y., Scott, G., Pelletier, P. and Adler, A. (2012) Method to evaluate pose variability in automatic face recognition performance. *International Journal of Biometrics*, 4, 373–387.

Received: February 03, 2016;  
 Accepted: July 14, 2016