MDTP: a novel multi-directional triangles pattern for face expression recognition

I. Michael Revina[1] · W. R. Sam Emmanuel[1]

Abstract

Face expression recognition is a key-subject of machine learning, and the primary issue on it is lack of accuracy. This paper proposes a novel face expression recognition method using Decision Based Rule-Oriented Median Filter (DBROMF) and Multi-Directional Triangles Pattern (FER-MDTP). The DBROMF is a novel noise reduction method which removes the impulse noise from the facial images. This FER method enriched with a new image descriptor (MDTP) which is structured via multi-directional triangle pattern to provide a superior image description. An unparalleled novel algorithm to locate the human face organ viz. lip and eyeball stuffed in this research by getting assistance with fuzzy edge strength and abbreviated as an MDTP-FES method. These landmarks of features extracted, and the histogram oriented features tailored with the classification division. The support vector neural network classifier (SVNN) is integrated to conduct the classification job. The JAFFE, CK, TFEID, and ADFES databases are linked to perform the simulation which telescoped with six face expressions. The proposed method lifts the accuracy to a significant range than the existing state-of-the-art methods.

Keywords MDTP descriptor · FER · SVNN · MDTP-FES · Triangle pattern

1 Introduction

Face Expression Recognition (FER) is an attractive field in pattern recognition and computer vision for the period of earlier days. Generally, an Automatic FER system includes three parts such as face tracking and detection, feature extraction and expression classification [1, 2]. In

✉ I. Michael Revina
michaelrevina09@gmail.com

W. R. Sam Emmanuel
sam_emmanuel@nmcc.ac.in

[1]   Department of Computer Science, Nesamony Memorial Christian College affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu 627012, India

the pattern recognition field, Face Expression Recognition is one of the challenging tasks. FER algorithms have some difficulties related to luminance which affects the accuracy. Thus the effective FER system is needed, to overcome these difficulties.

The various feature extraction methods are Adaptive Discriminative Metric Learning (ADML) [3], Scale Invariant Feature Transform (SIFT) [4], Laplacian of Gaussian (LOG), Local Binary Pattern (LBP) [5], etc. The various classifiers are Support Vector Machine (SVM) [6], Radial basis function neural network [7], Deep Neural Network (DNN) [4], Deep Belief Network (DBN), Time Delay NN (TDNN) [8], etc. Several exciting areas have some applications in Automatic Facial Expression Recognition such as robotics [9], telecommunications, video games, automobile safety, health care [10], behavioral science, etc. [1].

Munir et al. suggested that the Merged Binary Pattern Code (MBPC) can perform through zone based holistic manner. MBPC produces the two 8-bit codes such as HV-code and D-code. Before the feature extraction, image enhancement required. So the Contrast Limited Adaptive Histogram Equalization (CLAHE) is used in the frequency domain [11]. Zhang et al. proposed the FER with the fusion of Multi-signal Convolutional Neural Network (MSCNN) and Part-based Hierarchical Recurrent Neural Network (PHRNN). The PHRNN model is useful to extract the temporal features since the consecutive frames as well as MSCNN model is valuable to obtain the spatial features commencing the still frames [12].

Ding et al. offered the two descriptors for automatic FER such as Double Local Binary Pattern (DLBP) and Taylor Feature Pattern (TFP) for feature extraction. The DLBP with Logarithm Laplace (LL) domain efficiently detect the peak frames from the videos also reduces the detection time. The TFP extracts the useful discriminative information from the taylor feature map. [13]. Uddin et al. introduced the depth camera based approaches for proficient FER. The combined method of the Local Directional Rank Histogram Pattern (LDRHP) and the Local Directional Strength Pattern (LDSP) descriptors extracts the spatio-temporal features. These features applied to the Convolutional Neural Network (CNN) for classification of the expressions [14].

Yang et al. proposed the automatic FER via Weighted Mixture Deep NN (WMDNN) which uses the dual channel facial image it includes the grayscale facial image and its similar LBP image. The dynamic features are also extracted and tuning the VGG16 network model is based on ImageNet dataset [15]. Uddin, Hassan et al. proposed the Local Directional Positional Pattern (LDPP) for feature extraction in FER. Local Directional Position Pattern (LDPP) forms the 8-bit binary code for each pixel and extracts the high dimensional texture features. The feature dimensions are reduced using the Principal Component Analysis (PCA), and the robust features are made using Generalized Discriminant Analysis (GDA). Then these features are applied to the Deep Belief Network (DBN) classifier for expression recognition. [16].

Zeng et al. introduced the combination of Histogram of Oriented Gradient (HOG) and LBP descriptors which extracts the high dimensionality features in the form of a mixture of appearance [17] and geometric [18] facial features. These top dimensional features are reduced and fed to the Deep Sparse Autoencoder (DSAE), and it uses the forward propagation to recognize the expressions. Meena et al. suggested the Graphical Signal Processing (GSP) for feature vector dimensionality reduction. The features produced by Discrete Wavelet Transform (DWT) - HOG are high dimensional, and they reduced through GSP. Finally, the classification performed by using the k nearest neighbor classifier [19]. Cruz et al. introduced the Temporal Patterns of Oriented Edge Magnitudes (TPOEM) which based on the temporal and spatial derivatives. The adaptive weighted average procedure used with TPOEM which classifies the expressions [20].

In this paper, a novel DBROMF noise removal filter is used to remove the impulse noises from the input images in a practical way. Also, a novel MDTP descriptor which includes the influence of multi-directions triangle. The pattern is proposed to generate the description of the input face image with tolerance against brightness and luminance variations. The fusion process fused the left, bottom, top, and right direction oriented fuzzy edge strength to accomplish the face organ edge image. The lip and eyeball based features associated with the histogram feature model are extracted and fed as training input to the Support Vector Neural Network (SVNN) classifier to develop an effective classifier. The testing module of SVNN simulates or conducts a face expression detection task which incorporates with the recognizable face expressions viz. disgust, sad, smile, surprise, anger, and fear.

The rest of the paper is structured as follows, and Section 2 gives a brief description of the proposed method. Section 3 illustrates the experimental results and discussion of the proposed plan. The paper ends with a conclusion.

2 The proposed method

This paper proposes an FER method which is compared by lip and eyeball oriented features based on a novel image descriptor MDTP which is incorporated by triangle based window masks derived from multi-directions including the bottom, top, left and right directions.

This method is composed of four major phases.

- DBROMF based noise reduction
- MDTP descriptor image generation
- Fuzzy edge strength based Face-Organ-Edge image generation
- Feature extraction and classification

The input query image processed by DBROMF which removes the salt and pepper noises. The noise-free face image is handled by the four directional triangle patterns which cause edge images to form MDTP image descriptor. The fuzzy edge strength is formulated from each directions MDTP descriptor and the novel fusion process drawn out the landmarks of the lip and eyeball organs. The lip and eyeball oriented histogram features are extracted to serve as input for the SVNN classification process, which is of the category related to Neural Network for benefiting the face expression type recognition. Figure 1 denotes the architecture of the proposed FER method.

2.1 DBROMF based noise reduction

The proposed Decision-Based Rule-Oriented Median Filter (DBROMF) method removes the salt and pepper noises from the facial expression images. At first, the input face image is examined to know whether the image contains noisy pixels. Noisy pixel means the gray level value of the pixel is 0 or 255 if the gray level value of the pixel deceit among 0 or 255. If so the pixel is considered as noise free.

The working procedure of the DBROMF noise reduction described as follows. Firstly the input image is read along with the 3×3 window. After that, every pixel of the input image examined for the occurrence of salt and pepper noise. Imagine that the pixel N_{xy} considered if the pixel intensity value lies between 0 and 255, then the pixel is regarded as an unaffected,
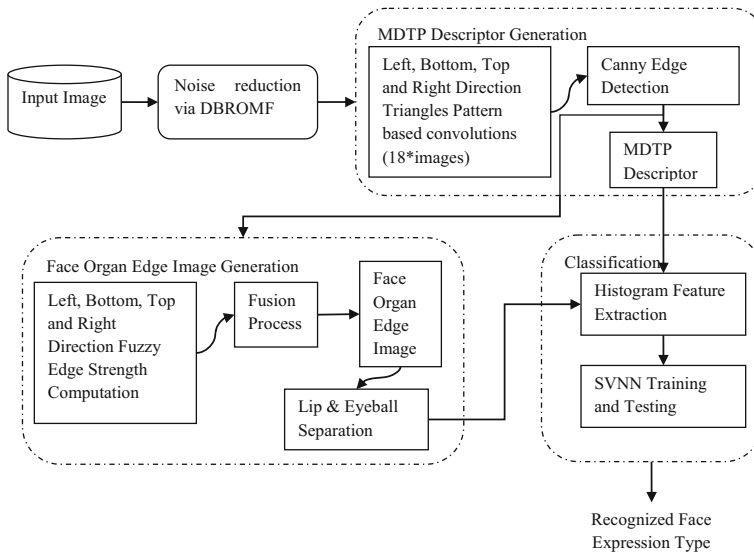
Fig. 1 The Architecture of the proposed FER-MDTP

and the pixel intensity value updated by the same value that means no change. If the pixel value is 0 or 255, the pixel considered as an affected pixel, and two cases are probable. In one case if every element of the chosen window holds 0's and 255's. The mean value should be discovered and update the 0's and 255's values by this mean value. In another case, if the entire elements are not of the chosen window which holds 0's and 255's but median values update some of the holds 0's and 255's and four cases are probable.

- Case 1: if the chosen window element's maximum majority strength is equal to $\frac{(windowsize)^2 - 1}{2}$ then the affected pixels are updated by median value. The median value calculated as follows, if it is single dominant instances, then the single dominant intensity values are taken as the median value. If the multi-dominant instances along with edge supported dominant intensity is available, then the edge supported dominant intensity values considered as median value otherwise average of the dominant intensities considered as the median value.

- Case 2: if the chosen window element's maximum majority strength is equal to $\left(\frac{(windowsize)^2 - 1}{2}\right) - 1$ then the affected pixels are updated by mean value. The median value calculated as follows, if it is single dominant instances, then the median value is calculated as same as in case 1. If the multi-dominant instances along with edge supported dominant intensity is available, then the edge supported dominant intensity values taken as median value otherwise first median value calculated and closest dominant intensity concerning the first median value is considered as the median value.

- Case 3: if the chosen window element's maximum majority strength is less than $\left(\frac{(windowsize)^2 - 1}{2}\right) - 1$ and greater than 1 then the median value is calculated, and it is taken as the median value.

- Case 4: if the chosen window element's maximum majority strength is equal to 1 then the mean value is calculated, and the mean value taken as the median value.

2.2 MDTP descriptor image generation

In conventional systems, the rectangle or square window models used for deriving the image descriptors. So the block level representations are produced for face images and the directional based descriptions and split part representations are missing. In proposed systems, the neighbor based triangle window models used, in which four directional oriented features derived. The multi-model triangles produce the multi-order oriented features and split part representations for face images which are efficient than the conventional systems.

A rectangular window of size 7×7 can subdivide as four triangular windows which reflect the representation of the four directions: bottom, top, left and right. These phenomenons depicted in Fig. 2.

The bottom directional triangle window is shown individually in Fig. 3 which is originated by [i,j] location and besides that the elements notations [a,b,c,…p], hints the addressing of triangle window entities. This bottom directional triangle window is used as a source to invent and organize the new triangle patterns (or sub triangle patterns) which are expressed in Fig. 4 to evaluate the MDTP-based image descriptor, through convolution process.

The concerned 18 patterns are convoluted with the input query image, and the outputs have undergone the canny edge detection process to yield edge detection output images. These edge output images generate an intermediate output which is known by I_{MDTP}, a partial descriptor associated with bottom direction triangle patterns using Eq. (1):

$$I_{MDTP}(i,j) = I_{MDTP}(i,j) + ED(i,j,p) \qquad (1)$$

where $i \in [0, H-1]$, $j \in [0, W-1]$, $p \in [0, P-1]$, I_{MDTP} is the MDTP based image descriptor, ED is the edge detection output from p^{th} triangle pattern related with bottom direction, p denotes the triangle pattern index, H is the image height, W is the image width, P is the whole triangle patterns.

The top directional triangle window of 18 triangle patterns formed that resembled with the bottom direction process. The top directional 18 patterns have undergone the convolution process with query image, and this output is processed by canny edge detection to grant edge detection outputs. This edge detection outputs applied with Eq. (1), and the resultant values projected over the same image matrix I_{MDTP}. The I_{MDTP} image continuously updated by the left and right triangle patterns with a similar mode, and finally, the MDTP image descriptor I_{MDTP} successfully obtained. The entire processing steps united with MDTP image descriptor formation is showcased in the algorithm inner-titled by step 1 and step 2. The MDTP oriented elements depicted in Fig. 7f.

Red   – Top Direction Triangle Window
Green   - Left Direction Triangle Window
Blue   - Right Direction Triangle Window
Yellow   – Bottom Direction Triangle Window

Fig. 2 Four Directional Triangle Windows

Fig. 3 Bottom Triangle Element Representation

				a		
		b	c	d		
	e	f	g	h	i	
j	k	l	m	n	o	p

2.3 MDTP fuzzy edge strength based face organ edge image generation

The bottom directional triangle pattern based edge outputs (totally 18) are used to generate the bottom directional edge image I_{BDE} to create the bottom directional edge image using Eq. (2),

$$I_{BDE}(i,j) = I_{BDE}(i,j) + ED(i,j,p) \qquad (2)$$

The top directional edge image I_{TDE}, the left directional edge image I_{LDE}, and right directional edge image care formed similarly to the I_{BDE} image formulation.

The fuzzy edge strength computation I_{FES} accomplished via Algorithm.1 and the edge strength computation $I_{FES}(i,j)$ which is built by a mnemonic value of the range 0 to 3 using the edge strength (ES) in Eq. (3),

$$I_{FES}(i,j) = \begin{cases} 0, 0 \leq ES \leq 4 \\ 1, 5 \leq ES \leq 9 \\ 2, 10 \leq ES \leq 14 \\ 3, 15 \leq ES \leq 18 \end{cases} \qquad (3)$$

Mask 0   Mask 1   Mask 2   Mask 3   Mask 4

Mask 5   Mask 6   Mask 7   Mask 8   Mask 9

Mask 10   Mask 11   Mask 12   Mask 13   Masks 14
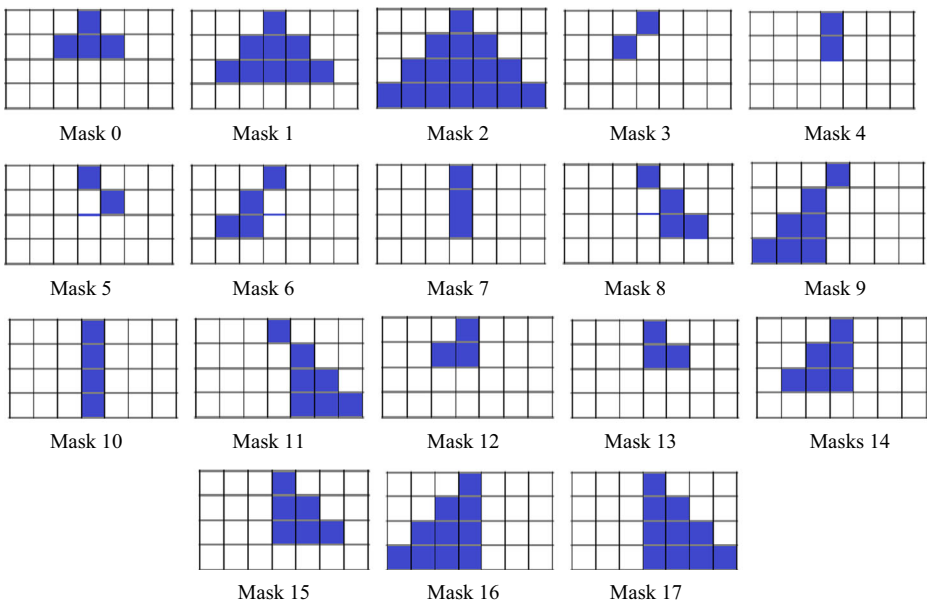
Mask 15   Mask 16   Mask 17

Fig. 4 Description of 18 triangle patterns united with the bottom direction

Subsequently, the same procedures are adapted to construct the bottom fuzzy edge strength image I_{BFES} top fuzzy edge strength image I_{TFES}, left fuzzy edge strength image I_{LFES} and right fuzzy edge strength I_{RFES}.

I_{TFES} is formed by the fusion of four directional edge strengths. The four directional fuzzy edge shapes fused into a single byte value based on Fig. 5 where each directional $[i,j]^{th}$ edge data fills two bits of the byte data of $[i,j]^{th}$ location. This reason, only the fuzzy edge value is limited into the range 0 to 3 which can be compacted by two bits. The algorithm steps 3 to 6 illustrate how to project each directional fuzzy edge data into a single byte which is the source to form the face organ edge image. Here the lip and eyeball areas are visible, and the other regions suppressed. It leads to the extraction of the eyeball and lip regions with ease and efficiency. This face-organ-edge image illustrated in Fig. 7g.

In existing descriptor representation the face organs are not easily represented because it contains unwanted edge artifacts. In the proposed MDTP descriptor representation the directional and multi-order oriented features are embedded, so the face organs like lip and eyeball easily extracted which is useful. Comparing with full face image features extraction the proper face organs feature extraction provides more efficient features for useful FER.

Algorithm 1

Input: MDTP descriptor image, I_{BDE}, I_{TDE}, I_{LDE}, I_{RDE}

Output: I_{BFES}, I_{TFES}, I_{LFES}, I_{RFES}

1. Read the MDTP descriptor image and initialize the Edge Strength (ES) to edge image.
2. Compare the edge strength with 18 patterns using the Eq.(3)
3. Compute fuzzy edge strength image for bottom directional edge image I_{BFES} using step 2, and it fills first two-byte locations.
4. Compute fuzzy edge strength image for top directional edge image I_{TFES} using step 2, and it fills two-byte location by right shifting 2 bytes (>> 2).
5. Compute fuzzy edge strength image for left directional edge image I_{LFES} using step 2, and it fills two-byte location by right shifting 4 bytes (>> 4).
6. Compute fuzzy edge strength image for right directional edge image I_{RFES} using step 2, and it fills two-byte location by right shifting 8 bytes (>> 6).

2.4 Feature extraction and classification

The histogram is the compact multi energetic information in which the feature extraction areas of lip and eyeball extracted, and histograms in associated with I_{MDTP} image descriptor are formulated to reflect the facial expression to feed training samples for SVNN classifier. The SVNN [21] classifier is a neural network based classifier which can successfully be adapt with FER system. SVNN trains the six face expressions and the query feature is used to conduct the SVNN testing process to draw out the resultant face expression type.
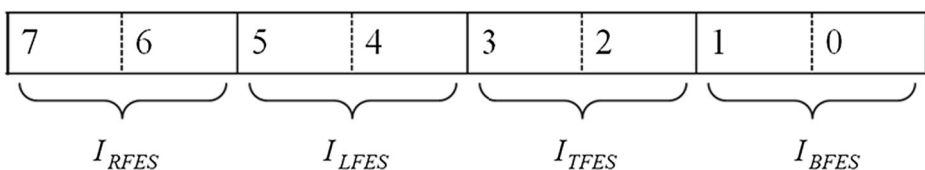

Fig. 5 Fuzzy Edge Strength Image

The architecture of the SVNN depicted in Fig. 6. The SVNN comprises of the input layer, hidden layer, and the output layer. The histogram features from the lip and eyeball areas fed to the input layers and the training are carried out. The score value provided by the SVNN ensures the recognition of the person. The features are corresponding to a person gaining the maximum score value authenticated.

The output of the SVNN described as $Output_{SVNN}$ in Eq. (4)

$$Output_{SVNN} = \left[z \times \log sig \left[\sum_{N=1}^{n} H_N \times We_N \right] + Weight \right] + Bias \qquad (4)$$

Input layer set,

$$We_N = \{We_1, We_2, ..., We_n\} \qquad (5)$$

Feature set,

$$H_N = \{H_1, H_2, ..., H_n\} \qquad (6)$$

where z denotes the bias value, We_N indicates the input layer defined by Eq. (5), H_N specifies the Nth feature described in Eq. (6), and n specifies the total number of features employed for facial expression recognition. The weight and bias of the output layer denoted as, $Weight$ and $Bias$.

The proposed FER system illustrated via Algorithm 2 which contains I_{MDTP} MDTP Descriptor Image, I_{BDE} Bottom Directional Edge Image, I_{TDE} Top Directional Edge Image, I_{LDE} Left Directional Edge Image, I_{RDE} Right Directional Edge Image, I_{FES} Fuzzy Edge Strength Image, I_{BFES} Bottom Fuzzy Edge Strength Image, I_{TFES} Top Fuzzy Edge Strength Image, I_{RFES} Right Fuzzy Edge Strength Image and I_{LFES} Left Fuzzy Edge Strength Image.

Algorithm 2

Input   : Face image
Output   : Face Expression Type.
   1. Read the input image I with H height, W width
   2. Reduce the noise by DBROMF method
   3. Initialize total directions to 4 and 18 patterns for each direction
   4. Compute MDTP descriptor image I_{MDTP} by convoluting the 18 patterns with the canny operator
   5. Compute bottom directional edge image I_{BDE} using bottom directional triangle patterns
   6. Compute top directional edge image I_{TDE} using top directional triangle patterns
   7. Compute left directional edge image I_{LDE} using left directional triangle patterns
   8. Compute right directional edge image I_{RDE} using right directional triangle patterns
   9. Compute MDTP based fuzzy edge strength I_{BFES}, I_{TFES}, I_{LFES}, I_{RFES} using Algorithm 1
   10. Separate the lip and eyeball facial organs
   11. Calculate the histograms for lip and eyeball regions.
   12. Perform expression recognition using SVNN training and testing
   13. Display the Face expression type

---

3 Experimental results

Japanese Female Facial Expression (JAFFE) [22] is an image database in which female facial expressions selectively gathered on account of 213 facial images referred with seven facial expressions like 256×256 pixel dimension. Cohn Kanade (CK) database [23] telescoped with 486 sequences of facial expression images connected by 97 subjects that have packed with
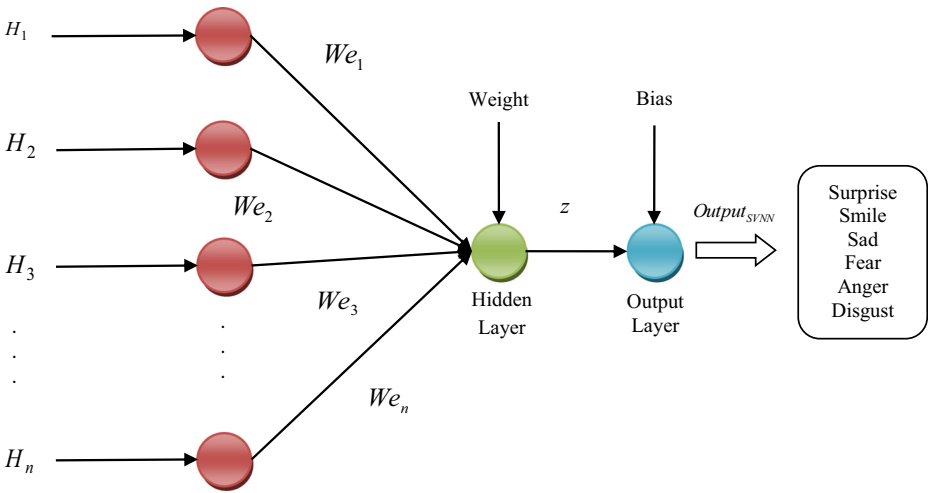
Fig. 6 The architecture of the SVNN

facial expressions belonging to the neutral frame to peak-expression frame. Taiwanese Facial Expression Image Database (TFEID) [24] includes 7200 stimulants which taken from 40 subjects. Amsterdam Dynamic Facial Expression Set (ADFES) [25] is a facial expression database, and it consists of the 648 stimuli of six basic expressions along with contempt, pride and embarrassment expressions.

Figure 7a to g shows the working principle of FER-MDTP system. Figure 7a denotes the original query image. Figure 7b illustrates the bottom directional edge of the original image. Figure 7c shows the top directional edge of the original image. Figure 7d denotes the left bottom directional edge of the original image. Figure 7e illustrates the right directional edge of the original image. Figure 7f indicates the MDTP descriptor image, and Fig. 7g shows the face organ structure of the query image.
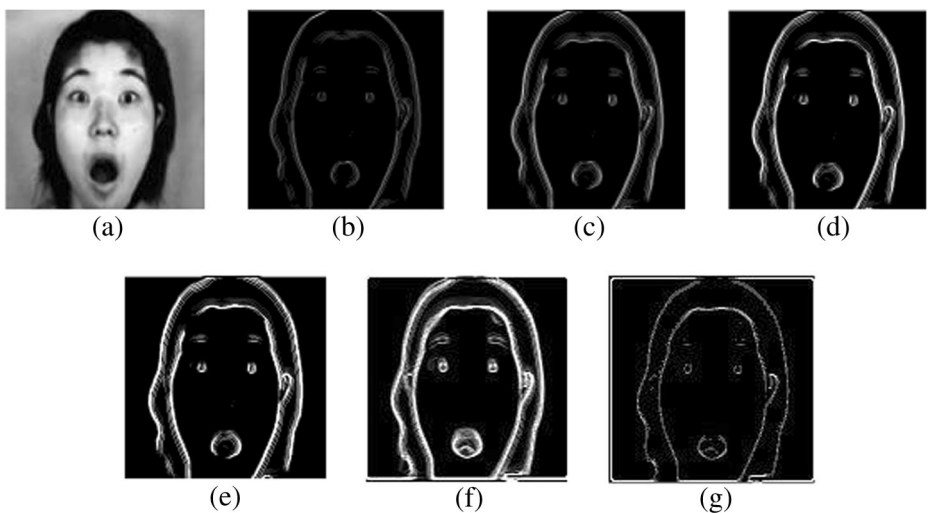

Fig. 7 **a** Original Image **b** Bottom Image **c** Top Image **d** Left Image **e** Right Image (f) MDTP descriptor image **g** Face organ structure

Table 1 Accuracy acquired by using various FER methods on the JAFFE database

Method	Anger (%)	Disgust (%)	Fear (%)	Smile (%)	Sad (%)	Surprise (%)
FER-LDN	85.27	87.71	82.56	96.32	84.17	95.08
FER-HOG	85.9	87.93	83.25	96.72	84.29	94.92
FER-LBP	86.71	88.97	83.93	96.81	84.96	95.81
FER-WLBI-CT	86.95	89.85	84.97	97.57	85.73	96.58
FER-HOG-DCT	87.52	90.02	85.5	97.9	85.9	96.92
Proposed	89.74	93.25	88.38	99.25	88.12	98.34

For evaluating the performance of the proposed method, two measures calculated that is recognition accuracy and confusion matrix using JAFFE, CK, TFEI, and ADFES databases. Also, the proposed method compared with various FER methods such as LDN [26], HOG [27], LBP [28], WLBI-CT [29], HOG-DCT [30]. For experimentation, images from JAFFE database, CK database, TFEI database, and ADFES database used which contains the six expressions such as anger, disgust, fear, smile, sad and surprise.

The accuracy analysis performed by using the accuracy formula, and it depicted in Eq. (7).

$$Accuracy = \frac{TruePositive}{TruePositive + FalsePositive} \tag{7}$$

Table 1 shows the recognition accuracy analysis using the JAFFE database for different expressions which include anger, disgust, fear, smile, sad and surprise. The anger expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 85.27, 85.9, 86.71, 86.95, 87.52 and 89.74 respectively. The disgust expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 87.71, 87.93, 88.97, 89.85, 90.02 and 93.25 correspondingly. The fear of expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 82.56, 83.25, 83.93, 84.97, 85.5 and 88.38 respectively. The smile expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 96.32, 96.72, 96.81, 97.57, 97.9 and 99.25 respectively. The sad expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 84.17, 84.29, 84.96, 85.73, 85.9 and 88.12 correspondingly. The surprise expression has the percentage of accuracy for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and Proposed are 95.08, 94.92, 95.81, 96.58, 96.92 and 98.34 respectively. The proposed method has high recognition accuracy than the existing method and also smile and surprise expression has the highest accuracy than the other expressions.

Table 2 Accuracy acquired by using various FER methods on CK database

Method	Anger (%)	Disgust (%)	Fear (%)	Smile (%)	Sad (%)	Surprise (%)
FER-LDN	83.34	86.53	82.47	94.52	82.57	93.46
FER-HOG	84.41	86.84	83.16	94.82	83.32	94.15
FER-LBP	85.27	87.77	84.26	97.13	84.12	95.26
FER-WLBI-CT	85.94	88.95	84.91	97.2	85.23	96.31
FER-HOG-DCT	86.15	89.04	85.12	97.43	85.42	96.45
Proposed	88.21	91.38	87.42	98.91	87.57	98.53

Table 3 Accuracy acquired by using various FER methods on ADFES database

Method	Anger (%)	Disgust (%)	Fear (%)	Smile (%)	Sad (%)	Surprise (%)
FER-LDN	80.36	83.04	79.49	91.45	79.03	89.55
FER-HOG	81.15	83.91	80.02	92.13	80.28	91.46
FER-LBP	82.3	84.84	81.47	93.3	81.16	92.37
FER-WLBI-CT	82.94	86.1	82.14	94.47	82.37	93.28
FER-HOG-DCT	83.3	86.38	82.31	94.59	82.49	93.41
Proposed	84.25	87.35	82.76	95.37	83.25	94.2

Table 2 shows the recognition accuracy analysis using the CK database for different expressions which include anger, disgust, fear, smile, sad and surprise. The FER-LDN method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 83.34, 86.53, 82.47, 94.52, 82.57 and 93.46 respectively. The FER-HOG method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 84.41, 86.84, 83.16, 94.82, 83.32 and 94.15 correspondingly. The FER-LBP method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 85.27, 87.77, 84.26, 97.13, 84.12 and 95.26 correspondingly. The FER-WLBI-CT method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 85.94, 88.95, 84.91, 97.2, 85.23 and 96.31 respectively. The FER-HOG-DCT method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 86.15, 89.05, 85.12, 97.43, 85.42 and 96.45 respectively. The proposed FER-MDTP method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 88.21, 91.38, 87.42, 98.91, 87.57 and 98.53 correspondingly. The smile and surprise expressions have the highest recognition accuracy than the other expressions and fear expression has the lowest accuracy. The proposed method gives the highest accuracy than the existing methods for difference six facial expressions.

Table 3 shows the recognition accuracy analysis using the ADFES database for different expressions which include anger, disgust, fear, smile, sad and surprise. The FER-LDN method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 80.36, 83.04, 79.49, 91.45, 79.03 and 89.55 respectively. The FER-HOG method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 81.45, 83.91, 80.02, 92.13, 80.28 and 91.46 correspondingly. The FER-LBP method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 82.3, 84.84, 81.47, 93.3, 81.16 and 92.37 correspondingly. The FER-WLBI-CT method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 82.94, 86.1, 82.14, 94.47, 82.37 and 93.28 respectively. The FER-HOG-DCT method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 83.3,

Table 4 Performance Analysis

| Method	Average Accuracy of JAFFE Database (%)	Average Accuracy of CK Database (%)	Average Accuracy of TFEI Database (%)	Average Accuracy of ADFES Database (%)	Average Time Taken (seconds)
FER-LDN	87.83	86.63	84.25	83.52	0.243
FER-HOG	88.25	87.45	85.2	84.65	0.251
FER-LBP	90.48	88.17	85.61	85.52	0.285
FER-WLBI-CT	92.73	91.53	89.15	88.35	0.353
FER-HOG-DCT	93.28	92.45	89.52	88.76	0.396
Proposed	97.23	95.78	92.54	90.98	0.422
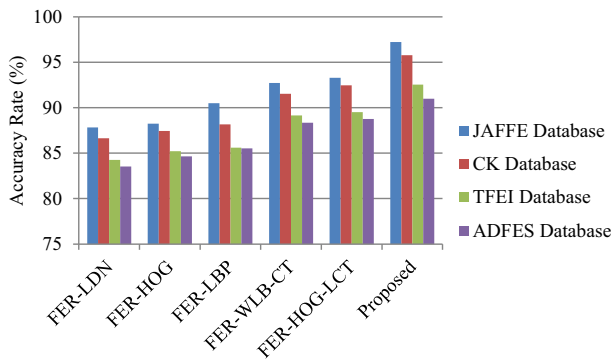
Fig. 8 Recognition accuracy analysis of JAFFE, CK, TFEI, and ADFES databases

86.38, 82.31, 94.59, 82.49 and 93.41 respectively. The proposed FER-MDTP method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 84.25, 87.35, 82.76, 95.37, 83.25 and 94.2 correspondingly. The smile and surprise expressions have the highest recognition accuracy than the other expressions and fear expression has the lowest accuracy. The proposed method gives the highest accuracy among the existing methods for difference six facial expressions.

Table 4 shows the performance analysis of the proposed method compared to the FER methods which are evaluated using the JAFFE, CK, TFEI and ADFES databases. The average percentage of accuracy using JAFFE database for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and proposed FER-MDTP are 87.83, 88.25, 90.48, 92.73, 93.28 and 97.23 respectively. The average percentage of accuracy using the CK database for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and proposed FER-MDTP are 86.63, 87.45, 88.17, 91.53, 92.45 and 95.78 correspondingly. The average percentage of accuracy using the TFEI database for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and proposed FER-MDTP are 84.25, 85.2, 85.61, 89.15, 89.52 and 92.54 correspondingly. The average percentage of accuracy using the ADFES database for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and proposed FER-MDTP are 83.52, 84.65, 85.52, 88.35, 88.76 and 90.98 respectively. The average seconds were taken to recognize the expressions for FER-LDN, FER-HOG, FER-LBP, FER-WLBI-CT, FER-HOG-DCT and proposed FER-MDTP are 0.243, 0.251, 0.285, 0.353, 0.396 and 0.422 respectively. The proposed method requires more processing time, but it achieves better recognition accuracy than the other compared methods. The proposed method gives 97.23% average accuracy for the JAFFE database which is the highest accuracy among the other methods.

Table 5 Accuracy Analysis in occurrence of 25% noise on CK database

Method	Anger (%)	Disgust (%)	Fear (%)	Smile (%)	Sad (%)	Surprise (%)
FER-LDN	84.62	86.91	83.67	94.96	82.64	93.29
FER-HOG	84.53	87.56	83.94	95.52	83.85	94.87
FER-LBP	86.41	88.47	84.97	96.81	84.42	95.82
FER-WLBI-CT	86.84	89.05	85.04	97.9	85.78	96.54
FER-HOG-DCT	87.56	89.73	85.77	98.06	86.31	96.89
Proposed	88.71	91.58	86.65	98.83	87.15	97.63

Table 6 The Confusion Matrix for 6-class Classification in FER using JAFFE database

	Anger	Disgust	Fear	Smile	Sad	Surprise
Anger	89.65	1.97	0	0	8.38	0
Disgust	1.83	92.27	0	0	5.9	0
Fear	8.06	0	87.95	0	3.99	0
Smile	0	0.5	0	99.5	0	0
Sad	2.04	7.02	2.47	0	88.47	0
Surprise	0	1.38	0	0	0	98.62

Figure 8 illustrates the recognition analysis of FER methods using JAFFE, CK, TFEI, and ADFES databases. Here the x-axis indicates the FER methods which include the compared methods and the proposed method. The y-axis shows the percentage of accuracy generated by the FER methods for the three databases. From this, it clearly understood that four databases used for analyzing the performances of the proposed FER-MDTP method. ADFES database obtains the accuracy is less than the other three databases such as JAFFE, CK, and TFEI as well as the proposed method FER-MDTP achieves better accuracy rate for the JAFFE database.

Table 5 shows the recognition accuracy analysis in the occurrence of 25% noise on the CK database for different expressions which include anger, disgust, fear, smile, sad and surprise. The FER-LDN method which has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 84.62, 86.91, 83.94, 94.96, 82.64 and 93.29 correspondingly. The FER-HOG method which has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 84.53, 85.56, 83.94, 95.52, 83.85 and 94.87 respectively. The FER-LBP method which has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 86.41, 88.47, 84.97, 96.81, 84.42 and 95.82 respectively. The FER-WLBI-CT method which has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 86.84, 89.05, 85.04, 97.9, 85.78 and 96.54 correspondingly. The FER-HOG-DCT method which has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 87.56, 89.73, 85.77, 98.06, 86.31 and 96.89 correspondingly. The proposed FER-MDTP method has the accuracy for anger, disgust, fear, smile, sad and surprise expressions are 88.71, 91.58, 86.65, 98.83, 87.15 and 97.63 respectively. The proposed method attains the maximum accuracy among the existing methods for various six facial expressions.

Tables 6, 7, 8 and 9 depicts the confusion matrix of the proposed method using JAFFE, CK, TFEI, and ADFES databases to understand the recognition behavior for individual expression. The fear expression is highly confused than the other expressions, and it is confused with anger and sad expressions. The anger expression also highly confused with sad and disgust expressions which achieve 89.65%, 87.53%, 85.07% and 84.10% accuracy for JAFFE, CK, TFEI, and ADFES databases respectively. The sad expression is confused with disgust, fear and

Table 7 The Confusion Matrix for 6-class Classification in FER using CK database

	Anger	Disgust	Fear	Smile	Sad	Surprise
Anger	87.53	3.57	0	0	8.9	0
Disgust	2.82	90.12	0	0	7.06	0
Fear	8.94	0	85.28	0	5.78	0
Smile	0	2.06	0	97.94	0	0
Sad	2.76	7.72	3.02	0	86.5	0
Surprise	0	2.95	0	0	0	97.05

Table 8 The Confusion Matrix for 6-class Classification in FER using TFEI database

	Anger	Disgust	Fear	Smile	Sad	Surprise
Anger	85.07	5.02	0	0	9.91	0
Disgust	3.59	88.35	0	0	8.06	0
Fear	10.84	0	83.67	0	5.49	0
Smile	0	3.48	0	96.52	0	0
Sad	2.39	8.91	3.87	0	84.83	0
Surprise	0	4.46	0	0	0	95.54

Table 9 The Confusion Matrix for 6-class Classification in FER using ADFES database

	Anger	Disgust	Fear	Smile	Sad	Surprise
Anger	84.10	5.01	0	0	10.89	0
Disgust	3.98	87.18	0	0	8.84	0
Fear	11.2	0	82.43	0	6.37	0
Smile	0	4.85	0	95.15	0	0
Sad	2.97	9.4	4.51	0	83.12	0
Surprise	0	5.95	0	0	0	94.05

anger expressions which achieve 88.47%, 86.5%, 84.83% and 83.12% accuracy for JAFFE, CK, TFEI, and ADFES databases respectively. The disgust expression is confused with sad and anger expressions which achieve 92.27%, 90.12%, 88.35% and 87.18% accuracy for JAFFE, CK, TFEI, and ADFES databases respectively. The surprise expression is slightly confused with disgust expression, and it attains 98.62%, 97.05%, 95.54% and 94.05% accuracy for JAFFE, CK, TFEI, and ADFES databases respectively. The smile expression has a little confusion with disgust expression, and it attains the highest accuracy rate of 99.5%, 97.94%, 96.52% and 95.15% accuracy for JAFFE, CK, TFEI, and ADFES databases respectively.

4 Conclusion

This paper comes forward with the three remarkable algorithms DBROMF, MDTP and MDTP-FES to efficiently extract the features at the locations of lip and eyeball for effective face expression recognition. The accuracy analysis of the JAFFE database exposes an evident for the greatness of the proposed method FER-MDTP while connected with its higher accuracy value of 97.23%. The CK, TFEI and ADFES databases analysis is also an essential proof for the successful accuracy rate of 95.78%, 92.54% and 90.98% in that order of the proposed method. The smile facial expression provides the most accuracy rate of 99.5% and surprise expression belong to the following accuracy rate as 98.62%. The proposed method outperforms the state-of-the-art FER methods through significantly improved accuracy.

References

1. Bettadapura V (2012) Face Expression Recognition and Analysis: The State of the Art. 1–27

2. Chu WS, Torre FD, Cohn JF (2017) Selective Transfer Machine for Personalized Facial Expression Analysis. IEEE Trans Pattern Anal Mach Intell 39:529–545. https://doi.org/10.1109/TPAMI.2016.2547397

3. Yan H, Ang MH, Poo AN (2012) Adaptive discriminative metric learning for facial expression recognition. IET Biometrics 1:160–167. https://doi.org/10.1049/iet-bmt.2012.0006

4. Zhang T, Zheng W, Cui Z et al (2016) A Deep Neural Network Driven Feature Learning Method for Multi-view Facial Expression Recognition. IEEE Trans Multimed 18:2528–2536. https://doi.org/10.1109/TMM.2016.2598092

5. Mistry K, Zhang L, Neoh SC et al (2017) A Micro-GA Embedded PSO Feature Selection Approach to Intelligent Facial Emotion Recognition. IEEE Trans Cybern 47:1496–1509

6. Hsieh C, Hsih M (2016) Effective Semantic Features for Facial Expressions Recognition using SVM. Multimed Tools Appl 75:6663–6682. https://doi.org/10.1007/s11042-015-2598-1

7. Yi J, Mao X, Chen L et al (2014) Facial Expression Recognition Considering Individual Differences in Facial Structure and Texture. IET Comput Vis 8:429–440. https://doi.org/10.1049/iet-cvi.2013.0171

8. Meng H, Bianchi-Berthouze N, Deng Y et al (2016) Time-Delay Neural Network for Continuous Emotional Dimension Prediction From Facial Expression Sequences. IEEE Trans Cybern 46:916–929. https://doi.org/10.1109/TCYB.2015.2418092

9. Liu Z, Wu M, Cao W, Chen L (2017) A Facial Expression Emotion Recognition Based Human-robot Interaction System. IEEE/CAA J Autom Sin 4:668–676

10. Muhammad G, Alsulaiman M, Amin SU et al (2017) A Facial-Expression Monitoring System for Improved Healthcare in Smart Cities. IEEE Access 5:10871–10881. https://doi.org/10.1109/ACCESS.2017.2712788

11. Munir A, Hussain A, Khan SA, Nadeem M (2018) Illumination invariant facial expression recognition using selected merged binary patterns for real world images. Opt - Int J Light Electron Opt 158:1016–1025. https://doi.org/10.1016/j.ijleo.2018.01.003

12. Zhang K, Huang Y, Du Y, Wang L (2017) Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks. IEEE Trans Image Process 29:4193–4203. https://doi.org/10.1109/TIP.2017.2689999

13. Ding Y, Zhao Q, Li B, Yuan X (2017) Facial Expression Recognition From Image Sequence Based on LBP and Taylor Expansion. IEEE Access 5:19409–19419

14. Uddin MZ, Khaksar W, Torresen J (2017) Facial Expression Recognition Using Salient Features and Convolutional Neural Network. IEEE Access 5:26146–26161

15. Yang B, Cao J, Ni R, Zhang Y (2018) Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images. IEEE Access 6:4630–4640

16. Uddin ZIA, Hassan MM (2017) Facial Expression Recognition Utilizing Local Direction-Based Robust Features and Deep Belief Network. IEEE Access 5:4525–4536

17. Mao Q, Rao Q, Yu Y, Dong M (2017) Hierarchical Bayesian Theme Models for Multi-pose Facial Expression Recognition. IEEE Trans Multimed 19:861–873. https://doi.org/10.1109/TMM.2016.2629282

18. Zeng N, Zhang H, Song B et al (2017) Facial Expression Recognition via Learning Deep Sparse Autoencoders. Neurocomputing 273:643–649. https://doi.org/10.1016/j.neucom.2017.08.043

19. Meena HK, Sharma KK, Joshi SD (2017) Improved Facial Expression Recognition Using Graph Signal Processing. Electron Lett 53:718–720. https://doi.org/10.1049/el.2017.0420

20. Cruz EAS, Jung CR, Franco CHE (2017) Facial Expression Recognition using temporal POEM features. Pattern Recogn Lett 114:13–21. https://doi.org/10.1016/j.patrec.2017.08.008

21. Ludwig O, Nunes U, Araujo R (2014) Eigenvalue Decay: a New Method for Neural Network Regularization. Neurocomputing 124:33–42

22. Michael L, Miyuki K, Jiro G (1999) Facial Expression Database: Japanese Female Facial Expression Database. http://www.kasrl.org/jaffe.html. Accessed 11 Nov 2018

23. (2000) Cohn-Kanade AU-Coded Expression Database. http://www.pitt.edu/~emotion/ck-spread.htm. Accessed 11 Nov 2018

24. (2007) TFEID: Taiwanese Facial Expression Image Database. http://bml.ym.edu.tw/tfeid/. Accessed 5 Feb 2019

25. Van Der SJ, Hawk ST, Fischer AH, Doosje B (2011) Moving Faces, Looking Places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). 11:907–920. doi: https://doi.org/10.1037/a0023853

26. Rivera AR, Castillo JR, Chae O (2013) Local Directional Number Pattern for Face Analysis : Face and Expression Recognition. IEEE Trans Image Process 22:1740–1752

27. Carcagnì P, Del Coco M, Leo M, Distante C (2015) Facial expression recognition and histograms of oriented gradients: a comprehensive study. Springerplus 4. https://doi.org/10.1186/s40064-015-1427-3

28. Kumar S, Bhuyan MK, Chakraborty BK (2016) Extraction of Informative Regions of a Face for Facial Expression Recognition. IET Comput Vis 10:567–576. https://doi.org/10.1049/iet-cvi.2015.0273

29. Khan SA, Hussain A, Usman M (2018) Reliable Facial Expression Recognition for Multi-scale Images using Weber Local Binary Image Based Cosine Transform Features. Multimed Tools Appl 77:1133–1165. https://doi.org/10.1007/s11042-016-4324-z
30. Nazir M, Jan Z, Sajjad M (2017) Facial expression recognition using histogram of oriented gradients based transformed features. Clust Comput. https://doi.org/10.1007/s10586-017-0921-5

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

I. Michael Revina received her M.E (Computer science and engineering) degree in 2015 from Anna University, Chennai. Now she is doing doctoral research under Manonmaniam Sundaranar University, Tirunelveli. Her interested research areas are image processing and pattern recognition.

W. R. Sam Emmanuel received his doctoral degree in computer science in 2012 from Vinayaka Missions University, Salem. He received his M. Phil (Computer Science) degree in 2002 from Manonmaniam Sundaranar University, Tirunelveli and MCA from Bharathidhason University, Tiruchirappalli. He is working as Associate Professor at Nesamony Memorial Christian College, Marthandam. His major research interests are Cryptography, Network Security, Segmentation and Classification.