RESEARCH ARTICLE

Ensemble Training Approach Based on Multivariate Empirical mode decomposition and Convolution Neural network for Periocular Recognition

Sheela R*, Smitha Rajagopal
Department of Computer Science and Engineering, Alliance University,
Chikkahagade Cross, Chandapura- Ankel Main Road, Bangalore-562 106, India.

ABSTRACT: The paper suggests an ensemble training method that utilizes a quick adaption of the Fast Adaptive Multivariate Empirical Mode Decomposition (FA-MVEMD) and a convolutional neural network (CNN) for periocular recognition. Initially, the periocular images go through pre-processing. After pre-processing, a single-level discrete wavelet transform decomposes the image, resulting in the LL, LH, HL, and HH coefficients. The LL coefficients undergo extra decomposition into K quantity of Intrinsic Mode Functions (IMFs) using the Fast Adaptive Multivariate Empirical Mode Decomposition (FA-MVEMD). The Intrinsic Mode Functions (IMF), along with the residual and HH coefficients, form K+2 ensemble images. During the training stage, the CNN model is trained with ensemble images. During the testing phase, the K+2 ensemble images are created on the test image. These K+2 ensemble images are classified as yielding K+2 outcomes. The actual classification result is obtained based on the maximum number of similar matching results. The evaluation was performed using metrics such as UBPIr, AR, CASIA Iris rank 1, rank 2, equal error rate (EER), and area under the curve (AUC). The proposed approach achieves rank-1 recognition accuracy of 93.47%, 98.23%, and 98.03% on UBPIr, AR, and CASIA Iris datasets, respectively.

KEYWORDS: Periocular recognition, Empirical mode decomposition, Discrete wavelet transform, Rank-1 recognition accuracy, Convolutional neural network.

https://doi.org/10.29294/IJASE.10.3.2024.3566-3580

1.0 INTRODUCTION

Identifying people with partial faces is one of the most difficult tasks in face recognition bio-metrics. In recent times, peri-ocular recognition has gained popularity in identifying people with partial faces. Peri-ocular features extracted around the eyes include features from the upper part of the brow, the cheekbone, the folds of the eye, the corners of the eye, and the skin texture. The face recognition algorithm performs poorly when the face is recognized by obstacles like eyeglasses, hair, face mask, helmet, etc. Peri-ocular images can be used for real-time application to identify the criminals wearing a face mask in a surveillance video. Fig 1 shows some real-time examples of peri-ocular images

Therefore, there is a need for high-performance periocular biometric recognition algorithms. As the face recognition algorithm takes into account the complete face region, the change of facial expression also affects the performance of face recognition system. As the upper half of face region is used for recognition, change of facial expression does not significantly affect the performance of periocular recognition algorithms. Other biometrics like iris recognition will perform better with partial faces and require a close focus on iris. Periocular recognition requires no close focus on the eye. The following are the contributions of the work: Ensemble training approach: The periocular images are first trained using a one-level DWT. The HH band image is then one of the ensembled images. Next, the algorithm uses FA-MVDEMD to further decompose LL band images into multiple IMFs and residuals.

The approach moreover employments a CNN calculation that can train the ensemble images, rather than preparing by coordinate periocular pictures. The approach too employments is a periocular classification component from the gathering classification results. The rest of the sections are orchestrated as mentioned below.

**Corresponding Author: sheelamohare@gmail.com*
Received: 05.01.2024 **Accepted:** 27.02.2024 **Published on**: 07.03.2024

Few of the related works are given in section 2. A point-by-point depiction of the proposed periocular calculation is given in the third section. The examinations of the proposed periocular calculation are given in the 4th section. Finally, the 5th section concludes the paper.

Fig 1: Few examples for periocular images [1, 2]

Thus, there is a necessity for high performance periocular-based biometric recognition algorithm. Since the face recognition algorithm considers the complete facial region, the change in facial expression also degrades the performance of the face recognition system. Since the upper half of the face region is used in recognition, the change in facial expression does not highly affect the periocular recognition algorithm. Other biometrics such as iris recognition will provide a better result with partial faces, and it needs a close focus on the iris. However, the periocular recognition does not require a close focus of the eye. The contributions of the work are as follows.

The paper proposes an ensemble training approach where several ensemble images are derived from the periocular images. The approach uses a single-level DWT that initially decomposes the periocular image into four bands, where the *HH* band image is used as one of the ensemble images. The algorithm also uses the FA-MVEMD that further decompose the *LL* band images into several IMFs and residual. The approach also uses a CNN algorithm that can train the ensemble image, instead of training by direct periocular images. The approach also uses a periocular classification mechanism from the ensemble classification results.

The rest of the sections are arranged as follows. Few of the related works are provided in section 2. A detailed description of the proposed periocular algorithm is given in section 3. The experimental results and analysis of the proposed periocular recognition algorithm are provided in section 4. Lastly, section 5 concludes the paper.

For the development of a high-performance periocular recognition algorithm, many researchers put forth different schemes based on local and global feature extraction. Algorithms such as local binary pattern (LBP), scale-invariant feature transform (SIFT) [3], phase intensive local pattern, and speeded up robust features (SURF) [4] are examples of local feature extraction algorithms that extract the features on sub-divided regions. Algorithms such as block histogram of gradient (HOG) [5], and color histogram are an example of global feature extraction algorithms that extracts the features of the complete region. Different classification algorithms are proposed that include distance-based classification approaches, machine learning, and deep-learning classification approaches. Algorithms such as Hamming distance, k-nearest neighbour, Bhattacharya distance chi-square, mean square error, and Euclidean distance are an example of distance-based classification approaches. Machine learning algorithms such as support vector machines [6], multi-layer perceptron [7], and Naïve Bayes classifiers [8] provide good recognition results for efficient feature extraction algorithms. The convolutional neural

network model plays a crucial role in deep learning classifiers in different applications [9, 10] where there is difficulty in extracting the features.

The shape of the eyebrow is used as a feature [11] which provides recognition accuracy of 85% and 76% for the MBGC dataset and AR dataset respectively. Part-based separation [12] is used to improve the performance of the recognition where the face image is matched after sub-dividing the facial image into different blocks. However, these approaches have the challenge of sub-dividing the unstructured facial images obtained from the real world. The deterministic annealing approach is introduced [13] that uses the geometric distribution that can align two-point sets. The deterministic annealing scheme uses geometric and texture information which is suitable for complete face recognition and degrades the performance in partial face recognition. The local binary pattern feature is extracted [14] from the periocular regions either in a counter clockwise or in a clockwise direction which uses a window so that the features are extracted with respect to the neighbourhood and the centre pixel. However, the performance of this algorithm is highly dependent on the rotation effect. Principle component analysis and linear discriminant analysis [15] is used as an optimization approach where the size of the features extracted can be highly minimized. Local binary patterns along with Zernike moment are used in texture feature extraction [16] along with the fusion of shapes. The author Chen et al. [17] used five different local features that include the features from iris color, tear duct texture features, inner eye corner intensities, upper eyelid, and eyebrow distance, and the texture features between the eyebrows.

The author Hugo et al., [18] utilized CNN which characterizes the locale of intrigued without utilizing the concealing approach that isolates the visuallocale from the periocular locale. They have also generated different samples by interchanging the ocular region and the periocular region. This approach provides an illumination rate of 85% for the FRGC dataset and 82% for the UBIRIS dataset. The author Zhao et al., [19] proposed a semantic-assisted CNN which recovers the comprehensive periocular information by including the explicit semantic features. The authors claimed that the semantic CNN approach can extract discriminative features without increasing the size of the training data. The authors Chul et al., [20] applied a random walker scheme to obtain a coarse segmented image. The coarse segmented image is further processed by a sequence of operations which can improve the recognition accuracy. The authors Zijing et al., [21] used an attention model that gives more weightage to the essential region of the periocular image. The essential region mainly includes the eye and eyebrow region within the periocular image, from which more discriminative features are extracted by deep CNN that can improve the recognition accuracy. Face intensive local pattern [22] was used to extract the features from the iris skin region which uses a collection of filter banks with various scales.

The authors Kumar et al. [23] used four equal non-overlapping sub-regions for extracting the non-overlapped interpolated LBP features. Binned histogram features along with intensive global pattern features are also extracted which improves the performance in periocular recognition. Deep feature fusion [24] with multiple attention mechanism that includes the process namely self-attention and co-attention are used in iris periocular recognition. The self-attention includes the processes such as spatial attention and channel attention to extract the features. Collaborative rendering of hand-crafted features and deep features are proposed by the author Ritesh et al. [25], which calculates the local statistical properties after sub-dividing the periocular region into several groups. The local statistical properties are used as the hand-crafted features whereas the convolutional neural network ResNet-101 is used in extracting the deep features. This paper proposes an ensemble image estimation mechanism for the training of CNN. This increases the recognition accuracy by increasing the number of training images that have different bands in each category. The next section depicts the proposed periocular recognition algorithm.

1. *Proposed periocular recognition algorithm*

The proposed periocular recognition algorithm has five stages namely (i). Pre-processing (ii). Single level discrete wavelet transforms, (iii). Ensemble image construction, (iv). Training, and (v). Testing. The block diagram representation of the proposed system is depicted in Fig 2.

(a) Pre-processing

Let the input periocular image be $I_1(x,y)$. The preprocessing includes the process namely cropping, scaling, and high pass filtering. The input image is cropped to obtain the periocular region. The cropped image is then scaled to a uniform size of $M \times N$. The resulting image is then passed to the high-pass filter to remove the blurring effect of the camera. Let the pre-processed image be represented as $I_2(x,y)$.

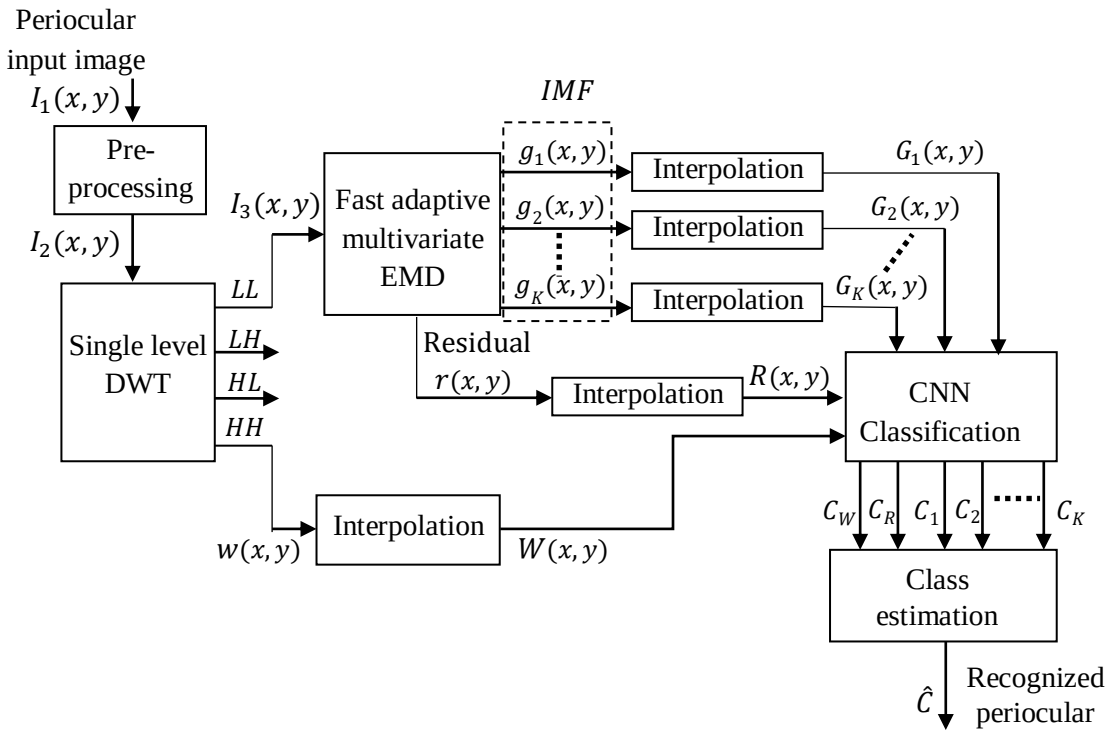


Fig 2: Block diagram representation of Proposed periocular recognition

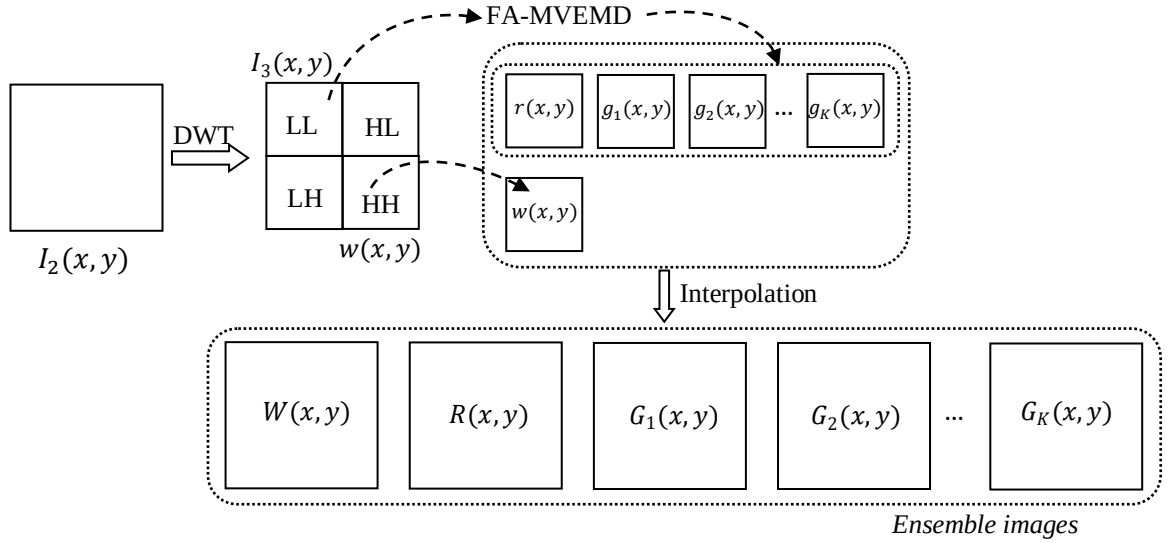


Fig 3: Representation of DWT decomposition and ensemble image formation

(b) Single level discrete wavelet transform

Ensemble training images can be constructed using the decomposition algorithms namely single level DWT, and fast adaptive multivariate empirical mode decomposition that can decompose the image to different bands. Let the

pre-processed image be $I_2(x,y)$. The pre-processed image $I_2(x,y)$ is decomposed using a single-level discrete wavelet transform [26] to obtain the bands LL, LH, HL, and HH. Let the LL band coefficients and HH coefficients are represented as $I_3(x,y)$ and $w(x,y)$ as depicted in Fig 3 whose size is $\left(\frac{M}{2}\times\frac{N}{2}\right)$. The HH band coefficient $w(x,y)$ is then interpolated [28] by 2 to obtain the interpolated HH band $W(x,y)$. The coefficient $W(x,y)$ has the size of $(M\times N)$.

(c) Ensemble training image construction

The band LL coefficients $I_3(x,y)$is again decomposed using a fast adaptive multivariate empirical mode decomposition (FA-MVEMD) [27] to obtain K number of IMF and the residual image. Let $I_3(x,y)$ be the image be decomposed by empirical mode decomposition. The FA-MVEMD algorithm can be summarized as follows

Algorithm: FA-MVEMD algorithm

Input: *LL band coefficients $I_3(x,y)$, Number of IMF 'K'*

Output: *IMFs $g_1(x,y)$, $g_2(x,y)$, $g_3(x,y)$..... $g_K(x,y)$ and the residual $r(x,y)$*

Initialize:$h_{10}(x,y)=I_3(x,y)$.

Step 1: Estimate the amplitude and position of local minima for the image $h_{10}(x,y)$. Similarly, estimate the amplitudes and position of local maxima for the signal $h_{10}(x,y)$.

Step 2: Using the amplitude and position of local minima estimate the lower envelope using the spline interpolation. Let the lower envelope be denoted by $F_{kl}^{-}(x,y)$. Similarly using the position and amplitude of local maxima, estimate the upper envelope using the spline interpolation. Let the upper envelope be denoted by $F_{kl}^{+}(x,y)$. Here l is the iteration number and $k=1,2\ldots\ldots K$.

Step 3: Estimate the average $\mu_{kl}(x,y)$ of the lower envelope and upper envelope for all the positions of (x,y) given by

$$\mu_{kl}(x,y)=\frac{F_{kl}^{-}(x,y)}{2}+\frac{F_{kl}^{+}(x,y)}{2} \qquad (1)$$

Step 4: The average envelope $\mu_{kl}(x,y)$ is subtracted from the previous signal given by

$$h_{kl}(x,y)=h_{k(l-1)}(x,y)-\mu_{kl}(x,y) \qquad (2)$$

Step 5: Check whether the signal $h_{kl}(x,y)$ is an IMF or not. This can be verified using the condition

$$|\mu_{kl}(x,y)|<\Delta\forall(x,y) \qquad (3)$$

Here Δ is the threshold value that decides the stopping condition of the iteration.

Step 6: If the condition in equation (3) is satisfied, then the iteration can be stopped. If the condition in the equation is not satisfied repeat steps 1 to 5 until the condition is satisfied

Step 7: As the stop condition is satisfied the IMF can be obtained from the recent result of $h_{kl}(x,y)$

$$g_k(x,y)=h_{kl}(x,y) \qquad (4)$$

Step 8: The residue can be obtained using the relation

$$\eta_k(x,y)=h_{kl}(x,y)-g_l(x,y) \qquad (5)$$

Step 9: The next IMF can be obtained by repeating the above steps with the residue $\eta_l(x,y)$ as the input signal

$$h_{(k+1)0}(x,y)=\eta_k(x,y) \qquad (6)$$

The final residue obtained after K number of IMF is given by $\eta_K(x,y)$. Therefore, the final residual image $\eta_K(x,y)$ and the IMFs can be related to the original image $I_3(x,y)$as

$$I_3(x,y)=\eta_k(x,y)+\sum_{j=1}^{k}g_j(x,y) \qquad (7)$$

Let the IMFs obtained by FA-MVEMD be given by $g_1(x,y)$, $g_2(x,y)$, $g_3(x,y)$..... $g_K(x,y)$ and the residual image is represented as $r(x,y)$ whose size is $\left(\frac{M}{2}\times\frac{N}{2}\right)$.. The IMFs and the residual are interpolated [28] by 2 so that the size of the IMFs and the residual remains $M\times N$. Let the interpolated IMFs and the residual be represented as $G_1(x,y)$, $G_2(x,y)$, $G_3(x,y)$..... $G_K(x,y)$ as depicted in Fig. 3.

(d). Convolutional neural network

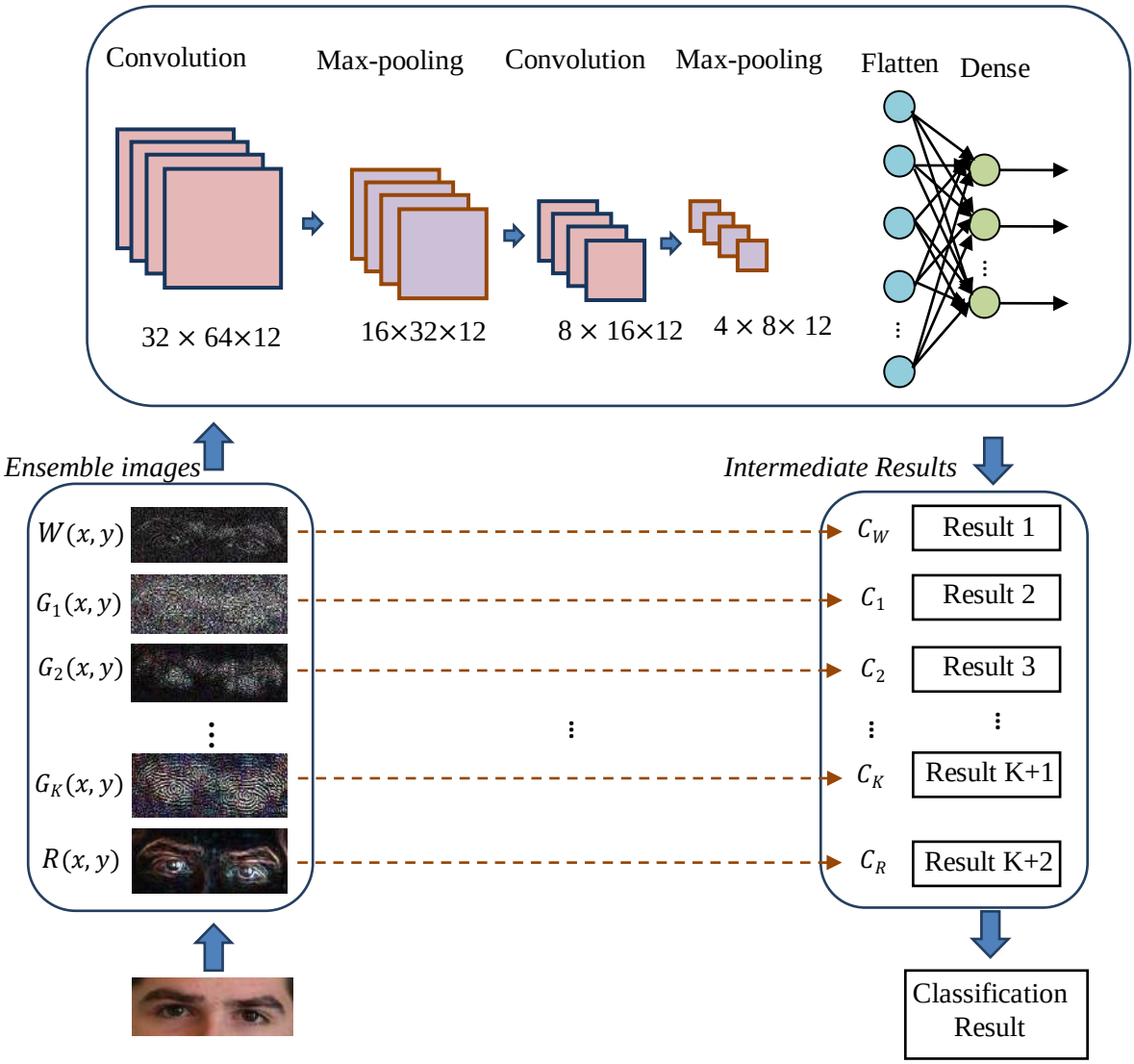


Fig 4: Ensemble image training in convolutional neural network

The ensemble images obtained by the DWT and the FA-MVEMD are converted to grayscale before training or testing using the CNN model. The input ensemble images to the CNN [29] have a size of $M \times N = 32 \times 64$. The convolution is performed between the weights and the input pixel values. We have used 12 convolutional filters therefore the output of the convolution layer has a volume of $32 \times 64 \times 12$. A sigmoid activation function is used at the output of the convolution layer so that the values are normalized between 0 and 1. The sigmoid activation does not change the size, so the size still remains $32 \times 64 \times 12$. The sigmoid activation function can be represented as

$$\sigma(n) = \frac{1}{(1+e^{-n})} \qquad (8)$$

The outputs of max-pooling result in the size of $16 \times 32 \times 12$with the use of a 2×2 window. The size of the features is $4 \times 8 \times 12$ at the end of the second max-pooling layer. The representation of the CNN model is depicted in Fig 4. Let the intermediate results obtained for $K + 2$ number of ensemble images be represented as

$$P_i = \{C_N, C_R, C_1, C_2, \ldots . C_K\} \quad \text{where} \quad P_i \in \{1, 2, \ldots . N_{class}\} \quad (9)$$

Where N_{class} is the number of classes that were trained by the model. The final result can be obtained from the intermediate result of the $K + 2$ ensemble image as

$$\hat{C} = \begin{cases} \max(P_i) if length(P_i == \max(P_i)) = 1 \\ No matching otherwise \end{cases} \quad (10)$$

3. Experimental Analysis

The dataset namely UBIPr, AR, and CASIA-Iris are used to evaluate the proposed periocular recognition scheme. Fig 5 shows a few sample images from these datasets. Performance metrics such as equal error rate (EER), the area under the ROC curve (AUC), rank1 recognition accuracy, and rank-5 recognition accuracy are used to evaluate the algorithm. The UBIPr dataset [30] consists of 10950 images from 261 subjects. The dataset contains 45.6% of female periocular images and the remaining 54.4% of male periocular images. The segmented version of face images (periocular region) of a single eye and the dual eye is available in the UBIPr dataset. The AR dataset has different facial expressions, occlusions, and illumination conditions of 56 women, and 70 men with a total of 4000 images. The CASIA-Iris dataset contains face images with dual eyes. The CASIA-Iris dataset consists of 6000 images collected from 200 subjects at a resolution of 1968×1024. The experimental evaluation was done with a different value of K. $K = 3, 4, 5, 6, \ldots .12$. The model was tested with ~30% of images in each subject and the remaining ~70% of images is used to train the model.

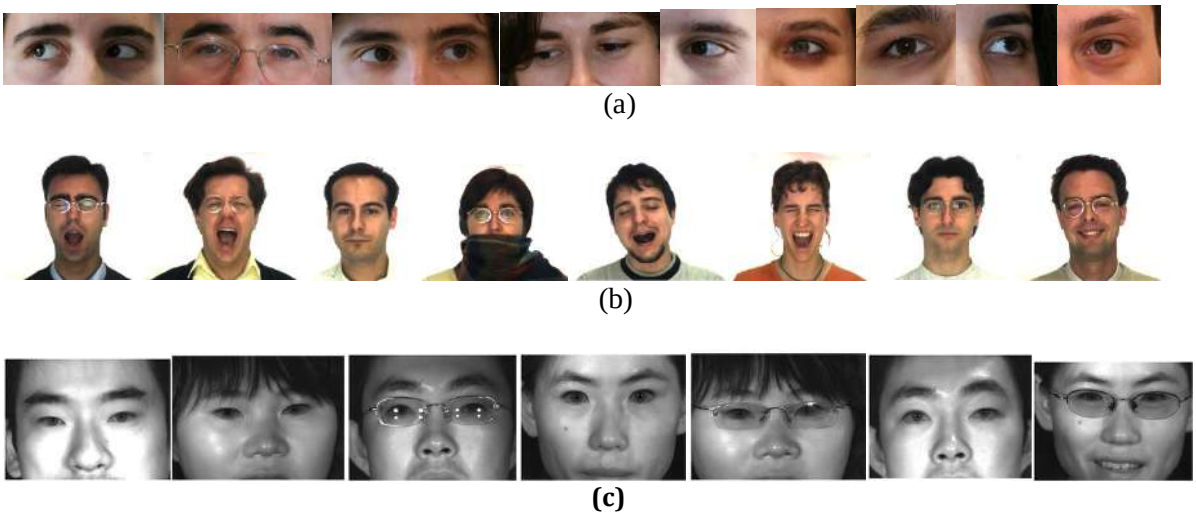
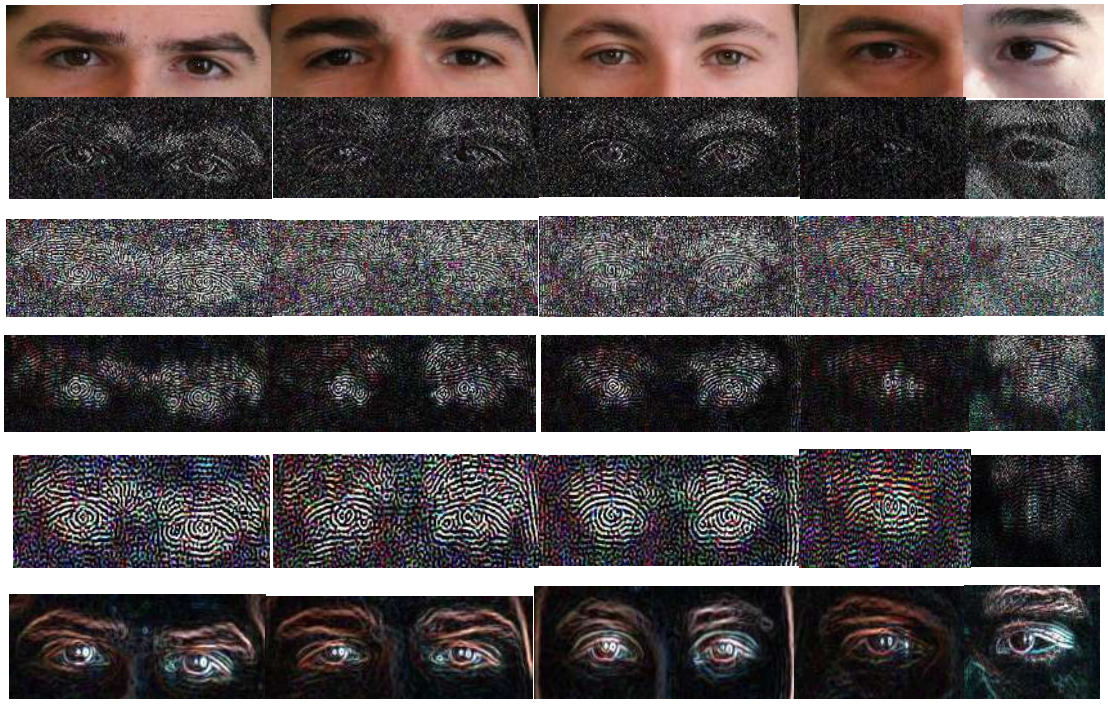


Fig 5: Sample test images from different datasets (a) UBIPr dataset (b) AR dataset (c) CASIA-Iris distance dataset

Few of the experimental results obtained from the proposed periocular recognition system for the UBIPr dataset are depicted in Fig 6. The evaluation was done in both single-eyed and dual-eyed periocular images. The results are obtained with the number of IMFs $K = 3$. The ensemble images are obtained from the HH band of DWT, three IMF images, and finally the residual as depicted in Fig 6.

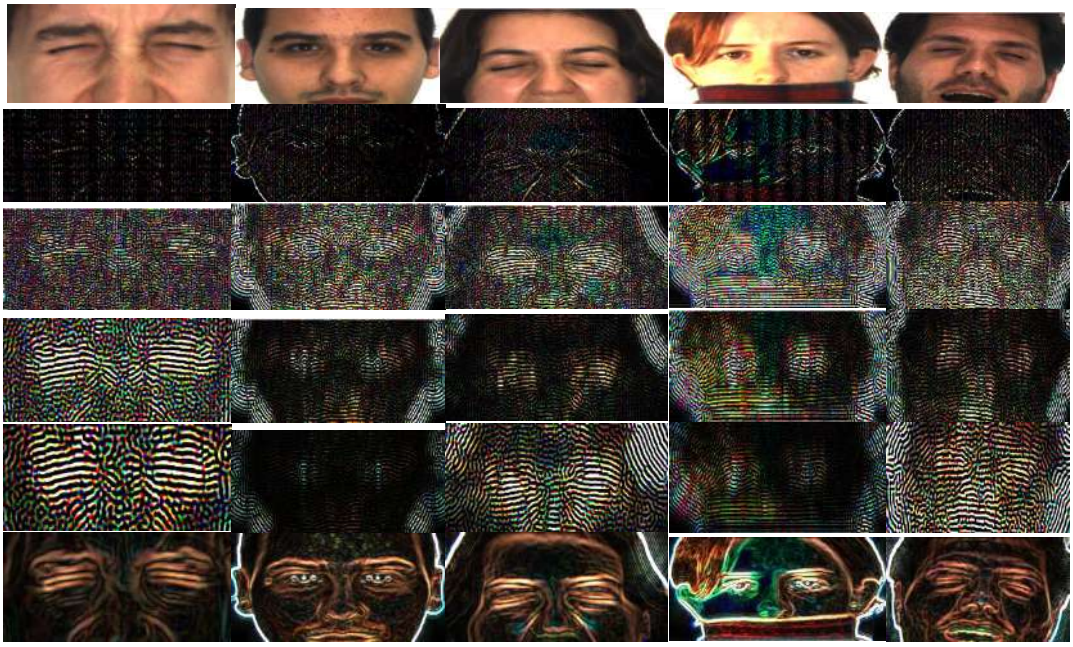
Few of the experimental results obtained from the proposed periocular recognition system for the AR and CASIA-Iris datasets are depicted in Fig 7 and Fig 8 respectively. The results are obtained with the number of IMFs $K = 3$. The ensemble images are obtained from the HH band of DWT, three IMF images, and finally the residual.

Fig 6, 7 and 8 depicts that the HH band of DWT holds the overall shape of the periocular region, while the three IMFs hold the periocular information related to different frequency bands. Finally the residual holds the overall facial information.



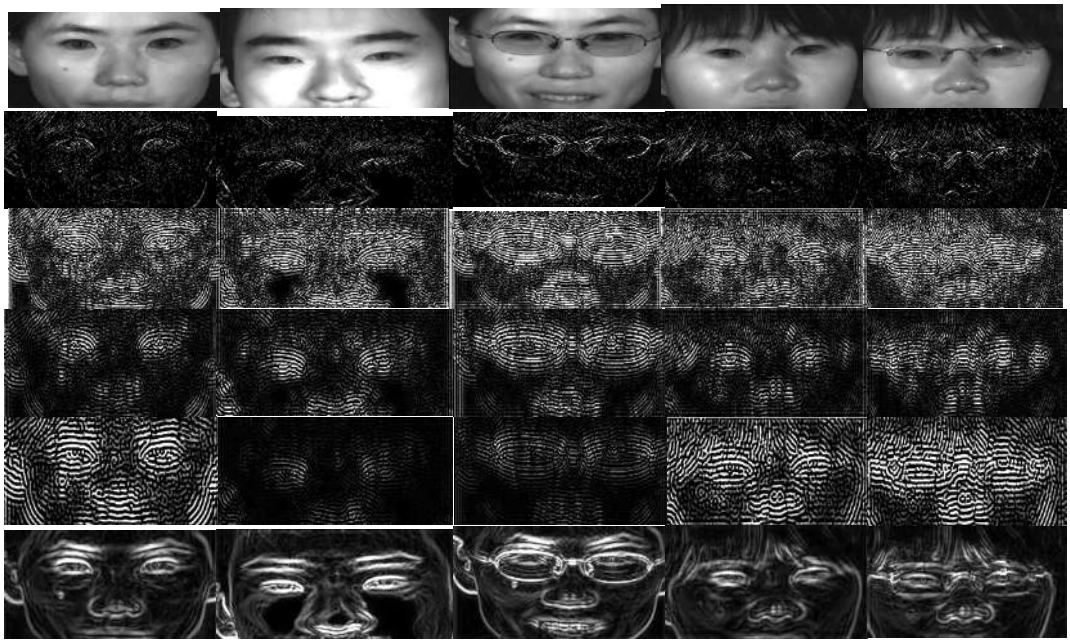
(a) (b) (c) (d) (e)

Fig 6 Experimental results of the proposed periocular recognition system for UBIPr dataset (row1) Input periocular image (row2) HH coefficients of single-level DWT (row3) IMF1 output (row4) IMF2 output (row5) IMF3 output (row6) Residual output



(a) (b) (c) (d) (e)

Fig 7 Experimental results of the proposed periocular recognition system for AR dataset (row1) Input periocular image (row2) HH coefficients of single-level DWT (row3) IMF1 output (row4) IMF2 output (row5) IMF3 output (row6) Residual output



(a) (b) (c) (d) (e)

Fig 8Experimental results of the proposed periocular recognition system for CASIA-Iris distance dataset (row1) Input periocular image (row2) HH coefficients of single-level DWT (row3) IMF1 output (row4) IMF2 output (row5) IMF3 output (row6) Residual output

Table 1: Comparison of EER and AUC for the proposed method with the traditional schemes for different dataset

	UBIPr		AR		CASIA-Iris	
Scheme	**EER (%)**	**AUC**	**EER (%)**	**AUC**	**EER (%)**	**AUC**
ImageNet [57]	7.11 ± 2.9	0.9805	14.53	0.9363	8.06 ± 5.3	0.9533
VGG16 [39]	4.38 ± 1.3	0.9892	7.69	0.9747	7.42 ± 1.7	0.9681
Clustering [35]	5.46 ± 1.5	0.987	9.4	0.9692	6.10 ± 2.2	0.9738
Iris recognition [32]	5.07 ± 2.2	0.9877	7.69	0.9751	7.51 ± 1.1	0.9674
Multilevel [37]	4.09 ± 2.1	0.9913	7.69	0.9756	8.69 ± 1.1	0.9594
LCNN [36]	6.34 ± 2.1	0.9849	9.39	0.9737	6.34 ± 1.6	0.9719
OC-LBCP [38]	4.21 ± 1.8	0.9898	6.83	0.9782	6.35 ± 0.5	0.976
Attention [24]	3.91 ± 1.3	0.9909	5.93	0.9834	5.86 ± 0.7	0.9813
I-LBP [23]	4.03 ± 1.3	0.9901	6.27	0.9792	6.08 ± 1.3	0.9798
EN-IR [25]	3.75 ± 1.4	0.9913	5.49	0.9869	5.06 ± 0.5	0.9878
Proposed	2.87 ±1.5	0.9961	4.71	0.9932	3.92 ±0.5	0.9925

The performance of the proposed method was compared with other recent periocular recognition schemes namely EN-IR [25], I-LBP [23], attention [24], OC-LBCP [38], LCNN [36], Multilevel [37], Iris recognition [32], Clustering [35], VGG16 [39] and ImageNet [31]. Table 1 shows the comparison of EER and AUC with other recent schemes with $K = 9$. The proposed method provides an EER of 2.87%, 4.71%, and 3.92% for the dataset UBIPr, AR, and CASIA-Iris datasets respectively. The EER of the proposed method is 0.88%, 0.78%, and 1.14% less than the scheme EN-IR. The proposed method has an AUC of 0.9961, 0.9932, and 0.9925 for the dataset UBIPr, AR, and CASIA-Iris which is higher than other recent schemes.

The recognition accuracy was evaluated using the metrics namely rank-1 and rank-5

recognition accuracy. The proposed approach provides a rank-1 recognition accuracy of 93.47%, 98.23%, and 98.03% for the dataset UBIPr, AR, and CASIA-Iris datasets respectively. The rank-1 recognition accuracy of the proposed ensemble training approach is 2.21%, 1.87%, and 1.32% higher than the EN-IR scheme for the dataset UBIPr, AR, and CASI-Iris datasets respectively. The Rank-5 recognition accuracy resembles the top-5 recognition prediction of the network. The proposed approach provides a rank-5 recognition accuracy of 98.01%, 99.31%, and 99.14% respectively. The rank-5 recognition accuracy is 0.02% less than the EN-IR scheme for the UBIPr dataset, however, it is 0.6% and 0.68% greater than the multilevel and EN-IR scheme for the AR and CASIA-Iris datasets respectively as depicted in Table 2. The graphical plot for rank-1 and rank-5 recognition accuracy for the dataset is UBIPr, AR, and CASIA-Iris dataset is depicted in Fig 9. From Fig 9, it is clear that the proposed scheme shows a better performance than other recent schemes in terms of rank-1 and rank-5 recognition accuracy.

Table 2: Comparison of recognition accuracy for the dataset UBIPr, AR and CASIA-Iris

Scheme	UBIPr		AR		CASIA-Iris	
	Rank-1	Rank-5	Rank-1	Rank-5	Rank-1	Rank-5
ImageNet [31]	84.88 ± 2.5	96.01 ± 1.8	93.59	96.75	95.00 ± 1.8	96.98 ± 2.5
VGG16 [39]	90.24 ± 1.4	97.09 ± 1.1	94.2	97.61	95.88 ± 0.1	97.99 ± 0.5
Clustering [35]	90.24 ± 1.4	97.36 ± 0.4	94.19	97.75	96.09 ± 2.1	98.10 ± 0.4
Iris recognition [32]	90.30 ± 1.2	97.41 ± 1.1	95.24	98.38	95.95 ± 2.1	98.15 ± 0.6
Multilevel [37]	90.75 ± 1.0	97.44 ± 0.3	96.07	98.71	95.81 ± 1.9	97.67 ± 1.0
LCNN [36]	90.28 ± 1.7	97.18 ± 0.7	94.27	97.52	96.01 ± 2.0	97.85 ± 0.9
OC-LBCP [38]	90.28 ± 1.2	97.39 ± 0.4	95.32	97.8	95.92 ± 1.3	98.12 ± 0.4
Attention [24]	91.02± 1.1	97.89 ± 0.8	96.17	98.01	96.32± 1.2	98.32± 0.6
I-LBP [23]	90.81 ± 1.3	97.53 ± 0.7	95.86	97.93	96.12± 1.5	98.01± 0.5
EN-IR [25]	91.26 ± 1.2	98.03 ± 0.7	96.36	98.17	96.71± 1.3	98.46± 0.7
Proposed	93.47±1.4	98.01±0.6	98.23	99.31	98.03±1.2	99.14±0.5

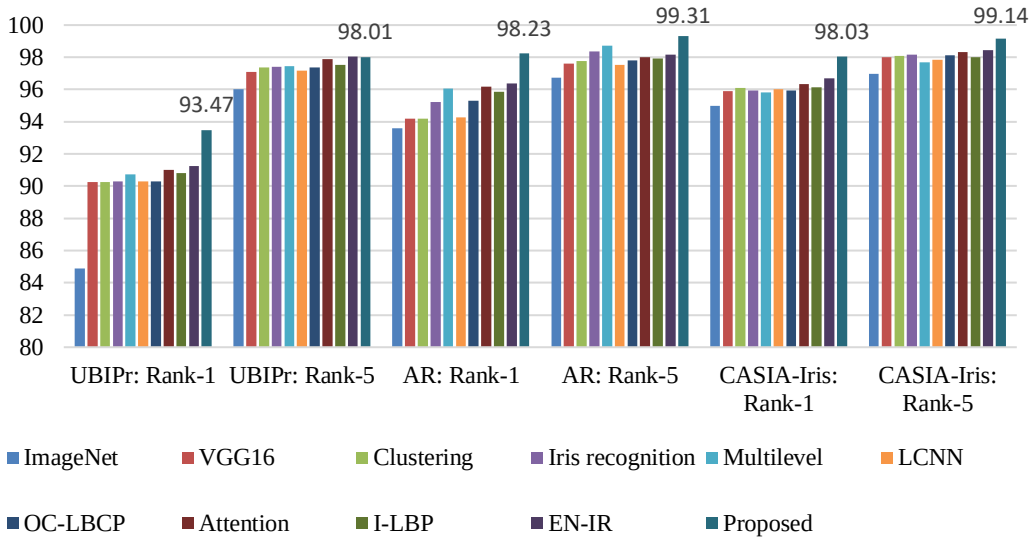


Fig 9: Graphical comparison of rank-1 and rank-5 recognition accuracy for different datasets.

The rank-1 and rank-5 recognition accuracy is also evaluated for different numbers of IMFs (K). The evaluation was done with $K = \{3,4,5,6,7,8,9,10,11,12\}$, which results in

5,6,7,8,9,10,11,12,13,14 ensemble images. As the value of K is increased from 3, the rank-1 and rank-5 recognition accuracy increases and it reaches a maximum at $K = 9$. For further increases in the value of K, the recognition accuracy gradually reduces. For example, the rank-1 recognition accuracy is 89.12% for the UBIPr dataset for $K = 3$. As the value of K is increased it reaches a maximum value of 93.47% at $K = 9$. For further increase in $K = 12$, the rank-1 recognition accuracy reduces to 93.04% as depicted in Table 3.

.

Table 3: Rank-1 and Rank-5 recognition accuracy for the different number of IMFs *K*

No. of IMFs K	Rank-1			Rank-5		
	UBIPr	AR	CASIA-Iris	UBIPr	AR	CASIA-Iris
3	89.12	93.79	92.87	93.66	94.87	93.98
4	90.24	94.17	93.76	94.78	95.25	94.87
5	90.76	94.92	94.19	95.3	96	95.3
6	91.37	95.88	95.76	95.91	96.96	96.87
7	92.42	96.72	96.91	96.96	97.8	98.02
8	93.36	98.01	97.83	97.9	99.09	98.94
9	93.47	98.23	98.03	98.01	99.31	99.14
10	93.21	98.08	97.91	97.75	99.16	99.02
11	93.17	97.92	97.68	97.71	99	98.79
12	93.04	97.79	97.52	97.58	98.87	98.63

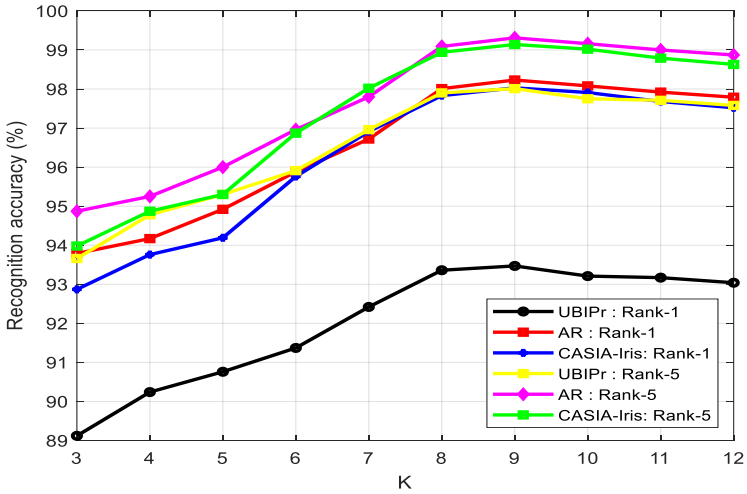


Fig 10: Graphical comparison of rank-1 and rank-5 recognition accuracy for different values of *K*

Fig 10 depicts the graphical comparison of rank-1 and rank-5 recognition accuracy for different values of K for the UBIPr, AR, and CASIA-Iris datasets. The maximum performance is achieved at $K = 9$. The ROC comparison for different schemes for the UBIPr, AR, and CASIA-Iris dataset is depicted in Fig 11. The time complexity of the proposed periocular recognition algorithm was evaluated on Matlab-2018a on Intel core i5 processor with a speed3.00GHz, 8GB RAM, Windows 10, 64-bit operating system. Let T_p, T_W, T_E, T_I, $T_{CNN,cl}$and $T_{CNN,tr}$represents the time of pre-processing, time of DWT estimation, time of EMD decomposition, time of interpolation, time of CNN classification, and time of CNN training respectively. The time of testing can be represented as:

$$T_{test} = T_p + T_W + T_E + T_I + T_{CNN,cl} \quad (11)$$

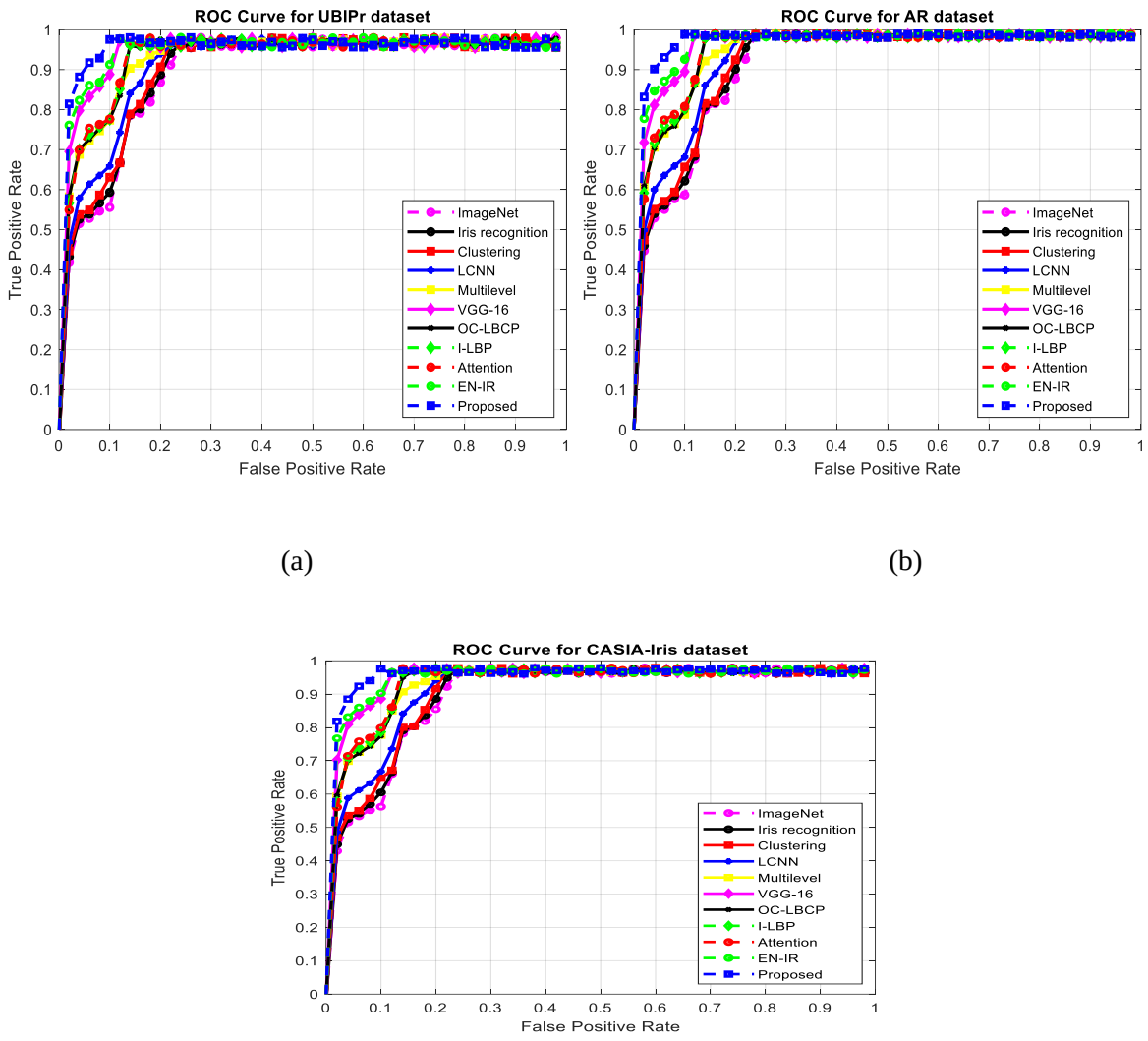


Fig 11: Comparison of ROC curve for different datasets (a) UBIPr dataset (b) AR dataset (c) CASIA Iris dataset

Table 4: Time complexity comparison for different numbers of IMFs *K*

Dataset	K	T_p (s)	T_W (s)	T_E (s)	T_I (s)	$T_{CNN,tr}$ (s)	$T_{CNN,cl}$ (s)	T_{test} (s)
UBIPr	3	0.561	0.478	0.21	0.213	2877	0.23	1.692
	5			0.86	0.326	4795	0.32	2.545
	7			1.21	0.458	6713	0.46	3.167
	9			1.51	0.567	8631	0.57	3.686
AR	3	0.568	0.482	0.28	0.213	1233	0.15	1.693
	5			0.92	0.326	2055	0.24	2.536
	7			1.41	0.458	2877	0.35	3.268
	9			1.54	0.567	3699	0.44	3.597
CASIA Iris	3	0.581	0.489	0.46	0.213	1644	0.18	1.923
	5			0.98	0.326	2740	0.27	2.646
	7			1.68	0.458	3836	0.38	3.588
	9			1.73	0.567	4932	0.49	3.857

Table 4 depicts the time consumption of different processes in the proposed periocular recognition algorithm for different values of K. As the value of K increases, the time of IMF estimation, interpolation, CNN training, and CNN classification also increases. Therefore, the time of testing increases as the value of K increases. For K=9, the UBIPr dataset, AR dataset, and CASIA Iris dataset provide a time of testing of 3.686s, 3.597s, and 3.857s respectively. Fig 12 depicts the graphical comparison of time of testing for different values of K. The next section shows the conclusion of the work.

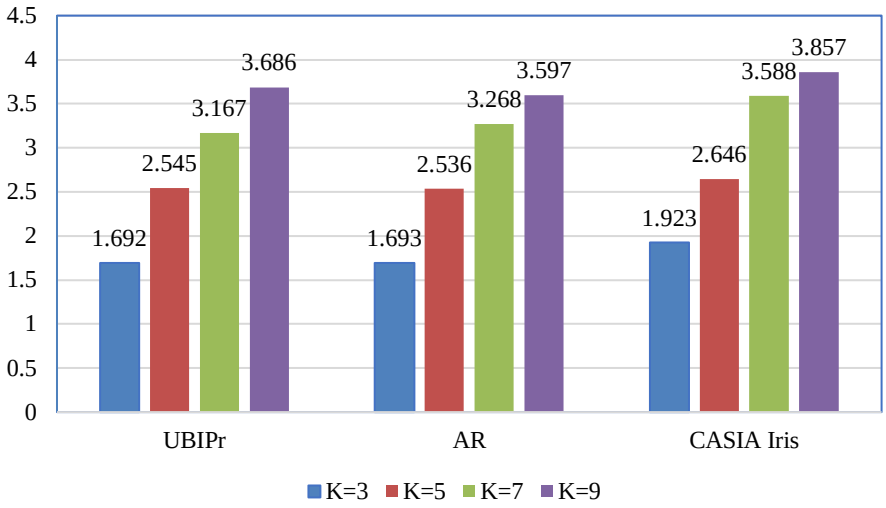


Fig 12: Graphical comparison of time of testing for different values of *K*

CONCLUSION

This paper proposed an ensemble training approach for the CNN classifier. The $(K+2)$ ensemble images are obtained from a single periocular image. The approach initially uses a single level DWT that decomposes the periocular image into four bands namely LL, LH, HL, and HH. The LL band is further decomposed using the FA-MVEMD algorithm into K-levels. The K-level IMF along with the residual and HH band of DWT is used as the ensemble image to obtain the ensemble images. The ensemble images are trained by the CNN model. In the testing phase, $K+2$ intermediate results are obtained from the ensemble test images. From the $K+2$ intermediate results, a final classification result is obtained. The evaluation was done using the datasets namely UBIPr, AR, and CASIA-Iris using the evaluation metrics such as EER, AUC, rank-1, and rank-5 recognition accuracy. The proposed approach provides a rank-1 recognition accuracy of 93.47%, 98.23%, and 98.03% for the datasets UBIPr, AR, and CASIA-Iris respectively. Also as the number of IMF increases the time complexity also increases, the proposed approach provides maximum recognition accuracy for $K=9$.

REFERENCES

[1] Uzair, Muhammad, Arif Mahmood, and Somaya Ali Al-Maadeed. 2017. Non-cooperative and occluded person identification using periocular region with visible, infra-red, and hyperspectral imaging. Biometric Security and Privacy. Springer, Cham, 223-251.

[2] Alahmadi, A., Hussain, M., Aboalsamh, H. and Azmi, A., 2020. ConvSRC: Smart Phone-based periocular recognition using deep convolutional neural network and sparsity augmented collaborative representation. Journal of Intelligent & Fuzzy Systems, 38(3) 3041-3057.

[3] Alonso-Fernandez, F., Bigun, J., 2014. Best regions for periocular recognition with nir and visible images. In: IEEE Intl. Conf. on Image Processing. pp. 4987 – 4991.

[4] Xu, J., Cha, M., Heyman, J. L., Venugopalan, S., Abiantun, R., Savvides, M., 2010. Robust local binary pattern feature sets for periocular biometric identification. In: IEEE Intl Conf. on Biometrics: Theory Applications and Systems (BTAS). pp. 1 – 8.

[5] Navneet Dalal and Bill Triggs,2005. Histograms of oriented gradients for

human detection, in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 1, 886–893.

[6] Kumari, Punam, and K. R. Seeja.2021.A novel periocular biometrics solution for authentication during Covid-19 pandemic situation. Journal of Ambient Intelligence and Humanized Computing 12(11) 10321-10337.

[7] Ramchoun, H., Ghanou, Y., Ettaouil, M., &Janati Idrissi, M. A. (2016). Multilayer perceptron: Architecture optimization and training.

[8] Ming Leung, K.. 2007. Naive Bayesian classifier. Polytechnic University Department of Computer Science/Finance and Risk Engineering, 1-16.

[9] Das, S., Kharbanda, K., M, S., Raman, R., & D, E.D. (2021). Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy. Biomed. Signal Process. Control., *68*, 102600.

[10] Sandhya kumari, S., Sandhya Rani, K.2022.Selection of MSER region based Ultrasound Doppler scan Image Big data classification using a faster RCNN network. International Journal of Computer Engineering In Research Trends (IJCERT), 9(10) 10,184-192.

[11] Le, T.H.N., Prabhu, U., Savvides, M., 2014. A novel eyebrow segmentation and eyebrow shape-based identification. In: IEEE International Joint Conference on Biometrics. IEEE, pp. 1–8.

[12] Ahonen, T., Hadid, A., Pietikainen, M., 2006. Face Description with Local Binary Patterns: Application to Face Recognition, in IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12) 2037-2041.

[13] Chui, H., Rangarajan, A. 2003. A new point matching algorithm for nonrigid registration, Computer Vision and Image Understanding, 89(2) 114–141.

[14] Mahalingam, G., Ricanek, K., Albert, A.M., 2014. Investigating the periocular-based face recognition across gender transformation. IEEE Trans. Inf. Forensics Secur. 9 (12), 2180–2192.

[15] Joshi, A., Gangwar, A., Sharma, R., Singh, A., Saquib, Z., 2014. Periocular recognition based on Gabor and Parzen PNN. In: 2014 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 4977–4981.

[16] Ambika, D.R., Radhika, K.R., Seshachalam, D., 2017. Fusion of shape and texture for unconstrained periocular authentication. World Acad. Sci., Eng. Technol., Int. J. Comp., Electr., Autom., Control Inf. Eng. 11 (7) 821–827.

[17] Chen, H., Gao, M., Ricanek, K., Xu, W., Fang, B., 2017. A novel race classification method based on periocular features fusion. Int. J. Pattern Recognit. Artif. Intell. 31 (08), 1750026.

[18] Proença, H., Neves, J. C. 2018. Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning Frameworks, in *IEEE Transactions on Information Forensics and Security*, 13(4) 888-896.

[19] Zhao, Z., Kumar, A. 2017. Accurate Periocular Recognition Under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network, in IEEE Transactions on Information Forensics and Security, 12(5) 1017-1030,

[20] Tan, C.,Kumar, A.2013. Towards Online Iris and Periocular Recognition Under Relaxed Imaging Constraints, in IEEE Transactions on Image Processing, 22(10) 3751-3765.

[21] Zhao, Z., Kumar, A. 2018. Improving Periocular Recognition by Explicit Attention to Critical Regions in Deep Neural Network, in IEEE Transactions on Information Forensics and Security, 13(12), 2937-2952.

[22] Bakshi, Sambit, Pankaj K. Sa, and Banshidhar Majhi. 2015. A novel phase-intensive local pattern for periocular recognition under visible spectrum. Biocybernetics and Biomedical Engineering 35(1) 30-44.

[23] Kumar, Gautam, Sambit Bakshi, Pankaj Kumar Sa, and Banshidhar Majhi. 2021.Non-overlapped blockwise interpolated local binary pattern as periocular feature. Multimedia Tools and Applications 80(11) 16565-16597.

[24] Luo, Zhengding, Junting Li, and Yuesheng Zhu. 2021. A deep feature fusion network based on multiple attention mechanisms for joint iris-periocular biometric recognition.*IEEE Signal Processing Letters* 28, 1060-1064.

[25] Vyas, Ritesh. 2022. Enhanced near-infrared periocular recognition through collaborative rendering of hand crafted and deep features. Multimedia Tools and Applications, 1-15.

[26] Heil, Christopher E., and David F. Walnut.1989. Continuous and discrete wavelet transforms. SIAM review 31(4) 628-666.

[27] Thirumalaisamy, Mruthun R., Phillip J. Ansell.2018. Fast and adaptive empirical mode decomposition for multidimensional, multivariate signals. IEEE Signal Processing Letters 25(10)1550-1554.

[28] Thévenaz, Philippe, Thierry Blu, and Michael Unser. 2000. Image interpolation and resampling. Handbook of medical imaging, processing and analysis 1(1) 393-420.

[29] Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi.2017. Understanding of a convolutional neural network. 2017 international conference on engineering and technology (ICET). IEEE.

[30] Chandrashekhar Padole, Hugo Proença, 2012. Periocular Recognition: Analysis of Performance Degradation Factors, in Proceedings of the Fifth IAPR/IEEE International Conference on Biometrics – ICB 2012, New Delhi, India, March 30-April 1,

[31] Krizhevsky, A.; Sutskever, I.; Geoffrey, H.2012. Imagenet classification with deep convolutional neural networks. In Proceedings of the International Conferences on Neural Information Processing Systems (NIPS), Lake Tahoe,NV, USA, 3–6 December; pp. 1097–1105.

[32] Gangwar, A., Joshi, A. DeepIrisNet, 2016. Deep iris representation with applications in iris recognition and cross-sensor iris recognition. In Proceedings of the International Conferences on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September, pp. 2301–2305.

[33] Schroff, F.; Kalenichenko, D.; Philbin, J.2015. FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June, pp. 815–823.

[34] Wu, X., He, R.; Sun, Z., Tan, T. 2018. A light CNN for deep face representation with noisy labels. IEEE Trans. Inf. Forensics Secur. 13, 2884–2896.

[35] Soleymani, S., Dabouei, A., Kazemi, H., Dawson, J.,Nasrabadi, N.M. 2018. Multi-level feature abstraction from convolutional neural networks for multimodal biometric identification. In Proceedings of the International Conferences on Pattern Recognition (ICPR), Beijing, China, 20–24 August, 3469–3476.

[36] Tiong, Leslie Ching Ow, Yunli Lee, and Andrew Beng Jin Teoh.2019. Periocular recognition in the wild: Implementation of RGB-OCLBCP dual-stream CNN. Applied Sciences 9(13) 2709.

[37] Parkhi, O.M.; Vedaldi, A.; Zisserman, A. 2015. Deep face recognition. In Proceedings of the British Machine Vision Conference, Swansea, UK, 7–10 September; pp. 1–12.