Human ear print recognition based on fusion of difference theoretic texture and gradient direction pattern features

Kawther Thabt Saleh, Raniah Ali Mustafa, Haitham Salman Chyad
Department of Computer Science, College of Education, Mustansiriyah University, Baghdad, Iraq

Article Info

ABSTRACT

Human ear recognition can be defined as a branch of biometrics that uses images of the ears to identify people. This paper provides a new ear print recognition approach depending on the combination of gradient direction pattern (GDP2) and difference theoretic texture features (DTTF) features. The region of interest (ROI), the gray scale of the ear print was cut off, noise removal by the median filter, histogram equalization, and local normalization (LN) are the first steps in this approach. After the image has been processed, it is used as input for the fusion of GDP2 and DTTF for extracting the features of ear print images. Lastly, the Gaussian distribution (GD) was utilized to compute the distance among fusion feature vectors (FV) for ear print images for recognizing ear print images for people using a set of images that had been trained and tested. The unconstrained ear recognition challenge (UERC) database, which comprises 330 subjects for ear print images, provides the approach that was suggested by employing ear print databases. Furthermore, experimental results on images from a benchmark dataset reveal that statistical-rely super-resolution methods outperform other algorithms in ear recognition accuracy, which was around 93.70% in this case.

Corresponding Author:

Raniah Ali Mustafa
Department of Computer Science, College of Education, Mustansiriyah University
Baghdad, Iraq
E-mail: rania83computer@uomustansiriyah.edu.iq

1. INTRODUCTION

The biometrics scheme is very secure compared to the traditional methods such as passwords and PINs, due they can be readily forgotten or tampered with. The three essential components of biometric schemes are input data, processing unit (authentication, identification, and verification), and output data. Input data are obtained from sensors. The behavioral parameters such as (fingerprint, earprint, speech, iris, and face) [1]. The authentication person has become a fundamental task for providing security for accessing restricted schemes and resources. The biometric-based authentication schemes have been utilized for this purpose [2], [3]. In the past few impressive years, human ear recognition (HER) had become very attractive in biometric authentications. The significant reasons behind human ear biometrics over another biometric modality are smaller in size; much-stabilized shape had been proven through clinical monitoring [4], [5].

The human ear is an interesting anatomic component of a passive; physiological biometrics scheme that relies on digital camera images. The human ear has several unique features which allow finding specific individuals. It can be utilized as efficient biometrics schemes, for instance, in crowd monitoring and terrorist identification in public places like airports and in controlling access to governmental offices [6]. Several biometric recognition approaches based upon gradient direction pattern (GDP2) and difference theoretic texture features (DTTF) have lately been investigated. Such as Mangayarkarsi *et al.* [7] suggested a detection

and recognition system for the human ear to utilize biometrics. In the first phase, the ear is detected and segmented from input image utilizing the contour detection algorithm (CDA). Scale invariant feature transform (SIFT) was utilized for segmented ear image in the second phase for extracting SIFT features. The model file is created from the training images utilizing the extracted features. Test images perform the same process, and finally phase utilized the Euclidean distance metric (ED), which computes the percentage of difference among the images, was utilized for ear recognition.

The value of this distance measure has thus been utilized to identify a certain person's ear. From low-resolution ear images, Zarachoff et al. [8] described a method for ear recognition depending on principal component analysis (PCA) and super-resolution algorithms. In this study, ear images were divided into an image query and a database, filtered and down-sampled, resulting in low-resolution ear images. The low-resolution images are subsequently extended to their original sizes with the use of a range of nearest neighbor classifier (NN) and statistical-based super-resolution approaches. The images are, after that, subjected to PCA to obtain their eigen values, utilized as matching features. According to experimental results on images from benchmark dataset, statistical-based super-resolution approaches, notably wavelet-based methods, outperform other algorithms regarding ear recognition accuracy.

Emersič et al. [9] present a study of a convolutional neural network (CNN) training problem based on (closed set) ear recognition models utilizing limited training data. In this work, various model training strategies were investigated and generated an outperformed model on better performing state-of-art methodology (which has been based upon histogram of oriented gradients (HOG) descriptors) compared to about 30% in terms of first-rank recognition rate (RR). The better model we have been able to generate has been based upon the SqueezeNet model, which has been fine-tuned with a limited set of 1,383 ear images of 166 classes which have been increased by a factor of 100 after learning parameters with ImageNet data.

Sarangi et al. [10] propose an excellent approach for representing ear images through the combination of the 2 most successful local feature descriptors, local directional patterns (LDP) and pyramid histogram of oriented gradients (PHOG). Use the PHOG to express spatial information of the shape and LDP to effectively encode local information of the texture. Because feature sets have a lot of high dimensionalities, the PCA has been utilized so as to lower dimensions before normalization and fusion in this study. Two sets of normalized heterogeneous features were combined after that in order to generate a single feature vector (FV). Lastly, using kernel discriminant analysis (KDA) approach, extract the relevant characteristics and effectively recognize them using the nearest neighbor classifier (NN). Experimentations on three standard datasets (Univ. of Notre Dame collection E, IIT Delhi versions (1 and 2) show that the suggested approach can provide adequate recognition performance when compared with existing successful techniques.

To enhance ear images, Sarangi et al. [11] proposed an automatic enhancement approach that has been based on metaheuristic optimization. The proposed algorithm for improving ear images in a few iterations by incorporating a mutation operator to a simple and new yet meta-heuristic optimization technicality is termed the enhanced Jaya algorithm. After that, using speeded-up robust features (SURF), a pose-invariant local feature extractor, extract local features. Lastly, the accuracy identification rate was calculated using the k-nearest neighbor (k-NN) classifier. Extensive tests are carried out on four standard datasets, with quantitative and qualitative metrics used to assess performance. Experimental results have clearly shown that the suggested enhancement method is competitive when compared with two conventional approaches contrast limited adaptive histogram equalization (CLAHE) and histogram equalization (HE), as well as two meta-heuristic algorithms, differential evolution (DE) and particle swarm optimization (PSO), all of which are based on an image-enhancing technicality. For the authentication technique, Thivakaran et al. [12] have presented a multimodel of fingerprint and ear biometrics. The feature extraction (FE) technique uses hybrid technicalities depending on fingerprint feature extraction (FE), like Minutiae and Singular point, whereas ear feature extraction uses SURF and binary robust invariant scalable keypoints (BRISK) technicalities. The feature-level fusion process is then used to obtain more reliable information concerning features.

Lastly, a matching operation has been carried out with the use of the registration and an affine transform with a similarity score. Ear biometrics appears to be an excellent solution for strengthening security requirements in numerous industries because the ears are noticeable and could be Taken simply, even without any awareness of an individual being checked. Results of the experiments have been based upon the IITDelhi database for ear images and the CASIA database for fingerprint images. The recommended approach could achieve 95.96% accuracy with low error rates of 0.19% false rejection rate (FRR) and 0.11% false alarm rate (FAR), compared to 0.17% FAR and 0.37% FRR for the present approach in-ear alone. Using different optimization technologies, the precision of the provided approach with more biometric images like palm print, face, and iris could be improved in future work.

Tariq et al. [13] proposed a unique approach for human ear recognition. There are three stages to the system. Preprocessing is the initial step (contrast enhancement (CE) and size normalization for ear image). The Haar wavelets are employed in the second step to extract features. Recognition utilizing quickly normalized

cross-correlation is the final phase. The method is being used on the lIT Delhi and University of Science and Technology Beijing (USTB) ear imaging databases. According to experimental results, the proposed method appears to attain an average precision of 97.2% and 95.2% on such databases.

Tian and Mu [14] suggested an ear recognition based upon the deep convolutional network provides an image for human ear recognition and suggests a deep CNN (DCNN). A total of three convolutional layers, a fully-connected layer, and a softmax classifier make up a CNN. The system employs the USTB ear database, suggesting that our proposed algorithm is simpler, more accurate, and superior to the standard algorithm in coping with partial occlusion. Ying *et al.* [15] suggested a DCNN-based solution to improve human ear recognition. A deep network structure-based CNN is used to challenge human ear recognition. The optimum activation function has been found to prevent network overfitting, and Dropout technology has been implemented in the final fully connected layer. The network model has been trained using a large number of samples of the human ear images in order to determine the learning rate, and a number of the feature graphs, in addition to other network characteristics. Furthermore, the human ear recognition test has been based upon a trained network model. The algorithm appears resistant to occlusion, lighting, and rotation in the comparison experiment, and the rate of human ear recognition has been much enhanced.

Zarachoff *et al.* [16] presented a two-dimensional wavelet relying on multi-band PCA (2DWMBPCA) technique, inspired through PCA based on technical for hyperspectral and multispectral images, which showed significantly higher performance than standard PCA. The suggested technique performs a non-destructive two-dimensional wavelet transformation on the input image to split the image into its sub-bands. After that, based on the coefficient values, it divides each resulting sub-band into several bands. The eigenvectors for the sub-bands are then extracted using standard PCA on each of the resulting bands, which are used after that to match the features. The suggested 2D WMBPCA outperforms traditional PCA as well as the eigenfaces techniques in images from two benchmark ear image datasets, according to experimental results. Jiddah and Yurtkan [17] contributed to the area of human ear recognition by including the texture and the geometrical features. This study employs the AMI ear database to extract the local binary pattern (LBP) features and run a laplacian filter (LF) on raw images to extract the geometrical characteristics. In order to discover the region of high importance in the images of the human ear, the ear database has been processed by dividing the ear image to 4 quarters and experimenting on every one of them separately. After that, the geometric and texture features are fused, and studies are conducted to verify the contribution of the fused features.

Al Rahhal *et al.* [18] presented a new ear recognition descriptor. Dense local phase quantization (DLPQ) is a suggested descriptor that has been based upon phase responses generated using a well-known LPQ descriptor. Local dense histograms have been generated from phase maps' horizontal stripes, succeeded by pooling procedure to account for changes in the viewpoint, and finally, an ear descriptor is concatenated. Although suggested DLPQ descriptor has been based upon the classic LPQ, we show that it achieves significant enhancements (over 20%) compared to the latter descriptor on two benchmark datasets. Othman *et al.* [19] offer a fully automated ear-based biometric system that does not require human interaction and can be utilized in real-time.

The suggested approach is designed to distinguish persons depending on their ear shape, which is retrieved from a profile facial image frequently partially obscured by earrings and/or hair. A cascaded classifier-based ear recognition method is used first to recognize ears in profile image depending on Haar-like characteristics. After that, a new ear detection method is used depending on the shape context descriptor. The results of testing the suggested method on a few standard datasets indicate encouraging results: 100% recognition was achieved for non-occluded images, whereas 57% accuracy was reached for images where the ear has been occluded by both earring and hair.

Khaldi and Benzaoui [20] suggest employing a tight region of interest (ROI) segmentation of an ear to avoid this and ensure that a classifier relies solely on ear pixels. This work used the Image-to-Image translation to generate ear ROI segmentation and remove the irrelevant pixels from the input images. Additionally, missing ear components owing to distortion or occlusion can be synthesized. To do this, we employed the Pix-2-Pix generative adversarial network (GAN) that has been trained on annotated web ears (AWE) dataset, which is a difficult ear dataset. The use of ear ROI segmentation improves the process of the classification and dramatically boosts the rate of recognition, according to the results of the experiments.

2. PROPOSED METHOD
2.1. Difference theoretic texture features (DTTF)
Texture identification and classification under diverse rotation, lighting and scale conditions is a difficult problem in pattern recognition, and grey level diversity statistics were widely used to solve it. DTTF presents a new set of rotation, lighting and scale-invariant texture classification features derived from correlated distributions of global and local grey level differences in texture image intensities. The authors examine the terms in the correlation formula to construct a difference-based feature set invariant and distinctive for a texture

class. The authors' studies use the nearest neighbor classifier, and the findings show that the proposed feature vector has great classification accuracy under diverse rotation, lighting, and scale conditions [21]; furthermore, because they contribute information regarding the same repeating pattern, a DTTF used in another system that color textures have a distinctive relation between their color levels. Therefore, the average information, or entropy, is considered redundant across color levels.

Additionally, which is focused on lowering the dimensionality of the features of the color texture while keeping high color texture identification accuracy through averaging the entropies across multi-dimensional color levels. Susan and Hanmandlu [21] used the mean operation for summarizing the features of the original 11-dimensional variance theoretic texture for the classification. In this work, they used the entropy of the characteristics spanning multiple color levels rather than the mean. The not-comprehensive entropy with Gaussian information gain has been utilized as a measure of the entropy in their investigations because it's nonlinear and a good predictor of the regular patterns in the textures. Concerning feature dimension reduction, comparisons to state-of-the-art reveal that their method is efficient and accurate [22]. It has also been used to recognize dynamic faces in videos using 3D-difference theoretic texture features (3D DTTF).

Along with current vertical, horizontal, and diagonal directions in the 2D DTTF, the 3 D-DTTF expands grey-level variation statistics along front (F), front-diagonal vertical (FDV), and front diagonal horizontal (FDH) directions. The new 3D characteristics are affine invariant, meaning they are identical to their 2D counterparts, which is useful for detecting faces in videos despite changes in facial expressions. On Cohn-Kanade facial expression video dataset, the recommended dynamic face recognition algorithm outperforms existing techniques [23].

2.2. Gradient direction pattern (GDP2)

To extract features (EF) from the grayscale image, only the eight pixels around each pixel were used. To get an eight-mask value, multiply the 3×3 pixels region by the Kirsch edge response mask in eight directions. Following the application of direction masks, the pixel's original color value has been substituted with the value of the corresponding mask. Figure 1 shows the obtaining GDP code from a 3×3 Region, where the corresponding mask value in north-west (NW) and south-east (SE) direction mask, corresponding mask value in N and S direction mask, corresponding mask value in north-east (NE) and south-west (SW) direction mask, corresponding mask value in W and E direction mask and the GDP is applied to that region to produce the GDP code illustrate in Figure 1(a), Figure 1(b), Figure 1(c), and Figure 1(d) respectively, which has higher stability in the noisy environments compared to the general GDP code. Furthermore, the GDP code is a four-bit binary pattern that could yield up to 16 different combinations [24].
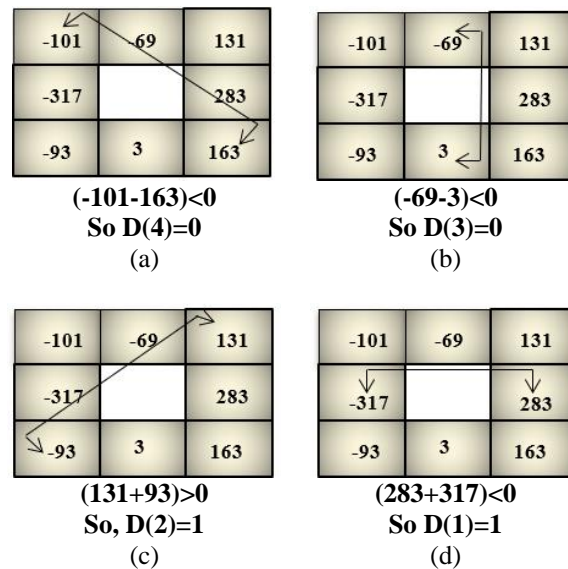

Figure 1. Obtaining GDP code from a 3×3 region corresponding mask value in (a) NW and SE direction mask, (b) N and S direction mask, (c) NE and SW direction mask, and (d) W and E direction mask (GDP code =0011=3(dec))

The grayscale image was separated into 81 equal-sized blocks, and the feature vector has been created by concatenating the histograms of GDP codes from every block. Figure 2 shows the obtaining mask value for a 3×3 region where the local region from the gray scale image, mask value for the pixel '20' using NW directional mask, mask value for the pixel '52' using N directional mask, corresponding mask value in eight directions and new representation of (a) using mask value are illustrated in Figure 2(a), Figure 2(b), Figure 2(c), Figure 2(d) and Figure 2(e) respectively. Only single transition patterns were considered while calculating GDP. As a result, every block's histogram length was 8, and the FV length for the entire image was 81x8= 648 [25].
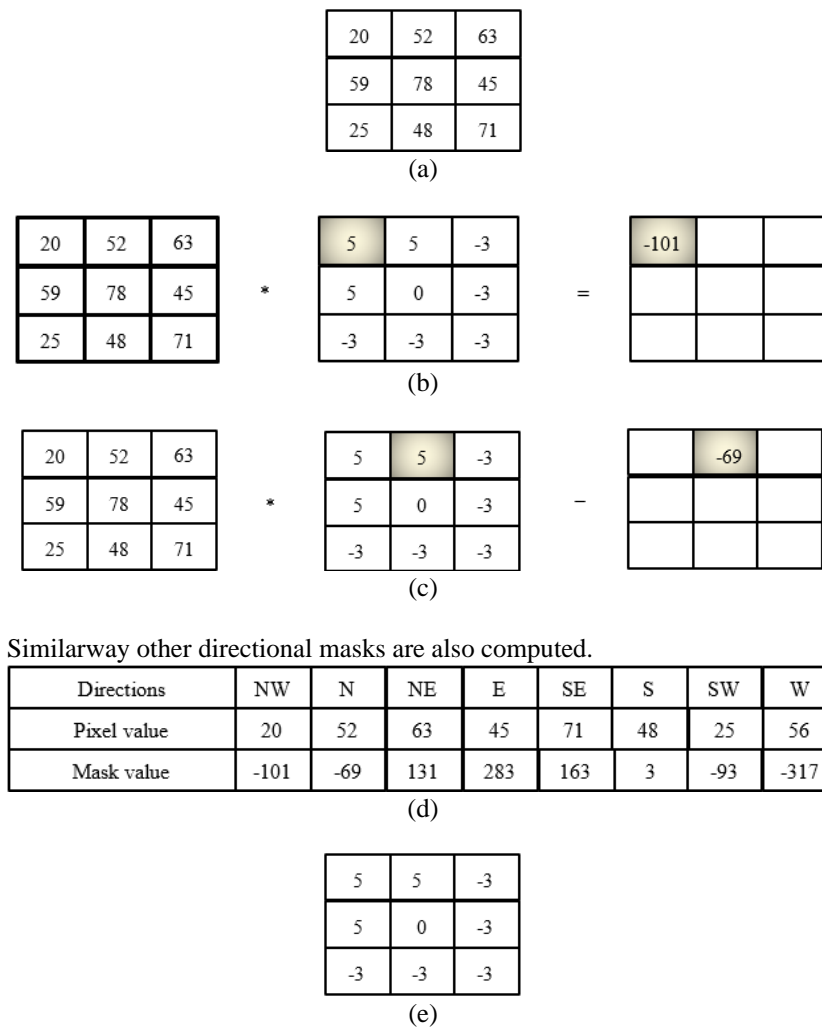
20	52	63
59	78	45
25	48	71

(a)

20	52	63
59	78	45
25	48	71

*

5	5	-3
5	0	-3
-3	-3	-3

=

-101		

(b)

20	52	63
59	78	45
25	48	71

*

5	5	-3
5	0	-3
-3	-3	-3

–

	-69	

(c)

Similarway other directional masks are also computed.

Directions	NW	N	NE	E	SE	S	SW	W
Pixel value	20	52	63	45	71	48	25	56
Mask value	-101	-69	131	283	163	3	-93	-317

(d)

5	5	-3
5	0	-3
-3	-3	-3

(e)

Figure 2. Obtaining mask value for a 3×3 region (a) local region from the gray scale image, (b) mask value for the pixel '20' using NW directional mask, (c) mask value for the pixel '52' using N directional mask, (d) corresponding mask value in eight directions, and (e) new representation of (a) using mask value

3.   PROPOSED SCHEME

The suggested method is divided into three phases: testing, training, and recognition. The phases of training are divided into 2 parts: the first one is the preprocessing, which includes several sub-phases (ROI, grayscale, remove noise, histogram equalization, and local normalization (LN)). The second phase involves extracting features for ear print images by combining DTTF and GDP2 and storing features for each class for each ear print image sample for an individual in the training database (TRDBF). The phase of the testing is involved as well in the scheme's critical phases (feature extraction phase and preprocessing phase). Lastly, Gaussian distribution (GD) is used in the recognition phase. Figure 3 depicts the suggested method's framework.
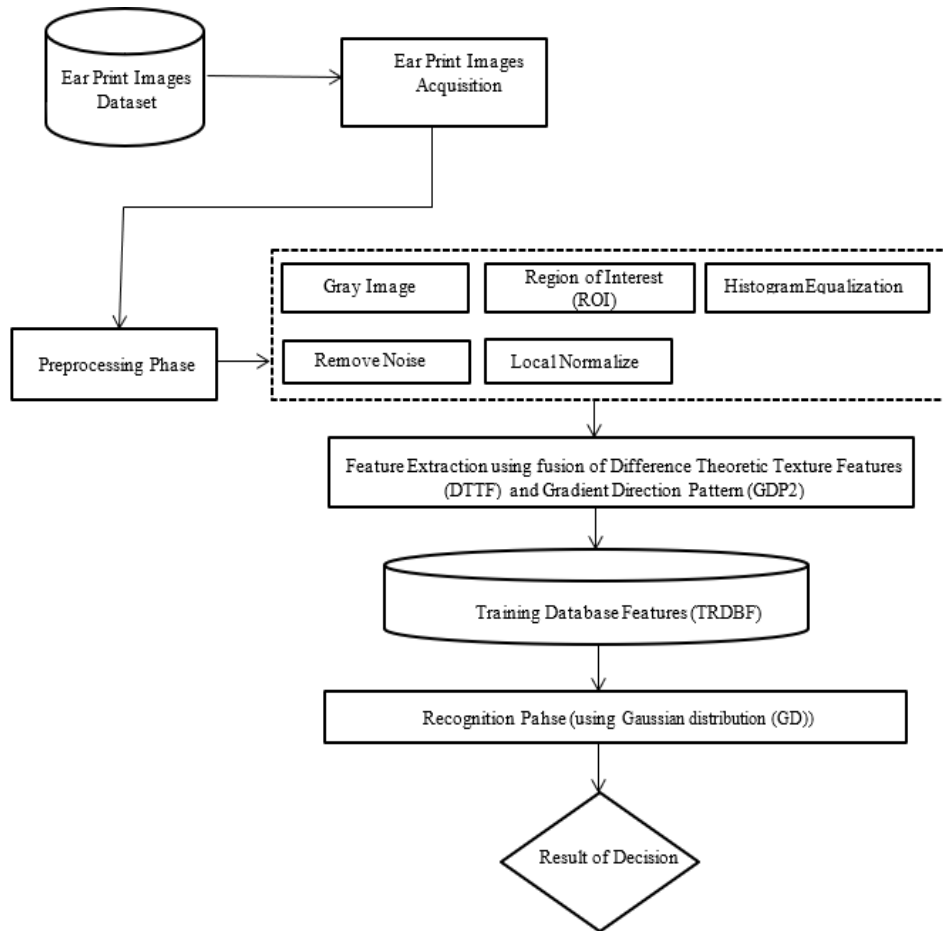
Figure 3. The general model of the proposed scheme

4.   RESULTS AND DISCUSSIONS

The results of recognition of ear print system, showing the results for the three phases: input images phase, preprocessing phase (convert to gray scale sub-phase, region of interest (ROI) sub-phase, histogram equalization (HE) sub-phase, median filter sub-phase, and local normalization (LN) sub-phase) and feature extraction phase in addition to a full explanation in this section.

4.1. Input images phase

This phase of proposed technique involves loading an image into the proposed scheme of recognition, after which it's made available for subsequent phases. The suggested approach can read an image with each extension (i.e., image format), and it used a BMP image format for ear print and color images. Figure 4 depicts the array.

No Pixels	Row/ Column (x,y)	Z (Preprocessing on x and y)							
		Value			HSV				
		Red	Green	Blue	H	S	V	SkinHSV	Contrast
1	1,1								
	1,2								
	...								
2	2,1								
	2,2								
	...								

Figure 4. Example of the 3D array contents

4.2. Pre-processing phase
In the preprocessing, the suggested scheme consists of five sub-phases as:
- Convert to gray scale sub-phase: in this sub-phase, after the input ear print image will convert the color ear print image to a grayscale ear print image by utilizing the appropriate formula. Figure 5 shows the conversion to grayscale. The original image has been illustrated in Figure 5(a) and the gray image has been illustrated in Figure 5(b). The (1) applies for conversion to grayscale.

$$Gray\ scale = 0.2989R + 0.5870G + 0.1140B \qquad (1)$$

(a)   (b)

Figure 5. Show conversion grayscale for (a) original image and (b) gray image

- Region of interest (ROI) sub-phase: After converting the earprint image to grayscale, will utilize logical indexing to extract the ROI. The idea is identical to indirect indexing, except that each index is logical value in this case. Assume that there is a matrix Mat of integer numbers generated randomly. Then we will use the lower bound 10 to substitute each value less than 10. The instructions idx < –Mat < 10 returns an array where each position obeys the true conditions. After, each value is substituted [26]. Figure 6 shows the logical indexing example.


```
1  Mat <- matrix(sample(25), nrow = 5)
2  idx <- Mat < 10
3  Mat[idx] <- 10
```

Figure 6. Logical indexing example

Another variation of indexing, logical indexing, is both beneficial and expressing. The matrix subscript is represented by a single logical array in logical indexing. MATLAB obtains the matrix elements that correspond to the logical array of nonzero values. The result is always presented as a column vector. Logical indexing (LI) is utilized to obtain and modify the ROI pixels. One good method to do that in MATLAB is to utilize logical indexing (LI). The expression A (B), if B is the logical and same size as A, selects every element in A that corresponds to B's real elements.

If you search the workspace's browser, you can see that the mask is a logical array. An important example (which is utilized here) is the mask = roipoly(I) function from MATLAB. Figure 7 depicts apply mask for image, where the gray image has been illustrated in Figure 7(a) and the mask image has been illustrated in Figure 7(b).
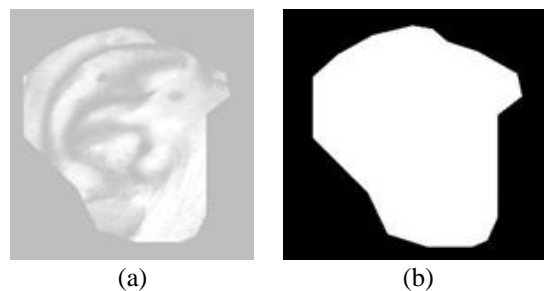

(a)   (b)

Figure 7. Show two images for (a) gray image and (b) mask image

Now, in the section show the results of the multiply the grayscale earprint image with the mask image to get the earprint image as a region of interest. Figure 8 depicts apply the region of interest. The gray image, mask image and region of interest are illustrated in Figure 8(a), Figure 8(b), and Figure 8(c), respectively.
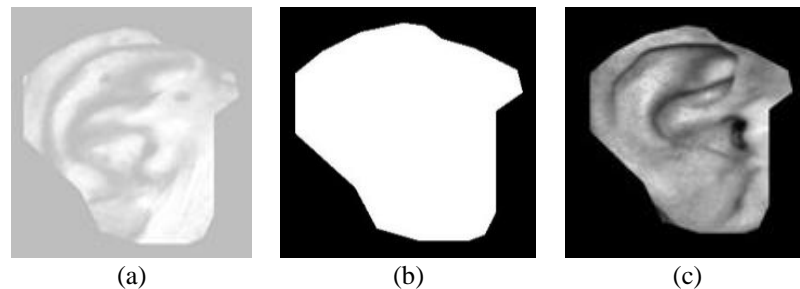

(a) (b) (c)

Figure 8. Show the three images for (a) gray image, (b) mask image, and (c) region of interest

- Histogram equalization (HE) sub-phase: in this sub-phase, after obtaining ROI of the earprint image, it will utilize HE, which is its method to adjust the contrast of the earprint image.

Suppose f be an image represented as an array r x c of integer pixel densities that range between 0 and L-1, where L represents number of gray level values in the image, typically 256. Suppose p represents normalized histogram (NH) of f. Figure 9 shows the HE, where the region of interest and histogram equalization are illustrated in Figure 9(a) and Figure 9(b), respectively. The (2) and (3) apply for HE [27], [28].

$$pn = \frac{\text{No.of pixels with intensity n}}{\text{total No.of pixels}} \quad (2)$$

HE images g will be determined through:

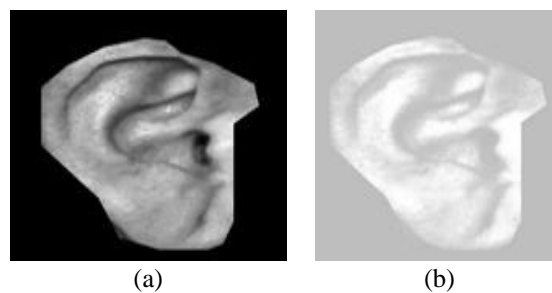$$gi, = ((L - 1)\sum f\ ni =, j0\ pn) \quad (3)$$

(a) (b)

Figure 9. Shows the two images for (a) region of interest and (b) histogram equalization

- Median filter sub-phase: In this sub-phase, after applies of the HE for the earprint image, it will utilize a median filter that reduces the blurring of edges. It may be better de-noising algorithms if they comprise not only noise but spatial characteristics of the image as well [21]. The median filter is a non-linear smoothing approach for reducing-edge blur. The substitution makes it of the current point in the image with a median of brightness in the neighborhood. Individual noise heights don't effectively affect the median brightness in the neighborhood; hence median smoothing efficiently removes impulsive noise [29]. According to performance, the median filter is the best filtering strategy that takes the least computational time. It smooths pepper and salt noises [30], [31]. In proposed scheme, the median filter is applied in the earprint image; note that the earprint image is enhanced better. Figure 10 shows the apply the median filter for the earprint image, where histogram equalization has been illustrated in Figure 10(a) and median filtered image has been illustrated in Figure 10(b).

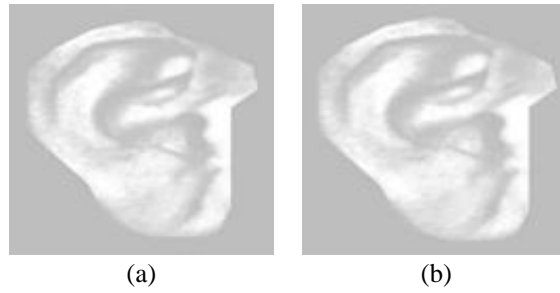(a)   (b)

Figure 10. Shows the two images for (a) histogram equalization and (b) median filtered image

- Local normalization (LN) sub-phase: in this sub-phase, after obtaining a better earprint image by applying the median filter, it will utilize LN, which is an important sub-phase in order to keep each beneficial information in a constant form of illumination. As it becomes easier to feature extraction and recognition for the earprint image, changing the illumination condition certainly leads to lower detection rates and may be removed through illumination normalization, normalization strategies must be well considered in the human earprint recognition scheme [32]–[34]. In the proposed scheme, the LN is applied to the earprint image; note that the better enhancement of the earprint image is obtained, the more useful it is in the next phase of suggested scheme. Figure 11 shows the local normalized earprint image, where the median filtered image has been illustrated in Figure 11(a) and the local normalized image has been illustrated in Figure 11(b).
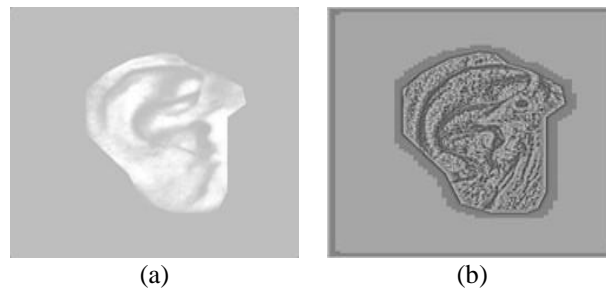

(a)   (b)

Figure 11. Shows the two images for (a) median filtered image and (b) local normalized image

4.3. Feature extraction phase (FEP)

After preprocessing, this phase extracts only important information from the ear print image. The proposed scheme utilizes the fusion vector of DTTF and GDP2 and stores features for all of the classes for each of the samples of ear print images for an individual in training database features (TRDBF). The feature extraction of two methods (DTTF and GDP2) for some samples of local normalized images are shown in Tables 1 and 2.

Table 1. Feature extraction of difference theoretic texture features for some sample

Class.	Difference theoretic texture features										
1	0.267679329208246	0.372102268291585	0.373155528504622	0.039911546489760	0.029621957283130	0.026663214110742	0.249497697705792	0.2844442872457	0.279073813402138	0.322059644119288	0.068702076712597
2	0.268419163783582	0.033778827679176	0.372708138931065	0.034935003203340	0.026600484573518	0.024643955170263	0.249969665913664	0.290122305460061	0.280439000086879	0.323381256381653	0.063740304906973
3	0.298187472560697	0.343838284179761	0.373639885974050	0.038954070065003	0.030740634030288	0.026205142606364	0.269145936394669	0.259168002618974	0.259766608069393	0.282192054496414	0.070587982734770
4	0.264549462696793	0.339970720365522	0.370013238135473	0.040701987286327	0.030585025875934	0.028962081453575	0.249679519633079	0.260958268212029	0.270198343236292	0.293685452361139	0.073906936543466
5	0.300405138847354	0.343813697857417	0.371273048826711	0.036427868934169	0.028321656643313	0.027095756152297	0.250538926150702	0.260046375508462	0.260806660245598	0.282407677486253	0.067668352479131

Table 2. Feature extraction of gradient direction pattern for some sample

Class	Gradient direction pattern							
1	9831	19069	14390	26776	12523	29393	12236	211851
2	10166	18582	13287	28082	13208	27674	12100	212932
3	9581	17175	13730	28218	14019	28576	11236	213420
4	10313	17994	13598	28064	13422	28278	12232	212360
5	10900	18769	12812	27216	13277	27920	11853	213586

4.4. Recognition phase

In this phase, after obtaining the fusion of DTTF and GDP2, it will utilize Gaussian distribution, which it utilizes for the calculation of distance among features for earprint recognition, where the distance utilizes to determine which class it belongs to. The distance by Gaussian distribution is shown in Table 3.

Table 3. Computation of the Gaussian distribution

Class	Gaussian distribution	
1	Man	73.2577105503673
	Variance	6910.98541538104
2	Man	120.149241739909
	Variance	160.732157314637
3	Man	51.6057133992513
	Variance	160.732157314637
4	Man	73.2577105503673
	Variance	89.4119000790942
5	Man	135.065880569067
	Variance	1813.01670162773

The results of the evaluation of the earprint recognition scheme were evaluated utilizing two measures: the recognition rate (RR) and false alarm rate (FAR). The (4) and (5) were used in order to calculate those two measurements [35]–[39]. The best recognition rate was achieved, as shown in Table 4.

$$FAR = \frac{Number\ of\ false\ recognition\ attempts}{Total\ number\ of\ attempts} * 100 \qquad (4)$$

$$RR = \frac{Number\ of\ correct\ attempts}{Total\ number\ of\ attempts} * 100 \qquad (5)$$

Table 4. RR that has been obtained for training and testing for all of the the classes

Recognition Rates		
Criteria of Evaluation		
	FAR	RR
Training	0.0	100%
Testing	14.6%	93.70%

In the proposed scheme, the 3-fold cross-validation of the dataset has been split into three equal portions, two of which have been used for training, and the third one has been used for testing. The comparison to another study which utilized the same UERC database earprint but various methods for recognition utilizing texture and geometric features. It must be noted that, in this work, better results have been obtained through utilizing the fusion of DTTF and GDP2.

Table 5 lists the dataset, preprocessing, feature extraction, methods/techniques, and RR used in previous works. Concerning the comparison, the analysis of the performance is noticed that the recognition ear prints most efficient due to the utilizing a set of sub-phases for pre-processing (convert to grayscale, a ROI, HE, median filter and LN) and because of several advantages for the fusion of DTTF and GDP2. Additionally, this work utilized the dataset of UERC and gave good results with a fusion of DTTF and GDP2.

Table 5. Comparison of the previous approaches

Ref	Year	Dataset	Pre-processing	Feature extraction	Recognition (methods/techinques)	Recognition rate (RR) %
[7]	2019	Ear images created of About 25 people have been collected irrespective of age. Dataset has both left and right images of human ear; other than this, the on-line dataset of human ear is also used.	The original image is resized to [492x702] pixels to be the same as the size of images in the database, the resized images are converted into grayscale images, Contour detection algorithm	SIFT algorithm	Euclidean distance measure approach	
[8]	2018	Top-1, Top-2, Top-4, Top-5 Top-5 dataset	Original image converted to gray image; The mean adjusted image X' is created	PCA	Euclidean distance	
[9]	2017	AWED and CVLED datasets	-	CNNs	CNNs	
[10]	2018	IIT Delhi-I, IIT Delhi-II, UND-E	Gaussian filter operation, CLAHE, Kirsc, ear edge image utilizing the Kirsch edge detection	PHOG, LDP, and PCA	KDA-NN	97.34
[11]	2019	IIT Delhi (IITD), Univ. of Notre Dame collection E (UNDE), collection J2 (UND-J2), AWE, low-resolution camera	localize ear shape, Enhanced Jaya algorithm (EJA).	SURF feature extraction technique	k-NN classifier	
[12]	2019	IITDelhi database	histogram standardization, ROI has been discovered with the use of the general protest discovery system	BRISK and SURF techniques	matching registration and similarity score process)	80.21%
[13]	2011	USTB ear image database and IIT Delhi ear image	Cropping image, normalization of cropped ear images, converted into grayscale, and contrast enhancement of gray-scale image	Haar wavelet transform (HWT)	fast normalized cross-correlation (NCC)	97.2% and 95.2%
[14]	2016	USTB ear	cropping, resizing, contrast maximization, and image means subtraction.	forward propagation and error back propagation.	deep convolutional neural network (DCNN)	98.27%
[15]	2018				DCNN	95% and up to 98%
[16]	2019	IITD II dataset	HE, wavelet decomposition (WD)	multiple-image generation (MIG), PCA	multi-band PCA (2DWMBPCA), Eigenvector	
[17]	2018	AMI ear database	laplacian filter (LF)	LBP, splitting ear images, histogram extraction	geometric and texture feature fusion	80%
[18]	2018	IIT Delhi-1 and IIT Delhi-2		local dense histograms (LDH)	dense local phase quantization (DLPQ)	
[19]	2018	WPUT and NCKU datasets	converted to gray image, HE, ear detection (ED) is using Haar feature-based cascade (HFBC)		shape context computation	100% for non-occluded images, 57% for occluded by both hair and earring
[20]	2020	AWE ear dataset	ROI segmentation, GAN	LBP, LPQ, BSIF	LBP, LPQ, and BSIF	

5. CONCLUSION

In this paper, a method for recognizing the human ear for biometrics has been proposed and implemented. Images of the human ear from the UERC database are used. Around 330 images were collected, of which 80% have been utilized for training and the rest 20% have been utilized for testing purposes. Initially, the ear is segmented from the input image with the use of the logical indexing method. The fusion vector of DTTF and GDP2 has been applied to the segmented ear image so as to extract the fusion features. The extracted features generate the model file from the training images. The test images carry out the same process, and for recognition of the ear, the distance by Gaussian distribution was employed to calculate the distance among

features for ear print recognition. Thus, the specific person's ear has been identified using the value of this distance measure. Further research will be on minimizing the feature values for better recognition.

ACKNOWLEDGEMENTS

REFERENCES

[1] V. S. Amritha and J. Aravinth, "Matcher performance-based score level fusion schemes for multi-modal biometric authentication system," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2020, pp. 79–85, doi: 10.1109/ICACCS48705.2020.9074446.
[2] W. Kabir, M. O. Ahmad, and M. N. S. Swamy, "A multi-biometric system based on feature and score level fusions," *IEEE Access*, vol. 7, pp. 59437–59450, 2019, doi: 10.1109/ACCESS.2019.2914992.
[3] A. Ross, K. Nandakumar, and A. K. Jain, "Introduction to multibiometrics," in *Handbook of Biometrics*, Boston, MA: Springer US, pp. 271–292.
[4] M. Arunachalam, "The human ear recognition based on phase-based matching algorithm," *3C Tecnología_Glosas de innovación aplicadas a la pyme*, pp. 141–157, Mar. 2020, doi: 10.17993/3ctecno.2020.specialissue4.141-157.
[5] G. N. Rutty, A. Abbas, and D. Crossling, "Could earprint identification be computerised? An illustrated proof of concept paper," *International Journal of Legal Medicine*, vol. 119, no. 6, pp. 335–343, Nov. 2005, doi: 10.1007/s00414-005-0527-y.
[6] A. Alemran and B. Rahmatullah, "Novel hybrid ear recognition framework in passive human identification," *International Journal of Academic Research in Business and Social Sciences*, vol. 9, no. 14, Oct. 2019, doi: 10.6007/IJARBSS/v9-i14/6505.
[7] N. Mangayarkarasi, G. Raghuraman, and A. Nasreen, "Contour detection based ear recognition for biometric applications," *Procedia Computer Science*, vol. 165, pp. 751–758, 2019, doi: 10.1016/j.procs.2020.01.016.
[8] M. Zarachoff, A. Sheikh-Akbari, and D. Monekosso, "Application of single image super-resolution in human ear recognition using eigenvalues," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, Oct. 2018, pp. 1–6, doi: 10.1109/IST.2018.8577134.
[9] Z. Emersič, D. Stepec, V. Struc, and P. Peer, "Training convolutional neural networks with limited training data for ear recognition in the wild," *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, IEEE, 2017, doi: 10.1109/fg.2017.123.
[10] P. P. Sarangi, B. S. P. Mishra, and S. Dehuri, "Fusion of PHOG and LDP local descriptors for kernel-based ear biometric recognition," *Multimedia Tools and Applications*, vol. 78, no. 8, pp. 9595–9623, Apr. 2019, doi: 10.1007/s11042-018-6489-0.
[11] P. P. Sarangi, B. S. P. Mishra, S. Dehuri, and S.-B. Cho, "An evaluation of ear biometric system based on enhanced Jaya algorithm and SURF descriptors," *Evolutionary Intelligence*, vol. 13, no. 3, pp. 443–461, Sep. 2020, doi: 10.1007/s12065-019-00311-9.
[12] T. Thivakaran, S. Padira, A. Kumar, and S. Reddy, "Fusion based multimodel biometric authentication system using ear and fingerprint," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 1, pp. 62–73, Feb. 2019, doi: 10.22266/ijies2019.0228.07.
[13] A. Tariq, M. A. Anjum, and M. U. Akram, "Personal identification using computerized human ear recognition system," in *Proceedings of 2011 International Conference on Computer Science and Network Technology*, Dec. 2011, pp. 50–54, doi: 10.1109/ICCSNT.2011.6181906.
[14] L. Tian and Z. Mu, "Ear recognition based on deep convolutional network," in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Oct. 2016, pp. 437–441, doi: 10.1109/CISP-BMEI.2016.7852751.
[15] T. Ying, W. Shining, and L. Wanxiang, "Human ear recognition based on deep convolutional neural network," in *2018 Chinese Control And Decision Conference (CCDC)*, Jun. 2018, pp. 1830–1835, doi: 10.1109/CCDC.2018.8407424.
[16] M. Zarachoff, A. Sheikh-Akbari, and D. Monekosso, "Single image ear recognition using wavelet-based multi-band PCA," in *2019 27th European Signal Processing Conference (EUSIPCO)*, Sep. 2019, pp. 1–4, doi: 10.23919/EUSIPCO.2019.8903090.
[17] S. M. Jiddah and K. Yurtkan, "Fusion of geometric and texture features for ear recognition," in *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Oct. 2018, pp. 1–5, doi: 10.1109/ISMSIT.2018.8567044.
[18] M. Mahmoud Al Rahhal, M. L. Mekhalfi, M. Guermoui, E. Othman, B. Lei, and A. Mahmood, "A dense phase descriptor for human ear recognition," *IEEE Access*, vol. 6, pp. 11883–11887, 2018, doi: 10.1109/ACCESS.2018.2810339.
[19] R. N. Othman, F. Alizadeh, and A. Sutherland, "A novel approach for occluded ear recognition based on shape context," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, Oct. 2018, pp. 93–98, doi: 10.1109/ICOASE.2018.8548856.
[20] Y. Khaldi and A. Benzaoui, "Region of interest synthesis using image-to-image translation for ear recognition," in *2020 International Conference on Advanced Aspects of Software Engineering (ICAASE)*, Nov. 2020, pp. 1–6, doi: 10.1109/ICAASE51408.2020.9380127.
[21] S. Susan and M. Hanmandlu, "Difference theoretic feature set for scale-, illumination- and rotation-invariant texture classification," *IET Image Processing*, vol. 7, no. 8, pp. 725–732, Nov. 2013, doi: 10.1049/iet-ipr.2012.0527.
[22] S. Susan and M. Hanmandlu, "Color texture recognition by color information fusion using the non-extensive entropy," *Multidimensional Systems and Signal Processing*, vol. 29, no. 4, pp. 1269–1284, Oct. 2018, doi: 10.1007/s11045-017-0502-z.
[23] S. Susan and R. Chakre, "3D-difference theoretic texture features for dynamic face recognition," in *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, Mar. 2016, pp. 227–232, doi: 10.1109/ICCTICT.2016.7514583.
[24] M. shahidul Islam, "Gender classification using gradient direction pattern," *Computational Science and Its Applications -- ICCSA 2015*, vol. 25, no. 4, pp. 797–799, Oct. 2013, doi: 10.48550/arXiv.1310.6808.
[25] B. Oluleye, "Zernike moments and genetic algorithm: tutorial and application," *British Journal of Mathematics & Computer Science*, vol. 4, no. 15, pp. 2217–2236, Jan. 2014, doi: 10.9734/BJMCS/2014/10931.
[26] O. A. Carmona Cortes and J. C. da Silva, "Unconstrained numerical optimization using real-coded genetic algorithms: a study case using benchmark functions in R from Scratch," *Revista Brasileira de Computação Aplicada*, vol. 11, no. 3, pp. 1–11, Sep. 2019, doi: 10.5335/rbca.v11i3.9047.

[27]   P. J. Acklam, "MATLAB array manipulation tips and tricks." 18th Oct. 2003, [Online]. Available: http://home.online.no/~pjacklam (accessed Sep. 11, 2022).

[28]   M. S. Khalefa, Z. A. Abduljabar, and Huda Ameer Zeki, "Fingerprint image enhancement by develop Mehtre technique," *Advanced Computing: An International Journal*, vol. 2, no. 6, pp. 171–182, Nov. 2011, doi: 10.5121/acij.2011.2616.

[29]   E. Chandra and K. Kanagalakshmi, "Noise elimination in fingerprint image using median filter," *International Journal of Advanced Networking and Applications*, vol. 955, no. 06, pp. 950–955, 2011.

[30]   K. Kanagalakshmi and E. Chandra, "Performance evaluation of filters in noise removal of fingerprint image," in *2011 3rd International Conference on Electronics Computer Technology*, Apr. 2011, pp. 117–121, doi: 10.1109/ICECTECH.2011.5941572.

[31]   E. Chandra and K. Kanagalakshmi, "Noise suppression scheme using median filter in gray and binary images," *International Journal of Computer Applications*, vol. 26, no. 1, pp. 49–57, Jul. 2011, doi: 10.5120/3064-4188.

[32]   Y. Tie and L. Guan, "Automatic face detection in video sequences using local normalization and optimal adaptive correlation techniques," *Pattern Recognition*, vol. 42, no. 9, pp. 1859–1868, Sep. 2009, doi: 10.1016/j.patcog.2008.11.026.

[33]   I. H. Witten, E. Frank, and M. A. Hall, *Data mining: practical machine learning tools and techniques*, 3rd ed. Elsevier, 2011.

[34]   O. Sudana, I. W. Gunaya, and I. K. G. Darma Putra, "Handwriting identification using deep convolutional neural network method," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 4, pp. 1934-1941, Aug. 2020, doi: 10.12928/telkomnika.v18i4.14864.

[35]   H. Salman Chyad, R. Ali Mustafa, and K. Thabt Saleh, "Study and implementation of resource allocation algorithms in cloud computing," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 591–594, 2018, [Online]. Available: www.sciencepubco.com/index.php/IJET.

[36]   H. S. Chyad, R. A. Mustafa, and R. A. Haleot, "An effective protocol and algorithmic approach for disaster management using wireless sensor networks," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 13 Special Issue, pp. 2405–2410, 2018.

[37]   H. S. Chyad, R. A. Mustafa, and D. N. George, "Cloud resources modelling using smart cloud management," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 11, no. 2, pp. 1134–1142, Apr. 2022, doi: 10.11591/eei.v11i2.3286.

[38]   R. A. Mustafa, H. S. Chyad, and J. R. Mutar, "Enhancement in privacy preservation in cloud computing using apriori algorithm," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 3, pp. 1747–1757, Jun. 2022, doi: 10.11591/ijeecs.v26.i3.pp1747-1757.

[39]   R. A. Mustafa, H. S. Chyad, and R. A. Haleot, "Palm print recognition based on harmony search algorithm," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 5, pp. 4113–4124, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4113-4124.

BIOGRAPHIES OF AUTHORS

Kawther Thabt Saleh received a B.Sc. in Computer Science from Mustansiriyah University - Iraq (2005), and M.Sc. in Computer Science from Al-Balqa' Applied University-Jordan (2011). She has nearly 16 years of experience in teaching in the Computer Science Department, College of Education, Mustansiriyah University. She has published many papers in international and national journals and conferences. Her fields of interest include image processing, pattern recognition, and artificial intelligence. She can be contacted at email: kawtherthabt@uomustansiriyah.edu.iq.

Raniah Ali Mustafa she was born in Baghdad, Iraq, in 1983. She received her M. Sc. Degree in computer science from Mustansiriyah University, Baghdad, Iraq, in 2016. she joined a programmer position in the laboratories of the Department of Computer Science at Al-Mustansiriya University for the period from 2005 to 2013 and after obtaining a master's degree, she became a teacher in the Department of Computer Science. She can be contacted at email: rania83computer@uomustansiriyah.edu.iq.

Haitham Salman Chyad He was born in Baghdad, Iraq, in 1979. He received her M. Sc. Degree in computer system from Osmania University, India, Hyderabad, in 2016. He joined a programmer position in the laboratories of the Department of Computer Science at Al-Mustansiriya University for the period from 2007 to 2013 and manager of the division of electronic calculator of the College of Education at Al-Mustansiriya University and in addition to teaching in the Department of Computer Science. He can be contacted at email: dr.haitham@uomustansiriyah.edu.iq.