Date of publication SEP-30, 2023, date of current version FEB-20, 2023.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.22.3.3236

Human Recognition based on Multi-instance Ear Scheme

INASS SH. HUSSEIN[1], NILAM NUR AMIR SJARIF[2]
[1]Middle technical university, Baquba technical institute, Baquba, Iraq
[2]Faculty Razak of Technology and Informatics, UTM, Kuala Lumpur, Malaysia

Corresponding author: Inass Sh. Hussein (inasshussin@mtu.edu.iq).

ABSTRACT Ear biometrics is one of the primary biometrics that is definitely standing out. Ear recognition enjoys special benefits and can make distinguishing proof safer and dependable along with other biometrics (for example fingerprints and face). Particularly as a supplement to face recognition schemes that experience issues in genuine circumstances. This is because of the extraordinary variety of a planar representation of a confusing object that is varied in shapes, illumination, and profile shape. This study is an endeavor to conquer these restrictions, by proposing scale-invariant feature transform (SIFT) calculation to extract feature vector descriptors from both left and right ears which is to be melded as one descriptor utilized for verification purposes. Likewise, another plan is proposed for the recognition stage, based on a genetic algorithm-backpropagation neural network as an accurate recognition approach. This approach will be tried by utilizing images from the Indian Institute of Technology Delhi's creation (IITD). The suggested system exhibits a 99.7% accuracy recognition rate.

KEYWORDS ear recognition, local features, SIFT, feature selection, GA, BP.

I. INTRODUCTION

IMPROVEMENTS in data innovation have assisted people with performing undertakings that were beforehand troublesome or physically did. By implication, these human advancements have helped by giving the quicker, more effective, and more useful execution of errands. There are various issues that make humankind use data innovation to obtain results that are more precise. The subject of safety is one issue looked at by all nations of the world. Biometric schemes are taken on as a more compelling answer for security infringement.

The investigation of biometrics depends on the physical or behavioural features of a person to confirm their character. Highlights like fingerprints, faces, and iris certainly stand out for quite a while. Colleagues consider fingerprints and iris to be more exact in biometric examination than the face, yet the face has different characteristics like being effectively gotten in genuine circumstances without client communication. Be that as it may, the face without anyone else isn't as adaptable as it ought to be because of brightening and demeanor changes [1]. On the other hand, ear recognition is a sort of new biometrics, the hypothesis and examination of ear recognition excite additional consideration from homegrown and unfamiliar researchers. Different specialists have confirmed that the ears are for sure exceptional enough to recognize an individual and they could have practical use as biometric features [2].

The human ear also differs greatly across individuals in terms of bending and geometric dimensions. Thus, ear structure is a promising characteristic. Although the idea of using the ear as a biometric has been known since the 1890s, there are currently no frameworks that are financially feasible to use ear data to identify persons. [3]. Ear images can be gotten with a similar methodology as the face, this situation proposes that it very well may be utilized as a supplement in a recognition framework.

Typically, ear recognition involves four stages.: detection, preprocessing, extraction, and recognition. Nonetheless, Additionally, some ear identification techniques require detection, and preprocessing, involve extraction, or even fall short of the previously mentioned steps. Prior to now, important methods of registration included facial and distinct fingerprint identification. However, despite advancements in science and technology, face and fingerprint identification have not been able to satisfy all ID requirements for the main reasons [3] listed below:

(1) Recognition rate is essentially diminished in the event that there is a significant obstruction to an image of a face.

(2) A change in facial expression is present because of terrifying or being harmed by a mishap bringing about erroneous ID.

(3) Even if it is possible to obtain distinct fingerprint images with a smartphone camera without the subject's consent, most of the time, distinguishing proof work cannot be done if the person doesn't provide a variety of fingerprints.

(4) Fingerprints are difficult to detect and take more time to scan in an overly damp or overly dry climate.

In light of the deficiencies referenced above, to accomplish precise recognizable evidence of ear recognition, a novel biometric because of the accompanying benefits:

(1) Accept before the age of eight and after the age of seventy, ear shape won't alter considerably throughout time.

(2) Ear shape doesn't change strongly when the face expression changes.

(3) Only simple instruments like a cell phone and camera can obtain ear information, without the need for subject coordination and in an unrestricted environment.

(4) Face and ear can be procured all the while, establishing the groundwork for face and ear multimodal recognition.

(5) By using infrared methods, the condition where the ear is obstructed by hair can be resolved.

Nonetheless, ear biometrics experiences issues in genuine circumstances. This is because of the extraordinary a planar representation of a confusing object with a variety of shapes, shifting illumination, and shifting profile shape [3].

Thusly, this study plans to fill these holes by offering an original approach that conquers the previously mentioned downsides through giving satisfactory decrease and enhancement execution improvement and working with the ear recognition. In addition, this article is to frame all significant ear recognition calculations and direct execution assessment. On the other hand, by applying a 2D entropy function for the binarization image and the maximum entropy as an adaptive threshold, this paper contributes to the segmentation of the ear image stage and designs a novel feature selection method employing backpropagation neural network with a genetic algorithm to aid in the recognition step (GABP).

The remainder of the article is described as joining in. A related work is introduced in Section 2. Section 3 introduces the material and the methodology. The proposed ear arrangement is explained in Section 4. Section 5 introduces the validation of the ear recognition technique, and Section 6 provides the article's conclusion.

II. RELATED WORK

Both geometric and textural aspects were used in [6]. In terms of its purposes and objectives, this analysis successfully investigates the AMI ear dataset by extracting features using the local binary pattern algorithm (LBP) and furthermore applying the Laplacian on the rough images separately to extract the geometric elements. In order to find regions of interest in human ear images, the ear dataset was processed by cutting the ear images into four quarters and investigating each one separately. Then, geometric and texture features are integrated, and tests are run to ensure that the combined pieces are reliable.

In work [7] LBP was used to take into account the ear recognition problem. Whereas, the LBP-like features define the spatial design image texture under the assumption that this texture contains two geographically correlative views and an example. The LBP-like features are appropriate for the ear recognition problem due to their high discriminative strength, invariance to monotonic gray scale changes, and computational productivity attributes. In order to determine how they might be employed for ear recognition, a few ongoing LBP versions provided in the literature as component extraction approaches are investigated. AMI, WPUT, AWE, IIT Delhi (I), and IIT Delhi (II) are five publicly available, mandatory, and unrestricted benchmark ear datasets pushing diverse imaging circumstances that are extensively used in the ear recognition proof and check tests. Nearly 100% of people recognized this work, but as the degree of distortion increases, it becomes more difficult to present.

A framework for ear recognition that uses a convolutional neural network (CNN) was proposed in [8] to identify a person from a reference image. When tested against clear photographs, the suggested technique performs similarly to other traditional methodologies. The researchers also outline a technique for enhancing the sliding window approach in order to increase the speed of a CNN applied to large information images.

On the other hand, authors in [9] proposed to extract the global NGTDM features as well as the local Haralick features. A final feature vector made up of these features is combined and given to SVM for recognition. IIT Delhi (I) and IIT Delhi (II) are two publicly accessible ear datasets on which ear recognition is briefly implemented. The suggested system has a 98.4% overall accuracy rate and an error rate of 1.825.

While in work[10] formulated by combining the concepts of random forest (RF) and histograms of oriented gradients (HOG). From images of ears, features are extracted using the HOG. The Indian Institute of Technology in Delhi provided the chosen ear images. The results of the experiments demonstrate that the accuracy of the suggested ear identification system is up to 99.69%.

In addition authors in [11] suggested image improvement and segmentation preprocessing. The feature vector is created by concatenating the same estimated statistical features from the sub-images at various levels with the mean, variance, and entropy for each block of the coarse image. Then, for ear recognition, features are taken from each network and supplied into a shallow classifier. For feature reduction, PCA is employed. First, an end-to-end test is conducted using three pre-trained nets: ResNet50, GoogleNet, and AlexNet. On two publicly available ear datasets, AMI and IIT Delhi (II), ear recognition is briefly used. The best outcomes for the IIT Delhi dataset were, respectively, 93.57 and 94.29 using ResNet50 and AlexNet.

III. MATERIAL AND METHOD

In general, ear recognition systems follow a number of rules, such as data collection, pre-processing, feature extraction, recognition of the specific object, and feature selection as an additional step. Fig. 1 shows a stream diagram of typical ear recognition. That will be completely explained in the next parts.

Figure 1. Typical ear recognition

A. FEATURE EXTRACTION

The components of an image can be divided into two categories: global features and local features. Global aspects continue to be as fascinating as they were ten years ago. "Global," mean concentrate the key elements from the entire image while taking each pixel into account [12]. Compressing the size of the representation and accelerating computation speed are two benefits of global. The drawbacks, meanwhile, are related to geometric distortion. As a result, relying solely on global features is ineffective. Particularly in the fields of pattern recognition and computer vision, local feature is thought to be the most effective representation of the image. The local features are not affected by changes in size, position, or rotation [13].

A few feature extraction algorithms have been trained to recognize inborn and discriminative local structures from images because feature extraction is the core of any recognition system. Texture descriptors, which encode and characterize the texture data of images, are one of these methods. The (LBP) addresses a unique texture descriptor that has demonstrated success in a number of PC vision applications, including object and face detection [14]. The LBP descriptor was used in [15] to extract ear features for evaluation using ear photos for 125 subjects from the IIT Delhi ear dataset. In contrast to principal component analysis, the authors found that LBP has excellent discriminative power, resistance to changes in global illumination, and minimal processing requirements (PCA). The recognition rate for this work is approximately 93%. The LBP descriptor and similarity (LSBP) were combined by Guo and Xu [16] to include extra information about the network of nearby pixels. On the USTB ear dataset, they found an identification rate of around 93%. In contrast, Benzaoui et al. [17] used an elliptical (ELBP) descriptor, a replacement for the fundamental LBP, to describe the finer details of ear pictures. The Discrete Wavelet Transform (DWT) is applied to each global histogram in order to reduce dimensionality and choose useful data. The IITD ear dataset, which contains 500 images for 100 people, is used to evaluate the identification performance, and the obtained recognition rate is close to 94%.

Additionally, some techniques depend on extracting features from 3D ear images [18] however, under uncontrolled imaging settings including scaling, rotation, different types of illumination, noise, and obstruction, these approaches fail [19].

In light of several local descriptors like LBP, local stage quantization, and binarized measurable image characteristics, Benzaoui et al. [20] provided a methodology for human verification evidence from his or her 2D ear image. The outcomes of local texture features in ear verification were confirmed overall by the results on the IIT Delhi-I, IIT Delhi-II, and USTB ear datasets.

On the other hand, ear biometrics methods based on local features are also suggested. These methods have proven successful in other biometric applications. These methods' resistance to partial occlusion and various affine transformations stand out as advantages [14].

Anwar, Ghany et al. [21] utilized SIFT, which initially recognizes a number of key points in images before determining a free descriptor for each key point, to extract the local features of the ear. Using two datasets, such as IIT Delhi and the AMI. Authors implemented a SIFT algorithm to recognize people and distinguish the ear picture. To obtain the Gaussian difference between scales and evaluate each pixel in DoG pictures, SIFT Extraction is used.

On the other hand, Ghoualmi et al. [22] used SIFT to extract features. The main steps are as follows: Identify keypoints, assign a direction course for each central issue, determine scale space boundaries, and represent central concerns. Additionally, SIFT has proven to be incredibly beneficial.

Due to their computational simplicity, procedures based on intensive descriptor calculation are proven to have good recognition performance. SIFT calculation is one such precise methodology utilized to extract local features [23]. We decided to apply the SIFT to the ear recognition problem because of the SIFT's advancement in these applications. The section that follows describes SIFT features.

B. SIFT LOCAL FEATURES

The SIFT idea was presented in 1999 by David Lowe, at the Computer Vision conference and settled in 2004 SIFT was expressed that is a strong and reliable image descriptor and detector [24]. The proposed work has been focused on SIFT because it has evidence to be powerful in accomplishing many challenging issues, especially with respect to the recognition of an object, number of the object including distortions such as detection of the scale, rotation, and partial obstruction in highly cluttered environments [25].

SIFT is a descriptive method of features that are strong and invariable scales. It has been widely used in the biometric system and used in matching, classification, and sewing images and used in many other fields [24].

This method's generation of several features that cover the image extensively over a range of scales and places is a key component. A typical 500x500 pixel image will have around 2000 stable features (although this number depends on both the content of the image and the options of the various parameters). The SIFT features are kept in a database after the extraction phase. By individually comparing each attribute of the new image with one that is kept in the database and determining the features of the candidates that match, the new image is matched [26].

On the other hand, the utilization of each feature in recognition is not definitely a good idea and gives the opportunity to increase the rating of errors. Consequently, choosing the right features from the full features is a good solution. Feature reduction or feature selection method are the common names for it. In the following section, we will make sense of the exhaustive feature selection method.

C. FEATURES SELECTION

By focusing on a select few crucial features that can deliver a respectable performance for the classification, the features selection process seeks to reduce the size of the data. The methods for feature selection will get rid of the smallest feature discriminator and leave some of the original features so that there is still enough information to tell one category from another. Filtering methods and wrapping methods are two categories into which function reduction algorithms can be divided. In order to analyze and choose subsets of features, the filtering techniques concentrate on the overall qualities of the data, excluding the chosen learning algorithm or classifier [27].

Principal component analysis (PCA), a well-known approach filtering technique, is being used by numerous researchers for the feature reduction stage of their studies [28]. The primary drawback of the PCA is that it does not work in conjunction with the classifier and does not significantly reduce the size of the dataset. Instead, it looks for features where the value of their sample has a higher variance than other features [29]. In contrast, wrapper approaches benefit from the selected subset's predictive performance being connected with a measure of relevance like a genetic algorithm (GA). GAs are a subset of evolutionary algorithms (EA), a class of methods with biological roots that employ a variety of strategies to mimic natural evolution. Problems involving feature subset optimization have been effectively solved using GAs [30].

In this manner, a genetic algorithm has been utilized to get rid of the PCA restrictions. A GA-based method takes into account the classifier's error rate of misclassification in its fitness function and attempts to decrease it. The section that follows contains GA's in-depth presentation.

D. GENETIC ALGORITHM

GAs are a component of the artificial intelligence evaluation techniques that are rapidly expanding. Evolutionary theory is the foundation of genetic algorithms. In other words, the issues that genetic algorithms have developed to answer. The algorithm has been calculated using a variety of biological techniques, such as mutation, crossover, and coding [31].

Finding a set of initial attributes that reduce f is the goal of the GA-based technique. Because researchers can alter GA's functional configuration to further enhance their results, it is well recognized that GA is a very adaptable and effective selection approach. As a result, the number of functions will be minimized by using a selection of GA-based features (a subspace or multiple projection approaches) [31].

The following are the G.A. methodology's means:

1. Represents the domain of problem variables such as fixed chromosome length, the crossover probability (pc), the size of a population of (N) chromosomes, and the probability of mutation (*pm)*;
2. Create a fitness function F(x) to assess an individual chromosome's performance or fitness inside the problem's domain;
3. Use randomness to create an initial population of *N* chromosomes with the values *x1, x2*,..., and *xn*;
4. Determine each chromosome's F(x) fitness, including *F(x1), F(x2),..., and F(xn)*;
5. Choose two chromosomes that came from the current population. A likelihood connected to their physical characteristics is used to choose the paternal chromosomes. Fewer chromosomes in shape are less likely to be selected for mating than high-fit chromosomes. Use the genetic operators crossover and mutation to produce a pair of offspring chromosomes;
6. The chromosomes of the newly born offspring are inserted into the new population;
7. Perform Step 5 while the initial population N's size is equal to the size of the new population's chromosome;
8. the new (offspring) population replaces the original (parent) population's chromosomes;
9. Continue performing step 4 until the termination criterion is complete.

E. RECOGNITION

Ear recognition methods can be partitioned into three, Statistics, Sparse Representation, and Neural Networks. In order to recognize ears, Kondappan et al.[2] used neural networks. The authors created a multi-layer feedforward network with unweighted linkages between nodes by combining ANN with fuzzy classifiers. On the other hand, backpropagation algorithm neural network (BP) is commonly used as a classifier in the recognition phase and provides strong findings based on ear recognition [3, 8]. As a result, this proposed work has employed BP to be integrated with a GA for both the selection and recognition phases in order to boost the accuracy rate of categorization. The following section fully explains how the BP algorithm functions.

F. BACKPROPAGATION NEURAL NETWORK

Backpropagation consists primarily of the gradient's descent; the gradient is a displacement along the gradient and a derivation vector [32]. In the main phase of the BP learning algorithm neural network, the signal is fed forward from the input layer until it reaches the output layer through the hidden layer. The signal back probes in the second phase from the output until it reaches the input through the concealed layer. Here, the weight of the neurons will vary based on their mistakes.

These are the mathematical formulas of the BP steps:

1. Compute the net weight:

$$x = \sum_{i=1}^{n} x_i \ w_i^- \tag{1}$$

where n = no. of input neurons, θ = threshold, x= input, w= weight, i= no. of layer

2. Proceed to the activation function:

$$Y^{sigmoied} = \frac{1}{1+e^{-x}} \tag{2}$$

3. Let:

i= no. of neuron in the input layer.
j= no. of the neuron in the hidden layer
K= no. of the neuron in output layer
Input signals = $x1$,$x2$……,xn
Error signals = $e1$,$e2$,……,en
Wij denotes the weight between i and j
Wjk denotes the weight between j and k

The definition of propagatinge error signal at the output k at iteration p is:

$$\boldsymbol{ek}(P) = y_{d,k}(p) - y_k(p) \tag{3}$$

where $y_{d,k}$ (p) is the desired output of neuron k is located into the output.

To update weight wjk as:

$$wjk(p+1) = \omega jk(p) + \Delta\omega jk(p) \tag{4}$$

where

$$\Delta\, wjk(p) \ = \ \alpha \ xy_j \ (p) \ * \ \delta_k \ (p) \tag{5}$$

where Δ wjk (p) is the weight correction, α=error gradient neuron k

IV. PROPOSED EAR RECOGNITION SCHEME

The proposed ear scheme has two main steps. First pre-processing step includes image acquisition and region of interest extraction (ROI) of ear images. Second, recognition step includes feature extraction, feature selection, and feature recognition. The proposed scheme steps are shown in Fig 2. The preprocessing step was carried out as follows:

1. Images of the left and right ears are acquired from the dataset;

2. Use binerzation to segment the ear images and identify the ROI by utilizing the 2D-entropy function and dilation morphology.

Implementation of the recognition step is as follows:

1. Using SIFT, to extract features from ROI from both the left and right ear images and integrate them into a single feature vector.
2. To choose the optimal feature combination, the GA algorithm is fed a features vector.
3. Based on the values of the fitness function, the probability of selecting each feature is determined.
4. Finally, the BP evaluates the selected features.

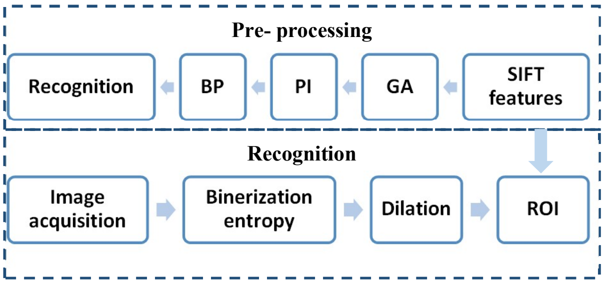

Figure 2. Proposed scheme for ear recognition

A. DATASET

The majority of the hands in this touchless ear image library were taken from IIT Delhi students and employees in India. Using a basic photographic setup, this database was obtained on the IIT Delhi campus between October 2006 and June 2007. Using a basic imaging setup, all of the photos are taken at a distance (touchless), and the imaging is done inside. The currently available dataset was collected from 121 different people, and each subject had a minimum of three photos of their ears. The respondents in the dataset are all between the ages of 14 and 58. With a number ID/number, the database of 471 images has been successively numbered for each user. a sampling of the IITD's ear images. With a number ID/number, the database of 471 images has been used to evaluate the proposed scheme. Fig. 3 displays an example of an ear image from the IITD dataset.

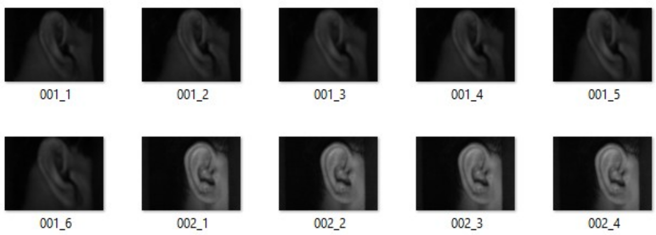

Figure 3. IITD ear images samples

B. EAR PREPROCESSING AND EXTRACT REGION OF INTEREST

All stages of image processing and analysis begin with ear preprocessing. In essence, the collected image from the dataset serves as the foundation for our research. Actions taken after the image was taken: first, convert the image to grayscale. Even in grayscale, occasionally, images are not clear because undesired areas can affect them or noise can be removed. Second, removing noise while preserving the edge details of the ear image with the median filter [33]. Then, using the Entropy function, the grayscale image is transformed into a binary image [1]. Finding the edges is supported by binary images. The edges can be used to calculate an outer line. The edges of the ear image are calculated using the dilation morphology [1]. The steps for ear ROI extraction are summarized in Fig. 4.

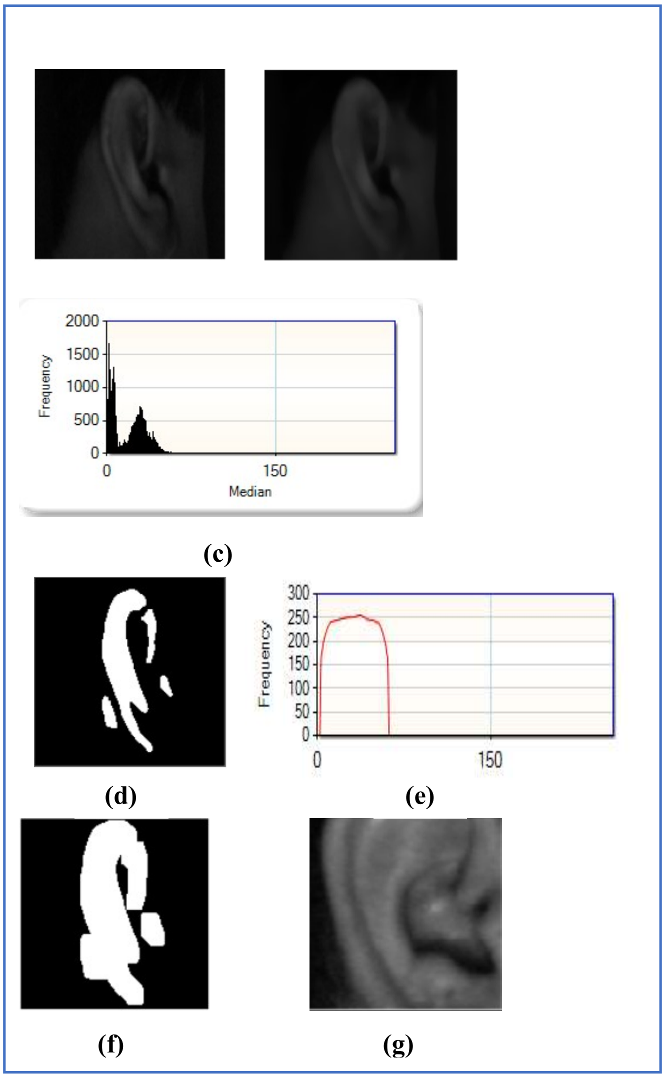

Figure 4. Preprocessing steps (a) original image; (b) image with median filter; (c) historam of the image; (d) binary image with entropy; (e) entropy histogram; (f) image with dilation; (g) ROI

C. EAR LOCAL FEATURES EXTRACTION

Each pixel in the image is processed by SIFT features with 128 descriptors depending on its characteristics, position, and correlations with nearby pixels in two distinct instance ear images. By using the technique described in Section SIFT features, SIFT features are extracted from the ROI ear to generate the feature matrix [i j]. For each image, a feature vector of 382 elements is used as a description. These properties are those of the input ear image. Utilizing feature-level fusion, the feature vectors are fused into one long vector via concatenation [27]. The result of a possible combination of these features is 3822 features. These features are passed to GA to find an appropriate combination of these features.

D. EAR FEATURES SELECTION

In order to select features, GA is applied from the feature matrix [i j]. A vector containing values of zero and one, referred to as the chromosome, is the result of the genetic algorithm calculation. If the value is "one," then the associated feature is selected, and if the value is "zero," then this feature is not selected. Different binary chromosomes are randomly created to create the initial population. The number of features that can be extracted equals the length of the chromosome. The fitness function measures how well a characteristic can be recognized using the matching chromosome and the features that are chosen. More ideal chromosomes are chosen to use GA operators to create a new population from the current population. The fitness function is used to calculate each chromosome's optimality. The selecting procedure has made use of the recognition accuracy function technique. The probability of choosing each chromosome is calculated based on its probability using the values of the fitness function. This probability is calculated using the formula below:

$$\boldsymbol{p_i} = \frac{f_i}{\Sigma_{i=1}^{n} f_i} \tag{6}$$

Where *fi* is the fitness function of chromosome *i, n* is the total number of chromosomes, and *pi* is the probability of selecting chromosome *i.* The greater *pi*, the greater the probability that (chromosome *i*) will be selected during the creation of a new population.

E. Ear Recognition using BP Neural Network

Subsequent selection features based on their probability function are sent to the BP in order to identify the user. To this end, the error rate between the input sample with all training samples is determined, and the training sample that has the smallest error will be selected as the identity of the individual. Datasets are separated into two categories: training and testing sets based on K-fold. In all experiments, the k value was selected based on the minimum number of existing images for each individual.

V. EAR RECOGNITION SCHEME VALIDATION

The experimental findings and performance assessment of the suggested method are presented in this section. Accuracy rate and receiver operating characteristic (ROC) curve were employed as metrics in the evaluation standard, and a confusion matrix was used for validation. By separating the dataset into a training set and a testing set according to K-fold, we were able to assess the proposed scheme. The analyses are implemented using Microsoft Visual Studio 2013 (Visual Studio C#). Fig. 5 shows the effect of increasing the number of epochs on the accuracy rate of ratio recognition. Figure 5 shows that when the number of epochs rises, performance rises as well. 7 epochs have thus been taken into account in the final implementation (according to our experiment, after 7 epochs, the rate is fixed). The best accuracy score is 99.7%.

Additionally, Fig. 6 ROC curves of the characteristics show the results and show how accurate the results are at recognizing objects. Finally, on the same dataset (IITD), the performance of the proposed scheme is compared with earlier techniques described in the literature. The findings are shown in Table 1.

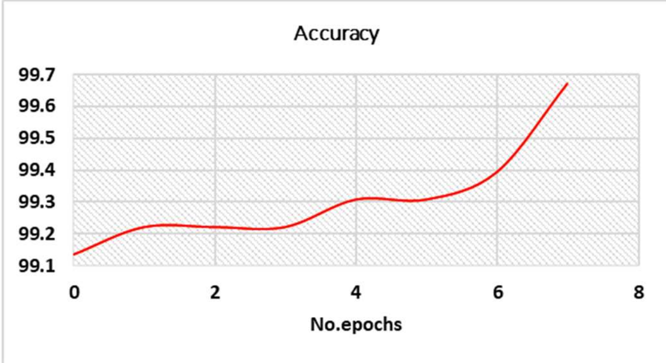

Figure 5. Accuracy rate

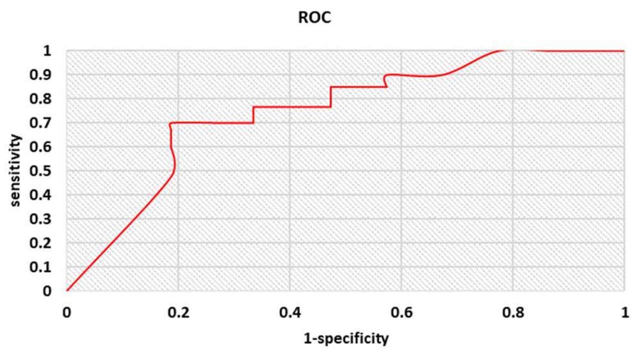

Figure 6. ROC representation

Table 1. Performance of proposed scheme with previous work

Work	Feature extraction	Recognition	Accuracy %
[14]	LBP	Chi-square	97.16
[11]	Resnet50 Alexnet	End-to-end End-to-end	93.57 94.29
[9]	Haralick features and NGTDM	SVM	98.4%
[10]	HOG	RF	99.69
Proposed scheme	SIFT	GABP	99.7

VI. CONCLUSION

In this article, two essential phases of the proposed ear recognition scheme are examined. The ear images are segmented to create the ROI in the first phase, preprocessing. Local features are extracted from both the left and right ROI ear images in the second phase of recognition, combined features are selected using GA, and the selected features are then transmitted to a BP network for evaluation. The proposed scheme yielded a substantial outcome with a 99.7% accuracy rate for the IITD dataset.

References

[1] H. Inass Shahadha, S. Shamsul Bin, N. Md Jan, and A. Nilam Nur, “The extract region of interest in high-resolution palmprint using 2D image histogram entropy function,” *Journal of Computer Science,* vol. 15, no. 5, 2019. https://doi.org/10.3844/jcssp.2019.635.647.

[2] S. B. Alagarsamy and S. Kondappan, “Ear recognition system using adaptive approach Runge–Kutta (AARK) threshold segmentation with ANFIS classification,” *Neural Computing and Applications,* vol. 32, no. 15, pp. 10995-11006, 2020. https://doi.org/10.1007/s00521-018-3805-6.

[3] Z. Wang, J. Yang, and Y. Zhu, "Review of ear biometrics," *Archives of Computational Methods in Engineering*, vol. 28, no. 1, pp. 149-180, 2021. https://doi.org/10.1007/s11831-019-09376-2.

[4] A. V. Iannarelli, *The Iannarelli System of Ear Identification*, Foundation Press, 1964.

[5] S. O. Berg, "Sherlock Holmes: father of scientific crime and detection," *J. Crim. L. Criminology & Police Sci.*, vol. 61, p. 446, 1970. https://doi.org/10.2307/1141973.

[6] S. M. Jiddah and K. Yurtkan, "Fusion of geometric and texture features for ear recognition," *Proceedings of the 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2018, pp. 1-5. https://doi.org/10.1109/ISMSIT.2018.8567044,

[7] L. Li, B. Zhong, C. Hutmacher, Y. Liang, W. J. Horrey, and X. Xu, "Detection of driver manual distraction via image-based hand and ear recognition," *Accident Analysis & Prevention*, vol. 137, p. 105432, 2020. https://doi.org/10.1016/j.aap.2020.105432,

[8] P. L. Galdámez, W. Raveane, and A. González Arrieta, "A brief review of the ear recognition process using deep neural networks," *Journal of Applied Logic*, vol. 24, pp. 62-70, 2017. https://doi.org/10.1016/j.jal.2016.11.014.

[9] R. Agarwal, "Local and global features based on ear recognition system," *Proceedings of the International Conference on Artificial Intelligence and Sustainable Engineering*, Springer Singapore, 2022, pp. 477-486. https://doi.org/10.1007/978-981-16-8546-0_39.

[10] M. H. Mutar, E. H. Ahmed, M. R. M. ALsemawi, H. O. Hanoosh, and A. H. Abbas, "Ear recognition system using random forest and histograms of oriented gradients techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 1, pp. 181-188, 2022. https://doi.org/10.11591/ijeecs.v27.i1.pp181-188.

[11] M. Sharkas, "Ear recognition with ensemble classifiers: A deep learning approach," *Multimedia Tools and Applications*, vol. 81, no. 30, pp. 43919-43945, 2022. https://doi.org/10.1007/s11042-022-13252-w.

[12] D. ping Tian, "A review on image feature extraction and representation techniques," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 4, pp. 385-396, 2013.

[13] I. S. Hussein, S. Sahibuddin, and N. N. A. Sjarif, "Shape feature extraction methods based pattern recognition: A survey," *Open International Journal of Informatics (OIJI)*, pp. 26-41, 2018.

[14] M. Hassaballah, H. A. Alshazly, and A. A. Ali, "Ear recognition using local binary patterns: A comparative experimental study," *Expert Systems with Applications*, vol. 118, pp. 182-200, 2019. https://doi.org/10.1016/j.eswa.2018.10.007.

[15] N. Boodoo-Jahangeer and S. Baichoo, "LBP-based ear recognition," *Proceedings of the 13th IEEE International Conference on BioInformatics and BioEngineering*, 2013, pp. 1-4. https://doi.org/10.1109/BIBE.2013.6701687.

[16] Y. Guo and Z. Xu, "Ear recognition using a new local matching approach," *Proceedings of the 2008 15th IEEE International Conference on Image Processing*, 2008, pp. 289-292.

[17] A. Benzaoui, A. Kheider, and A. Boukrouche, "Ear description and recognition using ELBP and wavelets," *Proceedings of the 2015 IEEE International Conference on Applied Research in Computer Science and Engineering (ICAR)*, 2015, pp. 1-6. https://doi.org/10.1109/ARCSE.2015.7338146.

[18] Y. Zhang, Z. Mu, L. Yuan, H. Zeng, and L. Chen, "3D ear normalization and recognition based on local surface variation," *Applied Sciences*, vol. 7, no. 1, p. 104, 2017. https://doi.org/10.3390/app7010104.

[19] L. Yuan, W. Liu, and Y. Li, "Non-negative dictionary based sparse representation classification for ear recognition with occlusion," *Neurocomputing*, vol. 171, pp. 540-550, 2016. https://doi.org/10.1016/j.neucom.2015.06.074.

[20] A. Benzaoui, A. Hadid, and A. Boukrouche, "Ear biometric recognition using local texture descriptors," *Journal of electronic imaging,* vol. 23, no. 5, p. 053008, 2014. https://doi.org/10.1117/1.JEI.23.5.053008.

[21] A. S. Anwar, K. K. A. Ghany, and H. ElMahdy, "Human ear recognition using SIFT features," *Proceedings of the 2015 IEEE Third World Conference on Complex Systems (WCCS)*, 2015, pp. 1-6. https://doi.org/10.1109/ICoCS.2015.7483254.

[22] L. Ghoualmi, A. Draa, and S. Chikhi, "An ear biometric system based on artificial bees and the scale invariant feature transform," *Expert Systems with Applications*, vol. 57, pp. 49-61, 2016. https://doi.org/10.1016/j.eswa.2016.03.004.

[23] G. Mawloud and M. Djamel, "Weighted sparse representation for human ear recognition based on local descriptor," *Journal of Electronic Imaging,* vol. 25, no. 1, p. 013036, 2016. https://doi.org/10.1117/1.JEI.25.1.013036.

[24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004. https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[25] E. A. Elgallad, N. Charfi, A. M. Alimi, and W. Ouarda, "Human identity recognition using sparse auto encoder for texture information representation in palmprint images based on voting technique," *Proceedings of the 2017 Sudan Conference on Computer Science and Information Technology (SCCSIT)*, 2017, pp. 1-8. https://doi.org/10.1109/SCCSIT.2017.8293060.

[26] A. Hassin, D. Abbood, "Machine learning system for human ear recognition using scale invariant feature transform," Artificial Intelligence & Robotics Development Journal (AIRDJ), vol. 1, issue 1, 2021. https://doi.org/10.52098/airdj.20217.

[27] I. S. Hussein, S. B. Sahibuddin, M. J. Nordin, and N. N. B. A. Sjarif, "Multimodal recognition system based on high-resolution palmprints," *IEEE Access*, vol. 8, pp. 56113-56123, 2020. https://doi.org/10.1109/ACCESS.2020.2982048.

[28] M. I. Ahmad, W. L. Woo, and S. Dlay, "Non-stationary feature fusion of face and palmprint multimodal biometrics," *Neurocomputing*, vol. 177, pp. 49-61, 2016. https://doi.org/10.1016/j.neucom.2015.11.003.

[29] S. Srivastava, "Accurate human recognition by score-level and feature-level fusion using palm–phalanges print," *Arabian Journal for Science and Engineering*, vol. 43, no. 2, pp. 543-554, 2018. https://doi.org/10.1007/s13369-017-2644-6.

[30] L. D. Vignolo and M. F. Gerard, "Evolutionary local improvement on genetic algorithms for feature selection," *Proceedings de XLIII CLEI–46 JAIIO,* pp. 1-8, 2017. https://doi.org/10.1109/CLEI.2017.8226467.

[31] M. Mahalakshmi, P. Kalaivani, and E. K. Nesamalar, "A review on genetic algorithm and its applications," *International Journal of Computing Algorithm*, vol. 2, no. 2, pp. 415-423, 2013.

[32] S. Skansi, *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*, Springer, 2018. https://doi.org/10.1007/978-3-319-73004-2.

[33] D. Yashomati, K. Madhuri, S. Pooja, and A. Varsha, "The human identification system using geometrical feature extraction of ear," *International Journal of Advanced Research in Computer Science and Software Engineering IJARCSSE*, vol. 6, issue 3, pp. 298-303, 2016.

***INASS SH.HUSSEIN** received the M.S degree in computer science from UKM, Malaysia, and a Ph.D. degree from the Faculty Razak of Technology and Informatics, UTM, Kuala Lumpur, Malaysia. She is currently a senior lecturer with Baquba Technical Institute, Middle Technical University, Baquba, Iraq. Her areas of interest include machine learning, biometric system, and pattern recognition.*

***NILAM NUR AMIR SJARIF** is a senior lecturer at Razak Faculty of Technology and Informatics, under the department of Advanced Infomatics, Universiti Teknologi Malaysia, Kuala Lumpur. She received her BSc in Computer Science (Software Engineering) in 2010 and PhD in Computer Science from Universiti Teknologi Malaysia in 2015, in the area of Human Action Recognition with the Geometrical Feature Representation for Video Surveillance. She is active in the research field Image and Video Processing, Pattern Recognition, Machine Learning, Deep Learning, Watermarking and Big Data Analytics.*

• • •