Ear Recognition Using Rank Level Fusion of Classifiers Outputs

https://doi.org/10.3991/ijoe.v19i03.36831

Resmi K. R.[1], Shijo M. Joseph[2], Raju G.[1], Debabrata Swain[3], Om Prakash Das[4], Biswaranjan Acharya[5](✉)
[1]CHRIST University, Bengaluru, India
[2]Mahatma Gandhi College, Iritty, India
[3]Department of Computer Science and Engineering, Pandit Deendayal Energy University, Gandhinagar, India
[4]GSK India Global Services Private Limited, Bangalore, India
[5]Department of Computer Engineering-AI, Marwadi University, Rajkot, India
biswaacharya@ieee.org

Abstract—An individual's authentication plays a vital role in our daily life. In the last decade, biometric-based authentication has become more prevalent than traditional approaches like passwords and pins. Ear recognition has gained attention in the biometric community in recent years. Researchers defined several features for the identification of a person from ear image. The features play a vital role in the success of classification models. This paper considers an ensemble of features for designing a new classification model. The features are assessed in isolation as well as through feature-level fusion. Subsequently, a rank-level fusion for classification is introduced. The experiments are conducted on both constrained and unconstrained ear datasets. The results are promising and open up new possibilities in machine learning-based ear recognition.

Keywords—ear biometrics, decision level fusion, rank level fusion, HOG, GLBP, BSIF

1 Introduction

Biometrics is a personal identification or authentication method relying on unique physiological or behavioural traits. Iris, palm print, and gait biometrics are well-researched topics with numerous practical systems. In the last decade, ear biometrics has become increasingly prominent. Ear Biometrics is the verification or identification of a person using a 2D or 3D ear image. As a physiological approach, an ear image has discriminative features to distinguish between individuals. Ear biometrics has broad applications, from authentication to forensics. Ear satisfies essential biometric requirements – universality, permanence, and uniqueness. After decades of study of "anthropometric measurement" of ear photos of thousands of people, it was concluded that no two ears are alike, even in the case of identical and fraternal twins [1].

Unlike the face, ear anatomy does not change much over time [1]. Furthermore, images of the ears are unaffected by makeup and emotion. Contact biometrics like iris and fingerprints require user cooperation while acquiring the biometric trait. However, ear images can be captured without any consent from the user. This makes ear biometrics useful in forensic investigation and surveillance. To improve security and performance, a human ear can be integrated with biometrics such as the face and iris to build a multimodal recognition system.

An ear recognition system with handcrafted features and machine learning algorithms includes ear segmentation/localization, feature extraction, and recognition/verification processes. The efficacy of an ear biometric system is influenced by the contrast, lighting, and different occlusion of ear images such as earrings and hair. Early ear biometrics research focused on images collected in controlled conditions, where the images are captured explicitly for research purposes using special hardware and indoor environments. Popular datasets like IIT Delhi, AMI, and USTB are among them. They have limited variability in rotation, occlusion, and illumination change. The recognition systems built with these datasets fail in many practical situations where the images are degraded by factors such as poor illumination, blur, noise, viewing angles, and occlusions. Hence researchers started using unconstrained datasets such as AWE, AWEx, and WebEar, where the recognition accuracy is a challenge. Though conventional machine learning-based approaches are successful with constrained datasets, they failed to give impressive results with unconstrained datasets. Deep learning-based models are proposed with relatively high accuracy for unconstrained datasets. Deep learning models generally require large datasets, and model building is resource-demanding.

Based on feature extraction methods, Ear recognition approaches are divided into four categories: geometrical, holistic/global, local, and hybrid [2]. Early Ear recognition works are based on geometric and global features [3] [4] [5], and they reported low recognition results. Recognition using local and hybrid descriptors [6] [7] [8] [9] reported high performance compared to other descriptors. In this work, three efficient local descriptors – Histogram of Oriented Gradients (HOG), Gradient Local Binary Pattern (GLBP), and Binarized Statistical Image Features (BSIF) are selected for feature extraction, and the goal is to evaluate the classifier performance using three approaches. The first approach is to independently find the recognition accuracy using the selected features. The second approach investigates whether combining features will improve recognition performance (feature-level fusion). The third approach is to build a novel rank-level fusion of classifier outputs to improve recognition accuracy. The Rank Level Fusion approach improves classification accuracy. Both constrained and unconstrained datasets are used in the study.

2 Related works

The French criminologist Mr Bertillon investigated the use of the ear to track criminals [10]. In 1964, Mr Iannarelli, an American police officer, analyzed about 10000 ear images and 12 manual measurements to prove the uniqueness of the ear [1]. He carried out experiments using twins and triplets to find individuality.

Early contributions in Ear recognition make use of geometrical shape and structure. These methods use edge information as attributes for classification. A graph matching technique, with an Adjacency Graph of Voronoi diagrams made from ear segments, is proposed in [3]. The authors provided a theoretical base for ear recognition. For occlusion due to hair, the authors suggested a thermogram. The paper brings out the ear biometrics challenges and offers possible research directions.

Mu et al. incorporated outer and inner ear structures for feature vector formation and Neural Networks for classification [11]. The paper reports 85% accuracy for the USTB II dataset. Choraś & Choraś defined a feature considering concentric circles for the centroid of an ear image centroid [12]. The authors created and experimented with their data set and reported 100% accuracy. Geometrical features, combining the outer contour's minimum and maximum ear height line of the outer contour, are used in [13]. The features include the maximum ear height line, minimum ear height line, a summation of height lines, the ratio of height lines, and the ratio of maximum and minimum ear height lines with the smallest distance from the centre of the maximum height line to the left. The experiment was conducted on the USTB subset 1 and IITD1 database and reported 98.33% and 99.6% accuracy, respectively. Geometrical methods are computationally simple but give a low recognition rate when the images are of low quality and occluded.

Holistic approaches like "Force Field", "Principal Component Analysis" (PCA), and "Linear Discriminant Analysis" (LDA) extract features from the whole ear image. Models based on these give high accuracy when the images are subjected to a few pre-processing steps, including size normalization. Hurley et al. used Force Field Transform approach that assumes pixels are mutually attracted to form a "Gaussian Force Field" [4]. The directional features derived from the Force Field are used to find potential wells and channels for matching. A rank-1 recognition rate of 99.2% is reported with the XM2VTS database. The Force Field techniques are also adopted in [14] [15] and [5]. The idea of Eigen space based on PCA to build a multimodal system, combining face and ear, is used in [5] and [16]. An observation from the above multimodal studies is that ear recognition accuracy is low compared to face recognition. The former reported a recognition rate of 40% on their dataset, while the latter projected a 71.6% rank-1 recognition rate on the UND-E dataset.

Zhang et al. [17] used Independent Component Analysis (ICA) and Radial Basis Function (RBF) for Ear recognition. The original image database is decomposed into several basic image combinations, and the combinations are fed to the RBF. They reported 94.1% accuracy with the CP dataset and 88.3% with their dataset. Analysis of works done using holistic approaches concludes that recognition accuracy decreases when unconstrained datasets are used.

Zarachoff et al. [18] used a multi-band two-dimensional PCA(2D-MBPCA) technique for ear recognition. This technique divides the given input image into multiple bands, and standard PCA is applied to each band to generate eigenvectors for feature matching. They reported 93.63% accuracy with the IITD II dataset and 96.11% with the USTBI dataset.

In local approaches, the ear image extracted is divided into regions, and local features are extracted from these regions. Nanni and Lumni used different colour space properties and ensemble-based ear matcher [19]. The method extracts Gabor features from different colour spaces. The experiments using the UND-E database resulted in 84% rank-1 recognition accuracy.

A study on "texture" and "surface" descriptors in Ear biometrics is presented in [20]. The authors used the "Local Binary Pattern" (LBP), "Local Phase Quantization" (LPQ), "Histogram of Oriented Gradients" (HOG), "Binarized Statistical Image Features" (BSIF), "Shape index", and "Curvedness". They experimented with feature-level fusion, introducing a novel histogram-based descriptor. LDA is used for generating feature subspace, and classification is done with different distance measures. The datasets used are UND-J2, IITK, and AMI. Extensive experiments with various features, transformers, and distance measures showed up to 100% accuracy.

Benzaoui et al. used the texture descriptors LBP, LPQ, and BSIF and the data sets IITD-I, IITD-II, and the USTB [21]. The authors reported a 97.3% accuracy, the highest, for IITD-II with the BSIF feature. The same authors proposed a new model that extracts local colour texture descriptors in the three-colour spaces, RGB, YCbCr, and HSV [7]. The training set includes one sample per person from the USTB-1 dataset, and the remaining images are used for testing. A rank-1 recognition rate of 96.53% is reported on the USTB-1 database with the RGB-BSIF combination.

Houcine et al. [22] defined a "multi-bags-of-features" histogram for ear recognition. The method computes histograms using a Bag of features and applies Kernel Discriminant Analysis (KDA). The experiments using IITD-I database resulted in 96.3% rank-1 accuracy. Hassaballah et al. used the Robust Local Oriented Patterns (RLOP) descriptor [6]. RLOP uses "edge directional information" to obtain a noise-tolerant and rotation-invariant descriptor. The dataset considered are IITD-I, IITD-II, AMI, and AWE. The paper presents details of extensive experiments carried out and reports an average accuracy of 98%.

Mahmoud et al. used a descriptor called "Dense Local Phase Quantization" (DLPQ) from LPQ [23]. DLPQ is calculated using phase responses generated from the LPQ descriptor. The experiment was conducted on IITD-I and IITD-II datasets and reported 98.49% and 98.4% recognition accuracy, respectively. Youbi et al. used a Multi-scale Local Binary Pattern (MLBP), a variant of LBP and city block distance [24]. The experiments were conducted with IITD-I, IITD-II, and USTB-I datasets and reported a rank-1 accuracy of 98.40%, 98.64%, and 98.33%, respectively.

Emersic et al. developed a model using LBP, LPQ, HOG, BSIF, and "Patterns of Oriented Edge Magnitudes" (POEM) [2]. The paper reports 98.5% and 48.4% accuracy on the IITD-II and the AWE datasets when BSIF is used as the descriptor. A new variant of BSIF – Multiscale Framework-BSIF [MS-BSIF] is proposed in [25]. A bank of BSIF filters is used for a given image to obtain response images, followed by histogram normalization and concatenation to get MS-BSIF histograms. "Whitened Linear Discriminant Analysis" is applied for dimensionality reduction and classification by KNN classifier with Chi-square distance. Rank-1 recognition rates of 98.08%, 97.72%, and 99.74% are obtained for IITD-I, IITD-II, and the USTB-1 datasets. Alshazly et al. used an approach based on Gradient or edge directions [26]. Experiments with Local Optimal Oriented Patterns (LOOP) feature descriptors and IITD-I, IITD-II, and the AMI datasets, report rank1 accuracy of 93.9%, 96.9%, and 70.2%, respectively.

Omara et al. [27] employed a hybrid approach by combining features. The authors used LBP, LPQ, HOG, BSIF, POEM, and Gabor features. "Dynamical Components

Analysis" (DCA) is applied for fusion and dimensionality reduction. SVM is chosen as the classifier. They reported 97.33% accuracy with the USTB and 96.25% with the IITD-II. Ear recognition by the fusion of two descriptors, "Pyramid Histogram of Oriented Gradients" (PHOG) and "Local Directional Pattern" (LDP), is proposed in [28]. The shape and texture information is extracted using the descriptors. Kernel Discriminant Analysis (KDA) and Nearest Neighbour classifier are combined for the classification task. Rank-1 recognition rate of 97.60%, 97.37%, and 96.83% is reported on IITD-I, IITD-II, and UND-E datasets.

Sajadi and Fathi employed a hybrid approach by combining global and local features [29]. Gabor Zernike is used as the global feature and LPQ is the local feature. The Genetic algorithm is used to select an optimal combination of the features. KNN classifier with Canberra Distance is adopted for classification. The rank-1 recognition rate is 100%, 99.2%, and 97.13% for USTBI, IITD-I, and IITD-II datasets.

Ear recognition based on shifted 1D-LBP is proposed in [30]. KNN is chosen as the classifier. The experiment using the USTB I and USTB II datasets resulted in 97.04% and 100% recognition accuracy. Muttasher used "Features from Accelerated Segment Test" (FAST) and HOG for feature extraction [31]. A probabilistic Neural Network (PNN) is used for the matching process. The experiments reported a 98% Recognition rate on the AMI database.

Hybrid approaches, which give better recognition results, are often computationally complex compared to holistic, local, or geometrical approaches. In the last decade, several Deep learning models have been proposed for ear recognition. Deep learning-based works are not included in the review as the focus is on handcrafted features and classification using machine learning algorithms.

3 Feature extraction methods

This section gives a brief overview of the local features that are included in the study.

3.1 Histogram of Oriented Gradients (HOG)

HOG is a popular and widely used feature descriptor for object recognition in computer vision [32]. HOG captures the structure or shape of an object. It determines the magnitude and direction of gradients for edges in an image.

3.2 Gradient Local Binary Pattern (GLBP)

The LBP is a texture descriptor introduced in [33]. LBP is simple, effective, and rotation and illumination invariant. For each pixel, the LBP code is computed, and the histogram of these codes is used as a texture descriptor.

LBP for pixel (x_c, y_c) is [33]

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

Where g_c is *the* value of the central pixel and g_p is the value of its 2D neighbours

In the literature, various LBP variations for feature extraction have been proposed. Of these, Gradient Local Binary Pattern (GLBP) is chosen for this work. GLBP utilizes pixel and gradient values to encode local texture information [34]. The LBP encoding technique encodes the pixel value to generate a binary code [34]. The gradient value is encoded as a single vector that provides gradient magnitude and direction. The binary code for gradient magnitude is obtained the same way as basic LBP (Eqn. 2), where gradient magnitude m is used instead of intensity.

$$t(s(m_0 - m_c), s(m_1 - m_c), \ldots\ldots s(m_{p-1} - m_c)) \tag{2}$$

For encoding the gradient direction, the centre pixel direction is compared with the directions in its 8 neighbourhoods. Two vectors have the same direction if both belong to the same range direction. The binary code for gradient direction is defined as

$$(f(d_0 - d_c), s(d_1 - d_c), \ldots\ldots s(d_{p-1} - d_c)) \tag{3}$$

Where d_c and d_p represent the centre pixel and eight surrounding pixels, respectively. The thresholding function $f(z)$ is defined as

$$f(z) = \begin{cases} 1, if\ |z| \le 20^\circ \\ 0, otherwise \end{cases} \tag{4}$$

The $GLBP_{P,R}$ operator is defined as follows:

$$GLBP_{P,R}(x_c, y_c) = \left[\sum_{p=0}^{P-1} (s(g_p - g_c \times 2^p)) \sum_{p=0}^{p-1} ((s(m_p - m_c) \vee f(d_p - d_c)) \times (128 + 2^p)) \right] \tag{5}$$

The binary code for the gradient is generated by ORing binary codes of the gradient magnitude and direction. This result is multiplied by pixel weights, and the sum is found. The final feature vector is obtained by concatenating the sum calculated using intensity values and the sum obtained from gradient values.

3.3 Binarized Statistical Image Features (BSIF)

BSIF is a local histogram-based texture descriptor reported in [35]. In BSIF, a binary code is generated by projecting the local image patches over filters of natural images. For an image patch X of size l × l pixels and a linear filter W_i of the same size, the filter response S_i is calculated using Eqn. 6.

$$s_i = \sum_{u,v} W_i(u,v) X(u,v) = w_i^T x \tag{6}$$

Where vectors w and x contain the pixels of W_i and X. The BSIF feature b_i is obtained by setting $b_i = 1$ if $s_i > 0$ and $b_i = 0$ otherwise. BSIF is a powerful descriptor that provides high ear recognition results under different conditions. Filters with size 11 × 11 are used in the present study.

4 Proposed rank level fusion approach

Features extracted from the segmented ear play a vital role in ear recognition. A set of local image features is chosen, and extensive experimentation is performed to identify their merits in ear recognition.

In classification, "Rank n" classification accuracy is one of the metrics used to project the performance (for example, the accuracy when the searched item is in the top five predicted outputs). This works well with classifiers like KNN. Popular Decision level fusion techniques include "majority voting, Borda count method, and weighted Borda count". In this paper, a decision-level fusion incorporating rank is introduced. The top ten predictions of three KNN classifiers are combined to generate the class label. HOG, GLBP, and BSIF features are extracted from IIT Delhi and AWE datasets. For each feature, a KNN model is created, adopting Euclidean distance. Euclidean distance outperforms all other distance measures for all the features. The KNN model predicts ten labels in rank order (distance).

Consider a test sample with actual label q_i, input to the three classifiers. For this input, for each classifier C, let $P(C)$, $C = 1$ to 3, be the set of predictions. Each $P(C)$ contains a set of ten class labels, ordered based on the distance, least to high. Let R (q_k, C) be the rank of a label in $P(C)$, $q_k \in$ labels in the training set.

Definition: Rank Weight (*RW*)

The rank weight $RW\,(q_k, C) = 10 - R\,(q_k, C) + 1$, if $q_k \in P(C)$ and RW $(q_k, C) = 0$, if $q_k \notin P(C)$, $C = 1$ to 3.

That is, the label at rank 1 gets a weight of 10, and the label at rank position 10 gets a weight of 1. For each label q_k, where q_k is in prediction sets *P©*, $C = 1$ to 3, calculate the Average Rank Weight, *ARW*:

$$ARW\,(q_k) = (RW\,(q_k, 1) + RW\,(q_k, 2) + RW\,(q_k, 3))/3. \tag{7}$$

The label q_k with the maximum ARW is taken as the final prediction, which is compared with the input qi for measuring prediction accuracy. If more than one class label has the same ARW, then the label with the minimum distance in BSIF is considered. Figures 1–3 illustrates the proposal.

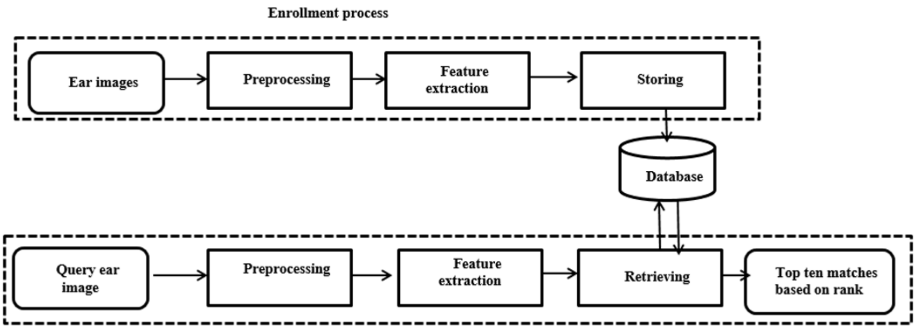

Fig. 1. Ranked output generated for a test image

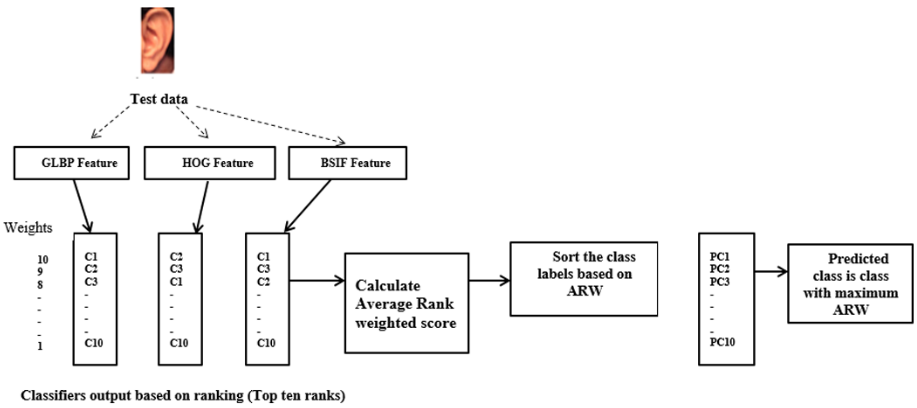

Fig. 2. Proposed classification based on ranking

The ranked output generated for a test image based on the rank level fusion of classifier output is illustrated with an example. Figure 3a shows the retrieved top three ranks and their corresponding class for a query image from three different classifiers. A feature vector is generated for each class id. For class id C1, the ranks are <1, 2, 1> respectively from three classifier outputs. The rank for each class is calculated as shown in Figure 3b and c. From Figure 3c, it is seen that Class C1, with the maximum score value, is the predicted output class for the given test image.

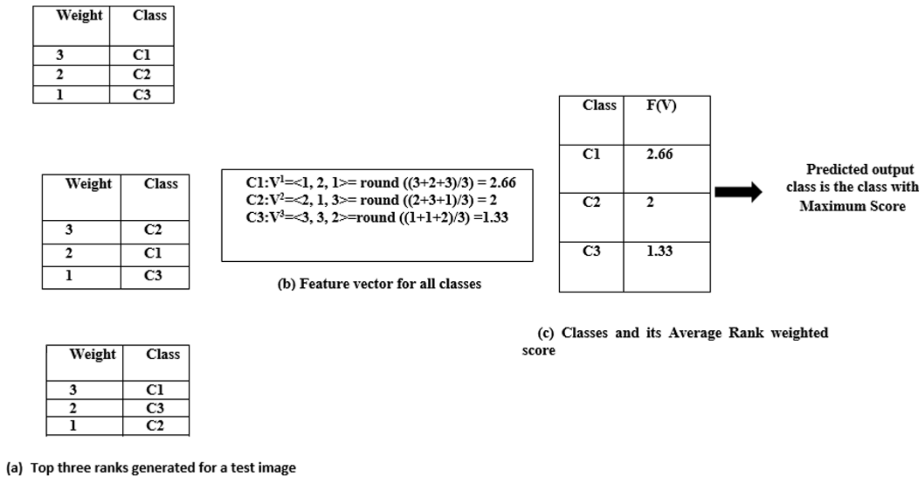

Fig. 3. An example of the rank-level fusion of the Top three classifier outputs for a test image

5 Experiments

Most ear biometrics research using traditional machine learning algorithms focuses on constrained datasets. Both constrained and unconstrained datasets are used in this study. Experiments using single features, feature-level fusion, and decision-level fusion are carried out on various datasets. This section describes the datasets, experimental setup, and results in detail.

5.1 Data sets

A brief overview of the datasets is presented below:

IIT Delhi. IIT Delhi ear database is a constrained dataset [36]. It consists of two sets of processed images, IIT-Delhi-I and IIT-Delhi-II, with 125 and 221 subjects. IIT-Delhi also contains a database with 493 images of 125 subjects.

All images are in greyscale ".bmp" format. The minimum number of images per sample is three, and the maximum is 6 in both databases. IIT-Delhi-I and IIT-Delhi-II contain cropped images of size 50 × 180 pixels. The size of images in the IIT-Delhi raw database is 272 × 204 pixels. Figures 4 and 5 show sample images of IIT Delhi raw and processed databases, respectively.

AWE. AWE is an unconstrained dataset [2]. AWE contains 1000 images, ten samples each of 100 subjects. The images, with different dimensions in ".png" format, are collected from the web. It contains images with different illumination, angle, size, and rotation. Sample images of a single person in the AWE database are shown in Figure 6.

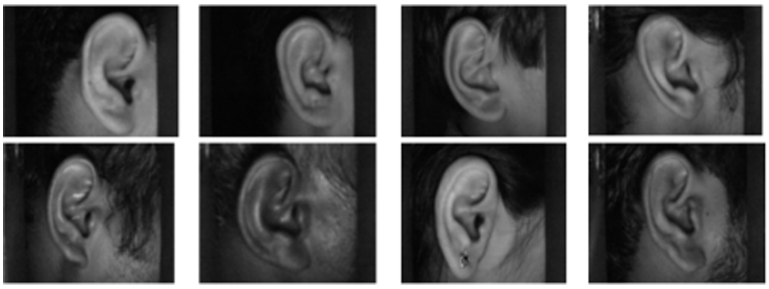

Fig. 4. Sample images from IIT Delhi raw database

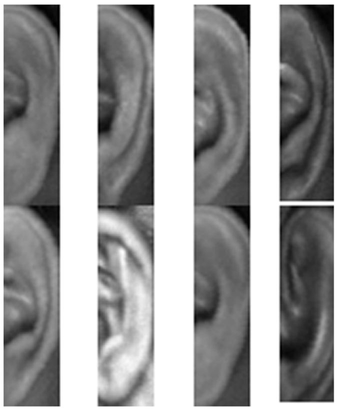

Fig. 5. Sample images from IITD I and IITD II database

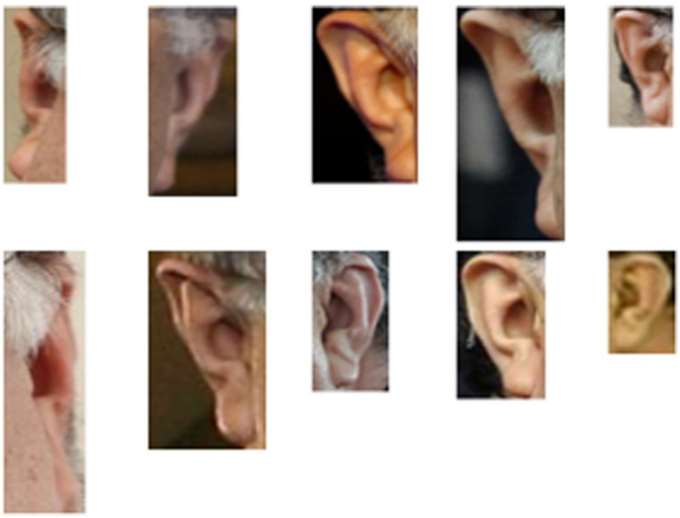

Fig. 6. Sample images of a single person AWE

Table 1 shows the databases used in the experiment with the number of persons and the total number of images in each dataset.

Table 1. Description of datasets used

Database	Number of Persons	Number of Images
IIT Delhi Raw database	125	493
IITD I	125	493
IITD II	221	793
AWE	100	1000

5.2 Experimental settings

Although the images in the IIT-Delhi datasets are size-normalized, AWE contains images of various sizes. As a result, AWE samples are resized to 100 × 100. All image samples are subjected to histogram equalization before feature extraction.

The LBP parameters are uniform LBP with a block size of 8 × 8 pixels and a radius of 2 pixels. The original image is divided into 8 × 8 cells for GLBP, and each cell's GLBP features are extracted. The parameters utilized in HOG are an 8x8 block size and a 16 × 16 block normalization of the histogram. The size of the filter window and the number of bits in the binary code are the most important factors in the BSIF filter. Different filter bank sizes are tested, and a filter bank size of 11 × 11 is chosen based on accuracy. Similarly, the binary code is limited to eight bits. Euclidean distance, City block, Mahalanobis, and Minkowski distance are among the distance measurements adopted for KNN. After experimenting with several K values, K is fixed as 10. Compared to different distance measures of KNN, Euclidean distance outperforms other distance measures. The training-test data split is fixed as 60:40. Fivefold cross-validation is adopted for training.

5.3 Results and discussion

Table 2 compares the performance of the proposed rank-level fusion with selected single local features and feature-level fusion. The BSIF feature shows the best recognition results among the selected single features in all databases. Experiments with feature-level fusion are carried out by considering the selected features and KNN with Euclidean distance. Feature level fusion of HOG, GLBP, and BSIF reported the highest recognition result. The dimensionality of various features used is HOG: 576; LBP: 3776; GLBP: 2304, and BSIF: 968. Feature fusion leads to high dimensionality, and hence PCA is applied.

The concepts of combining various classifiers or ensembles proved promising results across multiple classification problems. Rank-level fusion offers marginal improvement in recognition accuracy. The rank-level fusion approach is less complex and easy to implement than popular classifiers like Random Forest.

Table 3 shows a comparison with the existing state of art techniques. The maximum accuracy obtained in our studies is comparable to or better than the previous works.

For AWE, the unconstrained dataset, HOG + CNN, gives good accuracy but is a deep learning-based approach.

Table 2. Comparison of proposed rank level fusion method with selected features

Databases	Subjects	Feature	Classifier	Accuracy (Rank 1%)
IIT Delhi Raw database	125 subjects (493 images)	HOG LBP GLBP BSIF HOG+GLBP+BSIF Rank Level fusion	KNN (Euclidean)	80.8% 78.7% 80.7% 81.6% 83.6% 84.3%
IIT Delhi I	221 subjects (793 images)	HOG LBP GLBP BSIF HOG+GLBP+BSIF Rank Level fusion	KNN (Euclidean)	96.2% 95.7% 96.2% 96.4% 97.1% 97.9%
IIT Delhi II	125 subjects (493 images)	HOG LBP GLBP BSIF HOG+GLBP+BSIF Rank Level fusion	KNN (Euclidean)	97.3% 96.8% 97.3% 97.7% 97.8% 98.6%
AWE	100 subjects (1000 images)	HOG LBP GLBP BSIF HOG+GLBP+BSIF Rank Level fusion	KNN (Euclidean)	43.9% 43.5% 46.3% 48.4% 49.5% 49.8%

Table 3. Comparison of results of existing ear recognition methods

Reference	Features	Databases	Type	Recognition Accuracy (Rank 1)
[23]	DLPQ	IITDII	Local	98.4%
[24]	MLBP	IITDII	Local	98.64%
[25]	MS-BSIF	IITDII	Local	97.72%
[26]	LOOP	IITDII	Local	96.9%
[2]	LBP, LPQ, BSIF, HOG, POEM	IITDII AWE	Local	98.5% 48.4%
[40]	HOG+CNN	AWE	Hybrid	75.6%
[28]	PHOG and LDP	IITD1 IITDII	Hybrid	97.60% 97.37%
Proposed approach	HOG, GLBP, BSIF	IITD I IITDII AWE	Hybrid	**97.9%** **98.6%** **49.8%**

The proposed model can be improved by considering more features and fine-tuning feature parameters. Further, in the present study, only KNN classifier is used. Other classifiers shall be customized and adopted. Multiple classifiers can also be incorporated into the model. Finally, as the unconstrained data sets demand better models, the approach presented shall be extended to the deep learning domain. Several studies in deep learning report higher accuracy like [37] [38] [39].

6 Conclusion

In this paper, extensive experimentation on ear recognition using various local features and classifiers is presented. Features in isolation and feature-level fusion models are investigated. A novel rank-level fusion is proposed. The work identified the best combination of the features considered, the best classifier, and the best decision-level fusion strategy. The results obtained are promising. The proposed rank-level fusion is established as an efficient decision-level fusion approach. The study needs to be extended by considering more features and datasets.

7 References

[1] Iannarelli, Alfred V.: Ear Identification, Forensic Identification Series. Paramont Publishing Company, Fremont, California. (1989).
[2] Emeršič, Žiga, Vitomir Štruc, and Peter Peer.: Ear Recognition: More than a Survey. Neurocomputing. (2017). https://doi.org/10.1016/j.neucom.2016.08.139
[3] Burge, Mark, and Wilhelm Burger.: Ear Biometrics in Computer Vision. Proceedings – International Conference on Pattern Recognition. (2000). https://doi.org/10.1109/ICPR.2000.906202
[4] Hurley, David, J., Mark S. Nixon, and John N. Carter.: Automatic Ear Recognition by Force Field Transformations. In IEE Colloquium (Digest), (2000). https://doi.org/10.1049/ic:20000467
[5] Victor, Barnabas, Kevin Bowyer, and Sudeep Sarkar.: An Evaluation of Face and Ear Biometrics. Proceedings – International Conference on Pattern Recognition. (2002). https://doi.org/10.1109/ICPR.2002.1044746
[6] Hassaballah, M., Alshazly, H.A. and Ali, A.A.: Robust Local Oriented Patterns for Ear Recognition. Multimed Tools Appl 79, (2020) 31183–31204. https://doi.org/10.1007/s11042-020-09456-7
[7] Benzaoui, Amir, and Abdelhani Boukrouche.: Ear Recognition Using Local Color Texture Descriptors From One Sample Image Per Person. In 2017 4th International Conference on Control, Decision and Information Technologies, Codit (2017). https://doi.org/10.1109/CoDIT.2017.8102697
[8] Youbi, Zineb, Larbi Boubchir, and Abdelhani Boukrouche.: Human Ear Recognition Based on Local Multi-Scale LBP Features with City-Block Distance. Multimedia Tools and Applications. (2019). https://doi.org/10.1007/s11042-018-6768-9
[9] Sajadi, S. and Fathi, A.: Genetic Algorithm based Local and Global Spectral Features Extraction for Ear Recognition. Expert Syst. Appl. 159, 113639 (2020). https://doi.org/10.1016/j.eswa.2020.113639
[10] Bertillon, Alphonse.: The Bertillon System of Identification. Mcclaughry, Ed., Chicago, IL, 1896.

[11] Mu, Zhichun, Li Yuan, Zhengguang Xu, Dechun Xi, and Shuai Qi.: Shape and Structural Feature Based Ear Recognition. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), (2004). https://doi.org/10.1007/978-3-540-30548-4_76
[12] Choraś, Michał, and Ryszard S. Choraś.: Geometrical Algorithms of Ear Contour Shape Representation and Feature Extraction. In Proceedings – ISDA 2006: Sixth International Conference on Intelligent Systems Design and Applications, 2006. https://doi.org/10.1109/ISDA.2006.253879
[13] Omara, Ibrahim, Feng Li, Hongzhi Zhang, and Wangmeng Zuo.: A Novel Geometric Feature Extraction Method for Ear Recognition. Expert Systems with Applications, (2016). https://doi.org/10.1016/j.eswa.2016.08.035
[14] Dong, Jiyuan, and Zhichun Mu.: Multi-Pose Ear Recognition Based on Force Field Transformation. In Proceedings – 2008 2nd International Symposium on Intelligent Information Technology Application, IITA 2008. https://doi.org/10.1109/IITA.2008.325
[15] Abdel-Mottaleb, Mohamed, and Jindan Zhou.: Human Ear Recognition from Face Profile Images. In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2006. https://doi.org/10.1007/11608288_105
[16] Chang, Kyong, Kevin W. Bowyer, Sudeen Sarkar, and Barnabas Victor.: Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence, (2003). https://doi.org/10.1109/TPAMI.2003.1227990
[17] Zhang, Hai Jun, Zhi Chun Mu, Wei Qu, Lei Ming Liu, and Cheng Yang Zhang.: 2005. A Novel Approach for Ear Recognition Based on ICA and RBF Network. In 2005 International Conference on Machine Learning and Cybernetics, ICMLC 2005. https://doi.org/10.1109/ICMLC.2005.1527733
[18] Zarachoff, M.M., Sheikh-Akbari, A. and Monekosso, D.: Multi-Band PCA Based Ear Recognition Technique. Multimed Tools Appl, (2022). https://doi.org/10.1007/s11042-022-12905-0
[19] Nanni, Loris, and Alessandra Lumini.: Fusion of Color Spaces for Ear Authentication. Pattern Recognition, (2009). https://doi.org/10.1016/j.patcog.2008.10.016
[20] Pflug, Anika, Pascal Nicklas Paul, and Christoph Busch.: A Comparative Study on Texture and Surface Descriptors for Ear Biometrics. In Proceedings – International Carnahan Conference on Security Technology, 2014. https://doi.org/10.1109/CCST.2014.6986993
[21] Benzaoui, Amir, Abdenour Hadid, and Abdelhani Boukrouche.: Ear Biometric Recognition Using Local Texture Descriptors. Journal of Electronic Imaging, (2014). https://doi.org/10.1117/1.JEI.23.5.053008
[22] Houcine, Bourouba, Doghmane Hakim, Benzaoui Amir, and Boukrouche A. Hani.: 2015. Ear Recognition Based on Multi-Bags-of-Features Histogram. In 3rd International Conference on Control, Engineering and Information Technology, CEIT 2015. https://doi.org/10.1109/CEIT.2015.7232997
[23] Mahmoud Al Rahhal, Mohamad, Mohamed Lamine Mekhalfi, Mawloud Guermoui, Esam Othman, Baiying Lei, and Awais Mahmood.: A Dense Phase Descriptor for Human Ear Recognition. IEEE Access, (2018). https://doi.org/10.1109/ACCESS.2018.2810339
[24] Youbi, Zineb, Larbi Boubchir, and Abdelhani Boukrouche.: Human Ear Recognition Based on Local Multi-Scale LBP Features with City-Block Distance. Multimedia Tools and Applications, (2019). https://doi.org/10.1007/s11042-018-6768-9
[25] Doghmane, Hakim, Abdelhani Boukrouche, and Larbi Boubchir.: A Novel Discriminant Multiscale Representation for Ear Recognition. International Journal of Biometrics, (2019). https://doi.org/10.1504/IJBM.2019.096568

[26] Alshazly, H., Linse, C., Barth, E., and Martinetz, T.: Ensembles of Deep Learning Models and Transfer Learning for Ear Recognition. Sensors, 19, 4139, (2019). https://doi.org/10.3390/s19194139
[27] Omara, Ibrahim, Xiaoming Li, Gang Xiao, Khan Adil, and Wangmeng Zuo.: Discriminative Local Feature Fusion for Ear Recognition Problem. In ACM International Conference Proceeding Series, 2018. https://doi.org/10.1145/3180382.3180409
[28] Sarangi, Partha Pratim, Bhabani Shankar Prasad Mishra, and Satchidanand Dehuri.: Fusion of PHOG and LDP Local Descriptors for Kernel-Based Ear Biometric Recognition. Multimedia Tools and Applications, (2019). https://doi.org/10.1007/s11042-018-6489-0
[29] Sajadi, Shabbou and Fathi, Abdolhossein.: Genetic Algorithm Based Local and Global Spectral Features Extraction for Ear Recognition. Expert Systems with Applications, (2020), 159, 113639. https://doi.org/10.1016/j.eswa.2020.113639
[30] M. Benouis and M. Regouid.: Shifted 1D-LBP Based Ear Recognition System, 020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP), EL OUED, Algeria, 1–5, 2020. https://doi.org/10.1109/CCSSP49278.2020.9151650
[31] Raed Majeed Muttasher.: An Efficient Human Ear Recognition Scheme Based on Robust Feature Extraction and PNN, Solid State Technology, 63(5), 2020.
[32] Dalal, Navneet, and Bill Triggs.: 2005.: Histograms of Oriented Gradients for Human Detection. In Proceedings – 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR (2005). https://doi.org/10.1109/CVPR.2005.177
[33] Ojala, Timo, Matti Pietikäinen, and David Harwood.: A Comparative Study of Texture Measures with Classification Based on Feature Distributions. Pattern Recognition (1996). https://doi.org/10.1016/0031-3203(95)00067-4
[34] Tuan Le-Viet and Vinh Truong Hoang.: Local Binary Pattern Based on Image Gradient for Bark Image Classification. Proc. SPIE 11071, Tenth International Conference on Signal Processing Systems, 110710P, (2019). https://doi.org/10.1117/12.2522093
[35] Kannala, Juho, and Esa Rahtu.: BSIF: Binarized Statistical Image Features. In Proceedings – International Conference on Pattern Recognition. (2012).
[36] Kumar, Ajay, and Chenye Wu.: Automated Human Identification Using Ear Imaging. Pattern Recognition. (2012), 956–968. https://doi.org/10.1016/j.patcog.2011.06.005
[37] Salamh, A. B. S. and Akyüz, H. A New Deep Learning Model for Face Recognition and Registration in Distance Learning. International Journal of Emerging Technologies in Learning (iJET), (2022), 17(12), 29–41. https://doi.org/10.3991/ijet.v17i12.30377
[38] Muniasamy, A. and Alasiry, A.: Deep Learning: The Impact on Future eLearning. International Journal of Emerging Technologies in Learning (iJET), (2020), 15(01), 188–199. https://doi.org/10.3991/ijet.v15i01.11435
[39] Wang, C.: Emotion Recognition of College Students' Online Learning Engagement Based on Deep Learning. International Journal of Emerging Technologies in Learning (iJET), (2022), 17(06), pp. 110–122. https://doi.org/10.3991/ijet.v17i06.30019
[40] Hansley, Earnest, E., Maurício Pamplona Segundo, and Sudeep Sarkar.: Employing Fusion of Learned and Handcrafted Features for Unconstrained Ear Recognition. IET Biometrics, (2018). https://doi.org/10.1049/iet-bmt.2017.0210

8 Authors

Resmi K. R. obtained her degree in B.Sc. Physics in 2005 and M.C.A in 2008 from Mahatma Gandhi University Kerala, India. She completed her PhD in Computer Science in 2021 from the same university. She worked as a lecturer at Ilahia College, Muvattupuzha from 2009 to 2011 and currently working as an assistant professor in the Department of Computer Science, CHRIST University, Bengaluru, India. Her research interests include biometrics, pattern recognition, image processing and computer vision. She may be reached at resmykr@gmail.com.

Shijo M. Joseph obtained his degree in B.Sc. CMS (Computer Science Mathematics & Statistics) from CHRIST University, Bengaluru, in 1993, Master of Computer Science in 1996, and PhD in Computer Science from Kannur University Kerala, India. He is currently an Associate Professor in the Department of Computer Science at Mahatma Gandhi College, Kerala, India. Before this, he was a lecturer at Mary Matha College Manathavady, Kerala, from 1996 to 2001. His research interests include Signal Processing, Speech processing, image processing, and Coding. He published seven journal papers, one book, one book chapter, ten articles in international conference proceedings, and three articles in national conference proceedings all papers are related signal processing. He can be reached at shijomjose71@gmail.com.

Raju G. is currently working as a Professor in Computer Science, at CHRIST University, Yeshwanthpur Campus, Bengaluru, India. He has 30 years of teaching and 25 years of Research experience. He obtained his Ph. D. from the University of Kerala, India in 2003. He published more than 100 research articles and guided 20 PhD scholars. His area of interest includes Image Processing, Computer Vision, Machine and Deep Learning and Data Science. E-mail: raju.g@christuniversity.in

Debabrata Swain was born in Odisha, India. He obtained his B.Tech degree in Computer Science and Engineering from Bijupattnayak University of Technology, Odisha in 2008, and his M.Tech in Computer Science at Berhampur University, Odisha in 2011. He was awarded his Ph. D. in Computer Engineering from KIIT University, India in 2020. Now he is working as an Assistant Professor at Pandit Deendayal Petroleum University, Gandhi Nagar, India. Before he worked in different Univesity and engineering colleges like CHRIST Deemed to be University, Vishwakarma Institute of Technology, SRES College of Engineering Kopargaon, and Sharad Chandra Pawar College of Engineering at Pune, India. Dr Swain published more than 20 research articles in indexed journals. His research interest is Deep Learning and AI. E-mail: debabrata.swain7@yahoo.com

Om Prakash Das was born in Odisha, India. He obtained a Bachelor's in Engineering from Utkal University in 1999, MBA in Information Technology from Bharathidasan University in 2014 and an Executive Program in Business Analytics from IIM Kolkata. He is currently working for GSK as a Senior Manager in Digital and Technology. Before this, he worked as Analytic Manager at Wells Fargo and as a Manager Consulting Expert in CGI. He has extensive experience of around 18 years in the Information Technology Industry. His area of expertise is Machine Learning, Data Engineering and Business Intelligence. E-mail: ompdas@gmail.com

Biswaranjan Acharya (Senior Member IEEE) is a native of Odisha, India. He obtained his graduation in 2005 from Utkal University, Odisha, India. He received the M.C.A. degree from IGNOU, New Delhi, India, in 2009, and the M.Tech. degree in computer science and engineering from the Biju Pattanaik University of Technology (BPUT), Rourkela, Odisha, India, in 2012. He is currently pursuing a PhD degree in computer application with the Veer Surendra Sai University of Technology (VSSUT), Burla, Odisha, India. He is currently working as an Assistant Professor with the Department of Computer Engineering-AI, Marwadi University, Rajkot, Gujarat, India. He has a total of ten years of experience in academia at some reputed universities, such as Ravenshaw University and in the software development field. He has published more than 40 research articles in reputed international journals and served as a reviewer for many peer-reviewed journals. He has more than 50 patents on his credit.

Article submitted 2022-11-14. Resubmitted 2012-01-10. Final acceptance 2023-01-12. Final version published as submitted by the authors.