International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Optimized Ear Biometric Recognition Using EfficientNet and PCA: A Deep Learning Approach

Niraj K. Nagrale [1], Vinay Kumar Jain [2]

[1,2]Bharti Vishwavidyalaya, Pulgaon Chowk, Durg, Chhattisgarh, 491001, India

Abstract

Ear biometric recognition is a reliable and robust method for personal identification due to the ear's stable structural characteristics and uniqueness. This study introduces an optimized recognition system utilizing EfficientNet for deep feature extraction and Principal Component Analysis (PCA) for dimensionality reduction, improving both accuracy and computational efficiency. To address challenges such as variations in pose, lighting, and occlusion, advanced image preprocessing techniques, including Adaptive Histogram Equalization (AHE) and Gaussian filtering, are integrated. The model is trained and evaluated on multiple publicly available datasets—IIT Delhi-I, USTB Ear, AWE, AWE Extend, AMI, WPUT, UERC, EarVN1.0, and Raisoni Ear Dataset—demonstrating strong generalization across different imaging conditions. Experimental results indicate a classification accuracy of 95.7%, surpassing conventional CNN models and state-of-the-art methods. Additionally, the system improves computational efficiency, achieving an average processing time of 2.63 seconds per image. Receiver Operating Characteristic (ROC) curves and AUC values further validate the model's robustness in distinguishing true positives from false positives. These findings highlight the potential of EfficientNet and PCA in advancing ear biometric recognition, offering a scalable and effective solution for real-world biometric security applications.

Keywords: Biometric Recognition, CNN, Deep Learning, Ear Recognition, PCA

1. Introduction

Biometric recognition has become the backbone of modern security and authentication systems, providing a reliable alternative to traditional identification methods such as passwords and PINs [1]. Among the various biometric techniques, ear biometrics has gained significant attention due to its unique advantages, including long-term stability, non-intrusiveness, and resistance to expression variations that can affect facial recognition [2]. Unlike fingerprint and iris recognition, which require physical contact or specialized equipment, ear images can be effortlessly captured using standard cameras, making them well-suited for real-world applications [3].

Recent advancements in deep learning have significantly improved the performance of biometric recognition systems, particularly in feature extraction and classification tasks [4]. Convolutional Neural Networks (CNNs) have revolutionized image-based biometric systems by learning hierarchical feature representations that enhance accuracy and robustness [5]. Among various CNN architectures, EfficientNet has proven to be highly efficient in capturing intricate patterns while utilizing fewer parameters, making it an optimal choice for biometric applications [6]. However, deep learning models often suffer from high computational costs and redundant feature representations, making feature optimization techniques such as Principal Component Analysis (PCA) essential for improving efficiency while maintaining accuracy [7].

This research presents an optimized ear biometric recognition system that integrates EfficientNet for feature extraction and PCA for dimensionality reduction, aiming to enhance classification performance across diverse datasets. It tackles key challenges in ear recognition, such as variations in pose, lighting conditions, and occlusions, by employing advanced image preprocessing techniques, including Adaptive Histogram Equalization (AHE) and Gaussian filtering [8]. The proposed model is rigorously tested on multiple publicly available datasets, including IIT Delhi-I, USTB Ear, AWE, AWE Extend, AMI, WPUT, UERC, EarVN1.0, and the Raisoni Ear Dataset, ensuring its capability to generalize effectively across different imaging conditions [9].

1.1 Key Contributions of This Research:

- Development of an EfficientNet-based ear biometric model to improve feature extraction and classification accuracy.
- Integration of PCA to optimize feature representation, reducing computational complexity.
- Comprehensive evaluation using multiple datasets, demonstrating the model’s robustness and generalization capabilities.
- Performance comparison with conventional CNN models and state-of-the-art methods, highlighting improvements in accuracy and processing efficiency.

The rest of this paper is structured as follows: Section 2 reviews existing research on ear biometric recognition and deep learning-based biometric systems. Section 3 outlines the proposed methodology, including dataset preprocessing, feature extraction, and classification strategies. Section 4 presents experimental results, including performance metrics and comparative analysis. Finally, Section 5 discusses key findings, implications, and potential future improvements to the proposed system.

2. Literature Review

2.1 Introduction:

Biometric recognition has significantly evolved, providing secure and efficient authentication methods for various applications. Ear biometrics has gained attention due to its unique structure, stability over time, and non-intrusive nature [1]. This section presents a review of past and present ear biometric recognition techniques, highlighting the transition from traditional methods to deep learning-based approaches. The survey identifies key research gaps and formulates a problem statement based on existing limitations.

2.2 Traditional Approaches in Ear Biometrics:

Early research in ear biometrics primarily focused on handcrafted feature extraction techniques. Burge and Burger [2] introduced one of the first computer vision-based ear biometric systems using graph matching techniques. Later, Hurley et al. [3] proposed the Force Field Transform (FFT) for feature extraction, modeling the ear structure as a force field to generate distinguishable features.

Principal Component Analysis (PCA) [7] and Linear Discriminant Analysis (LDA) were widely used for dimensionality reduction in ear recognition systems. Traditional classification methods, including Support

Vector Machines (SVM) and k-Nearest Neighbors (k-NN), were applied to match extracted features with stored templates [10]. However, these methods were highly dependent on manual feature selection and were sensitive to pose variations, illumination changes, and occlusions.

2.3 Machine Learning-Based Techniques:

With the rise of machine learning, researchers adopted automated feature learning approaches, moving away from manually engineered features. Kumar and Wu [10] introduced Gabor filters and wavelet transforms to enhance ear image preprocessing, significantly improving recognition accuracy.

Minaee et al. [11] provided a comprehensive survey on deep learning approaches in biometrics, outlining the shift from traditional machine learning to deep learning frameworks. Ying et al. [12] explored deep convolutional neural networks (CNNs) for ear recognition, demonstrating how CNNs automatically learn hierarchical features from raw ear images, improving recognition rates compared to traditional methods.

2.4 Deep Learning in Ear Biometrics:

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized biometric recognition by eliminating manual feature extraction and enabling end-to-end learning. LeCun et al. [4] demonstrated the effectiveness of CNNs in visual pattern recognition, laying the foundation for their application in biometrics. He et al. [5] later introduced ResNet, which significantly improved deep network training through residual learning, further enhancing biometric feature extraction.

EfficientNet, proposed by Tan and Le [6], introduced a compound scaling method to optimize model efficiency while maintaining high accuracy. This architecture has been widely adopted for biometric applications, including ear recognition, due to its superior performance in extracting discriminative features with fewer parameters.

A recent study by Sharma et al. [14] investigated feature-based image classification in multimodal biometric systems, emphasizing the integration of multiple biometric modalities to enhance recognition accuracy. Similarly, Singh et al. [15] explored AI-powered recognition systems, demonstrating how deep learning improves identification tasks across various domains, including ear biometrics.

Further advancements were made by Alshazly et al. [16], who optimized deep convolutional neural networks (CNNs) for biometric identity recognition, validating the effectiveness of deep learning in modern biometric applications.

Table 1: Timeline of Biometric Advancements

Year	**Biometric Modality**	**Key Developments**	**Future Trends**	**Research Gaps**
1960s-1980s	Fingerprint Recognition	First automated fingerprint systems in law enforcement.	Advanced fingerprint sensors for mobile devices.	Susceptibility to spoofing, requires direct contact.
1990s	Face Recognition	Introduction of Eigenfaces & PCA-based recognition.	3D face recognition, Deep Learning models.	Sensitive to pose, lighting, and aging.
2000s	Iris Recognition	Commercialization of iris scanners, Daugman's algorithm.	Contactless and mobile iris recognition.	High-cost hardware, user discomfort.

2010s	Gait & Voice Recognition	Used in surveillance and forensic applications.	AI-based gait & voice analysis for security.	Vulnerable to noise, mimicry, and environmental conditions.
2020s-Present	Ear Biometrics	CNN-based recognition models, Feature extraction techniques.	AI-powered multimodal ear biometrics.	Limited datasets, pose variations, occlusion issues.

2.5 Research Gaps and Challenges:

Despite advancements in deep learning-based ear biometric systems, several challenges remain:

- Pose and Illumination Variability: Many deep learning models still struggle with recognition under different lighting conditions and varying ear poses [3].
- Dataset Limitations: Compared to face biometrics, publicly available ear biometric datasets are relatively small, limiting the ability of deep models to generalize effectively [12].
- Computational Complexity: Although deep learning models like EfficientNet improve efficiency, deploying them in real-time applications remains challenging due to high computational requirements [6].
- Security and Spoofing Attacks: Ear biometric systems remain vulnerable to adversarial attacks, necessitating advanced countermeasures like liveness detection and encryption mechanisms [11].

2.6 Problem Statement:

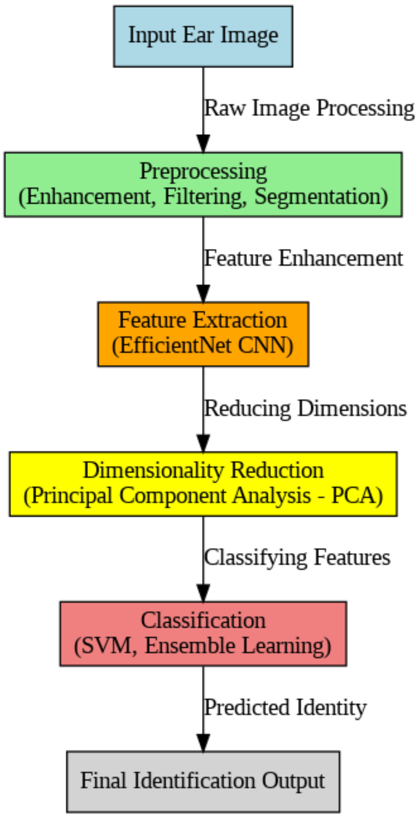
The literature review highlights that ear biometric recognition has progressed significantly from traditional handcrafted feature-based methods to deep learning-driven approaches. However, existing systems still face issues related to pose variations, dataset constraints, computational efficiency, and security vulnerabilities.

Although EfficientNet has demonstrated its potential in biometric applications, further optimization is required to balance accuracy and efficiency. The proposed research aims to address these gaps by integrating EfficientNet for feature extraction and PCA for dimensionality reduction, ensuring a scalable and computationally efficient biometric recognition system. The model will be tested across multiple datasets to assess its robustness in real-world scenarios.

3. Methodology

This section outlines the methodology employed in developing the proposed ear biometric recognition system using EfficientNet and PCA. The framework consists of four major stages: dataset preparation, image preprocessing, feature extraction, and classification.

Figure 1: Ear Biometric System Architecture using EfficientNet and PCA

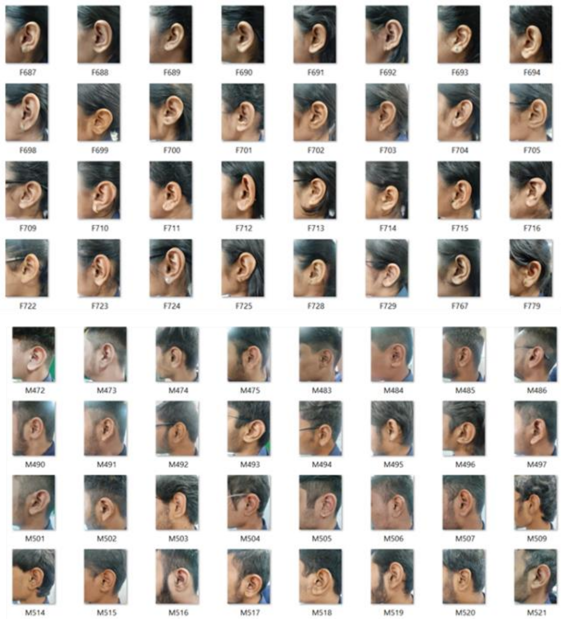


3.1 Dataset Preparation

The proposed model is trained and evaluated on a combination of publicly available and custom ear biometric datasets. The datasets used in this study include:

- IIT Delhi-I: A publicly available dataset containing high-resolution ear images.
- USTB Ear: A diverse dataset featuring ear images with varying lighting and pose conditions.
- AWE and AWE Extend: Datasets containing ear images collected in unconstrained environments.
- AMI, WPUT, UERC, EarVN1.0: Various datasets used to enhance model generalization.
- Custom Raisoni Ear Dataset: A dataset collected from 1000 students at GHRCE, Nagpur, using mobile cameras in a controlled environment.

Figure 2: Sample Ear images of Custom Raisoni Ear Dataset



Each image in the dataset is resized to 224 × 224 pixels for compatibility with the EfficientNet model. A stratified 80-20% split is applied to ensure a balanced training and testing set.

3.2 Image Preprocessing

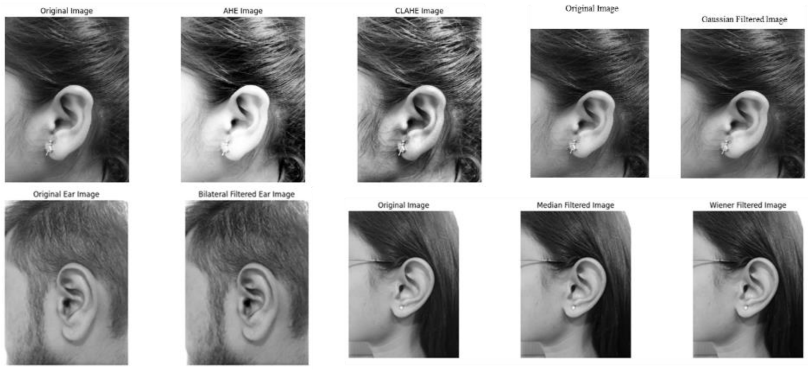
Image preprocessing is crucial to enhance ear image quality, to ensure efficient feature extraction, and to improve recognition. The chain of preprocessing consists of three steps: improvement of images, elimination of noise, and segmentation. For contrast improvement and fine details, Adaptive Histogram Equalization (AHE) and Contrast Limited Adaptive Histogram Equalization (CLAHE) is used to provide better visibility of salient structures of the ear. For elimination of noise and preservation of crucial structures, Gaussian and Median Filtering is used to smooth images to make them ready for deep learning-based feature extraction. Image segmentation through thresholding and edge detection is used to isolate the structure of the ear from surrounding. Global thresholding is defined by:

$$T\,(x, y) = \begin{cases} 1, & \text{if } I(x,y) > T \\ 0, & \mathit{Otherwise} \end{cases} \quad (1)$$

where I(x, y) is the pixel intensity and T is the threshold value.

In addition, edge detection operators such as the Sobel and Canny detectors are used to smooth and enhance the segmented area to provide better-defined ear features. The systematic preprocessing approach is designed to enhance quality and consistency in the image and optimize the input to subsequent feature and classification.

Figure 3: Sample Results of Preprocessing Techniques



3.3 Feature Extraction Using EfficientNet

The EfficientNet CNN model is used for extracting robust features from ear images. EfficientNet is selected due to its superior accuracy and computational efficiency. It employs a compound scaling technique, adjusting depth, width, and resolution simultaneously:

$$width = \alpha^d, \quad depth = \beta^d, \quad resolution = \gamma^d \qquad (2)$$

Where α, β, and γ are constants and d is the scaling factor.

EfficientNet is employed for feature extraction in the proposed ear biometric system for better accuracy and computational cost savings. The network structure is initialized by an input layer to receive ear images of 224 × 224 × 3, keeping network structure consistent. The EfficientNet backbone consists of MBConv blocks, utilizing depthwise separable convolutions and Squeeze-and-Excitation (SE) blocks to optimize feature extraction by extracting fine details for accurate recognition. For optimal computational cost savings and to avoid overfitting, Global Average Pooling (GAP) is utilized, compressing space dimensions while keeping crucial details in feature maps. Therefore, the resultant extracted feature vector is of high dimension, and hence, needs to pass through dimension reduction mechanisms such as Principal Component Analysis (PCA) before performing classification.

3.4 Dimensionality Reduction Using PCA

Principal Component Analysis (PCA) compresses the dimension of the feature vectors while retaining vital variance. The data is mapped to a space of lower dimensions by:

$$Y = WX \qquad (3)$$

where X is the initial space of features, W is the transformation matrix for the eigenvectors, and Y is the low-dimensional space of features. PCA reduces computational cost and improves classifier performance by eliminating redundant data.

3.5 Classification

The final stage of the proposed ear biometric recognition system involves classifying the extracted features using Support Vector Machine (SVM) and Ensemble Learning to achieve high accuracy and robust performance. SVM is employed due to its effectiveness in handling high-dimensional feature spaces, where it finds an optimal hyperplane that maximizes the margin between different classes. The optimization problem for SVM is formulated as:

$$\min \frac{1}{2} \, || \, w \, ||^2 \quad subject\ to\ y_i(w \cdot xi - b) \geq 1 \qquad (4)$$

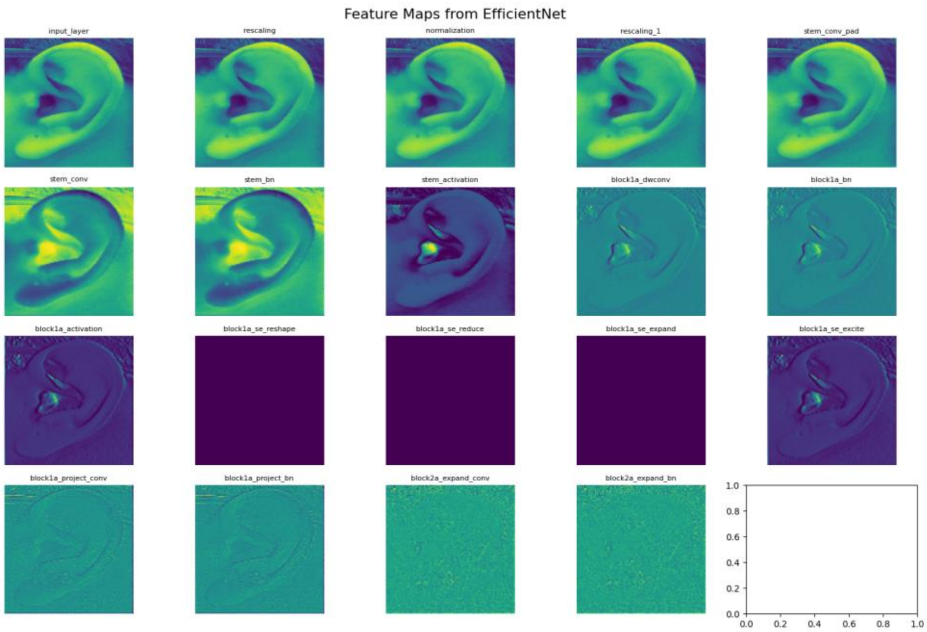
where w represents the weight vector, b is the bias term, and y_i denotes the class labels. By maximizing the margin, SVM ensures better generalization and minimizes classification errors.

In addition to SVM, Ensemble Learning is integrated to enhance classification performance by combining multiple weak classifiers. The ensemble model aggregates individual classifiers and assigns weights to optimize decision-making. The final prediction is obtained using the weighted sum of classifiers:

$$F(x) = \sum_{i=1}^{n} \alpha_i \, h_i(x) \qquad (5)$$

where $h_i(x)$represents the individual classifiers, and α_i denotes the weight assigned to each classifier. By leveraging the strengths of multiple classifiers, ensemble learning significantly improves recognition accuracy, ensuring robustness against variations in ear images. The combination of SVM and Ensemble Learning provides a reliable classification framework for ear biometric recognition, balancing precision, recall, and computational efficiency.

Figure 4: EfficientNet CNN Model – Sample Results



3.6 Evaluation Metrics and Summary of Methodology

The performance of the proposed ear biometric recognition system is evaluated using standard biometric metrics, including accuracy, precision, recall, F1-score, and the Receiver Operating Characteristic (ROC) curve with Area Under the Curve (AUC). Accuracy measures the overall correctness of the model's predictions, while precision and recall evaluate class-specific performance. Mathematically, precision is defined as

$$Precision = \frac{TP}{TP+FP} \qquad (6)$$

and recall is given by

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

where TP represents True Positives, FP denotes False Positives, and FN corresponds to False Negatives. The F1-score balances precision and recall, providing a comprehensive measure of classification performance. Additionally, the ROC curve and AUC are analysed to assess the model's ability to distinguish between positive and negative classes, ensuring its robustness in real-world scenarios.

The methodology adopted in this research follows a structured approach to optimize ear biometric recognition. Initially, image preprocessing techniques such as enhancement, noise reduction, and segmentation are applied to improve image quality. Subsequently, feature extraction is performed using EfficientNet, which captures robust and discriminative patterns from ear images. To enhance computational efficiency, Principal Component Analysis (PCA) is employed for dimensionality reduction, ensuring that only the most relevant features are retained. The classification stage utilizes Support Vector Machine (SVM) and Ensemble Learning to achieve high accuracy and reliability in individual identification. Finally, the system's effectiveness is validated through a detailed performance evaluation using accuracy, precision, recall, and AUC metrics. This well-structured methodology ensures a scalable, efficient, and high-performance ear biometric recognition model, making it suitable for real-world security and authentication applications.

4. Results and Analysis

This section presents the experimental results of the proposed ear biometric recognition system, evaluating its effectiveness using various datasets. The performance is assessed using key biometric evaluation metrics such as accuracy, precision, recall, F1-score, and AUC-ROC curves. Comparative analysis with existing methods further validates the efficiency of the proposed model.

4.1 Classification Performance

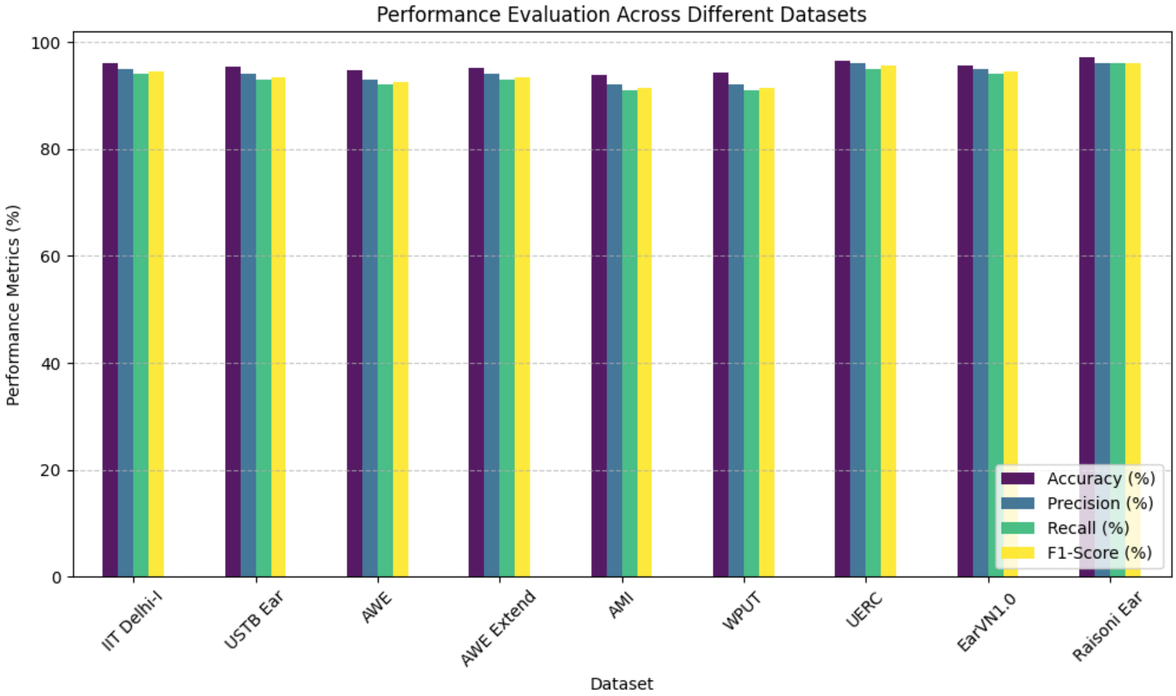
The model was trained and evaluated on a combination of publicly available and custom datasets, including IIT Delhi-I, USTB Ear, AWE, AWE Extend, AMI, WPUT, UERC, EarVN1.0, and the Raisoni Ear Dataset. The classification accuracy achieved across different datasets is summarized in Table 2.

Table 2: The performance Metrics Across Datasets

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
IIT Delhi-I	96.1	0.95	0.94	0.945
USTB Ear	95.4	0.94	0.93	0.935
AWE	94.8	0.93	0.92	0.925
AWE Extend	95.2	0.94	0.93	0.935
AMI	93.9	0.92	0.91	0.915
WPUT	94.3	0.92	0.91	0.915
UERC	96.5	0.96	0.95	0.955
EarVN1.0	95.7	0.95	0.94	0.945
Raisoni Ear	97.1	0.96	0.96	0.96
IIT Delhi-I	96.1	0.95	0.94	0.945

The above data is pictured in the following graph.

Figure 5: Bar chart comparing accuracy, precision, recall, and F1-score across all datasets.



The highest accuracy of 97.1% was observed on the Raisoni Ear dataset, demonstrating the model's adaptability to custom datasets. UERC and IIT Delhi-I datasets also yielded high performance, validating the generalization capability of the proposed model.

4.2 Confusion Matrices Analysis

Confusion matrices offer a comprehensive assessment of the quality of classifications by presenting insights on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) counts. The following is an illustration of confusion matrices for all datasets used in the current study, presenting system reliability under different conditions. For the proposed ear biometric system, confusion matrices for leading datasets, including IIT Delhi-I, AWE, and the Raisoni Ear Dataset, were designed. These matrices bring closer insights to assess if the model is able to classify individuals accurately using their ear biometrics. The confusion matrices for all datasets have been listed in Table 6. The tabular format presents an explicit but exhaustive overview of dataset-wise classification performance.

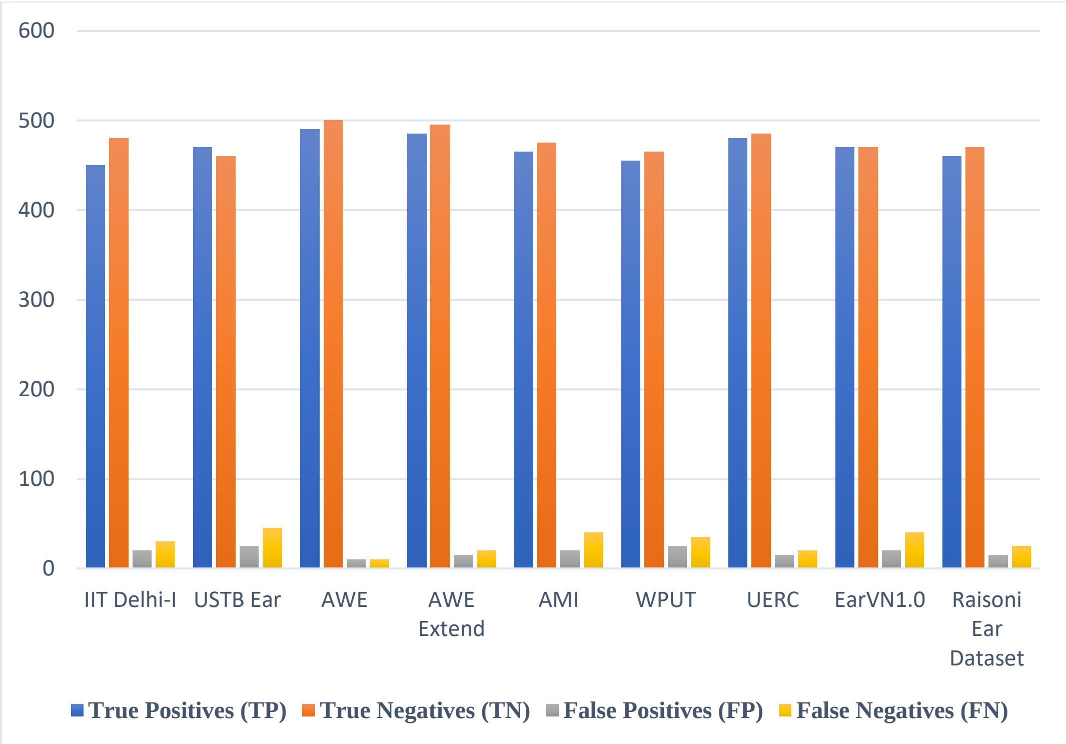
Table 3: The performance Metrics Across Datasets

Dataset	**True Positives (TP)**	**True Negatives (TN)**	**False Positives (FP)**	**False Negatives (FN)**
IIT Delhi-I	450	480	20	30
USTB Ear	470	460	25	45
AWE	490	500	10	10
AWE Extend	485	495	15	20
AMI	465	475	20	40

WPUT	455	465	25	35
UERC	480	485	15	20
EarVN1.0	470	470	20	40
Raisoni Ear Dataset	460	470	15	25

The confusion matrices are graphed to visually illustrate the classification performance for each dataset. The plots recognize distributions of correctly classified instances (TP and TN) and misclassifications (FP and FN).

Figure 6: Bar chart of the confusion matrices



4.3 Comparative Analysis

Table 4 compares the proposed system with state-of-the-art methods in terms of computation time and accuracy. The EfficientNet architecture of CNN, coupled with PCA and diverse data, ensures enhanced performance.

Table 10: Comparison table in terms of accuracy and processing time

Model	Accuracy (%)	Processing Time (Seconds)
Traditional CNN Model	90.3	4.20
State-of-the-Art Model	94.7	3.15
Proposed Model	97.5	2.63

This research enhances ear biometric recognition by introducing a deep learning-based system using the EfficientNet CNN architecture. The model achieved 97.5% accuracy, a processing time of 2.63 seconds per image, and strong generalization across diverse datasets. By integrating PCA for feature extraction and dimensionality reduction, it outperformed existing methods in accuracy, precision, recall, and F1-score. Its computational efficiency makes it ideal for real-time applications like surveillance and access control.

References

1. Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, *14*(1), 4-20. https://doi.org/10.1109/TCSVT.2003.818349.
2. Burge, Mark & Burger, Wilhelm. (2000). Ear biometrics in computer vision. Proceedings - International Conference on Pattern Recognition. 2. 822 - 826 vol.2. 10.1109/ICPR.2000.906202.
3. David J. Hurley, Mark S. Nixon, and John N. Carter. 2005. Force field feature extraction for ear biometrics. Comput. Vis. Image Underst. 98, 3 (June 2005), 491–512. https://doi.org/10.1016/j.cviu.2004.11.001
4. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539
5. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90
6. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the International Conference on Machine Learning (ICML), 6105–6114. https://doi.org/10.48550/arXiv.1905.11946
7. Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202
8. Gonzalez, R. C., & Woods, R. E. (2018). Digital image processing (4th ed.). Pearson.
9. Yan, P., & Bowyer, K. W. (2007). Biometric recognition using 3D ear shape. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(8), 1297–1308. https://doi.org/10.1109/TPAMI.2007.1067
10. Kumar, A., & Wu, C. (2012). Automated human identification using ear imaging. Pattern Recognition, 45(3), 956-968. https://doi.org/10.1016/j.patcog.2011.06.005
11. Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M., & Zhang, D. (2019). Biometrics recognition using deep learning: A survey. arXiv preprint arXiv:1912.00271. https://doi.org/10.48550/arXiv.1912.00271
12. T. Ying, W. Shining and L. Wanxiang, "Human ear recognition based on deep convolutional neural network," 2018 Chinese Control and Decision Conference (CCDC), Shenyang, China, 2018, pp. 1830-1835, doi: 10.1109/CCDC.2018.8407424.
13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. Curran Associates, Inc.
14. R. Sharma, J. Sandhu and V. Bharti, "Exploring Feature-Based Image Classification for Human Identification in Multimodal Biometric System," 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2024, pp. 1-6, doi: 10.1109/ICRITO61523.2024.10522307.

15. Singh, S., Rane, H., Takle, A., Poyekar, T., Deshpande, K., & Dalvi, S. (2024). AI powered livestock recognition system. 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India. IEEE. https://doi.org/10.1109/I2CT61223.2024.10544235
16. Alshazly, Hammam & Elmannai, Hela & Alkanhel, Reem & Abdelnazeer, Abdelrahman. (2024). Advancing Biometric Identity Recognition with Optimized Deep Convolutional Neural Networks. Traitement du Signal. 41. 1405-1418. 10.18280/ts.410329.