International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. **11** Special Issue **02 | 2026**

Advancements And Open Issues In Ear-based Biometric Identification

Ulka P. Patil[1], D. N. Besekar[2], Suhas G. Sapate[3], Seema V. Rane[4]
[1] Assistant Professor, Vidnyan Mahavidyalaya, Malkapur, Maharashtra, India
[2] Professor,Shri Shivaji College, Akola, Maharashtra, India
[3] Principal, Dr. Bapuji Salunkhe Institute of Engineering and Technology, Maharashtra, India
[4] Associate Professor, Smt. G.G. Khadse College, Muktainagar, Maharashtra, India

DOI: 10.5281/zenodo.19288581

ABSTRACT

Ear-based biometric identification has gained increasing attention as a reliable alternative to traditional biometric modalities such as face and fingerprint recognition, owing to the distinctive anatomical structure and long-term stability of the human ear. Unlike other biometric traits, ear features are less affected by aging and facial expressions, making them particularly suitable for passive and non-intrusive identification. This paper presents a structured and critical review of recent advancements in ear biometric recognition systems, with emphasis on preprocessing techniques, feature extraction methods, and classification models. Both conventional geometrical and handcrafted feature–based approaches, as well as modern deep learning–driven methods, are systematically examined and comparatively analyzed.

Recent research demonstrates significant performance improvements through the adoption of convolutional neural networks, transfer learning, and ensemble learning frameworks, which enable automatic and discriminative feature learning. Despite these advancements, the practical deployment of ear biometric systems in real-world scenarios remains challenging. Key limitations include sensitivity to occlusion caused by hair or accessories, variations in pose and illumination, limited availability of large-scale unconstrained datasets, and the high computational complexity of deep learning models. Furthermore, many existing approaches rely on multi-stage pipelines that separate ear localization and recognition, leading to error propagation and reduced system robustness.

This review highlights these challenges and identifies critical open research problems related to robustness, generalization, efficiency, and interpretability. Finally, potential future research directions are outlined, including the development of end-to-end lightweight models, occlusion-aware feature learning, cross-dataset evaluation strategies, and explainable artificial intelligence techniques, to support the design of robust, efficient, and generalizable ear biometric systems suitable for real-world applications.

Keywords:- Ear Biometrics, Biometric Identification, Feature Extraction, Deep Learning, Pattern Recognition

1. INTRODUCTION

Ear-based biometric identification has increasingly been studied within the broader domain of biometric security due to its inherent advantages such as structural stability, uniqueness, and non-intrusive acquisition. In comparison to traditional biometric traits including fingerprints, facial images, and iris patterns, ear biometrics are less affected by facial expressions and can be captured passively, making them suitable for surveillance and access-control applications.

This paper is structured as a systematic review and survey of ear-based biometric identification methods. It consolidates and critically analyzes prior research to highlight methodological trends, performance benchmarks, limitations, and open challenges. By synthesizing existing studies rather than proposing an immediate implementation, the paper aims to provide a comprehensive reference for researchers and practitioners. Furthermore, a conceptual methodology for a proposed ear recognition framework is presented to address the identified gaps and guide future research.

Biometric authentication has become a core component of modern security and surveillance systems, driven by the increasing demand for reliable and automated personal identification. Widely used biometric traits such as fingerprints, facial features, and iris patterns offer high accuracy; however, they often suffer from limitations including sensitivity to environmental conditions, vulnerability to spoofing, and the need for user cooperation. These challenges have motivated the exploration of alternative biometric modalities.

Among emerging traits, the human ear has attracted significant attention due to its distinctive anatomical structure, permanence over time, and minimal influence from facial expressions. Ear images can be acquired passively, making ear biometrics particularly suitable for surveillance and access-control scenarios [1]. A typical ear recognition system comprises ear detection and localization, image preprocessing, feature extraction, and classification stages.

Recent progress in machine learning, particularly deep learning, has substantially improved ear recognition performance by enabling automatic learning of discriminative features. Despite these advances, ear biometric systems still face unresolved challenges, especially under unconstrained conditions. This paper reviews existing ear recognition techniques, analyzes recent deep learning–based approaches, and highlights open research issues that motivate further investigation.

2. LITERATURE REVIEW

This section presents a comprehensive survey of ear biometric research, categorized into traditional feature-based methods, deep learning–based recognition approaches, hybrid and ensemble frameworks, and ear localization techniques. The reviewed studies are critically analyzed with respect to datasets used, feature extraction strategies, classification models, and reported performance.

Early research in ear biometrics primarily relied on handcrafted geometrical and structural features. Anwar et al. (2015) introduced a geometrical feature–based ear recognition method designed for passive identification [2]. Their approach employed active contour–based segmentation and extracted a compact set of shape-related features. High recognition accuracy was reported on a controlled dataset, demonstrating the effectiveness of carefully selected geometrical descriptors. However, the method showed limited robustness to occlusion and unconstrained imaging conditions.

With the advent of deep learning, convolutional neural networks (CNNs) have become the dominant approach in ear recognition. Chutani and Sharma (2025) proposed an ensemble framework that combines multiple lightweight CNN models to improve robustness and accuracy [3]. By leveraging ensemble fusion strategies, their method achieved superior performance on benchmark datasets, highlighting the benefits of model diversity. Nevertheless, the approach requires substantial computational resources and large training datasets.

Ear localization has been identified as a critical preprocessing step in ear biometric systems. Patil et al. (2024) addressed this issue by proposing a robust ear localization technique capable of operating under both controlled and uncontrolled environments [4]. While the localization accuracy was improved, recognition was not integrated into an end-to-end system, which limits real-world applicability.

Recent studies have also focused on optimizing CNN architectures and transfer learning strategies. Antonio et al. (2024) evaluated multiple CNN models and demonstrated that EfficientNet-based architectures provide improved performance under unconstrained conditions [5]. Mehta et al. (2024) proposed an ensemble-based hybrid transfer learning framework combining deep feature extraction with traditional classifiers, achieving high accuracy across benchmark datasets [6,7].

Hybrid deep learning approaches have further enhanced ear recognition performance. Mahajan and Singla (2024) proposed a CNN–BiLSTM framework to integrate spatial and contextual features, achieving strong results across multiple datasets [8]. However, increased model complexity and training cost remain key challenges.

Large-scale unconstrained datasets such as VGGFace-Ear introduced by Emersič et al. (2022) revealed significant performance degradation under real-world variations [9], emphasizing the necessity for robust, generalizable, and lightweight recognition models.

Early research in ear biometrics primarily relied on handcrafted geometrical and structural features. Anwar et al. (2015) introduced a geometrical feature–based ear recognition method designed for passive identification. Their approach employed active contour–based segmentation and extracted a compact set of shape-related features. High recognition accuracy was reported on a controlled dataset, demonstrating the effectiveness of carefully selected geometrical descriptors. However, the method showed limited robustness to occlusion and unconstrained imaging conditions.

With the advent of deep learning, convolutional neural networks (CNNs) have become the dominant approach in ear recognition. Chutani and Sharma (2025) proposed an ensemble framework that combines multiple lightweight CNN models to improve robustness and accuracy. By leveraging ensemble fusion strategies, their method achieved superior performance on benchmark datasets, highlighting the benefits of model diversity. Nevertheless, the approach requires substantial computational resources and large training datasets.

Ear localization has also been identified as a critical preprocessing step. Patil et al. (2024) addressed this issue by proposing a robust ear localization technique capable of operating under both controlled and uncontrolled environments. While their method improved detection accuracy, it did not integrate recognition into a unified framework, limiting its end-to-end applicability.

Several studies have focused on evaluating and optimizing CNN architectures for ear recognition. Antonio et al. (2024) conducted a comparative analysis of multiple CNN models and demonstrated that EfficientNet-based architectures provide improved performance under unconstrained conditions. Similarly, Mehta et al. (2024) introduced an ensemble-based hybrid transfer-learning framework that combines deep feature extraction with traditional classifiers, achieving high recognition accuracy on benchmark datasets.

Hybrid deep learning models have also been explored to capture complementary information. Mahajan and Singla (2024) proposed a CNN–BiLSTM framework that integrates spatial and contextual features for ear recognition. Their approach demonstrated strong performance across multiple datasets but involved increased model complexity and training cost.

Large-scale unconstrained datasets have further revealed the limitations of existing methods. Emersič et al. (2022) introduced a benchmark dataset for unconstrained ear recognition, showing that recognition accuracy drops significantly under real-world variations. This highlights the need for more robust and generalizable recognition models.

3. COMPARATIVE ANALYSIS OF EXISTING METHODS

A comparative examination of existing literature reveals a clear evolution from handcrafted feature–based methods to deep learning–driven frameworks. Traditional approaches offer low computational complexity but lack robustness, whereas deep learning models achieve higher accuracy at the expense of increased complexity and data dependency. Ensemble and hybrid methods further enhance performance but raise concerns regarding real-time deployment and scalability. Moreover, most studies evaluate performance on limited datasets, providing insufficient evidence of cross-dataset generalization.

The following table summarizes a comparative evaluation of existing ear recognition and localization approaches, emphasizing performance metrics, research gaps, and future research opportunities.

Table 1. Comparative Analysis of Ear Recognition and Localization Techniques with Research Gaps and Future Directions

Sr. No	Author (Year)	Methodology	Dataset	Feature Extraction	Classifier / Model	Accuracy / Performance	Research Gap	Future Directions
1	Anwar et al. (2015)	Geometrical feature–based ear recognition with active contour segmentation	IIT Delhi Ear Database	Mean intensity, centroid coordinates, Euclidean distance–based geometrical features	KNN with SAD distance	≈ 98%	Limited robustness to occlusion and unconstrained environments	Integration with deep learning and occlusion-aware feature extraction
2	Chutani & Sharma (2025)	Ensemble deep learning using multiple lightweight CNNs	IIT-D, UERC	CNN-based automatic feature learning	Weighted ensemble CNN	98.99% (IIT-D), 97.8% (UERC)	High computational complexity and large data dependency	Lightweight ensemble optimization for real-time deployment
3	Patil et al. (2024)	Robust ear localization in controlled and uncontrolled environments	Benchmark ear localization datasets	Structural and appearance-based localization features	Feature-based localization framework	Improved localization accuracy	Recognition stage not integrated	Unified end-to-end localization and recognition framework

4	Antonio et al. (2024)	CNN-based ear shape recognition with comparative model analysis	EarVN1.0	CNN-learned deep features	EfficientNetV2-S	92%	Lower robustness under occlusion	Attention mechanisms and occlusion-resilient learning
5	Mehta et al. (2024)	Hybrid ensemble transfer learning framework	Kaggle Ear, IITD-II	Deep features from VGG16 & VGG19	SVM with ensemble fusion	100% (Kaggle), 93.47% (IITD-II)	Limited cross-dataset generalization	Domain adaptation and cross-dataset validation
6	Mahajan & Singla (2024)	DeepBio: CNN + Bi-LSTM hybrid framework	IITD-I, IITD-II, AMI, AWE, EARVN1	CNN spatial + Bi-LSTM contextual features	CNN–BiLSTM	97.97%–99.37%	High model complexity and training cost	Model compression, pruning, and explainable AI
7	**Mehta et al. (2023)**	Hybrid transfer learning–based 2D ear recognition	Kaggle Ear Dataset (2600 images, 13 subjects)	Deep features extracted using fine-tuned VGG16	SVM classifier integrated with VGG16	99.23% validation accuracy	Limited evaluation on unconstrained, large-scale real-world datasets; no integrated localization module	Integration of ear localization, cross-dataset validation, and lightweight real-time deployment
8	**Emersič et al. (2022)**	Large-scale unconstrained ear dataset creation and benchmarking	**VGGFace-Ear Dataset** (extracted from VGGFace2; unconstrained images with pose, illumination, occlusion variations)	CNN-based deep features using pretrained models	Baseline CNN models (ResNet, VGG variants)	Baseline performance reported; recognition accuracy significantly lower under unconstrained conditions	Lack of robust recognition models capable of handling extreme unconstrained scenarios; dataset used mainly for benchmarking	Development of occlusion-aware, domain-adaptive, and large-scale generalizable ear recognition models

9	**Manvendra singh et al. (2022)**	Deep learning–based ear recognition using ResNeXt architecture	Benchmark ear datasets (controlled and semi-unconstrained)	Deep hierarchical features learned via ResNeXt blocks with grouped convolutions	ResNeXt CNN	High recognition accuracy reported; improved performance over traditional CNNs	Limited evaluation on large-scale fully unconstrained datasets; computational complexity remains high	Lightweight ResNeXt variants, cross-dataset validation, and integration with localization and attention mechanisms

4. RESEARCH GAPS AND OPEN ISSUES

Despite significant progress, several limitations persist in current ear biometric systems:
Robustness to Occlusion: Existing methods struggle with severe occlusion caused by hair, accessories, or masks.
Generalization Ability: Limited cross-dataset evaluation raises concerns about real-world applicability.
System Complexity: High-performing deep learning and ensemble models often require substantial computational resources.
Fragmented Pipelines: Ear localization and recognition are frequently treated as separate stages, increasing error propagation.
Lack of Explainability: Most deep learning–based systems operate as black boxes, reducing interpretability and trust.

5. PROBLEM STATEMENT

To develop an efficient and robust ear biometric recognition system capable of accurately identifying individuals under unconstrained conditions, including variations in pose, illumination, and occlusion, while maintaining low computational complexity and strong generalization across multiple datasets.

6. PROPOSED METHODOLOGY

Based on the identified research gaps, this section outlines a conceptual methodology for a proposed ear biometric recognition system. The methodology is designed to be lightweight, robust, and suitable for unconstrained environments. The proposed system is as shown in Fig-1.

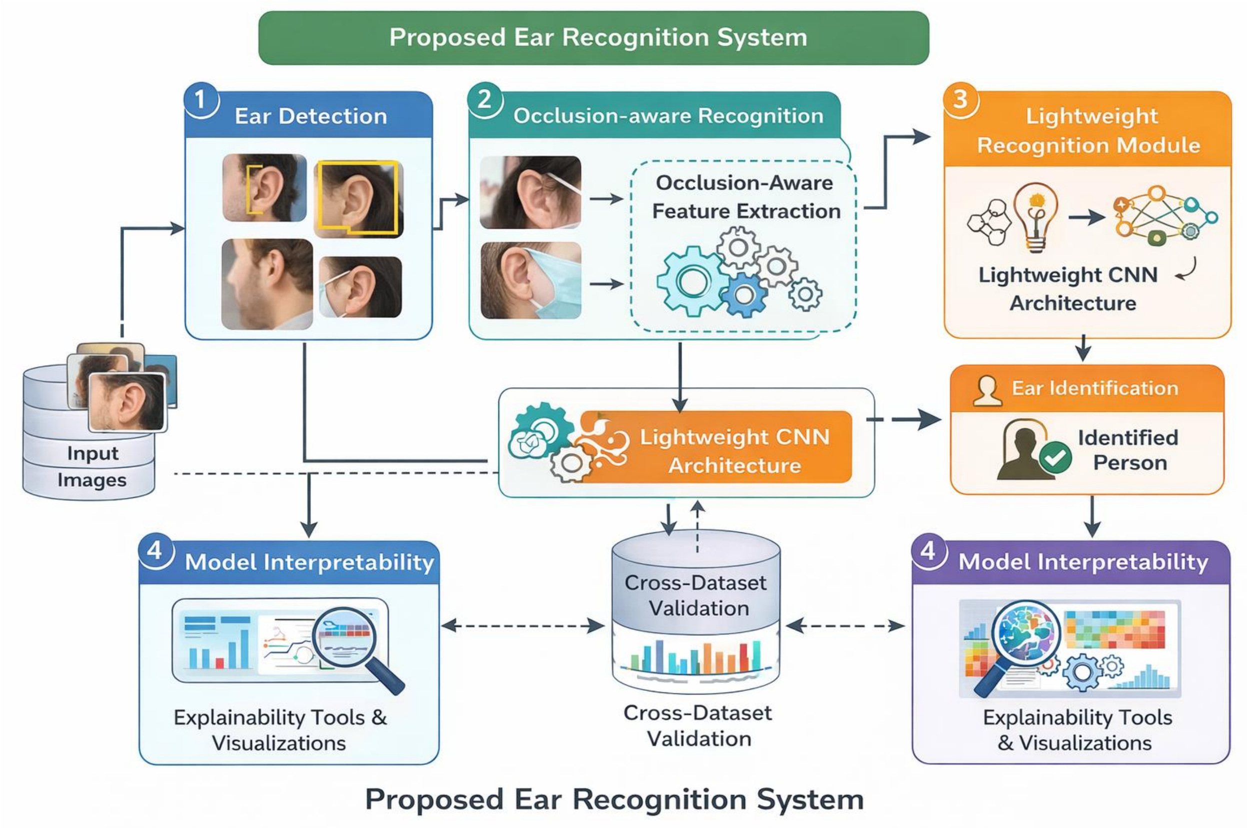

Fig -1: Proposed Ear Recognition System

System Overview

The proposed ear recognition system is an end-to-end, AI-driven biometric framework designed for accurate, robust, and deployable human identification under unconstrained conditions. The system integrates ear detection, occlusion-aware feature extraction, lightweight deep learning, and explainable AI to ensure high recognition accuracy while maintaining low computational complexity. Emphasis is placed on real-world applicability through cross-dataset validation and AI deployability, enabling seamless integration into edge devices, surveillance systems, and resource-constrained environments.

Methodology of the Proposed System

6.1 Input Image Acquisition

The system accepts input images containing human ear regions captured from multiple sources such as surveillance cameras, mobile devices, and biometric repositories. The acquisition stage accounts for variations in pose, illumination, scale, and background, ensuring robustness in unconstrained and real-world environments.

6.2 Ear Detection

An automated ear detection module localizes ear regions from side-face images. This module accurately isolates the region of interest while removing irrelevant background information, thereby reducing computational overhead and improving processing efficiency for subsequent stages.

6.3 Occlusion-Aware Ear Recognition

To handle partial occlusions caused by hair, masks, earrings, or headwear, the system incorporates an occlusion-aware recognition mechanism. This module extracts robust features by focusing on discriminative ear structures that remain visible under occlusion, significantly enhancing recognition reliability in surveillance scenarios.

6.4 Lightweight CNN-Based Feature Learning

The detected ear regions are processed using a lightweight Convolutional Neural Network (CNN) architecture. The CNN is optimized to extract high-level discriminative features with minimal parameters, low memory consumption, and reduced inference latency, directly supporting real-time AI deployment on edge and embedded platforms.

6.5 Ear Identification and Classification

The learned feature vectors are matched against enrolled ear templates using a recognition and classification module. This stage outputs the identity of the individual, enabling passive and non-intrusive biometric identification without requiring user cooperation.

6.6 Cross-Dataset Validation

The proposed system is evaluated using cross-dataset validation to assess generalization performance. Validation across multiple datasets ensures dataset independence, resistance to overfitting, and stability across demographic and environmental variations, confirming the system's robustness for real-world deployment.

6.7 Model Interpretability and Explainability

Explainable AI techniques are integrated to provide insights into the model's decision-making process. Visualization tools highlight salient ear regions, feature importance, and activation patterns, enhancing transparency, trustworthiness, and compliance with ethical AI requirements in security-sensitive applications.

6.8 AI Deployability and Practical Integration

The overall architecture is designed with AI deployability as a primary objective. The lightweight model, modular pipeline, and efficient inference enable scalable deployment across mobile devices, IoT nodes, edge systems, and surveillance platforms. Support for real-time operation and adaptability makes the system suitable for access control, smart security, and large-scale biometric applications.

6. OBSERVATIONS AND FINDINGS

The comparative analysis of existing ear recognition studies reveals that ear biometrics has evolved from traditional geometrical feature–based methods to advanced deep learning and ensemble frameworks, achieving consistently high recognition accuracy. However, the observation from the table clearly indicates that high accuracy alone is not sufficient for real-world biometric deployment. Most early approaches demonstrate strong performance only in controlled environments, while recent deep learning models, although more robust, introduce challenges related to computational complexity, generalization, and real-time applicability.
A critical observation is that ear recognition performance significantly degrades under unconstrained conditions, such as occlusion, illumination variation, pose changes, and background clutter. Although several studies employ powerful CNN architectures and ensemble learning strategies, they often rely on large datasets and high computational resources, making them unsuitable for deployment on edge devices or real-time surveillance systems. Additionally, most existing methods treat ear localization and recognition as independent tasks, which increases error propagation and reduces overall system reliability.
The inclusion of large-scale unconstrained datasets, such as VGGFace-Ear, further highlights a substantial gap between controlled experimental results and real-world performance, emphasizing the need for robust, generalizable models. Moreover, limited attention has been given to lightweight architectures, cross-dataset validation, and explainable learning mechanisms, which are essential for trustworthy and scalable biometric systems.
Therefore, this study is important because it addresses the practical limitations of existing ear recognition systems, focusing not only on accuracy but also on robustness, efficiency, generalization, and real-world applicability. By targeting these gaps, the proposed research contributes toward the development of an end-to-end, occlusion-aware, and computationally efficient ear recognition framework suitable for deployment in unconstrained and real-time biometric applications.

7. FUTURE SCOPE AND CONCLUSION

Future research should focus on developing lightweight and explainable ear biometric systems that operate reliably under unconstrained conditions. Emphasis should be placed on occlusion-aware learning, cross-dataset evaluation, and real-time deployment on edge devices. The integration of explainable AI will further support adoption in security-critical applications.
Future research should focus on developing integrated and lightweight ear recognition frameworks that combine localization and recognition. Incorporating attention mechanisms and occlusion-aware learning strategies can improve robustness under unconstrained conditions. Domain adaptation and cross-dataset evaluation are essential to enhance generalization. Additionally, integrating explainable AI techniques can improve system transparency and trust, facilitating adoption in security-critical applications.
This review and survey paper presented a structured analysis of advancements in ear-based biometric identification, covering traditional, deep learning, hybrid, and ensemble-based approaches. While recent methods demonstrate impressive recognition accuracy, their practical deployment is hindered by challenges related to robustness, computational complexity, and generalization. The proposed conceptual methodology outlines a promising direction toward addressing these challenges through an integrated, lightweight, and

explainable framework. This survey is expected to serve as a valuable reference for future research and development in ear biometric systems.

This paper presented a structured review of advancements in ear-based biometric identification, highlighting both traditional and deep learning–based approaches. While recent methods achieve high recognition accuracy, challenges related to robustness, generalization, and system complexity remain unresolved. Addressing these issues through integrated, efficient, and explainable frameworks represents a promising direction for future research in ear biometrics.

8. REFERENCES

[1] U. Patil , D. Besekar and S. Sapate, "Classification of Ear Images using Preprocessing Techniques and Gabor Feature Extraction", "Brazilian Journal of Biometrics", vol. 44 2026

[2] M. R. Anwar, M. R. Anwar, and S. M. Anwar, "Human ear recognition using geometrical features for passive biometric identification," Expert Systems with Applications, vol. 42, no. 22, pp. 8756–8766, 2015.

[3] P. Chutani and N. Sharma, "An ensemble deep learning-based approach for ear recognition," EVERGREEN – Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy, vol. 12, no. 1, pp. 399–422, 2025.

[4] S. Patil et al., "Method for human ear localization in controlled and uncontrolled environments," 2024.

[5] J. Antonio et al., "Human identification based on ear shape using convolutional neural networks," 2024.

[6] A. Mehta, R. Patel, and S. Shah, "A noble approach to 2D ear recognition system using hybrid transfer learning," 2023.

[7] A. Mehta, R. Patel, and S. Shah, "An ensemble-based hybrid transfer learning framework for effective 2D ear recognition," 2024.

[8] A. Mahajan and S. K. Singla, "DeepBio: A deep CNN and Bi-LSTM learning for person identification using ear biometrics," Computer Modeling in Engineering & Sciences, vol. 141, no. 2, pp. 1624–1649, 2024.

[9] V. Emirović, P. Štruc, and N. Pavešić, "VGGFace-Ear: An extended dataset for unconstrained ear recognition," in Proc. IEEE International Joint Conference on Biometrics (IJCB), 2022.

[10] M. Singh et al., "ResNeXt model for ear recognition and classification," 2022.