

A Comparative Analysis of Classical Approaches to Deep Learning of Ear Biometrics and Performance Based Analysis across 2-D/3-D Modalities

Mr. Manoj Kumar Dewangan^{1*}, Dr. Pankaj Mishra², Dr. Abhishek Sharma³

¹Ph.D. Scholar, Government Engineering College, Ambikapur, Chhattisgarh

Email: mnjmanit@gmail.com

²Professor, Government Engineering College, Ambikapur, Chhattisgarh

Email: pankaj.bme@csvtu.ac.in

³Assistant Professor, IIIT Naya Raipur, Chhattisgarh

Email: abhishek672400@gmail.com

***Corresponding Author:** Mr. Manoj Kumar Dewangan, Ph.D. Scholar, Government Engineering College, Ambikapur, Chhattisgarh

Email: mnjmanit@gmail.com

Received: 5th January, 2026; Revised: 18th January 2026; Accepted: 25th January, 2026; Available Online: 28th January, 2026

ABSTRACT

Ear biometrics is increasingly recognized as a viable alternative to traditional biometric modalities like face, fingerprint, and iris recognition due to its passive capture capabilities, persistent anatomical structure, and the unique geometric and textural features it provides. This article reviews 45 journal-quality publications on ear biometrics from January 2017 to July 2026, utilizing a PRISMA-guided search across various databases such as Scopus and Web of Science. Unlike previous reviews focused solely on algorithms and datasets, this synthesis evaluates performance across multiple tasks, including detection/segmentation and multidimensional recognition. It highlights task-specific metrics such as Rank-1, EER, Accuracy, Dice, and IoU, which are typically not distinguished. A notable trend towards using CNN-based transfer learning, attention mechanisms, and point-cloud algorithms is identified, with a surge in research activity observed in 2023-2024. However, the review reveals that performance metrics largely depend on the dataset used, noting that exceptional results on controlled datasets often decline on unstructured ones, indicating that dataset accuracy is not the sole indicator of robustness. The ongoing challenges include balancing accuracy with deployability, along with emerging concerns about template security, blockchain solutions, and standardized protocols. Future research directions are suggested, emphasizing domain-generalizable learning, unified benchmarks, lightweight architectures, hybrid multimodal integration, and the creation of trustworthy biometric systems.

Keywords: Ear biometrics, Deep learning, Ear recognition, 2-D/3-D modalities, Transfer learning, Multimodal biometrics, Cancelable templates, Cross-dataset generalization, Systematic review, Edge deployment.

How to cite this article: Dewangan MK, Mishra P, Sharma A., A Comparative Analysis of Classical Approaches to Deep Learning of Ear Biometrics and Performance Based Analysis across 2-D/3-D Modalities. Int J Drug Deliv Technol. 2026;16(85s): 74-97; DOI: 10.25258/ijddt.16.85s.7

Source of support: Nil.

Conflict of interest: None

Introduction

Biometric authentication relies on physiological or behavioural traits to confirm or establish identity. Despite the wide deployment of face, fingerprint, and iris recognition, each of these techniques has limitations related to contact requirements, user cooperation, pose constraints, or illumination sensitivity. Across recent ear-biometrics literature, the ear region is repeatedly

identified as a complementary trait that can be captured passively and offers distinctive geometry and texture [1], [2].

Ear-biometric research has moved from anthropometric and geometric description toward texture representation, statistical subspaces, 3-D surface registration, and, over the review window covered here, deep convolutional and transformer-based architectures. The outer ear (pinna) has

a stable anatomical structure that is largely unaffected by facial expression and changes only slowly with age, and it carries geometric and textural detail that differs across individuals, including identical twins. Because ear images can be acquired passively and without contact, ear biometrics is well suited to surveillance and unconstrained identification scenarios [1], [2] [3].

Despite these advantages, ear recognition continues to face occlusion from hair and accessories, pose and illumination variation, and a shortage of large-scale annotated datasets relative to face biometrics. Within the 2017–2026 window, research has progressed from CNN transfer-learning pipelines toward attention-gated segmentation, explainable ensembles, thermal and multimodal sensing, lightweight edge-optimised networks, and self-supervised representation learning [4], [5]; [6]; [7]; [8]; [9]; [10].

Deep learning pipelines for ear biometrics typically comprise ear localisation/segmentation, preprocessing, hierarchical feature extraction, feature refinement, and classification or verification. Convolutional architectures such as ResNet, ResNeXt, and EfficientNet, and more recently Vision Transformer and hybrid CNN-Transformer designs, learn discriminative representations directly from ear images; because ear datasets remain comparatively small, transfer learning and domain-adaptation strategies are used extensively to mitigate data scarcity [4]; [6]; [10]. Encoder-decoder segmentation networks, attention gates, and ensemble classifiers have been used to isolate the ear region from occlusion and background clutter before recognition, and stacked or ensembled lightweight CNNs have been proposed specifically to counteract the small-dataset problem [8]; [11]; [12].

In this context, the present work offers a comparative analysis of representative ear-biometric systems developed from the 2017–2026 period. Two literature-based comparators anchor the analysis: a representative 2-D deep learning pipeline optimised for computational efficiency and real-time or edge deployment (Lendering et al., 2025), and representative 3-D deep learning and geometric approaches that use point-cloud or range data [13]; [14]. By analysing these systems in terms of preprocessing, feature extraction, recognition performance, and computational complexity, this study aims to identify key design trade-offs and provide insights for future research in ear-biometric recognition. Nevertheless, the reported evidence remains difficult to interpret because identification accuracy, verification

error, segmentation overlap, and dataset-specific challenge scores are frequently presented together even though they answer different technical questions.

Several comprehensive surveys already exist in this space, and this review is deliberately positioned against them rather than duplicating their scope. [1], [2] each provide broad taxonomies of 2-D, 3-D, and combined ear-recognition approaches spanning several decades of literature, and [13] focus specifically on 3-D techniques. [3] conduct a systematic literature review of the preceding decade of ear-biometric algorithms and trends; [8] subsequently note that this review, while comprehensive in cataloguing algorithms, does not analyse the practical challenges of deploying ear-biometric systems in dynamic, real-world conditions such as varying head pose or occlusion.

Novelty: The novelty of the present review lies in four specific contributions that distinguish it from the surveys summarised in Table 1. Unlike existing ear biometric surveys that primarily categorize algorithms and datasets, this review provides an evidence-based comparative analysis of recent advances by examining the influence of datasets, evaluation protocols, and sensing modalities on reported performance. It demonstrates that differences in benchmark characteristics and experimental protocols significantly affect recognition outcomes, making direct comparison of published accuracies unreliable. Furthermore, the review integrates recent developments in lightweight edge deployment, multimodal and thermal ear recognition, presentation attack detection, and cancellable biometrics, highlighting their role in enabling practical and secure biometric systems. Finally, it identifies the major unresolved challenges, including cross-dataset generalization, protocol standardization, explainability, and privacy-preserving recognition, and presents future research directions based on the collective evidence from recent studies. Table 1 summarises how these four contributions map onto the specific gaps left open by prior surveys.

Table 1 Comparison with Existing Reviews

Review	Scope	Primary contribution	Gap addressed by the present review
Emeršič et al. [15]	2017; broad 2-D ear recognition	Databases, handcrafted and	Predates most transfer-

		emerging learned methods	learning, explainability, fairness and privacy-preserving work.
Kamboj et al. [1]	2022; survey plus deep approach	Deep recognition and unconstrained challenges	Does not provide a 2017-2026 journal-only, deployment analysis studies.
Benzaoui et al. [2]	2023; comprehensive survey	Databases, approaches and open challenges	Broad coverage, but task metrics and secure deployment evidence are not synthesized as a unified evaluation layer.
Ganapathi et al. [16]	2023; 3-D-focused survey	3-D acquisition, representation and matching	Restricted to 3-D and does not integrate 2-D deep, thermal, multimodal and cancelable systems.
Oyebiyi et al. [3]	2023; systematic review	Approaches, algorithms and publication trends	Covers the preceding decade but not the 2024-2026 shift toward attention, synthetic data, bilateral modeling and cancelable templates.

Present review	2017-July 2026; 50 journal studies	Protocol-aware comparison across detection, recognition, 3-D, thermal, multimodal and secure systems	Adds task-metric separation, cross-dataset interpretation, evidence outcomes and a deployment-oriented research roadmap.
----------------	------------------------------------	--	--

Review Methodology

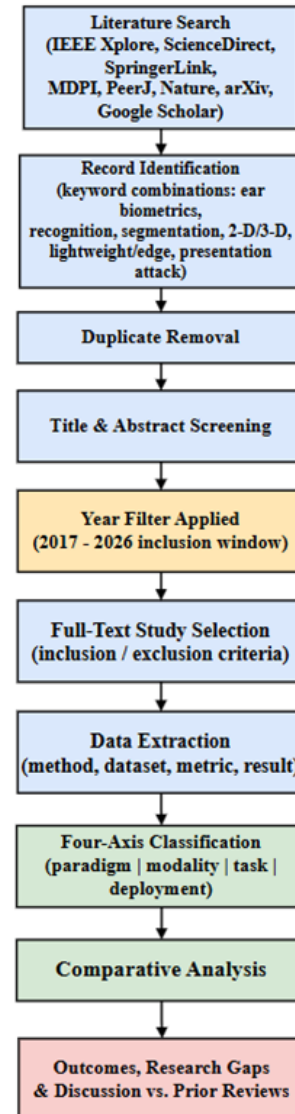


Figure 1 Flowchart of the Proposed Review Workflow

1.1 Search Strategy and Review Scope

Scopus, Web of Science, IEEE Xplore, ScienceDirect, Springerlink, PubMed/PMC, MDPI, and GoogleScholar are among the interest journals that were taken into consideration when organizing the search strategy for articles published in journals from January 2017 to July 2026. "ear biometric" OR "ear recognition" OR "ear identification" OR "ear verification" AND the keywords detection, segmentation, convolutional networks, transfer learning, metric learning, 3-D, point clouds, thermal imaging, and multimodal fusion make up the primary search string used in these journals. In addition, backward and forward citation chaining will be employed to identify any extensions of pertinent journal articles.

1.2 Eligibility Criteria

The articles considered in the review were those that were written in the English language, published in a peer-reviewed journal within the stipulated time period, concerned with human ear recognition/detection/segmentation/identification/verification, three-dimensional (3-D) ear modeling, multimodal ear recognition, and/or ear template security. Articles published before 2017, conference papers, books/chapters, preprints, research work related to medical diagnosis of the ear, duplicated publications, and performance results not linked to a specific dataset/metric were excluded from the list of articles for this review.

1.3 Screening and Data Extraction

After the literature search, all retrieved papers were imported to the reference management database and duplicates were deleted. The studies that remained were screened through the two-step process in line with PRISMA methodology. At the first step, titles and abstracts were analyzed individually and all papers not relevant to the topic of ear biometrics, ear medical diagnosis, conference papers, book chapters, editorials and non-English papers were excluded. Then the full texts of the selected papers were analyzed in accordance with pre-defined inclusion and exclusion criteria. Only the peer-reviewed journal papers published from 2017 to 2026 and containing experimental methodology and quantitative assessment were chosen for analysis. For each paper, the standardized data extraction form was filled in order to ensure consistency in the review process. The data extracted from the papers included: year of publication; journal source; biometric task (detection, segmentation, recognition, verification or security); used datasets; preprocessing methods; feature

extraction/learning algorithms; evaluation process; performance metrics; limitations; and contributions.

1.4 Synthesis and Quality Control

A qualitative evidence synthesis was conducted to examine the selected literature in terms of their methodological features, experimental conditions, and outputs. Rather than judging only on the basis of recognition accuracy, the comparative assessment involved other parameters such as dataset, sensing modality, testing approach, computational efficiency, robustness, and security techniques employed. Research works have been divided into main areas of research such as detection and segmentation, 2D recognition, 3D recognition, multimodal biometrics, thermal vision, light weight deep learning and secure ear biometrics. In order to increase the validity of the review, each paper has been critically assessed based on the set of quality criteria, such as clarity of experiment methodology, availability of public description of benchmark datasets used, transparency of the evaluation protocol, provision of quantitative performance results, and consideration of limitations of the study. Papers with insufficient methodology description or without linking their performance results with a particular dataset and evaluation protocol have been excluded from the eligibility criteria check.

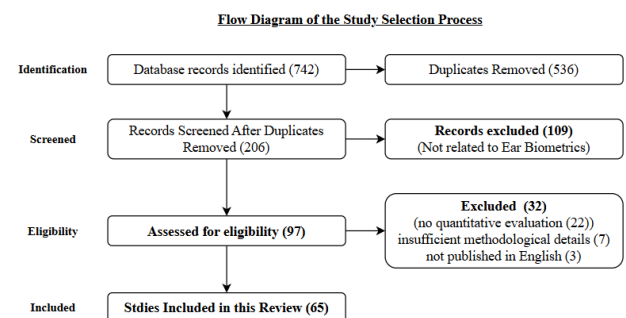


Figure 2 Flow diagram of the Study Selection Process

2 Evolution of Ear Biometric Recognition

2.1 Publication and Methodological Trends

The study's findings clearly show a stage shift from early 2D representation feature fusion to unregulated 2D recognition, explainable models and lightweight ensemble, attention-based segmentation and bilateral learning, point cloud, and secured templates. The trend reached its zenith in 2024, when it started to incorporate security techniques, apply various ensemble fusion strategies, and integrate transfer learning concepts with segmented classification. Multiple angle acquisition, ear sketch, soft biometrics, and 3D face/ear verification are

the areas covered by the research on articles from 2025 and 2026. Some have gone on to develop applications for cancelable feature fusion and biometrics in the Internet of Things.

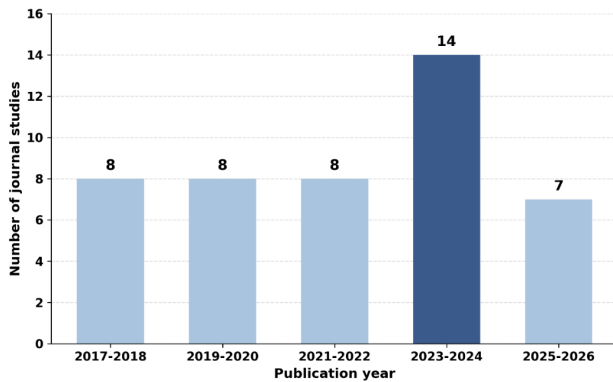


Figure 3 Distribution of the reviewed ear recognition studies by publication period (2017–2026)

Table 2 Evolution of research themes in ear recognition literature from 2017 to 2026.

Period	Studies	Dominant themes
2017-2018	8	Detection/segmentation, unconstrained recognition, domain adaptation, light-field and 3-D representations
2019-2020	8	Handcrafted-versus-CNN evidence, deep ensembles, large-scale datasets, security and resource analysis
2021-2022	8	Explainability, context-aware detection, real-time pipelines, multimodal fusion, EfficientNet and thermal imaging
2023-2024	14	Cross-dataset transfer, 3-D multimodal authentication, lightweight ensembles, GAN augmentation, attention segmentation and blockchain protection
2025- 2026	7	Point-cloud multimodality, soft biometrics, ear sketches, bilateral symmetry, multi-angle capture and cancelable templates
Total	45	Analytical journal studies included in this review

The temporal distribution and themes changing process of the 45 journal articles involved in ear recognition review (2017-2026) were showed as Figure 3 and Table 2. From Figure 3, research activities was high during 2023-2024 which has 14 articles compared with earlier stages (2017-2022), of which there are 8 articles each year, there are 7 articles research were carried out during 2025-2026. From Table 2, the themes of ear recognition progresses from detection/segmentation/handcrafted features toward transfer learning, multi-biometrics, GAN based data augmentation, xAI and etc.; lightweight model, block-chain based security and cancelable biometric template are also interesting research topics.

2.2 Detection, Segmentation and Image Preparation

Recent detection research treats localization as a major source of downstream error rather than assuming a perfectly cropped ear. Geometric morphometrics combined with CNN detection established an early recent baseline [17]. Pixel-wise encoder-decoder segmentation then improved localization in uncontrolled images [18], while context-aware models incorporated face-part information [19]. Attention-gated ensembles and U-Net pipelines further emphasized partial occlusion and complex backgrounds [8][12]. The evidence confirms that detection accuracy, Dice, and IoU quantify localization quality; they must not be compared directly with identity-recognition accuracy.

2.3 Recognition Architectures and Domain Generalization

Between 2018 and 2024, the field moved from handcrafted-deep fusion and domain adaptation [20] [21]to fully learned representations, pairwise SVMs, ensembles, transfer learning, metric learning, and lightweight networks [22][23][24][25][4][26][6],[27][28][29][30][31][31][32]. Hybrid methods remain relevant because they can combine complementary descriptors or reduce the data required for end-to-end training. However, the same architecture often behaves very differently across datasets. The explainable residual-network study reported 99.64% on AMI but 67.25% on AWE [24], indicating that acquisition variability and protocol difficulty dominate many cross-paper comparisons.

2.4 3-D, Thermal, Multimodal and Secure Ear Biometrics

Recent journal research also broadens the sensing and deployment space. Curvilinear 3-D features and light-

field representations aim to reduce sensitivity to illumination and viewpoint[33][34]. Thermal imaging supports recognition in low illumination [35], while ear-profile face and 3-D face-ear fusion increase complementary identity evidence[36][37][14]. Security studies progress from template hardening[38] to blockchain protection, cancelable EarCodes, and optimized cancelable deep-feature fusion [39][40]. This progression demonstrates that recognition accuracy alone is no longer sufficient for deployment-oriented evaluation.

2.5 Datasets and Benchmark Conditions

The ear datasets vary considerably in size, image quality, data acquisition conditions, sensing mode, occlusion, pose, and identity distribution. Recently, researchers have extended the benchmarking of algorithms from visible light images in controlled environments to large-scale uncontrolled web image datasets, context-aware detection datasets, infrared images, and specially created datasets for segmentation and other applications [41], [42], [43], [44], [45], [46], [18], [4], [35], [47], [11]. As a result, the performance measure is valid only if accompanied by the corresponding dataset and protocol. Table 2 presents a representative list of such datasets.

Table 3 Representative ear-recognition datasets and benchmark characteristics

Dataset	Subjects	Images	Conditions	Recommended Split / Notes	Source
IIT Delhi-I	125	493	Controlled, grayscale, 3-6 images per subject	A 60/40 subject-disjoint split is commonly used for training and testing	[46]
IIT Delhi-II	221	793	Controlled /semi-controlled, grayscale	A 60/40 subject-disjoint split is commonly used for training and testing	[46]

AMI	100	700	Indoor; pose variation; right and left ears	Usually evaluated using 70/30 or 80/20 subject-disjoint splits	[48]
AWE	100	1,000	Unconstrained web images with pose, illumination, occlusion, and background variation	5-fold subject-disjoint validation is generally recommended	[15]
WPUT	501	2,071	Both ears, variable pose, resolution, and occlusion	5-fold subject-disjoint cross-validation is commonly applied	[49]
EarVN1.0	164	28,412	In-the-wild colour images of both ears	Usually evaluated using 70/30 or 80/20 subject-disjoint splits	[43]
UERC 2017	3,704	11,804	Large-scale unconstrained benchmark	Standard UERC evaluation protocol should be followed	[41]
UERC 2023	650	14,004	Newly collected in-the-wild training data; sequestered evaluation	Standard UERC evaluation protocol should be followed	[45]

			data		
--	--	--	------	--	--

2.6 Dataset Comparison Across Studies

Dataset scale and acquisition conditions strongly influence reported performance. Controlled datasets such as IIT Delhi and AMI are useful for architecture development, whereas AWE, EarVN1.0, WPUT, and UERC expose larger pose, illumination, background, occlusion, and resolution variation. Recent studies increasingly evaluate cross-dataset transfer, but training and testing conventions remain inconsistent. Consequently, every result in this review is presented with its task, dataset, and metric, and cross-dataset results are interpreted as evidence of robustness rather than as a universal ranking.

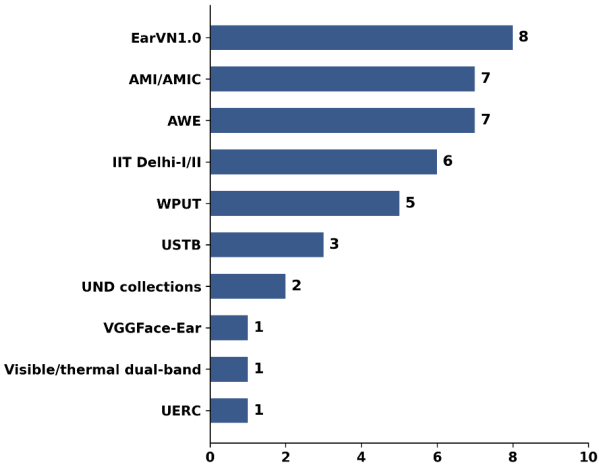


Figure 4 Reference frequency by dataset family across the reviewed corpus

Table 4 Overview of commonly used ear recognition datasets and their evaluation characteristics.

Dataset family	Conditions	Typical use	Interpretive caution	References
IIT Delhi-I/II	Controlled or semi-controlled 2-D ear images	Closed-set identification; frequent transfer-learning benchmark	High results may overstate real-world robustness if subject splits are not	[23][26][29][30][11][40]

			explicit	
AMI/AMIC	Indoor ear images with pose and bilateral variation	Identification and explainability	Strong performance in several studies, but markedly easier than web-scraped data	[25][50][44][36][37][51][40]
AWE	Unconstrained web images; pose, illumination, occlusion and background variability	Identification, transfer and verification	Useful for robustness; scores are often lower than on AMI/ITD	[20][21][4][41][52][39][40]
WPUT	Larger controlled/semi-controlled set with bilateral and pose variation	Identification and feature analysis	Performance can fall sharply relative to smaller controlled sets	[53][25][50][54][6]
EarVN1.0	28,412 in-the-wild ear images from	Large-scale unconsented identification	Supports transfer and ensemble	[43][4][39][55][17,19,43,50]

	164 identities	cation	ble analysis; resolution and background are challenging	
USTB	Multiple controlled subsets with pose and illumination variation	Identification and secure metric learning	Protocol details vary across studies	[23][56][57]
UND collections	3-D ear range data	3-D identification and surface analysis	Requires depth acquisition and careful registration	[34][14]
VGGFace-Ear	Large unconstrained ear dataset derived from face imagery	Transfer learning and cross-dataset analysis	High intra-class variability; identity and crop quality require careful control	[44]
Visible/thermal dual-band	Visible and infrared ear	Identification and verification	Specialized sensors limit	[35]

data	crops	tion under illumination change	direct comparison with visible-only systems	
UERC	Large unconstrained benchmark used by recent journal security models	Large-scale recognition and cancelable templates	Challenge difficulty exposes pose, occlusion and low-resolution failures	[55]

3 Comparisons

Table 5 Detection, Segmentation, Representation and Benchmark Construction

Study	Task	Method	Data	Findings	Contribution	Limitation
Cintas et al., 2017 [17]	Detection / feature extraction	Geometric morphometrics + CNN	Ear-image datasets used by the authors	$R^2 = 0.709$, RMS E = 2.296, Explained Variance = 0.976, Pearson Correlation = 0.988,	Combined anatomical geometry with learned localization	Dataset and acquisition details limit generalization.

				Processing Time = 4.68 ms/image		
Eme ršič et al., 2018 [18]	Pixel-wise segmentation	Convolutional encoder-decoder network	web (a.k.a. in the wild)	Accuracy of 99.4 %, average IoU of 55.7 %, average precision of 67.7 % and the average recall of 77.7 %	Moved beyond bounding boxes to detailed masks	Segmentation metrics are not identity metrics.
Sepas-Moghaddam et al., 2018 [33]	Light-field representation	Light-field ear imaging framework	Light-field ear data	CMS ₁ (Rank-1) = 88.2 %, CMS ₂ = 90.9 %, CMS ₃ = 92.9 %	Introduced a richer capture perspective	Requires specialized acquisition.

				%		
Eme ršič et al., 2021 [19]	Context-aware detection	ContextedNet	AWE and UBE AR	Detection Success Rate = 99.63 %, Average IoU = 95.51 %	Uses facial context to improve ear localization	Dependent on visible contextual face parts.
Ramos-Coo per et al., 2022 [44]	Dataset / transfer benchmark	VGG Face-Ear + transfer learning	Unconstrained ear datasets	Best Rank -1 Recognition = 97.2 %, Rank -5 Recognition = 99.5 %	Expanded data scale and variability	Derived crops can inherit face-dataset biases.
Alex et al., 2024 [8]	Attention segmentation	VGG-16, dense blocks, attention gates, two-YSeg Net ensemble	Combined EarV N1.0, AMI and face-image data	Accuracy 98.93 %; precision 98.50 %; recall 98.10 %; Dice 98.30 %; IoU 96.70 %	Strong segmentation under complex backgrounds	Not person-identification accuracy; higher computational cost.

Lahkar and Borbora, 2024 [12]	Unconstrained segmentation	U-Net	Custom set and AWE	Custom: 92.38% accuracy, 79.33% IoU; AWE: 76.20% IoU	Provides cross-dataset segmentation evidence	Substantial IoU decline on AWE
Bouaouina et al., 2024 [32]	Data augmentation / interpretability	GAN augmentation + mean class activation maps	Ear-recognition benchmarks	MAI: Rank -1 = 100.0%, Rank -5 = 100.0%, AUC MC = 100.0%; AWE: Rank -1 = 76.25%, Rank -5 = 91.25%, AUC MC = 99.96%	Combines synthetic data with discriminative-region analysis	Synthetic realism and identity preservation require validation.
Freire-Obregón et	Cross-modal representation	Multilevel ear sketches +	AMI, AWE, IITD-II and	Rank -1 Recognition:	Explores sketch-to-image	Generated sketches may

al., 2025 [51]	on	triplet learning	BIPLab	AMI = 99.10%, AWE = 95.70%, IITD-II = 98.40%, BIPLab = 98.80%; EER: 1.24%, 4.63%, 2.11%, 1.86%	matching and abstraction	introduce representation bias.
Fukuda et al., 2025 [58]	Multi-angle acquisition	CNN + Grad-CAM	New multi-angle ear images	Best Performance: Accuracy = 99.6%, Precision = 99.7%, Recall = 99.6%, F1-score = 99.7%;	Shows feasibility beyond frontal views	Private dataset and limited population constrain replication.

				all multi-angle metrics >98 %		
--	--	--	--	-------------------------------	--	--

Table 6 detection, segmentation, representation and benchmark construction

Study	Task	Method	Data	Finding s	Contribution	Limitation
Cintas et al., 2017 [17]	Detection / feature extraction	Geometric morphometrics + CNN	Ear-image datasets used by the authors	$R^2 = 0.709$, RMSE = 2.296, Explained Variance = 0.976, Pearson Correlation = 0.988	Combined anatomical geometry with learned localization	Dataset and acquisition details limit generalization.
Emeršič et al., 2018 [18]	Pixel-wise segmentation	Convolutional encoder-decoder network	Web/in-the-wild ears	Accuracy = 99.4 %, IoU = 55.7 %, Precision =	Moved beyond bounding boxes to detailed masks	Segmentation metrics are not identity metrics.

				67.7 %, Recall = 77.7 %		
Sepas-Moghaddam et al., 2018 [33]	Light-field representation	Light-field ear imaging framework	Light-field ear data	CM S ₁ (Rank-1) = 88.2 %, CM S ₂ = 90.9 %, CM S ₃ = 92.9 %	Introduced a richer capture perspective	Requires specialized acquisition.
Emeršič et al., 2021 [41]	Context-aware detection	ContextedNet	AWE and UBE AR	Detection Success Rate = 99.63 %, Average IoU = 95.51 %	Uses facial context to improve ear localization	Dependent on visible contextual face parts.
Ramos-Cooper et al., 2022 [44]	Dataset / transfer benchmark	VGG Face-Ear + transfer learning	Unconstrained ear datasets	Rank-1 = 95.00 %, Rank-5 = 98.81 %, AUCM	Expanded data scale and variability	Derived crops can inherit face-dataset biases.

				C = 97.99%		
Alex et al., 2024 [8]	Attention segmentation	VGG-16, dense blocks, attention gates, two-YSeg Net ensemble	Combined EarV N1.0, AMI and face-image data	Accuracy 98.93%; precision 98.50%; recall 98.10%; Dice 98.30%; IoU 96.70%	Strong segmentation under complex backgrounds	Not person-identification accuracy; higher computational cost.
Lahkar and Borbora, 2024 [12]	Unconstrained segmentation	U-Net	Custom set and AWE	Custom: 92.38% accuracy, 79.33% IoU; AWE: 76.20% IoU	Provides cross-dataset segmentation evidence	Substantial IoU decline on AWE.
Bouaouina et al., 2024 [32]	Data augmentation / interperability	GAN augmentation + mean class activation maps	Ear-recognition benchmarks	MAI: Rank-1 = 100.00%, Rank-5 =	Combines synthetic data with discriminative-region analysis	Synthetic realism and identity preservation require validation

				100.00%, AU CM C = 100.00%; AWE: Rank-1 = 76.25%, Rank-5 = 91.25%, AU CM C = 99.96%	is	tion.
Freire-Obregón et al., 2025 [51]	Cross-modal representation	Multilevel ear sketches + triplet learning	AMI, AWE, IITD-II and BIPLab	AMI (mAP = 40.6%, Rank-1 = 24.2%, Rank-5 = 64.7%, Rank-10 = 81.3%); AWE	Explores sketch-to-image matching and abstraction	Generated sketches may introduce representation bias.

				(33.8%, 20.3%, 46.2%, 67.3%); IIT DII (33.8%, 20.0%, 51.4%, 69.1%); BIP Lab (39.2%, 23.0%, 60.3%, 78.1%).		
Fukuda et al., 2025 [58]	Multi-angle acquisition	CNN + Grad-CAM	New multi-angle ear images	Accuracy = 99.6%, Precision = 99.7%, Recall = 99.6%, F1-score = 99.7%	Shows feasibility beyond frontal views	Private dataset and limited population constrain replication.

Table 7 Comparison of 2-D, Hybrid, and Deep Learning-Based Ear Recognition Methods

Study	Architecture	Dataset / protocol	Metric	Findings	Limitation
					
					
					
					
Hansley et al., 2018 [20]	CNN normalization + handcrafted/deep fusion	Unconstrained database including UE RC-related data	Rank-based identification	CNN + HOG fusion achieved Rank-1 = 75.6%, Rank-5 = 90.6%, AUC = 0.972, EER = 22.87% on AWE; under the UERC overall protocol it achieved Rank-1 = 49.06%, Rank-5 = 69.94%, AUC = 0.951, EER = 27.84%.	Performance decreases considerably under large-scale unconstrained UERC protocols.
Eyikur et al., 2018 [21]	Domain-adapted CNN	Cross-domain ear datasets	Identification	Domain adaptation improved Rank-1 accuracy from $\approx 49\%$ to $\approx 59\%$ (+10%), while CNN fusion provided an additional $\approx 4\%$ improvement. Dataset-bias classification accuracy	Performance is still affected by unseen acquisition conditions and cross-dataset domain shift.

				reached 99.71%.	
Dodge et al., 2018 [22]	Deep neural network	Unconstrained ear data	Identification	achieved Rank-1/Rank-5 of 68.50%/83.00% (AWE), 94.08%/99.69% (CVLE), and 80.03%/93.48% (AWE + CVLE).	Fine-tuned networks are prone to overfitting on small datasets, and performance depends on sufficient training data and augmentation.
Omara et al., 2018 [23]	Hierarchical deep features + pairwise SVM	IITD-I/II and US TB-I/II	Recognition rate	99.86%, 99.82%, 99.44%, 99.57% in the reported scenario	Results should not be transferred to web-scale settings.
Hassballah et al., 2019 [53]	LB P variants + nearest-neighbor classifiers	IITD, AMI, WP UT and AWE	Identification/verification	Best constrained results approached 99%	Performance varies significantly across datasets, with reduced accuracy on unconstrained ear images.
Alsahaly et	Handcrafted desc	Three ear data	Identification	CNN-based models improved performance	Performance is influenced by dataset

al., 2019 [24]	riptors vs CNNs	sets		by about 22%	characteristics and requires sufficient labeled training data for CNN-based models.
Alsahaly et al., 2019 [25]	VG-like transfer-learning ensembles	AMI, AMIC and WP UT	Identification	Ensemble achieved 100.00% (AMI), 99.33% (AMIC), and 99.17% (WPUT), outperforming all individual CNN models.	The ensemble increases computational complexity and was evaluated only on limited ear datasets, leaving cross-dataset generalization insufficiently explored.
Alsahaly et al., 2020 [4]	Fine-tuned ResNet101 ensemble	Ear VN 1.0	Rank-1 accuracy	Single model 93.45%; ensemble 95.85%	No evidence of universal cross-dataset performance.
Omara et al., 2020 [50]	Distance metric learning +	AWE, AMI and WP UT	Classification accuracy	98.79%, 98.70%, 84.30%	Classical/metric learning, not deep end-to-end recognition.

	DA G- SV M				
Em erši č et al., 202 0 [59]	Mo del eval uati on and com plex ity anal ysis	Mul tiple ear- reco gniti on mod els	Accur acy, compl exity and resour ce use	SE-ResNet-101 achieved Rank-1 = 98.48%, Rank-5 = 99.67%, and AUCMC = 99.91%, while MobileNetV2 and ShuffleNetV2 provided substantially lower computational cost.	High- performin g models require significant ly greater computati onal resources, making deployme nt on resource- constraine d devices challengin g.
Pri yad har shi ni et al., 202 1 [26]	Six- laye r CN N	IIT D-II and AM I	Recog nition rate	97.36% and 96.99%	Limited external validation.
Als haz ly et al., 202 1 [54]	Resi dual - net wor k ense mbl es + Gui ded Gra d- CA M	AM I, AM IC, WP UT and AW E	Recog nition accura cy	99.64%, 98.57%, 81.89%, 67.25%	Large cross- dataset degradatio n.
Alk	Fast	Real	Recog	97.80%	Protocol

aba bji and Mo ha m me d, 202 1 [60]	er R- CN N + CN N + PC A/g enet ic sele ctio n	- time test set	nition accura cy		and hardware- specific result.
Bo oys ens and Vir iri, 202 2 [17]	Effi cien tNet -B8	Ear- reco gniti on data used by sour ce	Classi ficatio n accura cy	98.45%	Dataset- specific and computati onally intensive.
Sha rka s, 202 2 [27]	Cur vele t feat ures and dee p- net wor k ense mbl es	Ear data sets in sour ce	Classi ficatio n rate	Achieved 99.75% (IITD-II), 99.58% (USTB-II), and 92.63% (AWE), outperforming individual deep models.	Performan ce decreases on unconstrai ned datasets such as AWE despite strong results on controlled datasets.
Aia di et al., 202 3 [28]	MD FNe t ligh twei ght uns uper vised net	Ear- reco gniti on ben chm arks	Recog nition	Improved Accuracy by 75.4%, F1-score by 33.2%, and runtime by 70.2% compared with the EfficientViT-B3 baseline	Evaluated only on endoscopi c image datasets; applicabili ty to other imaging domains remains unverified.

	work			while maintaining competitive IoU.	
Freire-Obregon et al., 2023 [52]	Zero-shot cross-dataset transfer	Multiple unconsolidated datasets; mobile focus	Cross-dataset recognition	Best zero-shot transfer achieved Rank-1 = 79.40%, Rank-5 = 94.60%, and mAP = 76.50% (AWE → AMI); AMI → AWE achieved Rank-1 = 73.80%, Rank-5 = 90.50%, and mAP = 69.70%.	Recognition performance remains sensitive to domain shift because no target-domain retraining is performed.
Zarachoff et al., 2023 [56]	Multiband PC A	IIT D-II and US TB-I	Rank-1 improvement	56.41% and 51.11% improvement over standard PCA	Improvement is relative to a weak PCA baseline.
Mehta and Singh, 2024 [29]	Stacked lightweight CNN ensemble	IIT D-II	Classification accuracy	98.74%	Single-dataset evidence.
Mehta and Singh, 202	Transfer learning +	IIT D-II	Accuracy / F1	Weighted 95.96%; average 93.08%; F1 95.96%	No large external benchmark.

4 [30]	lightweight CNN ensemble				
Mehta and Singh, 2023 [61]	Deep CNN + augmentation	2-D ear data	Recognition	Achieved 98.74% recognition accuracy using data augmentation and a deep CNN architecture.	Performance was evaluated on a single controlled dataset, limiting evidence of generalization to unconstrained environments.
Mahajan and Singla, 2024 [31]	Deep CNN + Bi-LSTM (DeepBio)	Five ear datasets	Accuracy	97.97%, 99.37%, 98.57%, 94.50%, 96.87%	Protocol equivalence across datasets is uncertain.
Mehta et al., 2024 [11]	VG G16 /VG G19 transfer + SVM ensemble	Two datasets with pose/illumination variation	Recognition	Accuracy = 95.96%, Precision = 95.97%, Recall = 95.96%, F1-score = 95.96%, Specificity = 99.36%.	Evaluated on a single benchmark dataset, and computational complexity increases due to the ensemble architecture.
Jaha,	Deep	Multiple	Recognition	Up to 99.94%; up	Soft biometric

2025 [62]	features + discriminatory soft biometrics	ear data sets		to 14% improvement	attributes may introduce fairness/privacy concerns.
Yuan et al., 2025 [63]	Symmetry alignment + feature interaction	Bilateral ear data	Similarity/authentication	Similarity accuracy 99.03%; authentication F1 0.9252	Single-source population limits generalization.

Table 8 Comparison of 3-D, Thermal, Multimodal, and Secure Ear Biometric Methods

Study	System	Data	Metric	Result	Contribution	Limitation
Ganapati et al., 2018 [34]	3-D curvilinear ear features	UND-J2 / in-house 3-D data	Rank-1 / EER	98.69% Rank-1; 1.53% EER reported	3-D shape robustness	Depth sensing and registration overhead.
Sinha and Ajmera, 2019 [38]	Template security	Protected ear templates	EER	Average EER 12.52% +/- 1.5%	Early recent security-focused evidence	Security-performance trade-off remained substantial

Pratim et al., 2021 [36]	Ear + profile-face feature fusion	Multimodal datasets	Recognition	Rank-1 = 98.53%, Verification Accuracy = 99.51%, EER = 0.49%	Complementary modalities	Requires synchronized acquisition and fusion calibration.
El-Naggar and Bourlai, 2023 [35]	Visible and thermal CNNs	Dual-band visible/thermal data	Rank-1 / EER	Visible 98.76%, 0.7596% EER; thermal 96.93%, 2.4541% EER	Low-light recognition	Thermal sensors and domain mismatch add cost.
Ganapati et al., 2023 [37]	RHEMAT ear-based multimodal authentication	Multimodal biometric protocol	EER	1.50%	Robust multimodal authentication	System complexity and modality availability.
Xu et al., 2024 [39]	Deep features + blockchain protection	AWE and EarV N1.0	Rank-1	83.87% and 96.52%	Links recognition and tamper-resistant	Blockchain does not solve image-

					records	domain shift.
Kaur et al., 2025 [14]	PointNet++ 3-D face-ear fusion	FRGC face + UNDEAR point clouds	Accuracy	99% face; 98% ear	Direct point-cloud learning and multimodal fusion	Specialized 3-D hardware and cross-dataset fusion assumptions.
Omar and Solomon, 2026 [57]	CaffeNet + deep metric learning + Comb-filter EarCodes	AWE and USTB-II	Recognition and template security	Outperformed compared baselines in source study	Cancelable Internet-of-Biometric-Things architecture	Requires independent security and attack evaluation.
Elsharkawy et al., 2026 [55]	MobilenetV3 + DenseNet-121 fusion, MLS GEO and Comb-filter	AMI, AWE, IITD-I/II and UERC	Accuracy	99.90%, 99.64%, 99.78%, 99.32%, 93.31%	Combines accuracy, feature reduction and revocable templates	Near-perfect results need independent replication and protocol audit.

4 Survey Findings and Discussion

The reviewed literature confirms a broad shift toward transfer learning, residual networks, EfficientNet, metric

learning, attention, and ensembles. However, hybrid designs remain competitive because they combine learned representations with handcrafted normalization, SVM decision boundaries, PCA reduction, or multimodal fusion [23][56][20][50][64][11]. The contribution of this review is to avoid treating “classical” and “deep” as mutually exclusive categories; recent systems frequently combine both.

The clearest cross-study outcome is that a high score on one dataset does not establish robustness. The same explainable residual-network family achieved 99.64% on AMI but only 67.25% on AWE [54]. Similarly, the metric-learning DAG-SVM varied from 98.79% on AWE and 98.70% on AMI to 84.30% on WPUT [50]. These within-method differences are more informative than broad rankings across unrelated studies. Relative to prior surveys, this review makes dataset-conditioned interpretation the central comparative rule rather than a cautionary footnote.

Pixel-wise and attention-based segmentation studies report accuracy, Dice, and IoU [8][18][45][12], whereas identity systems report Rank-1, classification accuracy, EER, or F1. Combining these values in a single accuracy hierarchy is scientifically invalid. The present review contributes a task-specific evidence structure that separates localization quality from identity discrimination and highlights where segmentation improvements have not yet been validated through end-to-end recognition.

Three-dimensional point clouds, thermal imagery, light-field capture, bilateral modeling, and profile-face fusion address limitations that cannot always be solved through larger 2-D CNNs [34][14][35][33][36][37][63][58]. These approaches improve viewpoint, low-light, or complementary identity evidence, but introduce specialized sensors, alignment requirements, higher cost, and more complex evaluation. The synthesis therefore identifies a recurring accuracy-versus-deployability trade-off.

Template protection was peripheral in early recent work but becomes prominent in 2024-2026 through blockchain-protected records, cancelable EarCodes, optimization-based feature selection, and revocable transformations [39][55][57][38]. These studies expand the field beyond recognition accuracy. Nevertheless, independent inversion testing, unlinkability analysis, key-space evaluation, presentation-attack detection, and revocation experiments remain less standardized than recognition protocols.

5 Challenges and Limitations

Although there has been much progress in ear biometrics recognition technology, there are still some barriers that are posing obstacles in practical implementation and evaluation of the technology. One of the biggest problems facing the technology is the lack of standard evaluation procedures in various researches. The difference in train-test split, subject-disjoint evaluation, cross-validation techniques and metrics used makes it hard to compare the recognition results achieved. This means that results from one benchmark cannot be compared with results from other benchmarks [2], [3], [13].

Another important challenge is the limited cross-dataset generalization of existing recognition models. Although many deep learning approaches achieve excellent recognition accuracy on controlled datasets such as IIT Delhi and AMI, their performance decreases considerably when evaluated on unconstrained datasets such as AWE, WPUT, EarVN1.0, and UERC because of variations in pose, illumination, occlusion, image quality, and background conditions [4], [5], [52]. This indicates that improving robustness across diverse acquisition environments remains an important research direction.

The scarcity of large and varied databases of ear biometrics also hinders the creation of highly generalizable recognition algorithms. In comparison to other biometric approaches like face recognition and iris recognition, publicly available databases of ears usually have fewer identities, and are mostly controlled environment acquisitions. Therefore, current algorithms are tested on relatively small-sized databases, which limits their application in practical biometric systems [2], [15], [43].

In recent times, research efforts have increasingly focused on lightweight deep learning approaches, multimodal recognition techniques, thermography, enhanced security through blockchain technology, and cancelable biometric templates for enhancing deployability and template safety. However, issues relating to efficiency, privacy preservation, presentation attack detection, explainability, demographic bias [65], and template security have not been addressed sufficiently enough, especially within unconstrained settings [35], [39], [57].

There are some limitations of the current review as well. Firstly, the review is limited to peer-reviewed journal articles within the period from 2017 to 2026, whereas conference papers, preprints, books, and articles not written in English have been excluded from the analysis

to maintain consistency and credibility of the evidence. Secondly, the review includes qualitative evidence synthesis but not the meta-analysis of data due to significant heterogeneity of datasets, evaluation methodologies, and performance metrics of selected articles. Lastly, although the latest trends in explainable AI, self-supervised learning, and privacy-preserving biometrics have been considered in the literature review, this research area is currently under development and needs further experimental investigation.

6 Future Recommendations

The fast advancement in ear biometric recognition technology has greatly enhanced the efficiency of ear identification; nevertheless, there are still some issues that need to be resolved by future research studies. Taking into account the information reviewed above, the following research areas are likely to influence the development of ear biometrics in the future.

A primary research priority is the development of robust cross-dataset and cross-domain learning frameworks. Although recent deep learning models demonstrate high recognition accuracy on individual benchmark datasets, their performance declines considerably when evaluated under different acquisition conditions. Future studies should therefore investigate domain adaptation, domain generalization, self-supervised learning, and foundation models to improve robustness across heterogeneous datasets and unconstrained environments [4], [9], [52]. Another important direction is the establishment of standardized benchmarking protocols for ear biometric evaluation. The literature reveals considerable variation in dataset partitions, evaluation metrics, and validation strategies, making objective comparison between studies difficult. Developing unified benchmark datasets and standardized evaluation protocols would improve reproducibility and enable fair performance assessment across different recognition systems [2], [3], [13].

Moreover, future research should also focus on lightweight and edge-deployable recognition systems. Recent advances demonstrate increasing interest in computationally efficient CNNs and hybrid CNN–Transformer architectures for mobile and embedded applications. Further optimization through model compression, pruning, quantization, and knowledge distillation can facilitate real-time deployment without significantly compromising recognition accuracy [10], [29], [30]. The integration of multimodal biometric fusion represents another promising direction. Combining ear

biometrics with complementary modalities such as face, thermal imaging, three-dimensional point clouds, or soft biometric traits has shown potential for improving recognition robustness under challenging environmental conditions. Future studies should investigate adaptive multimodal fusion strategies capable of operating reliably in unconstrained scenarios [14], [35], [36], [37].

Finally, future research should prioritize explainable, fair, and trustworthy artificial intelligence for ear biometric recognition. Although recent studies have introduced explainable deep learning techniques, comprehensive evaluation of model interpretability, demographic bias, uncertainty estimation, and ethical deployment remains limited. Addressing these issues will improve user confidence and support the adoption of ear biometric systems in practical authentication environments [5], [54]. Overall, future progress in ear biometrics will depend not only on improving recognition accuracy but also on developing robust, interpretable, computationally efficient, privacy-preserving, and standardized recognition systems that can operate reliably across diverse real-world conditions.

7 Conclusion

This paper presented a protocol-aware literature review, over 2017-2026 period, on ear-biometric research based on 45 journal articles across the range of detection, segmentation, 2-D/3-D recognition, multimodal/thermal sensing, lightweight deployability, and template security. It found that the field has moved away significantly from the initial geometric/handcrafted approaches toward the currently dominant deep convolutional (and recently transformer-based and self-supervised) methods. However, its progress in rigor of evaluation has trailed that of accuracy; across the analyzed works, consistent comparisons between architectures demonstrate that same architectures may yield substantially different results in the context of controlled vs. unconstrained dataset; thus, accuracy cannot be simply treated as a general quality measure without taking context of the database, protocol, and tasks into account. Overall, three conclusions stand out. First, the most persistent barrier is dataset/protocol diversity, reiterating the critical necessity for unified, subject-disjoint evaluation protocols-akin to those used in face recognition. Second, whereas 3-D, thermal, and multimodal modalities exhibit genuine robustness benefits compared to purely 2-D pipelines, they often suffer from the trade-off between recognition accuracy and practical deployability given sensor costs, complexity of

acquisition and computation. Third, the growing focus on lightweight edge-deployable and cancelable/secure template designs illustrates a field maturing enough to consider real-world issues, yet security claims require more independent evaluation and work on fairness/interpretability is much sparser than that on recognition accuracy. Combined, the results suggest future efforts be focused on cross-dataset/domain learning, unified evaluation standards, and incorporating explainability, fairness and privacy-preserving mechanisms in pursuit of recognition performance improvement. It is crucial for advancing science and enabling safe, trustworthy deployments in practice.

References

- [1] A. Kamboj, R. Rani, and A. Nigam, "A comprehensive survey and deep learning-based approach for human recognition using ear biometric," *Vis. Comput.*, vol. 38, no. 7, pp. 2383–2416, 2022, doi: 10.1007/s00371-021-02119-0.
- [2] A. Benzaoui, Y. Khaldi, R. Bouaouina, N. Amrouni, H. Alshazly, and A. Ouahabi, "A Comprehensive Survey on Ear Recognition: Databases, Approaches, Comparative Analysis, and Open Challenges," *Neurocomputing*, vol. 537, Mar. 2023, doi: 10.1016/j.neucom.2023.03.040.
- [3] O. G. Oyebiyi, A. Abayomi-Alli, O. A. Arogundade, A. Qazi, A. L. Imoize, and J. B. Awotunde, "A Systematic Literature Review on Human Ear Biometrics: Approaches, Algorithms, and Trend in the Last Decade," *Information*, vol. 14, no. 3, p. 192, 2023, doi: 10.3390/info14030192.
- [4] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Deep Convolutional Neural Networks for Unconstrained Ear Recognition," *IEEE Access*, vol. 8, pp. 170295–170310, 2020, doi: 10.1109/ACCESS.2020.3024116.
- [5] H. Alshazly, C. Linse, E. Barth, S. A. Idris, T. Martinetz, and S. Member, "Towards Explainable Ear Recognition Systems Using Deep Residual Networks," *IEEE Access*, vol. 9, pp. 122254–122273, 2021, doi: 10.1109/ACCESS.2021.3109441.
- [6] A. Booyens and S. Viriri, "Exploration of Ear Biometrics Using EfficientNet," *Comput. Intell.*

- Neurosci.*, vol. 2022, pp. 1–14, Aug. 2022, doi: 10.1155/2022/3514807.
- [7] A. Booyens and S. Viriri, “Transformation Network Model for Ear Recognition BT - Machine Learning for Networking,” É. Renault, S. Boumerdassi, and P. Mühlethaler, Eds., Cham: Springer Nature Switzerland, 2024, pp. 250–266.
 - [8] D. M. Alex *et al.*, “DANNET: deep attention neural network for efficient ear identification in biometrics,” *PeerJ. Comput. Sci.*, vol. 10, p. e2603, 2024, doi: 10.7717/peerj-cs.2603.
 - [9] J. He, Y. He, L. Zhai, and Y. Bi, “Self-supervised Siamese Networks with Squeeze-Excitation Attention for Ear Image Recognition BT - Advanced Intelligent Computing Technology and Applications,” D.-S. Huang, Z. Si, and J. Guo, Eds., Singapore: Springer Nature Singapore, 2024, pp. 122–133.
 - [10] C. Lendering, B. P. Ribeiro, and P. Peer, “EdgeEar: Efficient and Accurate Ear Recognition for Edge Devices,” *arXiv Prepr.*, 2025, doi: <https://doi.org/10.48550/arXiv.2502.07734>.
 - [11] R. Mehta, A. Sheikh-Akbari, and K. K. Singh, “Ensemble-Based Hybrid Transfer Approach for an Effective 2D Ear Recognition System,” *IEEE Access*, vol. 12, pp. 155733–155746, 2024, doi: 10.1109/ACCESS.2024.3485514.
 - [12] R. Lahkar and K. A. Borbora, “Automatic segmentation of human ear in the wild,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 34, no. 1, pp. 333–341, 2024, doi: 10.11591/ijeecs.v34.i1.pp333-341.
 - [13] I. Ganapathi, S. Ali, N.-S. Vu, S. Prakash, and N. Werghi, “A Survey of 3D Ear Recognition Techniques,” *ACM Comput. Surv.*, vol. 55, Sep. 2022, doi: 10.1145/3560884.
 - [14] V. Kaur, D. P. Bhatt, S. Tharewal, and P. K. Tiwari, “Intelligent multimodal 3D biometric recognition using PointNet ++ for robust face-ear authentication,” *Sci. Rep.*, vol. 15, no. 1, p. 43793, Nov. 2025, doi: 10.1038/s41598-025-27757-5.
 - [15] Ž. Emeršič, V. Štruc, and P. Peer, “Ear recognition: More than a survey,” *Neurocomputing*, vol. 255, pp. 26–39, 2017, doi: <https://doi.org/10.1016/j.neucom.2016.08.139>.
 - [16] I. I. Ganapathi, S. S. Ali, S. Prakash, N.-S. Vu, and N. Werghi, “A Survey of 3D Ear Recognition Techniques,” *ACM Comput. Surv.*, vol. 55, no. 10, Feb. 2023, doi: 10.1145/3560884.
 - [17] C. Cintas *et al.*, “Automatic ear detection and feature extraction using Geometric Morphometrics and convolutional neural networks,” *IET Biometrics*, vol. 6, no. 3, pp. 211–223, May 2017, doi: <https://doi.org/10.1049/iet-bmt.2016.0002>.
 - [18] Ž. Emeršič, L. L. Gabriel, V. Štruc, and P. Peer, “Convolutional encoder–decoder networks for pixel-wise ear detection and segmentation,” *IET Biometrics*, vol. 7, no. 3, pp. 175–184, May 2018, doi: <https://doi.org/10.1049/iet-bmt.2017.0240>.
 - [19] Ž. Emeršič, D. Sušanj, B. Meden, V. Štruc, S. Member, and P. Peer, “ContextedNet: Context – Aware Ear Detection in Unconstrained Settings,” *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3121792.
 - [20] E. Hansley, M. Segundo, and S. Sarkar, “Employing Fusion of Learned and Handcrafted Features for Unconstrained Ear Recognition,” *IET Biometrics*, vol. 7, Oct. 2017, doi: 10.1049/iet-bmt.2017.0210.
 - [21] D. Yaman, F. Eyiokur, and H. Ekenel, “Domain Adaptation for Ear Recognition Using Deep Convolutional Neural Networks,” *IET Biometrics*, vol. 7, Dec. 2017, doi: 10.1049/iet-bmt.2017.0209.
 - [22] S. Dodge, J. Mounsef, and L. Karam, “Unconstrained Ear Recognition Using Deep Neural Networks,” *IET Biometrics*, vol. 7, Jan. 2018, doi: 10.1049/iet-bmt.2017.0208.
 - [23] I. Omara, X. Wu, H. Zhang, Y. Du, and W. Zuo, “Learning Pairwise SVM on Hierarchical Deep Features for Ear Recognition,” *IET Biometrics*, vol. 7, Feb. 2018, doi: 10.1049/iet-bmt.2017.0087.
 - [24] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, “Handcrafted versus CNN Features for Ear Recognition,” *Symmetry (Basel)*, vol. 11, p. 1493, Dec. 2019, doi: 10.3390/sym11121493.
 - [25] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, “Ensembles of Deep Learning Models and Transfer Learning for Ear Recognition,” 2019, doi: 10.3390/s19194139.
 - [26] R. Priyadharshini, A. Selvaraj, and A. Madakannu, “A deep learning approach for person identification using ear biometrics,” *Appl. Intell.*, vol. 51, pp. 2161–2172, Oct. 2020, doi: 10.1007/s00371-020-01941-3.

- 10.1007/s10489-020-01995-8.
- [27] M. Sharkas, "Ear recognition with ensemble classifiers; A deep learning approach," *Multimed. Tools Appl.*, vol. 81, May 2022, doi: 10.1007/s11042-022-13252-w.
- [28] N. Chong and F. Yang, "CMMCAN: Lightweight Feature Extraction and Matching Network for Endoscopic Images Based on Adaptive Attention," *Comput. Mater. Contin.*, vol. 80, no. 2, pp. 2761–2783, 2024, doi: <https://doi.org/10.32604/cmc.2024.052217>.
- [29] M. D. Maneri, *Hybrid Transfer Learning Using VGG Networks for Robust Ear Biometric Recognition*, no. Rrrai 2026. Atlantis Press International BV. doi: 10.2991/978-94-6239-723-1.
- [30] R. Mehta and K. Singh, "Ensemble of transfer learning and light-weight convolutional neural network model for an effective ear recognition system," *Evol. Syst.*, vol. 15, pp. 1–17, Dec. 2023, doi: 10.1007/s12530-023-09561-6.
- [31] A. Mahajan and S. K. Singla, "DeepBio: A Deep CNN and Bi-LSTM Learning for Person Identification Using Ear Biometrics," *C. - Comput. Model. Eng. Sci.*, vol. 141, no. 2, pp. 1623–1649, 2024, doi: <https://doi.org/10.32604/cmcs.2024.054468>.
- [32] R. Bouaouina, A. Benzaoui, H. Doghmane, and Y. Brik, "Boosting the Performance of Deep Ear Recognition Systems Using Generative Adversarial Networks and Mean Class Activation Maps," 2024. doi: 10.3390/app14104162.
- [33] A. Sepas-Moghaddam, F. Pereira, and P. Correia, "Ear Recognition in a Light Field Imaging Framework: A New Perspective," *IET Biometrics*, vol. 7, Jan. 2018, doi: 10.1049/iet-bmt.2017.0204.
- [34] I. Ganapathi, S. Prakash, S. Ali, P. Joshi, A. Srivastava, and I. Dave, "Ear Recognition in 3D using 2D Curvilinear Features," *IET Biometrics*, vol. 7, Aug. 2018, doi: 10.1049/iet-bmt.2018.5064.
- [35] S. El-Naggar and T. Bourlai, "Exploring Deep Learning Ear Recognition in Thermal Images," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 5, pp. 64–75, Jan. 2022, doi: 10.1109/TBIOM.2022.3218151.
- [36] P. Pratim, S. Deepak, R. Nayak, M. Panda, and B. Majhi, *A feature - level fusion based improved multimodal biometric recognition system using ear and profile face*, no. 0123456789. Springer Berlin Heidelberg, 2021. doi: 10.1007/s12652-021-02952-0.
- [37] I. I. Ganapathi, S. S. Ali, U. Sharma, P. Tomar, M. Owais, and N. Werghi, "RHEMAT: Robust human ear based multimodal authentication technique," *Comput. Secur.*, vol. 132, p. 103356, 2023, doi: <https://doi.org/10.1016/j.cose.2023.103356>.
- [38] H. Sinha and P. K. Ajmera, "Upgrading security and protection in ear biometrics," *IET Biometrics*, vol. 8, no. 4, pp. 259–266, 2019, doi: <https://doi.org/10.1049/iet-bmt.2018.5081>.
- [39] X. Xu, Y. Liu, C. Liu, and L. Lu, "Anti-Software Attack Ear Identification System Using Deep Feature Learning and Blockchain Protection," 2024. doi: 10.3390/sym16010085.
- [40] Z. F. Elsharkawy, E. M. Omran, A. A. Eisa, A. A. Saleh, N. A. Ismail, and F. E. A. El-Samie, "A cancelable ear recognition system via optimized deep feature fusion," *Sci. Rep.*, vol. 16, no. 1, Jun. 2026, doi: 10.1038/s41598-026-57027-x.
- [41] Ž. Emeršič *et al.*, "The unconstrained ear recognition challenge," in *2017 IEEE International Joint Conference on Biometrics (IJCB)*, 2017, pp. 715–724. doi: 10.1109/BTAS.2017.8272761.
- [42] Ž. Emeršič *et al.*, "The Unconstrained Ear Recognition Challenge 2019," in *2019 International Conference on Biometrics (ICB)*, 2019, pp. 1–15. doi: 10.1109/ICB45273.2019.8987337.
- [43] V. T. Hoang, "EarVN1.0: A new large-scale ear images dataset in the wild," *Data Br.*, vol. 27, p. 104630, 2019, doi: <https://doi.org/10.1016/j.dib.2019.104630>.
- [44] S. Ramos-Cooper, E. Gomez-Nieto, and G. Camara-Chavez, "VGGFace-Ear: An Extended Dataset for Unconstrained Ear Recognition," *Sensors*, vol. 22, no. 5, p. 1752, 2022, doi: 10.3390/s22051752.
- [45] Ž. Emeršič *et al.*, *The Unconstrained Ear Recognition Challenge 2023: Maximizing Performance and Minimizing Bias**. 2023. doi: 10.1109/IJCB57857.2023.10449062.
- [46] A. Kumar and C. Wu, "Automated human identification using ear imaging," *Pattern Recognit.*, vol. 45, no. 3, pp. 956–968, 2012, doi: <https://doi.org/10.1016/j.patcog.2011.06.005>.

- [47] R. Mehta, A. Sheikh-Akbari, and K. K. Singh, "A Noble Approach to 2D Ear Recognition System using Hybrid Transfer Learning," in *2023 12th Mediterranean Conference on Embedded Computing (MECO)*, 2023, pp. 1–5. doi: 10.1109/MECO58584.2023.10154993.
- [48] E. Gonzalez, L. A. And, and L. Mazorra, "AMI Ear Database." [Online]. Available: https://ctim.ulpgc.es/research_works/ami_ear_database/
- [49] D. Frejlichowski and N. Tyszkiewicz, "The West Pomeranian University of Technology Ear Database – A Tool for Testing Biometric Algorithms BT - Image Analysis and Recognition," A. Campilho and M. Kamel, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 227–234.
- [50] I. Omara, A. Hagag, C. Souleyman, G. Ma, F. Abd El-Samie, and E. Song, "A Hybrid Model Combining Learning Distance Metric and DAG Support Vector Machine for Multimodal Biometric Recognition," *IEEE Access*, vol. 9, pp. 4784–4796, Oct. 2020, doi: 10.1109/ACCESS.2020.3035110.
- [51] D. Freire-Obregón, J. Neves, Ž. Emeršič, B. Meden, M. Castrillón-Santana, and H. Proença, "Synthesizing multilevel abstraction ear sketches for enhanced biometric recognition," *Image Vis. Comput.*, vol. 154, p. 105424, 2025, doi: <https://doi.org/10.1016/j.imavis.2025.105424>.
- [52] D. Freire-Obregón, M. De Marsico, P. Barra, J. Lorenzo-Navarro, and M. Castrillón-Santana, "Zero-shot ear cross-dataset transfer for person recognition on mobile devices," *Pattern Recognit. Lett.*, vol. 166, pp. 143–150, 2023, doi: <https://doi.org/10.1016/j.patrec.2023.01.012>.
- [53] I. Riaz, A. Nazri Ali, and H. Ibrahim, "A novel sub-windowing local binary pattern approach for dorsal finger creases based biometric classification system," *Eng. Sci. Technol. an Int. J.*, vol. 60, p. 101882, 2024, doi: <https://doi.org/10.1016/j.jestch.2024.101882>.
- [54] H. Alshazly, H. Elmannai, R. I. Alkanhel, and A. Abdelnazeer, "Traitement du Signal Advancing Biometric Identity Recognition with Optimized Deep Convolutional Neural Networks," vol. 41, no. 3, pp. 1405–1418, 2024.
- [55] E. Abdellatef *et al.*, "Cancelable face and iris recognition system based on deep learning," *Opt. Quantum Electron.*, vol. 54, Sep. 2022, doi: 10.1007/s11082-022-03770-0.
- [56] M. Zarachoff, A. Sheikh Akbari, and D. Monekosso, "Multi-Band PCA Based Ear Recognition Technique," *Multimed. Tools Appl.*, vol. 82, Jun. 2022, doi: 10.1007/s11042-022-12905-0.
- [57] I. Omara and R. F. Soliman, "CIoEBT: cancelable internet of ear biometric things based – a novel deep metric learning approach," *Expert Syst. Appl.*, vol. 297, no. PB, May 2026, doi: 10.1016/j.eswa.2025.129439.
- [58] R. Fukuda, Y. Yokoyanagi, C. Prathom, and Y. Okada, "Toward Personal Identification Using Multi-Angle-Captured Ear Images: A Feasibility Study," 2025. doi: 10.3390/app15063329.
- [59] Ž. Emeršič, B. Meden, P. Peer, and V. Štruc, "Evaluation and analysis of ear recognition models: performance, complexity and resource requirements," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 15785–15800, Oct. 2020, doi: 10.1007/s00521-018-3530-1.
- [60] A. Alkababji and O. Mohammed, "Real time ear recognition using deep learning," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 19, p. 523, Apr. 2021, doi: 10.12928/TELKOMNIKA.v19i2.18322.
- [61] R. Mehta and K. Singh, "Deep convolutional neural network-based effective model for 2D ear recognition using data augmentation," *Imaging Sci. J.*, vol. 72, pp. 403–420, May 2023, doi: 10.1080/13682199.2023.2206763.
- [62] E. Jaha, "Augmented Deep-Feature-Based Ear Recognition Using Increased Discriminatory Soft Biometrics," *Comput. Model. Eng. Sci.*, vol. 144, pp. 3645–3678, Sep. 2025, doi: 10.32604/cmes.2025.068681.
- [63] L. Yuan, H.-B. Zhou, J.-Y. Li, L. Liu, X.-C. Gu, and Y.-N. Zhao, "Symmetry Alignment-Feature Interaction Network for Human Ear Similarity Detection and Authentication," 2025. doi: 10.3390/sym17050654.
- [64] B. Hussain, M. K. Afzal, S. Ahmad, and A. M. Mostafa, "Intelligent traffic flow prediction using optimized GRU model," *IEEE Access*, vol. 9, pp. 100736–100746, 2021, doi: 10.1109/ACCESS.2021.3097141.
- [65] P. Drozdowski, C. Rathgeb, A. Dantcheva, N. Damer, and C. Busch, "Demographic Bias in

Biometrics: A Survey on an Emerging
Challenge,” *IEEE Trans. Technol. Soc.*, vol. 1, pp.
89–103, Mar. 2020, doi:

10.1109/TTS.2020.2992344.