Expert Systems With Applications 297 (2026) 129439

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

CIoEBT: cancelable internet of ear biometric things based – a novel deep metric learning approach

Ibrahim Omara [a,b], Randa F. Soliman [b,c,*]

[a] Department of Cybersecurity, College of Engineering and Information Technology, Buraydah Private Colleges, Buraydah, 51418, Saudi Arabia
[b] Machine Intelligence Deptartment, Faculty of Artificial Intelligence, Menoufia University, Shebin El-Kom 32511, Egypt
[c] Electrical and Computer Engineering Department, Tennessee Technological University, Cookeville, TN 38505, USA

ARTICLE INFO

Keywords:
Internet of Biometrics Things, EarCodes, Security, CaffeNet, Privacy Preservation, and Machine Learning

ABSTRACT

Biometric authentication is gaining widespread acceptance globally, driving the need for robust and secure systems. This study explores the use of outer ear images as a distinctive biometric modality. The human ear, much like the face, exhibits unique and permanent features, making it a promising candidate for biometric identification. However, ear biometrics, like facial recognition, face challenges such as variations in illumination, contrast, rotation, scale, and pose. To address these issues, this paper investigates the application of Convolutional Neural Networks (CNNs), a powerful tool in computer vision, for ear recognition. Specifically, we propose a hybrid approach that combines deep CNNs with metric learning. Using CaffeNet as a feature extractor and a novel Deep Effective Pairwise Constraints Metric Learning (DEP-ML) strategy, we encode ear images into secure representations called EarCodes. These codes are further protected using the Comb-filter algorithm, resulting in highly secure and reliable biometric templates. The proposed CaffeNet-DEP-ML framework is evaluated against well-known benchmarks like VGG-verydeep16 and VGG-S on two prominent ear image datasets, AWE and USTB II. Experimental results demonstrate that our method not only outperforms current state-of-the-art techniques but also benefits from fewer trainable parameters and faster processing times. This innovative approach shows strong potential for integration into Internet of Biometric Things (IoBT) environments, offering high accuracy while ensuring privacy preservation.

1. Introduction

Recent advancements in Wireless Body Area Networks (WBANs) and the Internet of Medical Things (IoMT) in smart healthcare, along with the broader Internet of Things (IoT) in smart environments, have heightened the demand for more secure and reliable user authentication systems (Huang, 2019; Mandal, 2020). Traditional authentication methods, such as passwords and ID cards, are increasingly vulnerable to theft, loss, or misuse by unauthorized individuals. In contrast, biometric authentication, which uses traits such as iris scans, DNA, facial recognition, and fingerprints, offers enhanced security due to the difficulty of replication (Victor et al., 2002). However, the growing complexity of biometric systems introduces new challenges in their design and implementation.

Despite being more secure than traditional methods, some biometric traits like voice, gait, signature, and facial expressions can still be mimicked by imposters. Even supposedly immutable traits like DNA and fingerprints are not immune to misuse, as they can potentially be collected and exploited without the user's knowledge. Moreover, most biometric identifiers are non-revocable, making recovery from data breaches more difficult and raising further concerns about long-term privacy and security. Considering these challenges, research has increasingly highlighted the potential of the human ear as a robust biometric trait (Victor et al., 2002; Camile et al., 2502). The ear possesses several key advantages: its shape is largely stable over time and resistant to changes due to emotions, mental states, or external influences. Ears can also be captured discreetly from a distance, even under constrained or sensitive conditions such as in national security or

Abbreviations: AWE, Annotated Web Ears; CaffeNet, Convolutional Architecture for Fast Feature Embedding; CNN, Convolutional Neural Network; DEP-ML, Deep Effective Pairwise Constraints Metric Learning; IoMT, Internet of Medical Things; ROC, Receiver Operating Characteristics; SVM, Support Vector Machine; USTB II, University of Science and Technology in Beijing; VGG, Visual Geometry Group; WBAN, Wireless Body Area Network.

* Corresponding author.
E-mail addresses: i_omara84@ai.menofia.edu.eg, i_omara84@yahoo.com (I. Omara), c-rsoliman@tntech.edu, randa_soliman@ai.menofia.edu.eg (R.F. Soliman).

https://doi.org/10.1016/j.eswa.2025.129439
Received 15 November 2024; Received in revised form 11 August 2025; Accepted 18 August 2025
Available online 21 August 2025

forensic applications. These unique properties make ear recognition a compelling solution for tasks like forensic image analysis and intelligent surveillance.

Unlike certain biometric features such as the face, which can be affected by changing expressions, lighting conditions, or the use of makeup, the human ear offers a more stable and consistent structure. Despite sharing common anatomical components across individuals, the ear remains uniquely identifiable due to its distinctive features. These include the anti-helix, which runs parallel to the helix; the tragus, a uniquely shaped hairpin bend above the earlobe; and the concha, the central region of the ear named for its resemblance to a seashell (Emersic et al., 2017; Mariam Mahmoud Mohammed, 2025). This combination of features contributes significantly to the uniqueness and reliability of ear-based biometric identification. However, despite these advantages, ear-based authentication systems often lack adequate security measures, raising significant concerns about data privacy and system vulnerability, especially within IoT environments. Without proper protection, raw ear data can be intercepted or exposed, leaving authentication systems susceptible to spoofing attacks and unauthorized access. Moreover, if the biometric data is compromised or corrupted, the ear's unique information cannot be revoked or changed, leading to permanent loss of security for the affected user.

Another critical issue involves compliance with international data privacy regulations. Various countries and regions have enacted laws aimed at safeguarding personal data, such as the California Privacy Rights Act (CPRA) in the U.S., the Personal Information Protection Law (PIPL) in China, and the General Data Protection Regulation (GDPR) in the European Union. These frameworks demand stringent protection of biometric data, emphasizing the need for secure implementation of ear-based authentication in IoT systems. To address these challenges, two major categories of ear template protection strategies have emerged, Cryptography-based approaches and cancelable ear template techniques. Cryptographic methods, such as symmetric, asymmetric, and homomorphic encryption, offer strong security by transforming raw data into protected formats. These techniques allow for secure, flexible template storage but are often computationally intensive, making them less suitable for resource-constrained IoT devices (Kim et al., 2020; Barni et al., 2010; Bohan et al., 2013). As such, striking a balance between security, efficiency, and privacy compliance remains a key challenge in deploying ear biometrics in real-world IoT applications.

An alternative method for securing biometric data is the use of cancelable biometrics (Ratha et al., 2007; Ratha et al., 2001; Punithavathi and Geetha, 2019; Kaur and Khanna, 2020; Jiang et al., 2020; Patel et al., 2015). This approach involves permanently transforming the original ear template into an altered version that cannot be reversed. By doing so, even if the transformed template is compromised, the original biometric data remains protected. Cancelable ear templates thus offer a strong defense against privacy breaches while enhancing the overall security and reusability of biometric systems. However, in resource constrained IoT environments, using overly simplified transformation techniques may lead to significant drops in authentication accuracy. This performance degradation stems from the trade-off between maintaining security and accommodating the limited computational capabilities of IoT devices (Oyebiyi and Abayomi-Alli, 2023). Therefore, the design of cancelable biometric systems must carefully balance privacy protection, system performance, and computational efficiency.

To address the growing need for secure and revocable biometric templates, this study proposes a novel approach within the domain of ear biometrics. The human outer ear has attracted increasing interest due to its notable advantages, such as distinctiveness, resilience, accessibility, and acceptability comparable to facial biometrics (Benarous et al., 2017). Moreover, ear recognition has been successfully integrated into various platforms, including Android (Boczek, 2017) and IOS applications (Bargal et al., 2015), reinforcing its practicality and real-world applicability.

Furthermore, it is crucial to emphasize that various past metric learning methods, including ITML, encounter limitations due to their utilization of fixed pairwise constraints during training, in contrast, these constraints are established during the preprocessing phase. Also, the use of biometric systems poses security risks and the potential for misuse, primarily due to the storage of original templates in databases. Moreover, the lack of revocability in traditional biometric systems, the computational overhead of cryptographic methods, and the performance degradation in simplified cancelable approaches. In response to these limitations of existing techniques, this paper presents a robust, cancelable, and privacy-preserving authentication framework. The proposed methodology is designed to be both efficient and secure, particularly suitable for resource constrained IoT environments. The main notable contributions of this paper can be summarized as follows: Firstly, the acquisition of testing images using IoT devices, followed by the extraction of deep features from the human ear employing CaffeNet as the feature extractor. The novel Deep Effective Pairwise Constraints Metric Learning (DEP-ML) technique is then employed for the classification task within the proposed joint CaffeNet-DEP-ML framework. Subsequently, a cancelable biometric scheme based on intelligent agents and utilizing Comb-filters is implemented. Lastly, the testing images are matched against the stored template database on the servers. By combining deep learning, metric learning, and template cancelability, this methodology effectively addresses the challenges of template security, system performance, and compliance with privacy regulations, paving the way for scalable and secure ear authentication in IoT-based applications.

The remainder of this paper is organized as follows: Section 2 presents a comprehensive review of the related work in secure biometrics and ear biometrics. Section 3 details the proposed methodology for achieving secure and cancelable ear recognition. Section 4 provides an evaluation of the system's performance, while Section 5 offers a thorough security analysis of the proposed approach. Finally, Section 6 concludes the paper by summarizing the key findings and contributions.

2. Preliminaries

2.1. Ear biometrics

The field of ear recognition began with the foundational work of Iannarelli (Iannarelli, 1989), who used a manual technique involving 12 distinct measurements taken from the center of the ear. A significant advancement was later made by Burge and Burger (Burge, 2000), who introduced a method for localized ear detection using deformable shape models. Their approach employed a Gaussian pyramid transformation to process ear images for identification. A breakthrough came with Moreno et al. (Moreno and Sanchez, 1999), who developed the first fully automated ear recognition system. Their method leveraged an ensemble of neural classifiers along with a compression network to extract macro-level features from ear images. Recognizing the limitations of working with unprocessed ear images, Ghoualmi et al. (Ghoualmi et al., 2016) highlighted the inaccuracy of several biometric techniques under challenging conditions. To address these shortcomings, Omara et al. (Omara et al., 2016) and Anwar et al. (Anwar et al., 2015) proposed more robust image preprocessing pipelines. These improvements aim to increase recognition accuracy by effectively handling variations in scale, pose, and lighting conditions.

In the field of ear biometrics, significant research efforts have focused on developing effective methods for feature extraction from ear images. Notable approaches include wavelet-based techniques (Sana et al., 2007; Wang et al., 2008) and filter-based methods (Jamil et al., 2014; Meraoumia et al., 2015), both of which have shown promising results. Among these, the force-field transformation has emerged as a particularly effective method, achieving a recognition accuracy of 99.2 % on the XM2VTS dataset (Hurley et al., 2005). Additionally, the recommendations of LeCun et al. (LeCun et al., 1989) emphasized the importance of feature selection based on classifier performance,

demonstrating substantial improvements in recognition outcomes.

A notable shift from traditional techniques is evident with the adoption of Convolutional Neural Networks (CNNs). Unlike conventional methods that rely on manual feature extraction, CNNs are capable of processing raw images directly, eliminating the need for preprocessing steps. By generating feature vectors automatically and using the backpropagation algorithm to train a posterior classifier, CNN-based approaches have shown superior performance compared to traditional classifiers that depend on handcrafted features.

2.2. Cancelable biometrics

Extensive research has been devoted to advance secure biometric authentication, with scholars generally pursuing two primary approaches:

Fuzzy techniques for constructing biometric-based cryptosystems, commonly known as biometric salting.
Non-invertible transformations applied to captured biometric data to prevent reconstruction.

These approaches, outlined by researchers, are indicative of the substantial interest in enhancing the security of biometric authentication (Jain et al., 2008).

2.3. Biometric-based cryptosystems

The implementation of biometric-based cryptosystems has been explored through two primary approaches: Combining a user-specific key with a biometric sample and directly generating a key from the biometric template, resulting in a randomized and distorted template (Rathgeb and Uhl, 2011). Common techniques in this domain include fuzzy commitment, fuzzy vault, and fuzzy extractor. Additionally, Nagar et al. (Nagar et al., 2010) proposed a hybrid scheme to enhance the security of fingerprint templates. The fuzzy commitment method, introduced by Juels and Wattenberg (Juels and Wattenberg, 1999), combines error-correcting codes and cryptography to create a hybrid approach. It involves using a biometric sample to decommit a key. This technique has been applied in various recognition systems, including iris (Rathgeb and Uhl, 2012; Hao et al., 2006; Bringer et al., 2008) and fingerprint recognition (Tong et al., 2007). Despite their utility, these systems face several challenges, including unrealistic assumptions, limited error correction capacity, and restricted key lengths.

Expanding on fuzzy commitment schemes, fuzzy vault schemes (Juels and Sudan, 2006) have been applied in various contexts, including smart cards (Charles Clancy et al., 2003), fingerprint recognition (Uludag et al., 2005; Nandakumar et al., 2007; Tams, 2016), handwritten signature recognition (Dong and Tan, 2008; Freire-Santos et al., 2006), face recognition (Dong and Tan, 2008), and iris recognition (Lee et al., 2007). Fuzzy vault schemes are valued for their reliability, security, and revocability. However, they can be vulnerable if crucial parameters, such as polynomials and chaff points, are not properly protected. Furthermore, the original biometric data may be inferred by correlating values across different vault systems. For instance, Chang et al. (Chang et al., 2006) demonstrated how statistical analysis could exploit the non-uniformity of biometric data from a set of chaff points, potentially compromising security.

The fuzzy extractor scheme (Li et al., 2008) authenticates a user by extracting a key from the biometric sample. This method ensures that the extraction remains consistent even in the presence of noise. During the enrollment phase, a helper key is generated to further enhance security. The effectiveness of this approach depends on the extractor's ability to consistently reproduce the same string, or one that is sufficiently close, according to a predefined distance metric, even in uncontrolled environments.

2.4. Non-invertible transformation

The potential for compromising the original biometric sample through manipulation of biometric cryptosystems has been acknowledged (Cappelli et al., 2007). In response, feature transformation techniques have been developed to modify the biometric sample in a way that makes it non-invertible. This transformation ensures that the original biometric data cannot be retrieved, and the comparison of templates occurs within this transformed space. The concept of non-invertible transformations was first introduced by Ratha et al. (Ratha et al., 2007), who applied Cartesian, polar, and functional transformations to fingerprint data. Building on this, Leng et al. (Leng et al., 2012) proposed a bidirectional 2D random projection for secure face and palm-print recognition.

A comprehensive survey on cancelable biometrics is provided in (Rathgeb and Uhl, 2011, 2011,; Soliman et al., 2020; Randa et al., 2018; Soliman et al., 2018; Soliman et al., 2018). More recently, the focus has shifted toward incorporating deep learning techniques into secure biometric authentication. In this context, deep learning methods generate a transformed feature vector that not only uniquely represents the biometric data but also minimizes correlations between the original data and the transformed vector. This approach has shown promise in applications such as face recognition (Pandey et al., 2016; Pandey et al., 2017), palm-print recognition (Meraoumia et al., 2017), and iris recognition (Omran et al., 2019; Abdellatef et al., 2022; Essam et al., 2020; Omran et al., 2020). In secure biometrics, ear samples are increasingly relevant in multimodal recognition systems, often combined with face recognition or other soft biometric traits. For instance, Paul and Gavrilova (Paul et al., 2012) demonstrate multimodal recognition by combining features extracted from both the face and the ear, creating cancelable templates using random projection and the fuzzy commitment scheme.

Previous studies have examined biometric authentication systems in IoT environments (Yang et al., 2019; Habib et al., 2014; Kantarci et al., 2015; Maček et al., 2016; Shahim et al., 2016; Dhillon and Kalra, 2017; Yang et al., 2019). Habib et al. (Habib et al., 2014) developed a pioneering authentication framework that combines biometric modalities with wireless device radio fingerprinting for IoT healthcare applications. This innovative approach not only ensures that health data originates from the correct patient but also protects its integrity. Similarly, Kantarci et al. (Kantarci et al., 2015) proposed a cloud-centric architecture for biometric identification, enhancing the security of mobile applications by integrating biometric and context-aware techniques to prevent unauthorized access. Macek et al. (Maček et al., 2016) introduced a multimodal authentication method, using high-quality cameras from devices like laptops, smartphones, and tablets to generate face and iris templates. However, they acknowledged privacy concerns related to the storage of these templates. Shahim et al. (Shahim et al., 2016) developed an authentication system based on hand geometry scans and a sequence of gestures, implemented on a Raspberry Pi platform. Dhillon and Kalra (Dhillon and Kalra, 2017) proposed a lightweight, multifactor remote user authentication technique that relies on a computationally efficient hash function and XOR operations. Yang et al. (Yang et al., Mar. 2019) introduced a privacy-preserving, lightweight fingerprint system designed for resource-constrained IoT devices. Their approach uses a block logic-based algorithm to reduce the size of the fingerprint templates. Additionally, Yang et al. (Yang et al., Jul. 2019) applied Binary Decision Diagrams (BDD) to a deep learning-based finger-vein system, using BDD to irreversibly transform the original finger-vein template. This transformed template is then processed by a multilayer extreme learning machine. A comparative evaluation against existing finger-vein recognition methods, both machine learning and non-machine learning-based, demonstrated the competitive performance of the proposed cancelable finger-vein template.

The methodology proposed in this study uses CaffeNet to extract feature vectors that remain consistent for classification. The training

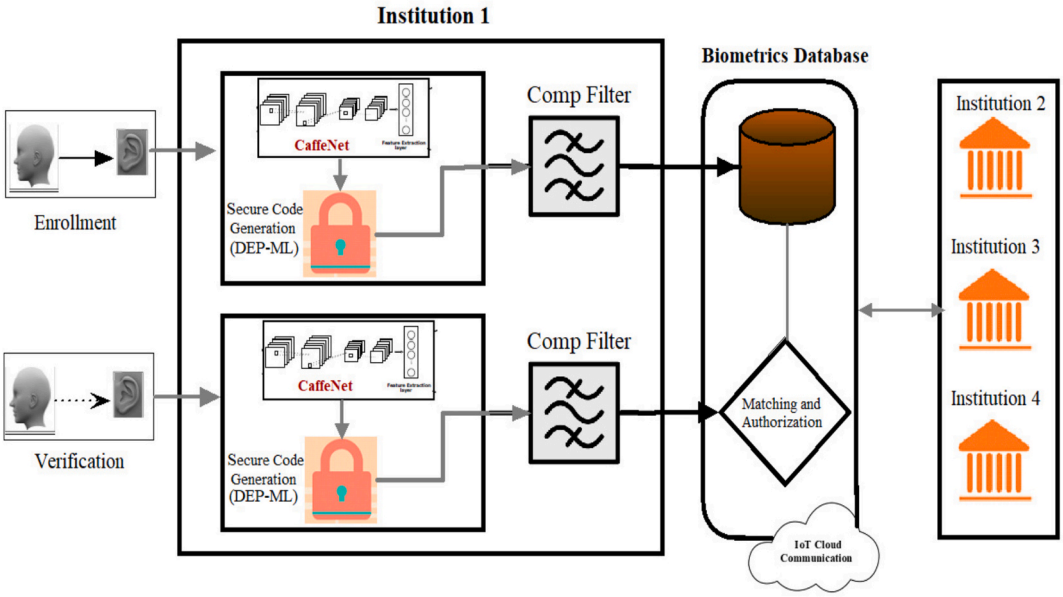

Fig. 1. Proposed authentication system based on secure cancelable ear biometrics.

process incorporates an innovative approach to metric learning, specifically the Deep Effective Pairwise Constraints Metric Learning (DEP-ML) technique, which is used to develop a discriminant function. This function helps map the extracted features from the CNNs. To enhance data security, the resulting EarCodes are encrypted using the Comb-filter algorithm (Ferreira and Wu, 2018). As a result, the proposed method effectively addresses challenges such as the non-uniformity of biometric data and difficulties in acquiring data from uncontrolled environments.

3. Proposed methodology

This study presents an innovative approach to ear biometric recognition, delivering both high accuracy and robust security without making assumptions about pose variations, changes in illumination, or potential security threats. The proposed method leverages CaffeNet as a feature extraction module, which processes images of the outer ear and outputs corresponding feature vectors. Due to its architecture, CaffeNet can be trained to produce features that are invariant to pose, lighting, and noise conditions. It consists of multiple convolutional and fully connected layers, with the penultimate layer, often referred to as bottleneck features (BNFs), serving as a reliable source of generic features for classification tasks.

Research has shown that descriptors from CaffeNet's bottleneck layer are highly effective for classification. Building on this, the study employs a novel classification strategy called Deep Effective Pairwise Constraints Metric Learning (DEP-ML) to classify the extracted BNFs with enhanced precision. Unlike conventional biometric systems that store original samples, thereby risking the exposure of Personally Identifiable Information (PII) or Sensitive Personal Information (SPI), this approach addresses those vulnerabilities directly. To enhance security, the proposed method incorporates the Comb-filter technique, which generates output templates (referred to as ear prints) by applying the Comb-filtering process. These templates are then distorted and securely stored for authentication.

The primary objective of the distortion process is to generate cancelable feature patterns that uniquely characterize everyone's ear while supporting template revocability. This is achieved by modifying the order of the Comb-filter, allowing the system to produce new cancelable patterns as needed, thereby enhancing both privacy protection and system adaptability. The overall architecture of the proposed framework is depicted in Fig. 1, which illustrates the integrated use of CaffeNet, DEP-ML, and the Comb-filter within a unified mechanism. The framework is organized into three key components: CaffeNet, serving as the feature extractor to generate deep representations of ear images, DEP-ML, acting as the classifier to evaluate feature similarity, and Comb-filter, used for secure and cancelable template generation.

During an authentication attempt, a test sample is processed through the trained model to compute an EarCode, which is then matched against the stored codes in the database. These EarCodes are non-invertible, meaning the original biometric data cannot be reconstructed, ensuring strong privacy protection. In the event of a security breach or compromise, a new set of EarCodes can be quickly generated by altering the Comb-filter settings, thus embedding cancelability into the system by design. The subsequent sections provide a detailed explanation of each component within the proposed framework.

3.1. Feature extraction based on transfer learning

3.1.1. CaffeNet

CaffeNet, developed by the Berkeley Vision and Learning Center (BVLC) (Jia et al., 2014), is a deep convolutional neural network known for its efficiency and simplicity compared to other neural architectures. Built on the Caffe deep learning framework, it comes with pre-trained models specifically optimized for various image classification tasks. CaffeNet's straightforward architecture makes it easy to understand and implement, while offering several practical advantages such as low memory consumption, high accuracy on benchmark datasets, and built-in image preprocessing tools. These features collectively make CaffeNet a reliable and efficient choice for a wide range of image recognition and understanding applications. Due to its proven effectiveness in visual recognition, CaffeNet was selected as the core feature extraction module in our proposed deep ear biometric recognition framework.

CaffeNet, a CNN-based architecture, excels at training and deploying general-purpose convolutional neural networks with impressive efficiency. It utilizes an N-dimensional data structure known as a "Blob" to manage data storage and flow efficiently within the CPU. Data is imported from the CPU into the network for computation. CaffeNet operates through a structured sequence of layers, including convolutional, pooling, and ReLU activation layers, while employing optimization losses such as SoftMax for classification tasks. Each layer of CaffeNet serves dual functions: the forward pass employs inputs to generate outputs, while the backward pass leverages gradients in conjunction

with outputs to compute gradients based on parameters and inputs. These gradients are then propagated backward to earlier layers in a process known as back-propagation. The architecture begins with a data layer that handles input loading from storage and ends with a loss layer that defines the objective function for classification, customized to the specific problem at hand.

The CaffeNet architecture consists of seven layers, beginning with an input layer and culminating in a SoftMax layer. The first five layers, including convolutional, pooling, and normalization layers, are dedicated to extracting detailed and discriminative features from the input image. The final two layers, a fully connected layer and the SoftMax layer, are responsible for classifying the ear image based on the extracted features.

4. Loss function

The choice of a loss function is critically important in optimizing any neural network-based model. In our proposed approach, we use the Cross-Entropy loss function for training. This loss function is particularly effective for multi-class classification problems, where the class label is expected to be an integer in the range $[0, \boldsymbol{C}-1]$, with $\boldsymbol{C}$ representing the number of classes.

The Cross-Entropy loss is defined in Eq. (1), where $\boldsymbol{x}$ is the input (typically the raw output or logits of the model), and class is the index of the correct target class.

$$loss\,(x,\ class) = -\log\left(\frac{\exp(x[class])}{\sum_j \exp(x[j])}\right) = -x[class] + log\left(\sum_j \exp(x[j])\right) \quad (1)$$

If a weight matrix is applied to handle class imbalance or importance, the weighted loss function is given in Eq. (2):

$$loss(x, class) = weight[class] \times \left(-x[class] + log\left(\sum_j \exp(x[j])\right)\right) \quad (2)$$

4.1. Parameter tuning

When working on a deep learning task involving an image dataset, the initial instinct is often to train a model from scratch. However, this approach can be problematic, especially when dealing with convolutional neural networks (CNNs), which typically contain millions of learnable parameters. Training such a large model on a small dataset increases the risk of overfitting and may hinder overall performance. To address this issue, a common and effective strategy is fine-tuning. This involves adapting a pre-trained CNN, originally trained on a large dataset, to a new, smaller dataset by updating its parameters.

In essence, fine-tuning retrains part of the network using the new dataset over several additional training iterations. In this study, we fine-tune the CaffeNet model. Specifically, we replaced the final fully connected layer with a new one, where the number of output units matches the number of classes in the target dataset. While this new layer is trained from scratch, the remaining layers of the network are kept frozen, meaning their weights remain unchanged during training. The goal of fine-tuning is to tailor the pre-trained model to the unique characteristics of the new dataset while preserving the valuable feature representations learned from the original, larger dataset. By training only the updated final layer, the model can effectively recognize patterns specific to the new task, ultimately leading to improved accuracy.

4.2. Proposed dynamic pairwise constraints metric learning (DEP-ML)

The proposed method integrates CaffeNet as a feature extractor with a novel metric learning technique, DEP-ML, to perform classification. This combination forms the CaffeNet-DEP-ML framework, which enables the mapping of ear images into compact representations called EarCodes. Initially trained for classification, CaffeNet becomes class-specific after training. However, retraining it to accommodate new classes is both challenging and dependent on collecting additional user images. To address this, the method leverages CaffeNet's robust feature extraction capabilities by using bottleneck features (BNFs), the activations from its penultimate layer, which serve as general, noise-resistant descriptors of biometric samples.

A key aspect of this approach is using DEP-ML to classify these BNFs. DEP-ML is designed to support incremental learning, allowing new users to be added to the system at any time without retraining the entire network. This flexibility is central to the adaptability of the proposed method. The remainder of this section explores the DEP-ML technique in detail. The training process uses a nearest neighbor strategy to dynamically update pairwise constraints, which are based on the Mahalanobis distance metric. This metric is optimized to learn meaningful distances between deep features extracted from the network. For classification and ear matching tasks, a K-Nearest Neighbors (KNN) approach is employed to make final predictions.

It is important to highlight that many traditional metric learning methods, such as Information-Theoretic Metric Learning (ITML) (Davis et al., 2007), are limited by their use of fixed pairwise constraints determined during the preprocessing stage. These static constraints can restrict the model's adaptability during training. To address this limitation, our proposed method, DEP-ML, introduces a dynamic strategy that seeks to optimize the kernel matrix using linearly effective pairwise constraints, implemented through the $\boldsymbol{n}$ cyclic projection method.

To clarify, consider a training set with $\boldsymbol{n}$ samples represented as $\{(a_i, a_j)|i = 1, 2, \cdots, n\}$, where each $\boldsymbol{a_i} \in \boldsymbol{R^d}$. In the first cycle, the model creates initial sets of similar and dissimilar sample pairs. In each subsequent cycle, these sets are updated dynamically to refine the learning process. At the start, the matrix $\boldsymbol{A}$ is initialized as the identity matrix $\boldsymbol{I}$. During training $\boldsymbol{x_t}$ and $\boldsymbol{y_t}$ represents maximum and minimum distances between sample pairs, respectively. In the $\boldsymbol{m^{th}}$ cycle, the model calculates the distance matrix $\boldsymbol{D_A}$. The previously learned metric $\boldsymbol{A_{t-1}}$ is used to compute Mahalanobis distances between all training samples.

To maximize the effectiveness of pairwise constraints, the model is trained over many cycles ($\boldsymbol{m}$). In each cycle, the nearest neighbor technique is used to iteratively refine and update these constraints. This approach allows for the integration of a greater number of pairwise relationships into the training process, thereby enhancing the learned distance function. By enriching the training phase with additional and dynamically updated pairwise constraints, our proposed method significantly improves the performance of metric learning. As a result, it contributes to increased recognition accuracy across a broad range of computer vision and machine learning tasks.

The characterization of our problem function can be framed as the LogDet divergence projection function denoted as $\boldsymbol{P_{ld}}(\boldsymbol{A}, \boldsymbol{A}_0)$ between two matrices $\boldsymbol{A}$ and $\boldsymbol{A}_0$, given by:

$P_{ld}(A, A_0) = tr\left(A, A_0^{-1}\right) - logdet\left(A, A_0^{-1}\right) - N$, where tr(.) represents the trace of the matrix, and $\boldsymbol{N}$ is the dimension of the training samples. As a result, the formulated classification problem can be expressed as shown in Equation (3):

$$\underset{A}{arg\,Min}\ P_{ld}(A, A_0)\ \text{s.t.}\, d_A\left(a_i, a_j\right) \le u\,,\ (i,j) \in S\ \&\ d_A\left(a_i, a_j\right) \ge l\,,\ (i,j) \in D \quad (3)$$

Here, $\boldsymbol{u}$ and $\boldsymbol{l}$ denote the upper and lower thresholds respectively, while $\boldsymbol{S}$ and $\boldsymbol{D}$ denote the sets of similar and dissimilar subjects. For a positive definite matrix $\boldsymbol{A}$, the function $\boldsymbol{d_A}$ defines the Mahalanobis distance between two points, $(\boldsymbol{a_i}, \boldsymbol{a_j})$. Both the proposed method and ITML are fast and scalable, as they operate without the need for semi-definite programming or eigenvalue computations. However, ITML relies on fixed constraints, which can lead to slower convergence and may

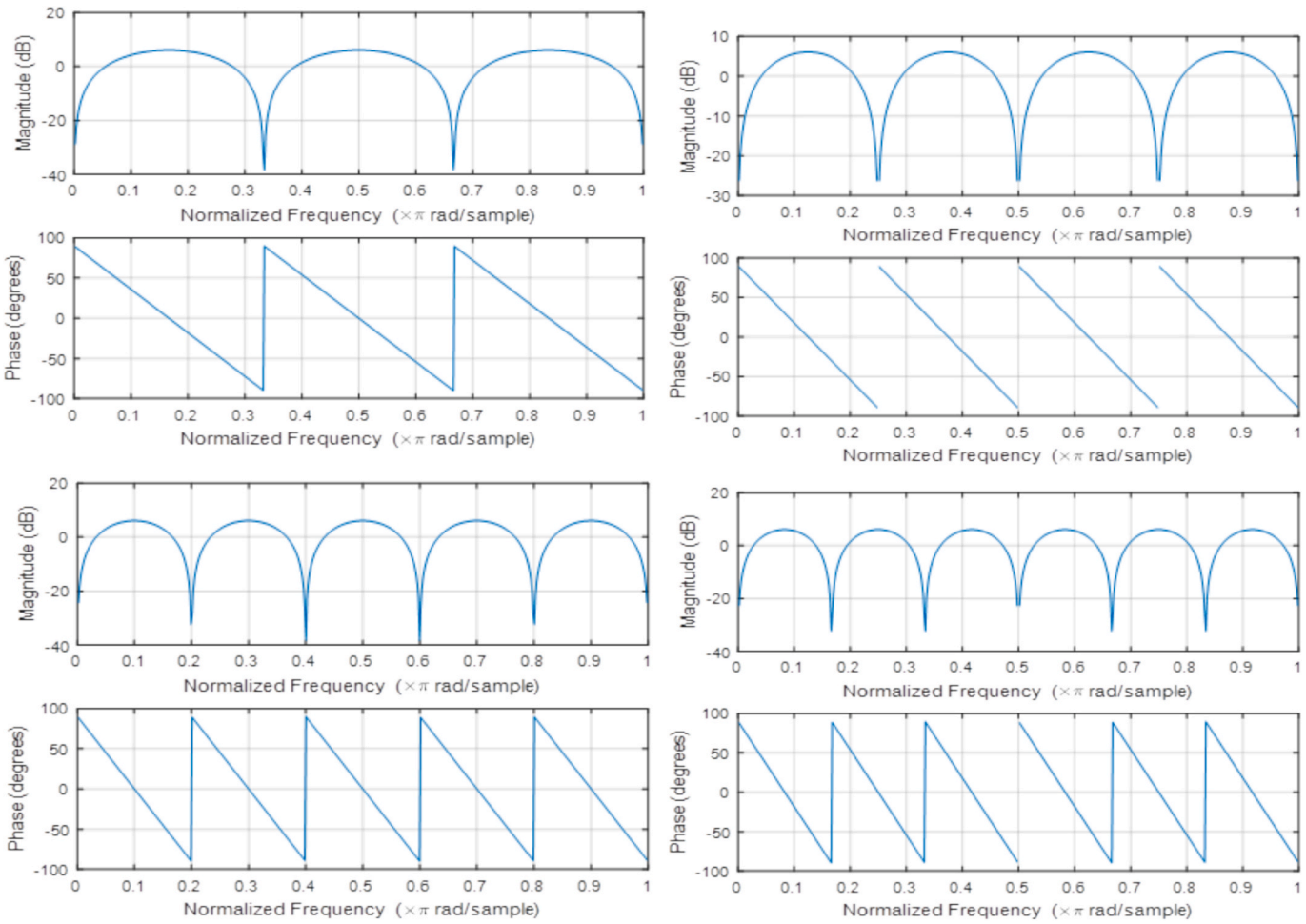

Fig. 2. The comb-filter magnitude and phase responses for order 6, 8, 10, and 12 respectively.

require more iterations. As a result, ITML may struggle to fully exploit the fixed pairwise constraints designed to improve KNN classifier performance. In contrast, our proposed method effectively overcomes these limitations by using multiple cycles to establish more flexible and effective constraints for distance learning, all without additional learning iterations. This leads to better performance and higher precision for the KNN classifier.

The proposed prediction algorithm can be summarized as follows:

First, we consider the training set $\{(x_i, y_i)|i = 1, 2, \cdots, n\}$ where i represents the number of subjects. The learning distance metric $\boldsymbol{d_A}$ defined by Eq. (4), serves as the targeted output:

$$d_A(a_i, a_j) = (a_i - a_j)^T A (a_i - a_j) \tag{4}$$

In the next step, we initialize $\boldsymbol{A}_0$ as the identity matrix $\boldsymbol{I}$, using it to compute the initial Mahalanobis distance $\boldsymbol{d_A}$. The process then iterates from **1** to the total number of cycles $\boldsymbol{m}$. For each cycle, the similarity and dissimilarity between every pair of training samples are computed using the distance metric $\boldsymbol{d_{Am}}$. This step involves identifying the nearest neighbors for each sample $\boldsymbol{a_i}$, both for similar and dissimilar instances.

Moving to the third step, the algorithm focuses on refining the distance metric. In each cycle, new pairwise constraints are generated, progressing from $\boldsymbol{P_l}$ to $\boldsymbol{P}_{(l+1)}$, based on the sets $\boldsymbol{S}$ and $\boldsymbol{D}$. (similar and dissimilar pairs). Following the ITML framework (Davis et al., 2007), the distance metric is updated through iterative Bregman projections, as shown in Eq. (5):

$$A_m = A_m + \beta_{ij} A_m (a_i - a_j)\ (a_i - a_j)^T A_m \tag{5}$$

where $\boldsymbol{a_i} - \boldsymbol{a_j}$ denotes a training pair in either $\boldsymbol{S}$ or $\boldsymbol{D}$, and $\boldsymbol{\beta_{ij}}$ is the corresponding Lagrange multiplier. Additionally, the algorithm monitors convergence by comparing the prediction and target metrics using a threshold $\boldsymbol{\eta}$ as defined in Eq. (6):

$$D_{A_{l+1}}(a_i, a_j) - D_{A_l}(a_i, a_j) < \eta \tag{6}$$

In this context, two possible scenarios arise based on the value of the threshold $\boldsymbol{\eta}$. When the convergence factor meets the $\boldsymbol{\eta}$ threshold, the proposed algorithm designates the Mahalanobis distance as the desired metric for the training model. Conversely, if the convergence factor falls short of the $\boldsymbol{\eta}$ threshold, the algorithm takes measures to enhance both the pairwise constraints and the Mahalanobis distance, as explained earlier, with the aim of achieving the desired convergence factor.

Finally, for the classification of similar and dissimilar subjects $(\boldsymbol{a_i}, \boldsymbol{a_j})$; the proposed approach employs the Lagrange multiplier $\boldsymbol{\beta_{ij}}$, a regularization parameter $\boldsymbol{\gamma}$ that balances the regularization function $\boldsymbol{P_{ld}}(\boldsymbol{A}, \boldsymbol{A}_0)$, initial distance $\boldsymbol{P_{ld}}(\boldsymbol{\xi}_0, \boldsymbol{\xi})$ $\boldsymbol{t}$, and the desired distance $\boldsymbol{\xi_{ij}}$, as demonstrated in Eq. (7).

$$\xi_{ij} = \gamma\, \xi_{ij} / (\gamma + \xi_{ij}) \tag{7}$$

4.3. Biometric template protection

In traditional biometric authentication systems, features extracted from a biometric sample are typically stored as a template in a database. However, this approach carries a significant risk: if the database is breached, it can lead to identity theft and serious privacy violations. To

mitigate this vulnerability, it is crucial to store templates in a way that prevents any direct connection to the original biometric data. This ensures that even if a template is compromised, it would be extremely difficult to impersonate the user.

One effective solution is to intentionally distort the extracted features using a multi-pass band Comb-filter, producing what are known as deformed EarCodes, which are then used for authentication. Incorporating a Comb-filter offers two major advantages: Its multi-band characteristics deliberately distort the pattern of ear features, significantly reducing the correlation between the stored template and the original biometric data. If a template is compromised, a new random code can easily be assigned as the user's identifier, introducing the critical concept of cancelability. An additional benefit is that the distortion introduced by the Comb-filter naturally varies across users, as the initial inputs to the filters differ.

In the proposed system, during an authentication attempt, the ear image sample first undergoes feature extraction via CaffeNet. These extracted features are then passed to DEP-ML for classification. The EarCode generated for the test sample is compared against the stored codes to verify the subject's identity. Importantly, neither the original biometric data nor the EarCode itself needs to be stored in the database, eliminating the risk of data leakage. Instead, non-invertible templates are saved for comparison during authentication. These non-invertible EarCodes protect user privacy by making it infeasible to reconstruct the original biometric sample.

4.4. EarCodes generation using Comb-filter

The proposed cancelable ear recognition system generates transformed features from the original ear characteristics. These modified features must retain enough discriminatory power to distinguish between different ear images. To achieve this, a Comb-filter is applied during the output phase of the CaffeNet model in the recognition system. To better understand this transformation, it is essential to first explain the principles behind the Comb-filter.

At its core, a Comb-filter can be thought of as a series of notch filters, where the points of minimum response (notches) occur periodically across the frequency spectrum. This pattern is reflected in both the magnitude and phase responses, as shown in Fig. 2 (Ferreira and Wu, 2018). The Comb-filter is considered a specialized type of multi-band filter, featuring multiple passbands and stopbands. Its distinctive feature is the equally spaced zeros in its frequency response, giving the magnitude response a Comb-like appearance.

Mathematically, the difference equation of a basic Comb-filter is given by (Kuo et al., 2006):

$$y(n) = x(n) - x(n - L) \tag{8}$$

where the integer $\boldsymbol{L}$ represents the delay length. The delay is $\boldsymbol{L} = \boldsymbol{f_S}/\boldsymbol{f_M}$ where $\boldsymbol{f_S}$ is the sampling frequency and $\boldsymbol{f_M}$ is the fundamental frequencies of the periodic null (Ferreira and Wu, 2018). The transfer function of this Finite Impulse Response (FIR) Comb-filter is:

$$H(z) = 1 - z^{-L} = \frac{z^L - 1}{z^L} \tag{9}$$

Therefore, at the origin, the Comb-filter has $\boldsymbol{L}$ poles, while having $\boldsymbol{L}$ zeros equally spaced on the unit circle at:

$$L \tag{10}$$

For example, when $\mathbf{L} = 8$, the zeros are located at:

$$z_l = 1,\ e^{\pi/4},\ e^{\pi/2},\ e^{3\pi/4}\ ,\ e^{\pi} = -1,\ e^{5\pi/4},\ e^{3\pi/2},\ e^{7\pi/4} \tag{11}$$

4.5. Secure storage of assigned EarCodes

This section presents a detailed explanation of the strategy developed to create secure biometric templates. The common security concerns associated with traditional biometric systems are effectively addressed using a Comb-filter. The Comb-filter plays a critical role by intentionally transforming arbitrary data blocks into multi-pass band sequences that are non-invertible. This transformation ensures that the stored templates bear no resemblance to the original biometric samples, thereby enhancing security.

The proposed use of the Comb-filter is an innovative and previously unexplored approach in the field of cancelable ear recognition. One major advantage of this method is that the Comb-filter, functioning as a multi-band filter, selectively removes specific frequencies at the null bands from the acquired feature map. This selective frequency removal leads to significant distortion of the patterns, which enhances user differentiation. Importantly, because the distortion varies with each user's unique input, pattern repetition is avoided. As a result, the system achieves high user discrimination and improved detection accuracy. The key attributes of the proposed approach can be summarized as follows:

The Comb-filter's multi-band nature deliberately disrupts ear feature patterns.
The altered versions of the ear patterns are used for subject authentication.
The distortions are unique to each user due to different initial inputs into the Comb-filters.

In this system, features extracted from ear images using CaffeNet are first mapped into EarCodes. These EarCodes are then intentionally deformed by applying the multi-pass band Comb-filter during the final stage of enrollment, ensuring secure storage. The use of the Comb-filter is based on its proven capability to enhance security, as it can selectively pass or suppress specific frequencies and their harmonics. This property makes it highly suitable for our cancelable ear recognition framework.

Finally, the system guarantees the cancelability of EarCodes in the event of database breach, thanks to the modifiability of the filter order, allowing new, secure templates to be easily generated.

5. Experimental setup

5.1. Datasets

Our proposed model was tested on two distinct ear image databases: the University of Science and Technology Beijing (USTB-II) (Mu et al., 2004) and the Annotated Web Ear (AWE) dataset (Emersic et al., 2017). The AWE dataset is particularly challenging because it contains ear images captured in unconstrained environments. It includes 1,000 images from 100 subjects, with 10 ear images per subject.

Similarly, the USTB-II dataset consists of 380 facial profile images collected from 77 individuals, covering a wide range of illumination conditions. Each subject is represented by four images: The first image is a frontal ear view captured under standard lighting. The second and third images depict the ear rotating at $+30^{\circ}$ and -30°, respectively. The fourth image is captured under low-light conditions.

To evaluate the performance of the proposed methodology, several key metrics were used, including the Receiver Operating Characteristic (ROC) curve (Benzaoui et al., 2014), accuracy, and the confusion matrix. The ROC curve is generated by plotting the True Positive Rate (TPR) against the False Negative Rate (FNR). These metrics together provide a comprehensive assessment of the model's effectiveness, with the ROC curve's average performance calculated across all classes.

5.2. Training

For the input image preprocessing, the proposed scheme applies linear normalization of the gray scale images by the following Eq. (12):

Table 1
Recognition rates (%) for unprotected EarCodes databases (KNN = 1, 2, and 3).

Unprotected EarCodes proposed methods		Databases for training ratio (60 %)					
		AWE			USTB II		
VGG – verydeep16	KNN	71.80	71.80	73.90	90.08	90.57	90.57
		1.55	1.55	1.22	4.78	5.00	5.17
	ITML distance Classification	71.90	66.90	70.30	96.43	86.04	82.14
		1.25	3.09	1.12	3.75	2.35	1.22
	Mahalanobis distance Classification	72.88	72.88	73.55	90.08	90.16	90.16
		1.88	1.88	2.16	4.16	3.95	3.95
	Proposed Metric Learning	**70.93**	**70.93**	**74.15**	**92.85**	**92.85**	**92.60**
	(DEP-ML)	**1.86**	**1.86**	**2.15**	**3.65**	**3.65**	**3.56**
VGG-S	KNN	70.13	70.13	71.68	92.93	93.09	93.10
		2.42	2.42	2.73	5.66	5.68	5.73
	ITML distance Classification	68.5	67.20	69.57	94.16	81.49	77.27
		1.34	2.34	3.45	2.45	4.04	2.34
	Mahalanobis distance Classification	71.55	71.55	73.23	93.98	94.15	94.15
		1.46	1.46	1.72	5.49	5.36	5.36
	Proposed Metric Learning	**70.80**	**70.80**	**72.40**	**92.68**	**92.68**	**92.44**
	(DEP-ML)	**1.88**	**1.88**	**1.97**	**4.49**	**4.49**	**4.62**
CaffeNet	KNN	72.25	72.25	74.35	92.96	92.96	93.65
		1.85	1.85	1.98	3.24	3.24	3.68
	ITML distance Classification	73.23	73.23	72.98	96.98	92.33	88.75
		2.23	2.23	2.42	3.12	2.45	2.56
	Mahalanobis distance Classification	72.96	72.96	73.85	94.12	94.12	95.33
		2.45	2.45	2.52	2.38	2.38	3.24
	Proposed Metric Learning	**72.18**	**73.65**	**74.43**	**93.33**	**93.33**	**93.17**
	(DEP-ML)	**1.35**	**1.43**	**1.58**	**3.27**	**3.27**	**2.93**

$$Img_N = (Img - Min)\frac{NMax - NMin}{Max - Min} + NMin \tag{12}$$

where $\boldsymbol{Img_N}$ is the normalized image, ***Img*** is the original image, ***NMax*** and ***NMin*** are the desired pixels range intensity. Subsequently, considering the varying sizes of available and public ear databases, such as USTB-II and the AWE datasets, all images need to undergo a conversion process to become three-channel images by resizing them to a consistent dimension of $224x224x3$ pixels. Additionally, alignment and color enhancement are performed before the feature extraction phase to ensure compatibility with the pre-trained CaffeNet model.

For performance evaluation, each database is divided into a training set (60 %) and a testing set (40 %). In the case of the AWE dataset, this split is already provided by the AWE toolbox (Emersic et al., 2017). Following a closed-set experimental protocol like that used in (Emersic et al., 1711; Dodge et al., 2018), identification experiments are conducted, where the model predicts the correct class for each input image.

To assess the effectiveness of the proposed architecture, key performance metrics are used, including the ROC curve, accuracy, and the confusion matrix.

6. Experimental results

The experiments were conducted in two distinct phases. In the first phase, all tests were performed without applying any data protection. In the second phase, the datasets were protected using Comb-filters with orders 6, 8, 10, and 12, which effectively expanded the original training data by a factor of 10, to support further experimentation. Two datasets and their various combinations were used for these experiments.

To explain the motivation behind securing the ear recognition system, we introduce three models that employ VGG-very deep16, VGG S, and CaffeNet for extracting features from ear images. These models were paired with different metric learning techniques such as KNN, Mahalanobis distance, and ITML for ear classification. Additionally, a novel approach known as DEP-ML is explored for ear classification. DEP-ML involves learning the Mahalanobis distance and KNN classifier using VGG-very deep16, VGG S, and CaffeNet, while also incorporating cancelability through the application of Comb-filters with orders 6, 8, 10, and 12 to the extracted ear features.

The experimental results, across both protected and unprotected EarCodes, showed that our DEP-ML approach consistently outperformed traditional metric learning methods (KNN, Mahalanobis, and ITML) when paired with all three feature extractors. For instance, in the case of unprotected EarCodes, the recognition rates for VGG 16-DEP-ML, VGG S-DEP-ML, and CaffeNet-DEP-ML on the AWE dataset are 74.15 ± 2.15, 72.40 ± 1.97, and 74.43 ± 1.58, respectively. Similarly, on the USTB II dataset, the unprotected EarCodes recognition rates for VGG 16-DEP-ML, VGG S-DEP-ML, and CaffeNet-DEP-ML are 92.60 ± 3.56, 92.44 ± 4.62, and 93.17 ± 2.93, respectively. CaffeNet offers several advantages over VGG in cancelable biometric schemes, primarily due to its lower intra-class variance. It also preserves essential spatial details needed for effective ear recognition without the over-parameterization seen in VGG. Additionally, CaffeNet requires less training data to achieve convergence, making it more suitable for small biometric datasets where VGG's data-intensive nature increases the risk of overfitting. Its simpler architecture and significantly fewer parameters make CaffeNet less computationally demanding and easier to train. These qualities also make it well-suited for deployment in resource-constrained environments, such as IoT systems. Further, CaffeNet retains sufficient spatial details for ear recognition while avoiding the over-parameterization seen in VGG, making it less dependent on large datasets for convergence. This efficiency makes CaffeNet more suitable for small biometric datasets, where VGG's data-intensive nature increases the risk of overfitting. Moreover, in cancelable biometric schemes, CaffeNet features demonstrate lower intra-class variance compared to VGG (Hahn and Marcel, 2022; Yang et al., 2024). Regarding ear classification, ITML's limitations stem from its use of fixed pairwise constraints established during training. In contrast, our proposed method, DEP-ML, addresses this by optimizing the kernel matrix using linearly effective pairwise constraints determined through the n-cyclic projection method during preprocessing.

When cancelability was introduced using Comb-filters, it was observed that the Comb-filter of order 10 produced the best recognition accuracy. Therefore, the protected EarCodes recognition rates using the Comb-filter of order 10 for VGG 16-DEP-ML, VGG S-DEP-ML, and CaffeNet-DEP-ML on the AWE dataset are 74.70 ± 2.06, 71.88 ± 2.36, and 74.85 ± 1.68, respectively. Similarly, on the USTB II dataset, the protected EarCodes recognition rates using the Comb-filter of order 10 for VGG 16-DEP-ML, VGG S-DEP-ML, and CaffeNet-DEP-ML are 92.68

Table 2
Recognition rates (%) for protected EarCodes databases using Comb-filter of order 6 (KNN = 1, 2, and 3).

Protected EarCodes	Proposed Methods	Databases					
		AWE			USTB II		
VGG – verydeep16	KNN	71.60	71.60	74.60	89.78	89.19	89.19
		2.44	2.44	2.51	4.52	4.30	4.30
	ITML distance Classification	72.10	63.80	70.50	91.88	76.95	79.55
		1.71	2.65	1.51	1.02	2.41	2.34
	Mahalanobis distance Classification	67.60	67.60	70.18	86.18	86.67	86.67
		2.01	2.01	2.12	4.32	4.49	4.49
	Proposed Metric Learning	**71.73**	**71.73**	**74.53**	**91.14**	**91.14**	**90.82**
	(DEP-ML)	**2.50**	**2.50**	**2.41**	**5.75**	**5.75**	**5.81**
VGG-S	KNN	70.15	70.15	72.00	90.89	90.89	90.73
		2.09	2.09	1.90	4.38	4.38	4.31
	ITML distance Classification	70.5	63.50	66.80	99.03	84.74	86.36
		2.85	4.25	3.17	3.15	3.49	3.65
	Mahalanobis distance Classification	65.72	65.72	67.98	89.19	89.59	89.59
		3.14	3.14	2.74	3.87	3.31	3.31
	Proposed Metric Learning	**70.25**	**70.25**	**71.98**	**91.79**	**91.79**	**91.79**
	(DEP-ML)	**2.28**	**2.28**	**2.72**	**3.72**	**3.72**	**3.72**
CaffeNet	KNN	71.88	71.88	73.90	91.46	91.46	91.46
		1.41	1.41	1.11	2.40	2.40	2.43
	ITML distance Classification	72.89	72.28	72.15	95.54	91.45	87.82
		1.92	1.72	1.89	2.71	2.24	3.17
	Mahalanobis distance Classification	72.28	72.28	72.40	94.01	94.01	93.98
		2.23	2.23	2.00	3.12	3.12	3.19
	Proposed Metric Learning	**71.88**	**73.53**	**74.58**	**92.31**	**92.31**	**92.15**
	(DEP-ML)	**1.89**	**1.75**	**1.91**	**3.07**	**3.07**	**3.09**

Table 3
Recognition rates (%) for protected EarCodes databases using Comb-filter of order 8 (KNN = 1, 2, and 3).

Protected EarCodes	**Proposed Methods**	**Databases**					
		AWE			**USTB II**		
VGG – verydeep16	KNN	71.56	71.56	74.20	91.46	91.46	91.46
		3.28	3.28	3.16	3.99	3.99	4.11
	ITML distance Classification	74.00	65.70	69.6	96.75	86.36	75.00
		2.45	2.15	2.65	2.12	1.89	3.66
	Mahalanobis distance Classification	68.43	68.43	71.10	87.72	87.72	87.89
		1.93	1.93	1.74	5.51	5.51	5.97
	Proposed Metric Learning	**72.00**	**72.00**	**75.13**	**91.79**	**91.79**	**91.54**
	(DEP-ML)	**1.15**	**1.15**	**1.17**	**2.98**	**2.98**	**2.90**
VGG-S	KNN	70.40	70.40	71.65	93.98	94.23	94.23
		1.61	1.61	1.86	1.51	1.51	1.72
	ITML distance Classification	70.50	62.90	66.30	96.10	82.14	79.22
		3.18	2.77	3.15	2.85	5.23	4.15
	Mahalanobis distance Classification	66.58	66.58	78.18	89.75	89.92	89.92
		2.19	2.19	2.39	3.87	3.86	3.86
	Proposed Metric Learning	**70.40**	**70.40**	**72.23**	**92.28**	**92.28**	**92.11**
	(DEP-ML)	**2.59**	**2.59**	**2.40**	**4.88**	**4.88**	**5.01**
CaffeNet	KNN	71.85	71.85	73.23	93.14	93.14	93.89
		2.44	2.44	2.31	2.26	2.26	2.07
	ITML distance Classification	73.01	72.58	72.21	95.68	91.45	88.18
		1.08	1.39	2.23	2.81	2.14	2.87
	Mahalanobis distance Classification	72.38	72.38	73.53	94.09	94.09	94.17
		2.97	2.97	2.54	3.14	3.14	3.21
	Proposed Metric Learning	**72.28**	**74.03**	**75.00**	**93.26**	**93.26**	**93.10**
	(DEP-ML)	**2.62**	**2.65**	**2.61**	**2.74**	**2.74**	**2.05**

$\pm$ 5.54, 92.36 $\pm$ 4.62, and 95.04 $\pm$ 2.93, respectively.

The proposed cancelable CaffeNet-DEP-ML approach demonstrates a significantly stronger impact compared to both unprotected and other cancelable EarCodes models, particularly under KNN = 3 distance testing repeated over 10 trials, as detailed in Tables 1–5. The results clearly show that incorporating Comb-filters into the CaffeNet-DEP-ML model notably enhances ear recognition performance while simultaneously ensuring privacy preservation.

Figs. 3 and 4 present the ROC curves for both unprotected and cancelable EarCodes across the two datasets. As observed, the proposed CaffeNet-DEP-ML approach consistently outperforms the other two methods on both the AWE and USTB-II datasets. This superiority is due to the combined power of CaffeNet's deep learning architecture and the novel DEP-ML metric learning strategy. Using CNN-based feature extraction, which generates highly discriminative feature vectors, proves to be more effective than traditional handcrafted features. In this context, CaffeNet surpasses manually designed feature extraction pipelines, resulting in better classification performance. Consequently, the proposed CaffeNet-DEP-ML method outshines both VGG 16-DEP-ML and VGG S-DEP-ML models for both protected and unprotected EarCodes, even with the integration of Comb-filters of orders 6, 8, 10, and 12.

Figs. 5 and 6 further support these findings by showcasing the confusion matrices for both unprotected and protected EarCodes on the two datasets. The confusion matrix, a standard tool for evaluating classification models, illustrates the relationship between predicted and

Table 4
Recognition rates (%) for protected EarCodes databases using Comb-filter of order 10 (KNN = 1, 2, and 3).

Protected EarCodes Proposed Methods		Databases					
		AWE			USTB II		
VGG − verydeep16	KNN	71.48	71.48	74.83	90.16	90.41	90.41
		1.91	1.91	2.38	3.60	3.51	3.51
	ITML distance Classification	70.20	64.20	71.00	98.05	83.77	83.77
		0.95	3.84	2.32	3.24	2.51	2.20
	Mahalanobis distance Classification	68.95	68.95	71.53	86.35	86.35	86.67
		2.60	2.60	2.84	5.01	5.01	5.11
	Proposed Metric Learning	**71.50**	**71.50**	**74.70**	**92.60**	**92.60**	**92.68**
	(DEP-ML)	**1.78**	**1.78**	**2.06**	**5.66**	**5.66**	**5.54**
VGG-S	KNN	71.45	71.45	72.93	94.55	94.87	94.87
		1.55	1.55	1.93	3.49	3.49	3.39
	ITML distance Classification	71.3	63.50	67.30	95.45	83.77	82.79
		2.21	2.18	2.01	1.25	1.18	1.33
	Mahalanobis distance Classification	68.00	68.00	69.65	91.55	92.20	92.20
		1.57	1.57	1.88	2.85	3.05	3.05
	Proposed Metric Learning	**70.25**	**7.25**	**71.88**	**92.44**	**92.44**	**92.36**
	(DEP-ML)	**2.67**	**2.67**	**2.36**	**4.62**	**4.62**	**4.62**
CaffeNet	KNN	72.60	72.60	74.88	93.76	93.76	94.34
		2.30	2.30	2.43	1.88	1.88	1.78
	ITML distance Classification	73.14	73.14	72.73	95.78	91.69	88.38
		2.92	2.92	3.05	2.91	3.74	3.68
	Mahalanobis distance Classification	72.20	72.20	74.05	94.33	94.33	95.25
		2.39	2.39	2.36	2.32	2.32	2.43
	Proposed Metric Learning	**72.68**	**73.93**	**74.85**	**94.88**	**94.88**	**95.04**
	(DEP-ML)	**1.83**	**1.72**	**1.68**	**3.02**	**3.02**	**2.93**

Table 5
Recognition rates (%) for protected EarCodes databases using Comb-filter of order 12 (KNN = 1, 2, and 3).

Protected EarCodes proposed methods		Databases					
		AWE			USTB II		
VGG − verydeep16	KNN	72.13	72.13	74.98	91.30	91.63	91.63
		2.23	2.23	2.64	3.80	3.49	9.49
	ITML distance Classification	72.80	63.80	71.60	91.88	83.12	84.42
		1.33	2.39	1.29	3.25	2.20	2.85
	Mahalanobis distance Classification	69.12	69.12	71.35	88.37	88.46	88.46
		2.23	2.23	2.30	4.55	4.56	4.56
	Proposed Metric Learning	**71.68**	**71.68**	**74.75**	**92.60**	**92.60**	**92.20**
	(DEP-ML)	**2.12**	**2.12**	**2.68**	**3.82**	**3.82**	**3.87**
VGG-S	KNN	72.00	72.00	73.30	93.17	93.41	93.41
		2.54	2.54	2.26	4.17	4.31	4.31
	ITML distance Classification	67.00	63.40	66.10	93.18	84.74	80.52
		2.45	3.55	2.12	1.31	2.58	3.98
	Mahalanobis distance Classification	66.75	66.75	68.60	88.53	88.37	88.37
		2.92	2.92	2.87	5.21	5.26	5.26
	Proposed Metric Learning	**70.63**	**70.63**	**71.90**	**92.20**	**92.20**	**91.79**
	(DEP-ML)	**3.20**	**3.20**	**2.50**	**3.68**	**3.68**	**3.86**
CaffeNet	KNN	72.45	72.45	74.55	93.65	93.65	94.25
		2.15	2.15	2.42	2.12	2.12	2.78
	ITML distance Classification	73.04	73.04	72.63	96.34	92.45	88.28
		3.25	3.25	3.14	2.61	3.24	4.87
	Mahalanobis distance Classification	72.66	72.66	73.76	93.45	93.45	93.98
		3.42	3.42	3.37	3.89	3.89	4.15
	Proposed Metric Learning	**72.80**	**73.93**	**74.68**	**93.17**	**93.17**	**93.09**
	(DEP-ML)	**2.43**	**2.62**	**2.53**	**3.56**	**3.56**	**3.74**

actual labels. The proposed CaffeNet-DEP-ML outperforms VGG 16-DEP-ML and VGG S-DEP-ML, for both unprotected and cancelable EarCodes, confirming the effectiveness of the proposed architecture across different protection levels.

In summary, a comprehensive comparison among the three models shows that CaffeNet-DEP-ML achieves the highest accuracy for both unprotected and cancelable EarCodes. This conclusion is supported by both the numerical results in the tables and the graphical representations in the figures. Notably, when using Comb-filters of order 10, the proposed cancelable CaffeNet-DEP-ML model delivers superior recognition accuracy while effectively preserving privacy. The success of the proposed methodology lies in its integration of three key components: A deep feature extraction module based on the CaffeNet architecture, A novel DEP-ML metric learning technique for classification, and the use of Comb-filters (orders 6, 8, 10, and 12) for template protection. This combination results in outstanding performance in both recognition accuracy and template security. The application of Comb-filters ensures that the protected templates maintain diversity and non-invertibility, strengthening privacy without compromising performance. Ultimately, the impressive recognition results achieved by the cancelable framework highlight the crucial importance of developing secure ear recognition systems that maintain a strong balance between high accuracy and robust privacy preservation.

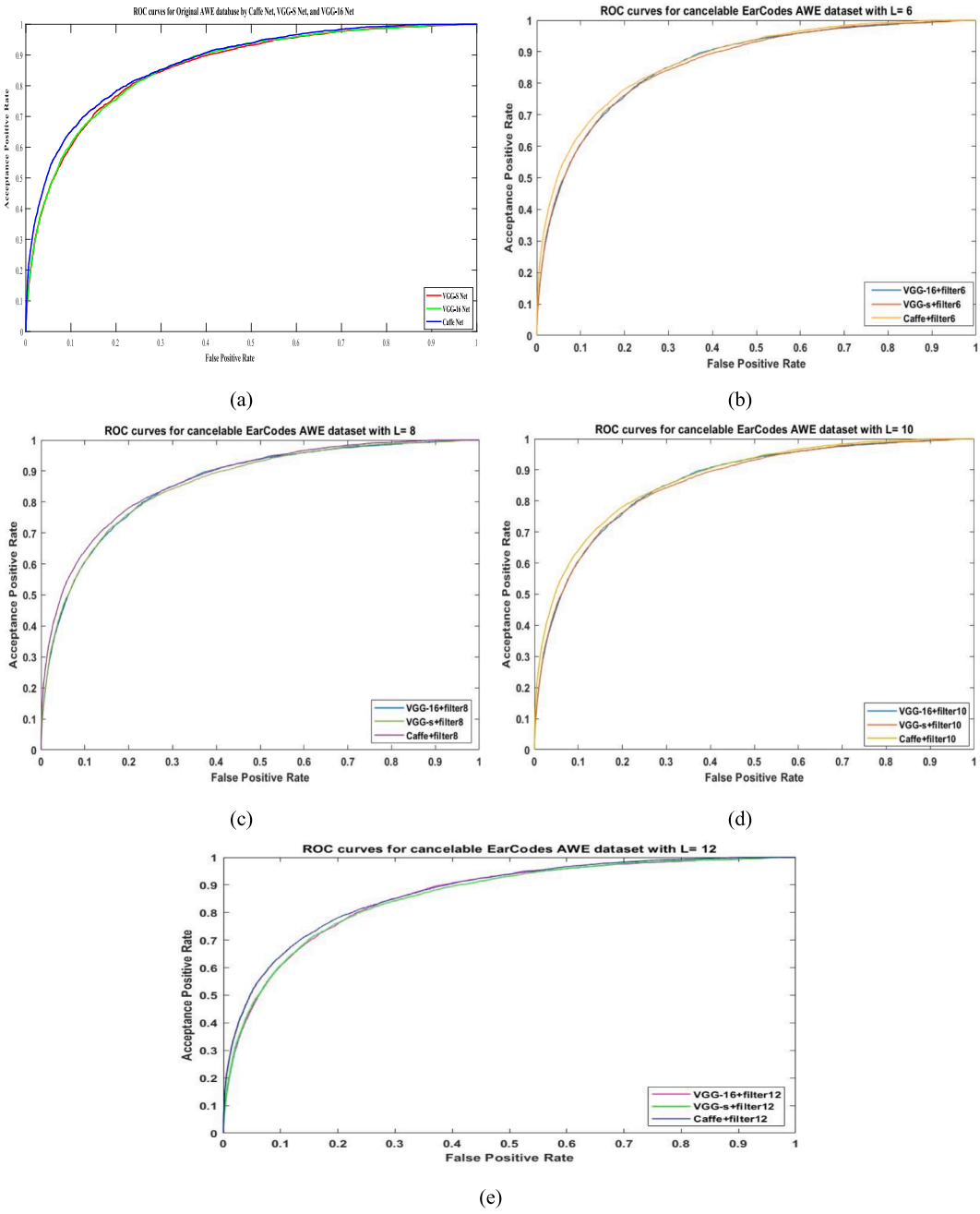

Fig. 3. ROC curves for various deep features of the AWE dataset of unprotected and the proposed cancelable EarCodes schemes (a) Unprotected EarCodes, (b) L = 6, (c) L = 8, (d) L = 10, (e) L = 12.

6.1. *Security analysis*

This section discusses the security aspects of the proposed biometric authentication framework:

Result analysis: The accuracy metrics presented in Tables 1–5 demonstrate strong recognition performance, both before and after applying cancelability. These results confirm the effectiveness of the proposed method in achieving high recognition accuracy.

Suitability for IoBT assessment: The ear, as a stable and distinctive biometric trait, shows great potential for biometric monitoring systems aligned with Industry 4.0 initiatives, specifically within the framework of the IoBT (Castiglione et al., 2017). In the IoBT model, users can easily utilize cost-effective biometric devices connected to cloud-based platforms. This setup not only strengthens security and maintains privacy but also allows enterprises to monitor biometric systems across multiple locations. Experimental results validate that the proposed cancelable deep learning model is highly compatible with the IoBT architecture, particularly when applied to ear biometrics.

Privacy preservation analysis: Traditional privacy-preserving biometric systems often rely on cryptographic techniques to secure templates within databases, protecting against unauthorized access or misuse. In contrast, the proposed cancelable ear recognition system takes a different approach: instead of storing original image templates, it saves only the concealed EarCodes for classification purposes. The integration of a multi-passband Comb-filter is central to this privacy strategy. By intentionally distorting ear features, the Comb-filter minimizes any direct correlation between the biometric sample and the stored EarCodes, significantly enhancing template

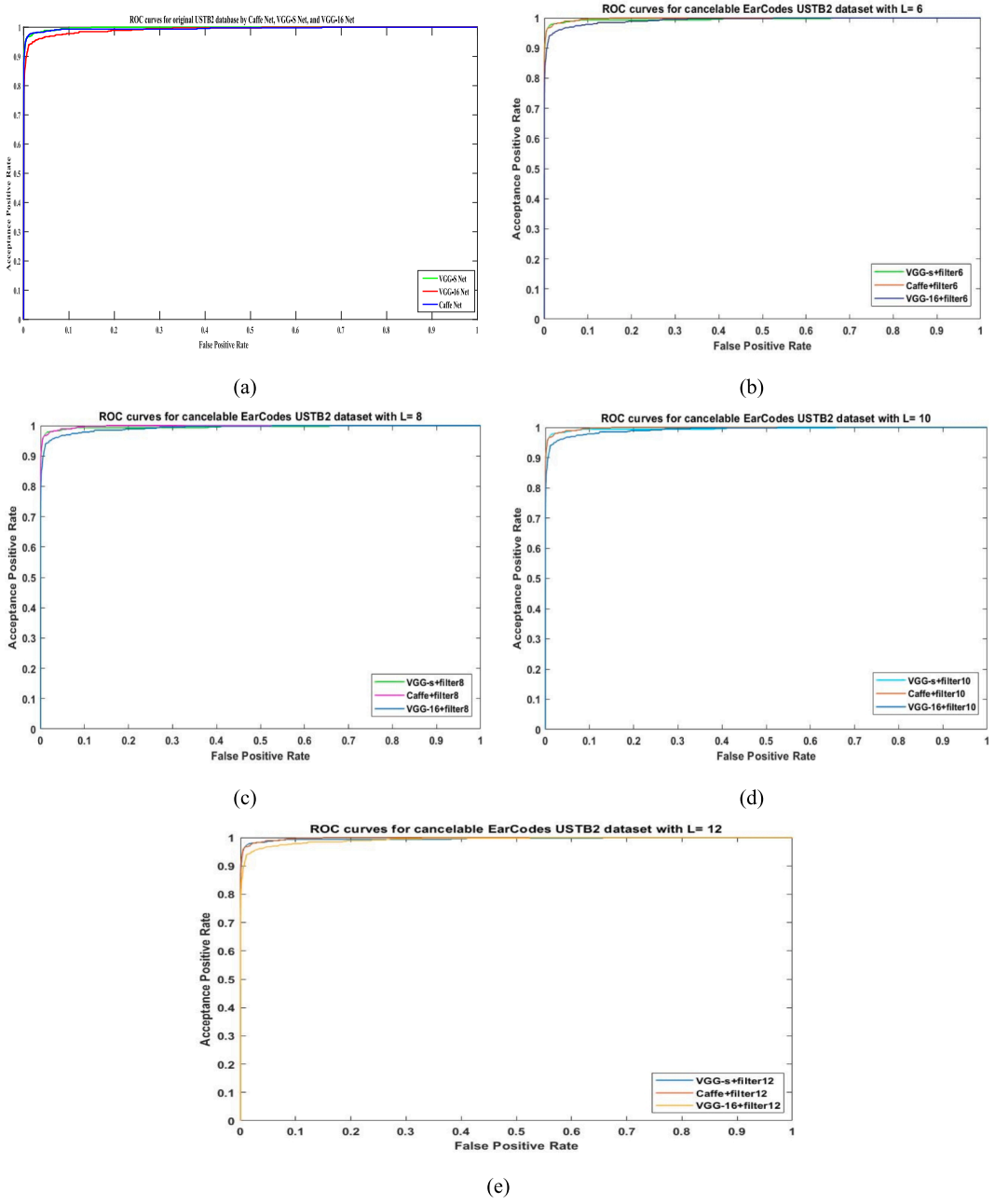

Fig. 4. ROC curves for various deep features of the USTB II dataset of unprotected and proposed cancelable EarCodes schemes (a) Unprotected EarCodes, (b) L = 6, (c) L = 8, (d) L = 10, (e) L = 12.

security. In case a template is compromised, a new random code can be easily generated, ensuring cancelability. Additionally, because the distortion process varies depending on the initial input into the Comb-filters, each user's protected template is uniquely different. This further reinforces both the security and privacy preservation capabilities of the proposed system.

7. Performance comparison

This section compares the performance of our method with several well-known techniques that have used the AWE dataset. In particular, Hansley et al. (Hansley et al., 2018) and Emersic et al. (Emersic et al., 2017) investigated various approaches ranging from holistic and handcrafted features to learned representations to tackle the challenge of ear recognition in unconstrained settings. Fig. 7 compares the performance of various feature extraction methods such as Local Binary Patterns (LBP) (Hansley et al., 2018), Dense SIFT (DSIFT) (Hansley et al., 2018), Patterns of Oriented Edge Magnitudes (POEM) (Emersic et al., 2017), and Convolutional Neural Networks (CNN) (Hansley et al., 2018) on the AWE dataset. It also shows the proposed framework, which outperforms all other techniques. The proposed framework achieves Equal Error Rates (EER) of 15.25 % for unprotected EarCodes while the proposed cancelable framework achieves the lowest EER of 15.03 % for cancelable EarCodes. It is obvious that the proposed scheme can generate multiple unique ear patterns for each application and can create new patterns if the database is compromised. Importantly, it preserves the original recognition performance while effectively preventing the extraction of sensitive information from the transformed patterns.

As demonstrated by Lecun et al. (LeCun et al., 1989), selecting

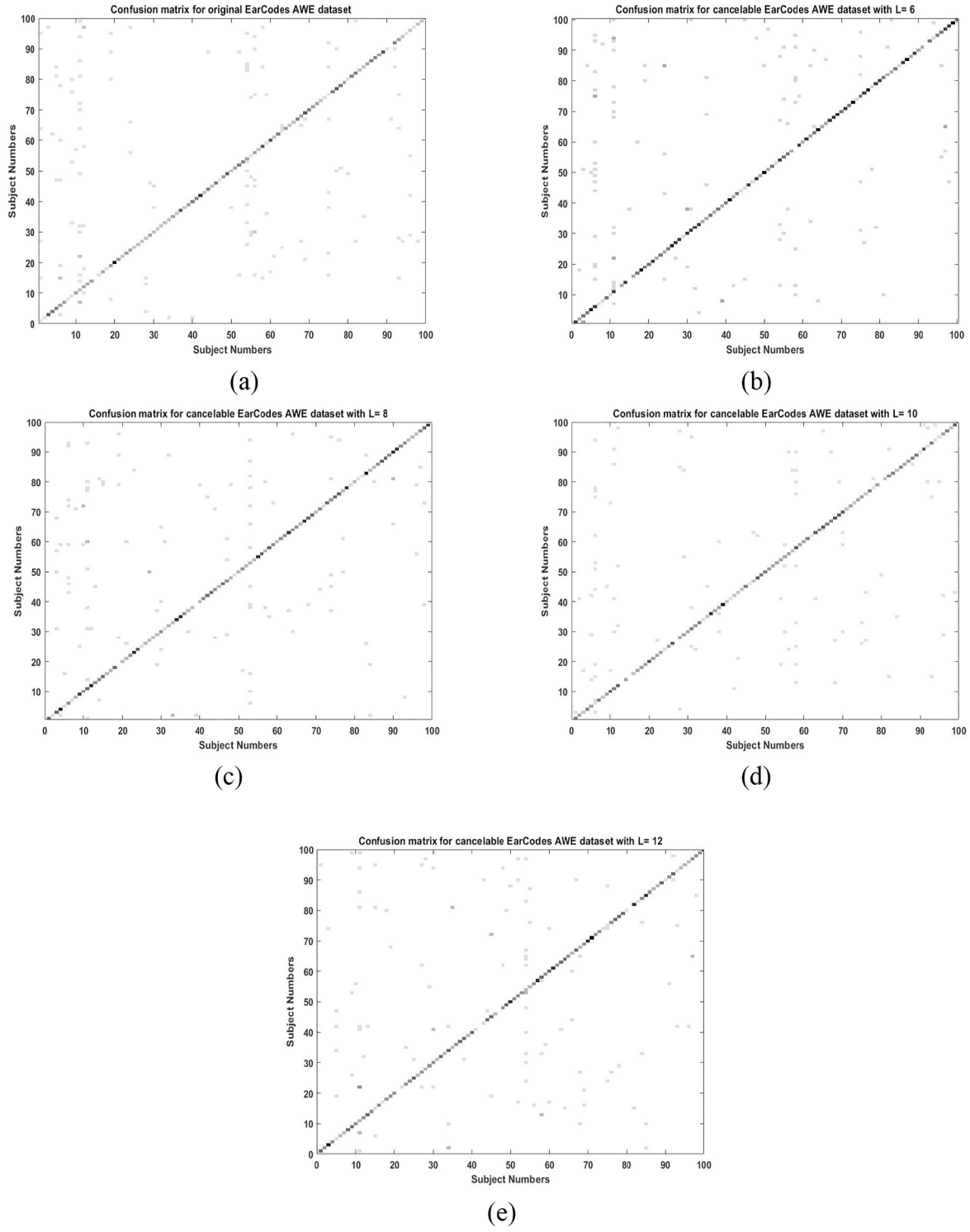

Fig. 5. Confusion matrix for the AWE dataset of unprotected and the proposed cancelable EarCodes schemes (a) Unprotected EarCodes, (b) L = 6, (c) L = 8, (d) L = 10, (e) L = 12.

features based on classifier performance is a more effective approach. The proposed CaffeNet-DEP-ML model benefits from the combination of CaffeNet's deep learning architecture and our proposed DEP-ML metric learning method. This integration allows for the extraction of highly discriminative features using CNNs, which significantly outperforms traditional handcrafted methods. In this regard, CaffeNet proves more effective than manually designed feature extraction pipelines, resulting in improved classification accuracy. When using Comb-filters of order 10, the proposed cancelable CaffeNet-DEP-ML model not only delivers superior recognition performance but also ensures strong privacy preserving.

8. Conclusions

This paper presents a novel cancelable ear biometric authentication system that effectively combines three key components: CaffeNet for feature extraction, DEP-ML for metric learning and class ranking, and the Comb-filter for generating secure, non-invertible EarCodes.

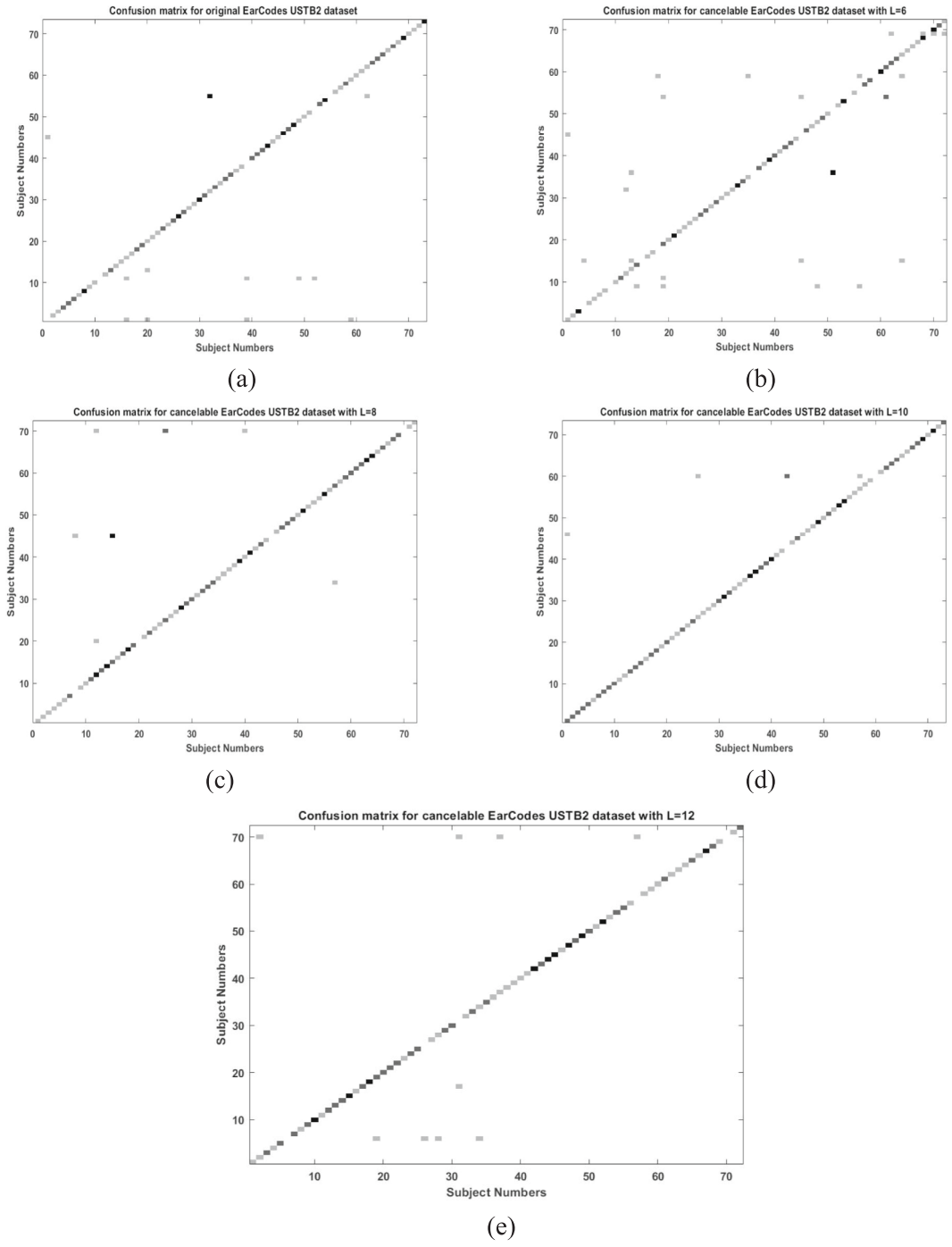

Fig. 6. Confusion matrix for the USTB II dataset of unprotected and the proposed cancelable EarCodes schemes (a) Unprotected EarCodes, (b) L = 6, (c) L = 8, (d) L = 10, (e) L = 12.

Together, these elements create robust and privacy preserving ear biometric authentication framework. Extensive evaluations and experiments on the proposed cancelable CaffeNet-DEP-ML system demonstrate its outstanding performance in both recognition accuracy and privacy preserving. The key contributions can be concluded as follows:

(i) Feasibility of ear biometrics: The experiments validate the use of the outer ear image as a reliable biometric trait, achieving high recognition rates and confirming its potential for biometric authentication
(ii) Security through non-invertible templates: The proposed cancelable CaffeNet-DEP-ML approach incorporates Comb-filter for template generation and storage, the system ensures that stored templates are non-invertible, effectively mitigating the risk of unauthorized intrusions or template reconstruction
(iii) Enhanced security via cancelability: The system guarantees the cancelability of EarCodes, even in scenarios where the database might be compromised, adds an extra layer of security even in the event of database compromise. This flexibility is further enhanced by the ability to adjust the filter order, allowing enterprises to fine-tune the security level according to their specific needs

Our long-term goal is to improve the versatility and effectiveness of

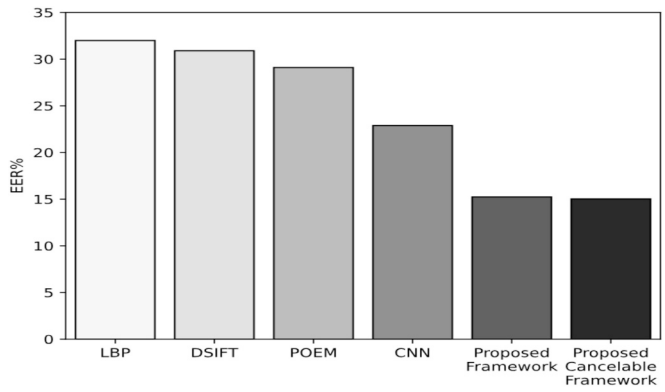

Fig. 7. Illustrative comparative performance of the proposed framework.

the proposed ear biometric authentication system, ensuring it meets the evolving demands of real-world security and usability. Although the system already demonstrates strong performance, several promising research directions could further enhance its robustness and applicability. Incorporating additional biometric modalities, such as facial or fingerprint recognition, could improve overall accuracy and resilience against spoofing attacks. Exploring advanced cryptographic methods or dynamic key-based transformations may strengthen template security while maintaining revocability. Additionally, evaluating the system on larger and more diverse ear image datasets would help assess its generalizability across various demographic and environmental conditions.

Declarations

Availability of data and material

The datasets used and/or analyzed during the current study available on reasonable request.

CRediT authorship contribution statement

Ibrahim Omara: Conceptualization, Data curation, Methodology, Investigation, Writing – original draft, Writing – review & editing. **Randa F. Soliman:** Conceptualization, Data curation, Methodology, Investigation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Randa F. Soliman reports was provided by Menoufia University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Essam Abdellatef, Randa F. Soliman, Eman M. Omran, Nabil A. Ismail, Salah E. S. Abd Elrahman, Khalid N. Ismail, Mohamed Rihan, Mohamed, Amin, Ayman A. Eisa, Fathi E. Abd El-Samie, 2022. Cancelable Face and Iris Recognition Systems Based on Deep Learning", Optical and Quantum Electronics 54, Article number: 702 (2022).

Anwar, A. S., Ghany, K. K. A., & Elmahdy, H. (2015). Human ear recognition using geometrical features extraction. *Procedia Comput. Sci., 65*, 529–537.

Bargal, S.A., Welles, A., Chan, C.R., et al., 2015. 'Image-based ear biometric smartphone app for patient identification in field settings'. VISAPP (3), Berlin, Germany, 2015, pp. 171–179.

M. Barni, T. Bianchi, D. Catalano, M.D. Raimondo, R.D. Labati, P. Failla et al., A privacy-compliant fingerprint recognition system based on homomorphic encryption and fingercode templates, in: IEEE International Conference on Biometrics: Theory, Applications and Systems, 2010, pp. 1–7.

Benarous, L., Kadri, B., Bouridane, A., 2017. A survey on cyber security evolution and threats: biometric authentication solutions. In Jiang, R., Al-maadeed, S., Bouridane, A., et al. (Eds.): 'Biometric security and privacy' (Springer, Cham, Switzerland, 2017), pp. 371–411.

Benzaoui, A., Hadid, A., & Boukrouche, A. (2014). Ear biometric recognition using local texture descriptors. *J. Electron. Imaging, 23*(5), Article 053008.

Boczek, M., 2017, Ear biometric capture, authentication, and identification method and system'. US Patent 9,613,200, 4 April 2017.

Z. Bohan, W. Xu, Z. Kaili, Z. Xueyuan, Encryption node design in internet of things based on fingerprint features and cc2530, in: IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, 2013, pp. 1454–1457.

Bringer, J., Chabanne, H., Cohen, G., et al. (2008). Theoretical and practical boundaries of binary secure sketches. *IEEE Trans. Inf. Forensics Sec., 3*(4), 673–683.

Burge, M., Burger, W., 2000. 'Ear biometrics in computer vision'. Proc. 15th Int. Conf. Pattern Recognition, Barcelona, Spain, 2000, Art. no. 826830.

Camile Lendering, Bernardo Perrone Ribeiro, Žiga Emeršič, Peter Peer, EdgeEar: efficient and accurate ear recognition for edge devices. arXiv preprint arXiv:2502.07734 (2025).

Cappelli, R., Maio, D., Lumini, A., et al. (2007). Fingerprint image reconstruction from standard templates. *IEEE Trans. Pattern Anal. Mach. Intell., 29*(9), 1489–1503.

Castiglione, A., Choo, K.-K., Nappi, M., & Narducci, F. (2017). Biometrics in the cloud: challenges and research opportunities. *IEEE Cloud Comput., 4*(4), 12–17.

Chang, E.-C., Shen, R., Teo, F.W.: 'Finding the original point set hidden among chaff'. Proc. ACM Symp. Information, Computer and Communications Security, Taipei, Taiwan, 2006, pp. 182–188.

Charles Clancy, T., Kiyavash, N., Lin, D.J., 2003. Secure smartcard-based fingerprint authentication. In: Proc. ACM SIGMM Workshop Biometrics Methods and Applications, Berkeley, California, USA; 2003. pp. 45–52.

J.V. Davis, B. Kulis, P. Jain, S. Sra, I.S. Dhillon, "Information-theoretic metric learning, in: Proceedings of the 24th international conference on Machine Learning. ACM, 2007, pp. 209–216.

Dhillon, P. K., & Kalra, S. (2017). A lightweight biometrics based remote user authentication scheme for IoT services. *J. Inf. Secur. Appl., 34*, 255–270.

Dodge, S., Mounsef, J., & Karam, L. (2018). Unconstrained ear recognition using deep neural networks. *IET Biom., 7*(3), 207–214.

Dong, J., Tan, T., 2008. Security enhancement of biometrics, cryptography and data hiding by their combinations. In: 5th International Conference on Visual Information Engineering (VIE 2008), Xian China; 2008. pp. 239–244.

Emersic, Z., Stepec, D., Struc, V., Peer, P., 2017. Training convolutional neural networks with limited training data for ear recognition in the wild; 2017. arXiv preprint arXiv: 1711.09952.

Emersic, Z., Struc, V., & Peer, P. (2017). Ear recognition: more than a survey. *Neurocomputing, 255*, 26–39.

Essam Abdellatef; Nabil A. Ismail; Salah Elden S. E. Abd Elrahman; Khalid N. Ismail; Mohamed Rihan; Fathi E. Abd El-Samie; Eman Omran; Randa Soliman; Ayman Eisa, "Fusion of deep-learned and hand-crafted features for cancelable recognition systems", Soft Computing (2020), Springer, https://doi.org/10.1007/s00500-020-04856-1.

Ferreira, J. L., Wu, Y., & Aarts, R. M. (2018). Enhancement of the comb-filtering selectivity using iterative moving average for periodic waveform and harmonic elimination. *J. Healthcare Eng.*

Freire-Santos, M., Fierrez-Aguilar, J., Ortega-Garcia, J., 2006. Cryptographic key generation using handwritten signature. In: Flynn, P.J., Pankanti, S. (Eds.): 'Biometric technology for human identification III', vol. 6202 (International Society for Optics and Photonics, SPIE, USA, 2006), p. 62020N.

Ghoualmi, L., Draa, A., & Chikhi, S. (2016). An ear biometric system based on artificial bees and the scale invariant feature transform. *Expert Syst. Appl., 57*, 49–61.

Habib, K., Torjusen, A., Leister, W., 2014. A novel authentication framework based on biometric and radio fingerprinting for the IoT in eHealth. In: Proc. 3rd Int. Conf. Smart Syst., Devices, Technol., 2014, pp. 32–37.

Hahn, V. K., & Marcel, S. (2022). Biometric template protection for neural-network-based face recognition systems: a survey of methods and evaluation techniques. *IEEE Trans. Inf. Forensics Secur., 18*, 639–666.

Hansley, E. E., Segundo, M. P., & Sarkar, S. (2018). Employing fusion of learned and handcrafted features for unconstrained ear recognition. *IET Biom., 7*(3), 215–223.

Hao, F., Anderson, R., & Daugman, J. (2006). Combining crypto with biometrics effectively. *IEEE Trans. Comput., 55*(9), 1081–1088.

Huang, P., et al. (Oct. 2019). Practical privacy-preserving ECG-based authentication for IoT-based healthcare. *IEEE Internet Things J., 6*(5), 9200–9210.

Hurley, D. J., Nixon, M. S., & Carter, J. N. (2005). Force field feature extraction for ear biometrics. *Comput. Vis. Image Underst., 98*(3), 491–512.

Iannarelli, A.V., 1989. 'Ear identification' (Paramont Publishing Company, Fremont, California, 1989).

Jain, A. K., Nandakumar, K., & Nagar, A. (2008). Biometric template security. *EURASIP J. Adv. Signal Process., 2008*, 113.

Jamil, N., AlMisreb, A., & Halin, A. A. (2014). Illumination-invariant ear authentication. *Proc. Comput. Sci., 42*, 271–278.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia, MM '14, Association for Computing Machinery; 2014. p. 675–678.

Jiang, X., Xu, K., Liu, X., Dai, C., Clifton, D. A., Clancy, E. A., Akay, M., & Chen, W. (2020). Cancelable hd-semg-based biometrics for cross-application discrepant personal identification. *IEEE J. Biomed. Health Inform., 25*(4), 1070–1079.

Juels, A., Wattenberg, M., 1999. A fuzzy commitment scheme. In: Proc. 6th ACM Conf. Computer and Communications Security, Singapore, 1999, pp. 28–36.
Juels, A., & Sudan, M. (2006). A fuzzy vault scheme. *Des. Codes Cryptogr., 38*(2), 237–257.
Kantarci, B., Erol-Kantarci, M., Schuckers, S., 2015. Towards secure cloudcentric internet of biometric things. In: Proc. 4th Int. Conf. Cloud Netw, 2015, pp. 81–83.
Kaur, H., & Khanna, P. (2020). Privacy preserving remote multi-server biometric authentication using cancelable biometrics and secret sharing. *Futur. Gener. Comput. Syst., 102*, 30–41.
Kim, T., Oh, Y., & Kim, H. (2020). Efficient privacy-preserving fingerprint-based authentication system using fully homomorphic encryption. *Secur. Commun. Netw., 2020*, Article 4195852.
Kuo, S.M., Lee, B.H., Tain, W., 2006. Real-Time signal processing, implantations and applications. John Wiley & Sons Ltd, 2nd.
LeCun, Y., Boser, B., Denker, J. S., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comput., 1*(4), 541–551.
Lee, Y.J., Bae, K., Lee, S.J., et al., 2007. Biometric key binding: fuzzy vault based on iris images. Proc. Int. Conf. Biometrics, Seoul, Korea; 2007, pp. 800–808.
Leng, L., Zhang, S., Bi, X., et al., 2012. 'Two-dimensional cancelable biometric scheme. In: Proc. Int. Conf. Wavelet Analysis and Pattern Recognition, Xian, China, 2012, pp. 164–169.
Li, Q., Guo, M., Chang, E.-C.: 'Fuzzy extractors for asymmetric biometric representations. In: Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops, Anchorage, Alaska, USA; 2008, pp. 1–6.
Maček, N., Franc, I., Bogdanoski, M., Mirković, A., 2016. Multimodal biometric authentication in IoT: Single camera case study. In: Proc. 8th Int. Conf. Bus. Info. Secur., Belgrade, Serbia, 2016, pp. 33–37.
Mandal, S., et al. (Apr. 2020). Certificateless-signcryption-based three-factor user access control scheme for IoT environment. *IEEE Internet Things J., 7*(4), 3184–3197.
Mariam Mahmoud Mohammed, Abdul Monem Rahma, Ayad Al-Adhami, A new ear recognition system based on moments analysis, AIP Conf. Proc. 3169 (2025) 030023.
Meraoumia, A., Chitroub, S., Bouridane, A.: 'An automated ear identification system using Gabor filter responses. In: Proc. IEEE 13th Int. New Circuits and Systems Conf., Grenoble, France; 2015. pp. 1–4.
Meraoumia, A., Kadri, F., Bendjenna, H., et al., 2017. Improving biometric identification performance using PCANet deep learning and multispectral palmprint. In: Jiang, R., Al-maadeed, S., Bouridane, A., et al. (Eds.): Biometric security and privacy (Springer, Cham, Switzerland, 2017, pp. 51–69.
Moreno, B., Sanchez, A., Vélez, J.F., 1999. On the use of outer ear images for personal identification in security applications. In: Proc. IEEE 33rd Annual Int. Carnahan Conf. Security Technology, Madrid, Spain, 1999, pp. 469–476.
Mu, Z., Yuan, L., Xu, Z., Xi, D., Qi, S., 2004. Shape and structural feature-based ear recognition. In: Advances in biometric person authentication. Springer; 2004. p. 663–670.
Nagar, A., Nandakumar, K., & Jain, A. K. (2010). A hybrid biometric cryptosystem for securing fingerprint minutiae templates. *Pattern Recogn. Lett., 31*(8), 733–741.
Nandakumar, K., Jain, A. K., & Pankanti, S. (2007). Fingerprint-based fuzzy vault: implementation and performance. *IEEE Trans. Inf. Forensics Sec., 2*(4), 744–757.
Omara, I., Li, F., Zhang, H., et al. (2016). A novel geometric feature extraction method for ear recognition. *Expert Syst. Appl., 65*, 127–135.
Omran, E.M., Soliman, Randa F., salah, M.M., napoleon, S.A., el-rabaie, E.-S.M., abdeelnaby, M.M., el-samie, F. ABD, 2020. Noisy Iris Recognition Based on Deep Neural Network. Menoufia J. Electron. Eng. Res. (MJEER), 29(2), 64–69.
Oyebiyi, Oyediran George, Adebayo Abayomi-Alli, Oluwasefunmi 'Tale Arogundade, Atika Qazi, Agbotiname Lucky Imoize, Joseph Bamidele Awotunde, A systematic literature review on human ear biometrics: approaches, algorithms, and trend in the last decade Informat. 14(3) (2023) 192.
Pandey, R.K., Zhou, Y., Kota, B.U., et al., 2016. Deep secure encoding for face template protection. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 2016, pp. 77–83.
Pandey, R.K., Zhou, Y., Kota, B.U., et al., 2017. Learning representations for cryptographic hash-based face template protection. In: Bhanu, B., Kumar, A. (Eds.): 'Deep learning for biometrics' (Springer, Cham, Switzerland, 2017), pp. 259–285.
Omran, E. M., Soliman, Randa, F., Eisa, A. A, Ismail, N. A., & Abd el-samie, F. E. (2019). Cancelable iris recognition system with pre-trained convolutional neural networks. *Menoufia J. Electron. Eng. Res. (MJEER), 28(ICEEM2019-Special Issue)*, 95–101.
Patel, V. M., Ratha, N. K., & Chellappa, R. (2015). Cancelable biometrics: a review. *IEEE Signal Process Mag., 32*(5), 54–65.
Paul, P.P., Gavrilova, M., 2012. Multimodal biometric approach for cancelable face template generation. In: Braun, J.J. (Ed.): 'Multisensor, multisource information fusion: architectures, algorithms, and applications' vol. 8407, (International Society for Optics and Photonics, SPIE, USA, 2012), p. 84070H.
P. Punithavathi, S. Geetha, Partial dct-based cancelable biometric authentication with security and privacy preservation for iot applications, Multimedia Tools Appl. 78 (18) (2019) 487–25.
Randa F. Soliman. Mohamed Amin. Fathi E. Abd El-Samie, A modified cancelable biometrics scheme using random projection, Springer, Annals of Data Science, 2018, pp. 1–14.
Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J., 40*, 614–634.
Ratha, N. K., Chikkerur, S., Connell, J. H., & Bolle, R. M. (2007). Generating cancelable fingerprint templates. *IEEE Trans. Pattern Anal. Mach. Intell., 29*(4), 561–572.
Rathgeb, C., & Uhl, A. (2011,). 'A survey on biometric cryptosystems and cancelable biometrics. *EURASIP J. Inf. Secur.*, 3.
Rathgeb, C., & Uhl, A. (2012). Statistical attack against fuzzy commitment scheme. *IET Biom., 1*(2), 94–104.
Sana, A., Gupta, P., Purkait, R., 2007. Ear biometrics: a new approach. In: Pal, P. (Ed.): 'Advances in pattern recognition' (World Scientific, India, 2007. pp. 46–50.
Shahim, L.-P., Snyman, D., du Toit, T., Kruger, H., 2016. Cost-effective biometric authentication using leap motion and IoT devices. In: Proc. 10th Int. Conf. Emerg. Secur. Inf., Syst. Technol., Nice, France, 2016, pp. 24–28.
Soliman, R. F., Amin, M., & Abd El-Samie, F. E. (2018). A double random phase encoding approach for cancelable iris recognition. *Springer Opt. Quant. Electron., 50*(326), 1–12.
Randa F. Soliman, Ghada M. El banby, Abeer D. Algarni, Mohamed Elsheikh, Naglaa F. Soliman, Mohamed Amin, and Fathi E. Abd El-Samie, Utilization of double random phase encoding for cancelable face and iris recognition, App. Opt. (OSA) Optical Society of America Vol. 57, No. 35, pp. 10305-10316, 2018.
Randa F. Soliman. Noha Ramadan. Mohamed Amin. HossamEldin H. Ahmed. Said El-Khamy. Fathi E. Abd El-Samie "Efficient cancelable iris recognition scheme based on modified logistic map". Proceedings of the National Academy of Sciences Published, 2018.
Randa F. Soliman. Mohamed Amin. Fathi E. Abd El-Samie, Cancelable iris recognition system based on comb-filter", Multimedia Tools and Applications, Springer 79 (2020) 2521–254, https://doi.org/10.1007/s11042-019-08163-2.
Tams, B. (2016). Unlinkable minutiae-based fuzzy vault for multiple fingerprints. *IET Biom., 5*(3), 170–180.
Tong, V.V.T., Sibert, H., Lecoeur, J., et al., 2007. 'Biometric fuzzy extractors made practical: a proposal based on fingercodes'. In: Proc. Int. Conf. Biometrics, Seoul, Korea, 2007, pp. 604–613.
Uludag, U., Pankanti, S., Jain, A.K., 2005. 'Fuzzy vault for fingerprints. In: Proc. Int. Conf. Audio-and Video-Based Biometric Person Authentication, New York, USA; 2005. pp. 310–319.
Barnabas Victor, Kevin Bowyer, Sudeep Sarkar, An evaluation of face and ear biometrics, in: ICPR'02, vol.1, 2002, pp.429–432.
Wang, Y., Mu, Z.-C., Zeng, H., 2008. 'Block-based and multi-resolution methods for ear recognition using wavelet transform and uniform local binary patterns. In: Proc. 19th Int. Conf. Pattern Recognition, Tampa, Florida, USA; 2008. p. 1–4.
Yang, W., Wang, S., Zheng, G., Yang, J., & Valli, C. (Mar. 2019). A privacy-preserving lightweight biometric system for internet of things security. *IEEE Commun. Mag., 57*(3), 84–89.
Yang, W., Wang, S., Hu, J., Zheng, G., Yang, J., & Valli, C. (Jul. 2019). Securing deep learning based edge finger vein biometrics with binary decision diagram. *IEEE Trans. Ind. Informat., 15*(7), 4244–4253.
Yang, W., Wang, S., Jiankun, Hu., Tao, X., & Li, Y. (2024). Feature extraction and learning approaches for cancellable biometrics: a survey. *CAAI Trans. Intell. Technol., 9*(1), 4–25.