Convolution Neural Network Based Method for Biometric Recognition

Adel Jalal Yousif

University of Diyala, Diyala, Iraq

Abstract: *In recent years, there has been increasing interest in the potential of precisely identifying individuals through ear images within the biometric community, owing to the distinctive characteristics of the human ear. This paper introduces deep neural network architecture for ear recognition. The suggested method incorporates a preprocessing stage that enhances significant features in ear images through contrast-limited adaptive histogram equalization. Subsequently, a classifier with deep convolutional neural network is employed to recognize the preprocessed ear images. Experimental results demonstrate a remarkable testing accuracy of 97.92% for the proposed recognition system.*

Keywords: *Ear recognition, machine learning, features extraction, convolutional neural networks, neural network.*

1-Introduction

In modern society, biometric systems have gained immense significance due to the escalating demand for security measures across various fields. These systems find extensive applications in identity management, law enforcement, surveillance, forensics, and more [1]. A biometric system operates as a pattern recognition system that verifies individual authenticity based on unique biometric characteristics. Conventional identification methods, such as passwords and identification cards, have proven to be increasingly unreliable in the face of advancing cyber-attacks, and they can easily be forgotten or stolen. In contrast, biometric systems have emerged as the most trustworthy and accurate means of identity authentication [2]. Biometric characteristics can be broadly categorized into two types: physiological biometrics and behavioral biometrics. Physiological biometrics pertain to physical attributes of the human body, such as fingerprints, ear prints, face, iris, retina, and hand geometry. On the other hand, behavioral biometrics are related to patterns of human behavior, including gait, voice, signature, and keystrokes. Biometric techniques utilize image acquisition devices, such as scanners or cameras for physiological biometrics, and platens for behavioral biometrics [3].

To differentiate between users, the most reliable and distinct features are selected and converted into a biometric reference. These extracted features are then stored in a database or repository and used for comparative processes when pattern recognition is needed. Among the various biometric characteristics, ear recognition holds significant importance in the field of human recognition. The human ear possesses unique biological characteristics that remain consistent over time, making it an ideal biometric identifier compared to other attributes [4]. Ear recognition offers several advantages, such as stability over time, insensitivity to emotional feelings, and ease of capture from a distance. Consequently, the image of the

ear provides a rich source of biometric information, making it suitable for the development of an effective recognition system [1]. The process of features extraction is a crucial phase in biometric recognition. In the early stages of ear recognition techniques, handcrafted feature engineering approaches were used to extract features from images. These selected features were then utilized to train conventional classifiers, enabling them to learn specific patterns from the extracted features. However, the effectiveness of these recognition approaches heavily depends on the classifier performance and the strength of the feature extractor. There are two main challenges associated with these traditional methods. Firstly, the effectiveness of the system diminishes as the visual variation level in the processed images increases. Secondly, the process of manually extracting correlated features from images proves to be quite time-consuming task that requires significant expertise in the specific domain [1, 5]. To overcome these limitations, recent advancements in Convolutional Neural Networks (CNNs) have proven successful in many computer vision applications such as image classification [9-11], biometric recognition [6-8], and object detection [12-15]. The focus of this work is to develop an ear recognition system based on a deep convolutional neural network classifier. This approach aims to harness the power of CNNs to improve the accuracy and efficiency of ear recognition.

2- Related Works

In this section, a brief summary of various ear recognition studies conducted in recent years is presented. Taertulakarn et al. [16] proposed an ear recognition system based on geometric features extracted from 2D ear images for the 3D surface of the human ear. Principal components analysis (PCA) was employed for feature extraction from the ear structure. The system required a combination of a 3D scanner and a camera for image capturing, and it achieved an accuracy rate of 92% in ear image recognition. M. Chowdhury et al. [17] suggested an ear-based biometric recognition technique that utilized local image features and artificial neural networks. They used an AdaBoost-based detector to detect the ear region in profile images. Fuzzy filters were applied for pre-processing ear images to remove holes and spikes from the region of interest.

Jiddah et al. [18] developed a recognition technique based on the fusion of texture and geometric features of the ear using the AMI dataset. They used Laplacian filters for extracting geometric features and Ojala operators for texture features. After feature fusion, a k-nearest neighbor classifier was employed for ear pattern classification, achieving an accuracy rate of 90% in six iterations. N. Petaitiemthong et al. [19] adopted CNN based scheme for recognizing side view and front view images of the human ear. The system achieved a correct classification rate of 80% for side view images and 84% for front view images. S. Nikose and H. Meena [20] developed an ear biometric identification scheme based on a convolutional neural network. They preprocessed dataset images using Gaussian filters and Canny operators to enhance the recognition rate. The proposed system achieved an impressive accuracy rate of 93.3% in correct ear recognition.

These related works highlight the ongoing efforts to enhance ear recognition systems, with approaches ranging from geometric features and artificial neural networks to convolutional neural networks, all contributing to improved accuracy and reliability in ear biometric identification. Hamdany et al. [21] proposed an earprint recognition model based on deep learning techniques. The method utilized Adam optimization to determine the best parameters for the convolutional neural network. The suggested model achieved an impressive accuracy rate of 94% using the IIT Delhi ear dataset. This study further contributes to the advancement of ear recognition systems, demonstrating the effectiveness of deep learning approaches in achieving high accuracy in earprint recognition.

3- Convolutional Neural Network

Convolutional Neural Network (CNN) is a widely used architecture of multilayer perceptron deep neural networks, particularly in the domains of image processing and computer vision. Its popularity stems from several essential characteristics that make it highly effective in these fields [22, 23]. These characteristics include low complexity, tolerance to visual variations, local perception, pooling sampling, weight sharing, and the integration of automated feature extraction with the classification task. One of the key advantages of CNN is its ability to be trained using small image patches, such as (32 x 32) pixels, which reduces the computational time and complexity. The typical structure of CNN consists of three main types of layers: convolutional layers, pooling layers, and fully connected layers [24-26], as illustrated in figure 1. Each of these layers serves a specific purpose in the network.

Convolution Layer: The convolutional layer is responsible for extracting feature from input images. It applies a set of learnable filters (also known as kernels) to the input image, convolving them across the image to produce feature maps. These feature maps highlight specific patterns and features present in the input.

Pooling Layer: The pooling layer decreases the spatial dimensions of the feature maps obtained from the convolutional layers. It achieves this by down sampling the feature maps, reducing the number of parameters and the computational burden while preserving significant information.

Fully Connected Layer: This layer is responsible for the classification task. It takes the features extracted from the previous layers and performs a classification based on the learned representations, determining the class or category to which the input image belongs.

In summary, the CNN's structure enables it to automatically learn hierarchical representations of features from input images, making it a powerful tool for image-related tasks. The next sections will provide further details on each layer [19].
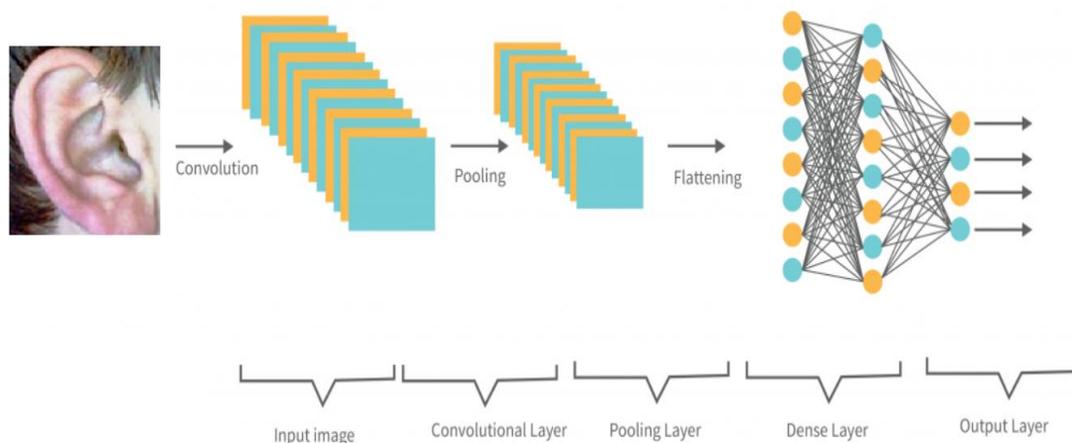

Fig. 1: General CNN architecture

4- Proposed Method

The proposed ear recognition system is divided into two main stages: the pre-processing stage and the automated features extraction stage. The system is designed to enhance the quality of the input image using the Contrast-Limited Adaptive Histogram Equalization (CLAHE) technique during the pre-processing stage. Subsequently, a convolutional network is utilized for automated features extraction. In the pre-processing stage, the input ear image undergoes the CLAHE technique, which enhances the

image's contrast and improves its overall quality. This step is crucial for ensuring that the CNN receives well-prepared input data.

In the automated features extraction stage, the pre-processed ear image is fed into the CNN. The CNN uses its convolutional layers to extract relevant features from the image. These layers employ sliding filters to capture patterns and distinct characteristics present in the ear image. The pooling layers then reduce the dimensionality of the extracted features, further refining the information for classification. Finally, the fully connected layers combine the learned features to generate a final output for classification, representing the identification or recognition of the input ear image. The proposed system design leverages the CLAHE pre-processing technique and the power of CNNs for automated feature extraction, providing an effective and accurate ear recognition system. Figure 2 illustrates the general structure of the proposed method, outlining the flow from pre-processing to features extraction and classification.
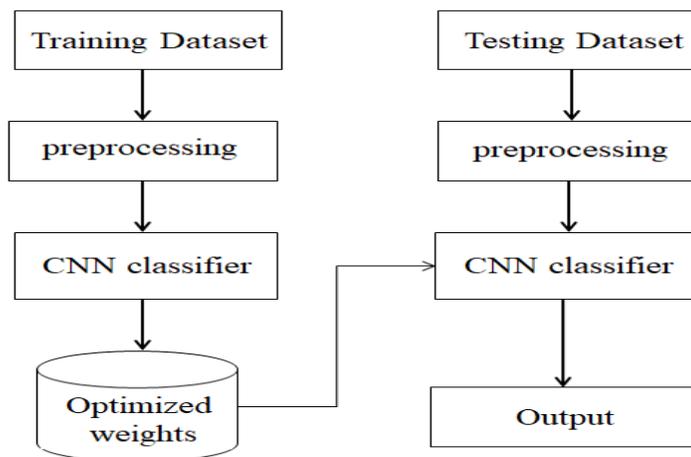

Fig. 2: General structure of the proposed method

A. Contrast-Limited Adaptive Histogram Equalization

Contrast-Limited Adaptive Histogram Equalization (CLAHE) [27, 28]is a powerful image processing technique used for enhancing the contrast of an image. It addresses the limitations of traditional adaptive histogram equalization, which can lead to noise amplification in homogeneous areas of the processed image. CLAHE operates on small local regions of an image rather than the entire image, which helps to limit noise amplification. The result of CLAHE enhancement is illustrated in figure 3, where the left column shows samples of the original ear image dataset, and the corresponding enhanced images using CLAHE are displayed in the right column. The steps of adopted CLAHE technique can be summarized as follows [29, 30]:

1. Image Partitioning: The input image is divided into non-overlapping blocks with equal size.

2. Histogram Calculation: the histogram of pixel intensities is calculated for each block. which represents the frequency distribution of pixel values within the block.

3. Determining the Clip Limit: A clip limit is determined based on the desired level of contrast stretching. This clip limit sets a threshold for limiting the extent of contrast enhancement in each block.

4. Histogram Equalization: The pixel intensities in each block's histogram are redistributed or equalized to stretch the contrast locally. This involves adjusting the pixel values to achieve a more balanced distribution of intensities within the block.

5. Clipping the Histograms: To ensure that the contrast enhancement does not lead to excessive noise amplification, the heights of the histograms are clipped or limited based on the determined clip limit. This prevents any extreme intensities from dominating the enhancement process.

6. Cumulative Distribution Functions (CDF): The cumulative distribution functions (CDF) are computed for the histograms after the clipping step. The CDF represents the cumulative probabilities of the pixel intensities in the histograms.

By following these steps, the CLAHE technique ensures that the contrast enhancement is performed locally on small blocks of the image, limiting noise amplification and providing improved image quality. The resulting contrast-enhanced images are then used for feature extraction in the subsequent stages of the proposed ear recognition system.
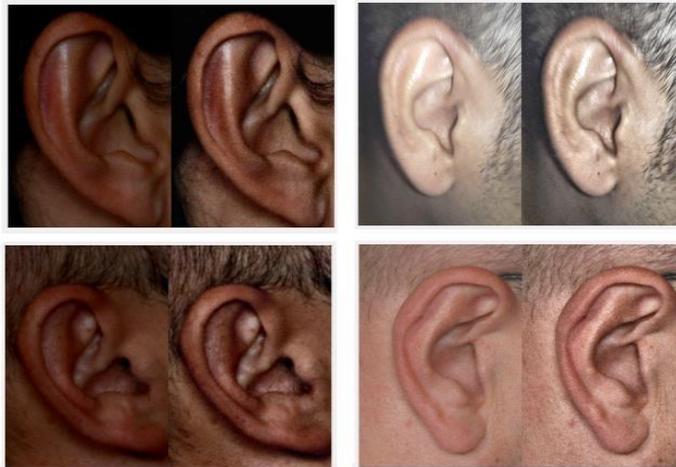

Fig. 3: Sample of images after CLAHE enhancement

B. Convolutional Neural Network Architecture

The proposed Convolutional Neural Network (CNN) model comprises three convolutional layers, each followed by a max-pooling layer with a filter size of 2x2. Additionally, there are two dense fully connected layers in the network. The input to the CNN consists of ear images with a size of (227x227x3). Specifically, the first convolutional layer applies 32 filters of size 3x3, the second convolutional layer applies 64 filters of size 3x3, and the third convolutional layer applies 128 filters of size 3x3. To make the input suitable for the fully connected layers, the image is flattened to a vector with dimensions of $(n^2 \times 1)$, where n represents the dimensions of the image. This vector is then fed as input to the first fully connected layer, which comprises 130 neurons. The final layer is a fully connected layer with neurons equal to the number of output classes. To introduce non-linearity and enhance the fitting capability, Rectified Linear Unit (ReLU) activation function is applied after each convolutional layer and the first fully connected layer. For obtaining the probabilities of each class, the last fully connected layer uses the Softmax activation function. To address the overfitting issue, a dropout function with a 20% dropout ratio is implemented at the first convolutional and fully connected layers. The suggested CNN architecture is summarized in Table 1 and is implemented using the Keras library within the Python environment. The overall proposed ear recognition system is illustrated in figure 4. By employing this CNN architecture, the system aims to achieve accurate and efficient ear recognition, benefiting from the power of deep learning and the advantages of using ReLU activation and dropout functions.

Table 1. Proposed CNN architecture

| Layer	Output shape	Parameter #
conv2d_3 (Conv2D)	(None, 225, 225, 32)	896
max_pooling2d_3	(None, 112, 112, 32)	0
dropout_2 (Dropout)	(None, 112, 112, 32)	0
conv2d_4 (Conv2D)	(None, 110, 110, 64)	18496
max_pooling2d_4	(None, 55, 55, 64)	0
conv2d_5 (Conv2D)	(None, 53, 53, 128)	73856
max_pooling2d_5	(None, 26, 26, 128)	0
flatten_1 (Flatten)	(None, 86528)	0
dense_2 (Dense)	(None, 130)	11248770
dropout_3 (Dropout)	(None, 130)	0
dense_3 (Dense)	(None, 10)	1310
Total parameters: 11,342,149 , Trainable parameters: 11,342,149		
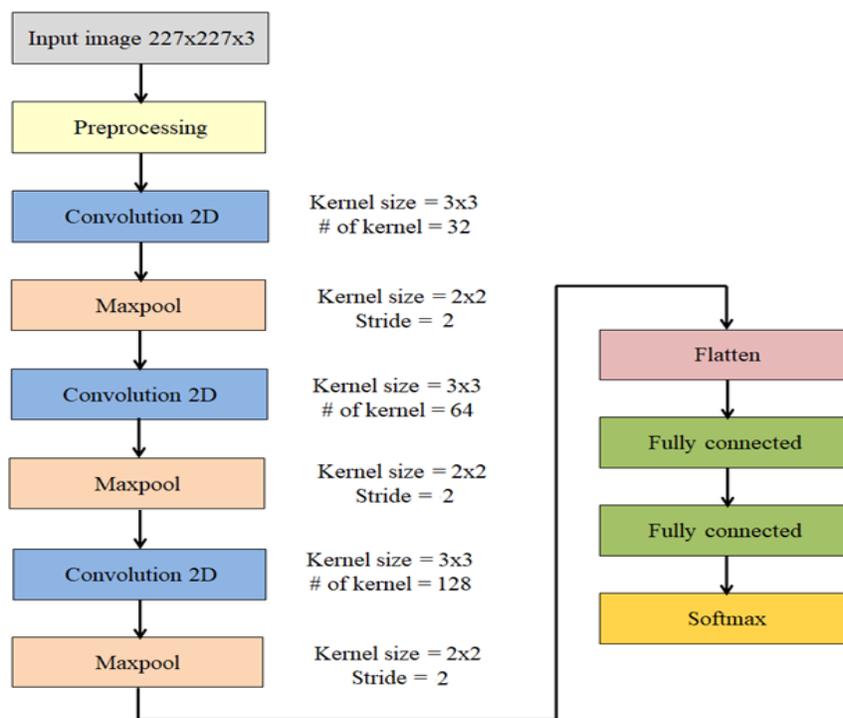
Fig. 4: Architecture of the proposed ear recognition

Fig. 5: Sample of Dataset-1 for four different classes

5- Experimental Results

For evaluating the performance of the suggested ear recognition system, experiments using the publicly available ear dataset known as Dataset1 is conducted. This dataset is widely used in the biometric community and is characterized by containing images with high variability, presenting various levels of visual variation within the same class. This variation poses a significant challenge for biometric recognition algorithms. Figure 5 illustrates the variability of images in Dataset1, showcasing several aspects such as the angle of capturing, intensity of illumination, capturing distance, ear position, and presented ear size. These variations make the dataset representative of real-world scenarios and enable a thorough assessment of the ear recognition system's robustness and accuracy. For the experiments, CNN architecture with Contrast-Limited Adaptive Histogram Equalization (CLAHE) is used for pre-processing the input ear images. The CNN was trained using Dataset1 and the Adam optimization technique for determining the best parameters of the network. The performance of the system was evaluated based on recognition accuracy, which represents the percentage of correctly identified ear images compared to the total number of images in the dataset. The proposed method is compared with the modern ear recognition approaches to measure the effectiveness of the suggested system.

The dataset used in this work consists of 1600 images of right and left ears, each with a size of (227x227x3), spread across 10 different classes. Each class contains 160 images. To train and evaluate the proposed ear recognition system, the dataset is divided into two subsets: a training set (80%) and a testing set (20%). The proposed CNN model, as described in Section III, is employed for the ear recognition task. During the training phase, the network learns from the training dataset and builds its intelligence based on the provided ear images and their corresponding classes. Once the training is completed, the model's accuracy is evaluated by applying the learned intelligence to the testing dataset. The results of the CNN's training and testing, including accuracy and loss values across different epochs, are presented in Table II. Additionally, the performance of the suggested system, in terms of accuracy and loss, is illustrated in figure 6.

The proposed ear recognition method achieved an impressive overall testing accuracy of 97.92%. This high accuracy demonstrates the effectiveness of the CNN model in accurately identifying and classifying ear images, even in the presence of significant variability and challenges posed by the Dataset1. The low loss values during training also indicate that the CNN model effectively learned the representations and features necessary for ear recognition. The results confirm that the proposed system, with the integration of the CLAHE pre-processing technique and the CNN architecture, can successfully handle the complexities of ear recognition, making it a reliable and robust solution for biometric identification tasks involving ear images.

Table 2. Accuracy and loss of proposed method

Epoch	Loss	Accuracy (%)
5	1.5914	0.4750
10	0.8419	0.6833
15	0.4531	0.8625
20	0.2699	0.9000
25	0.2380	0.9250
30	0.2330	0.9208
35	0.2018	0.9458
40	0.2279	0.9625
45	0.1254	0.9708
50	0.1803	0.9625
55	0.1316	**0.9792**
60	0.1525	0.9667
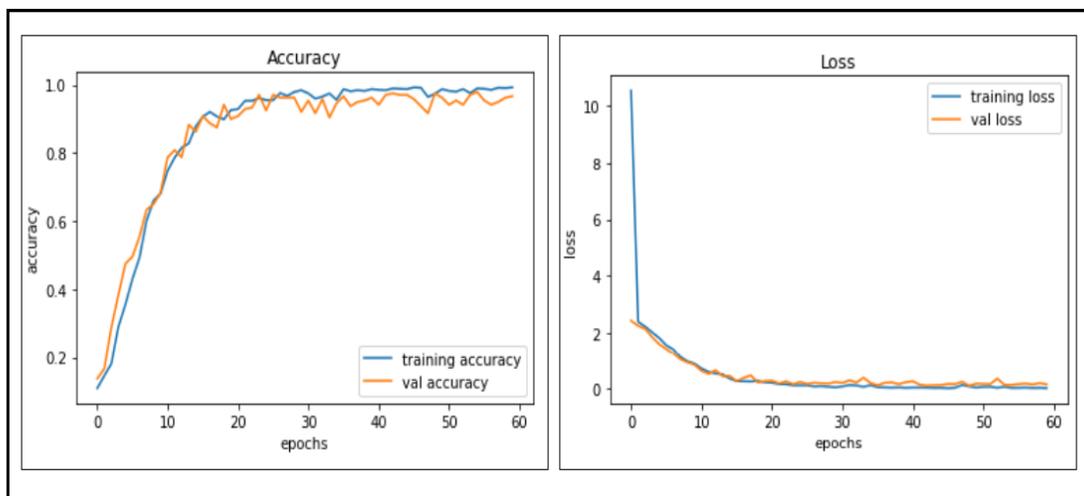
Fig. 6: Accuracy and loss of training/validation sets

6- Conclusion

In this paper, a deep learning-based human ear recognition method have been presented that incorporates a preprocessing stage using the contrast-limited adaptive histogram equalization (CLAHE) technique. The aim of this preprocessing stage is to enhance the features of the ear images before they are fed into the CNN classifier. The proposed CNN approach has proven to be highly effective, achieving remarkable accuracy in ear image classification. Specifically, the system achieved an impressive overall testing accuracy of 97.92%. This high accuracy demonstrates the success of the proposed method in accurately identifying and classifying ear images, even in the presence of variability and challenges encountered in the Dataset1. The combination of the CLAHE preprocessing technique and the power of CNNs has resulted in a robust and reliable ear recognition system. This system can serve as an important tool in various biometric applications, offering a highly accurate and efficient means of human ear recognition. In conclusion, the proposed method represents a significant advancement in the field of ear recognition and demonstrates the potential of deep learning techniques in biometric identification tasks. The achieved accuracy showcases the system's effectiveness in handling the complexities and variations present in ear images, making it a promising solution for real-world biometric recognition scenarios.

REFERENCES

1. H. Alshazly, C. Linse, E. Barth and T. Martinetz, "Deep Convolutional Neural Networks for Unconstrained Ear Recognition," in IEEE Access, vol. 8, pp. 170295-170310, 2020.

2. I. Boucherit, M. Zmirli, H. Hentabli, and B. Rosdi, "Finger vein identification using deeply-fused Convolutional Neural Network," Journal of King Saud University - Computer and Information Sciences, 2020.

3. B. Pandya, G. Cosma, A. A. Alani, A. Taherkhani, V. Bharadi and T.M. McGinnity, "Fingerprint classification using a deep convolutional neural network," 2018 4th International Conference on Information Management (ICIM), pp. 86-91, 2018.

4. C. Chen, J. Ding and C. Huang, "Advanced Ear Detection Algorithm Using Faster R-CNN, Refocus Filters, and the Gradient Map," 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), pp. 1-5, 2018.

5. H. Alshazly et al. "Handcrafted versus cnn features for ear recognition,". Symmetry 11(12), 2019.

6. S. Jia et al., "3-D Gabor Convolutional Neural Network for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022.

7. H. Yu, H. Zhang, Y. Liu, K. Zheng, Z. Xu and C. Xiao, "DualChannel Convolution Network With Image-Based Global Learning Framework for Hyperspectral Image Classification," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022.

8. M. S. Patil and P. B. Mane, "Fused Image Classification using Pre-trained Deep Convolutional Neural Networks," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 1215-1221.

9. Fadhil Kadhim Zaidan et al. "Ear Recognition System Based on CLAHE and Convolution Neural Network", International Journal of Computer Science & Wireless Security (IJCSWS), Vol. 09(03), May-June 2022, pp.04-09.

10. J. Shen et al., "Finger Vein Recognition Algorithm Based on Lightweight Deep Convolutional Neural Network," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-13, 2022.

11. M. K. Roy and U. Keshwala, "Res2Net based Text Independent Speaker recognition system," 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 612-616, 2022.

12. M. Reza Obeidavi and K. Maghooli, "Tumor Detection in Brain MRI using Residual Convolutional Neural Networks," 2022 International Conference on Machine Vision and Image Processing (MVIP), pp.1-5, 2022

13. M. Ponika, K. Jahnavi, P. S. V. S. Sridhar and K. Veena, "Developing a YOLO based Object Detection Application using OpenCV," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 662-668.

14. M. Prakash, M. Janarthanan and D. Devi, "Multiple Objects Identification for Autonomous Car using YOLO and CNN," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023, pp. 597-601.

15. B. Bai, W. Fu, T. Lu and S. Li, "Edge-Guided Recurrent Convolutional Neural Network for Multitemporal Remote Sensing Image Building Change Detection," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-13, 2022.

16. S. Taertulakarn, C. Pintavirooj, P. Tosranon and K. Hamamoto, "The preliminary investigation of ear recognition using hybrid technique," 2016 9th Biomedical Engineering International Conference (BMEiCON), pp. 1-4, 2016.

17. M. Chowdhury, R. Islam and J. Gao, "Robust ear biometric recognition using neural network," 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1855-1859, 2017.

18. S. M. Jiddah and K. Yurtkan, "Fusion of Geometric and Texture Features For Ear Recognition," 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1-5, 2018.

19. N. Petaitiemthong, P. Chuenpet, S. Auephanwiriyakul and N. TheeraUmpon, "Person Identification from Ear Images Using Convolutional Neural Networks," 2019 9th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 148-151, 2019.

20. S. Nikose and H. K. Meena, "Ear-biometrics for human identification," 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), pp. 8-13, 2020.

21. A. Hamdany, A. Ebrahem, and A. Alkababji, "Earprint recognition using deep learning technique," TELKOMNIKA Telecommunication, Computing, Electronics and Control, Vol. 19, No. 2, pp. 432-437, April 2021.

22. A. Ajit, K. Acharya and A. Samanta, "A Review of Convolutional Neural Networks," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), pp. 1-5, 2020.

23. Phung, V.H.; Rhee, E.J. "A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets". Appl. Sci. 2019.

24. K. Singh, A. Seth, H. S. Sandhu and K. Samdani, "A Comprehensive Review of Convolutional Neural Network based Image Enhancement Techniques," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2019, pp. 1-6, doi: 10.1109/ICSCAN.2019.8878706.

25. B. Biswas and R. K. Yadav, "A Review of Convolutional Neural Network-based Approaches for Disease Detection in Plants," 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2023, pp. 514-518, doi: 10.1109/IDCIoT56793.2023.10053428.

26. Palak and A. L. Sangal, "Traffic Signs Classification using Convolutional Neural Networks: A review," 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC), Jalandhar, India, 2021, pp. 450-455, doi: 10.1109/ICSCCC51823.2021.9478172.

27. S. O. Oppong, F. Twum, J. B. Hayfron-Acquah and Y. M. Missah, "Hybrid Approach to Medicinal Plant Leaf Image Enhancement Using Contrast Limited Adaptive Histogram Equalization and Homomorphic Filter," 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 2021, pp. 1-6, doi: 10.1109/ICECET52533.2021.9698460.

28. M. S. Hitam, E. A. Awalludin, W. N. Jawahir Hj Wan Yussof and Z. Bachok, "Mixture contrast limited adaptive histogram equalization for underwater image enhancement," 2013 International Conference on Computer Applications Technology (ICCAT), Sousse, Tunisia, 2013, pp. 1-5, doi: 10.1109/ICCAT.2013.6522017.

29. A. Harsoyo, M. C. Rezi and P. H. Rusmin, "Design of face recognition system using local binary pattern and CLAHE on Smart Meeting Room System," 2013 IEEE 3rd International Conference on System Engineering and Technology, pp. 341-345, 2013.

30. P. Musa, F. A. Rafi and M. Lamsani, "A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition," 2018 Third International Conference on Informatics and Computing (ICIC), pp. 1-6, 2018.