Print ISSN 2710-0952 Electronic ISSN 2790-1254

A Comprehensive Review of CNN-Based Human Recognition Using Ear Shape Images

Mohammed Abdulsattar Abdulghani Abdullah[1]
Batool Abdulsatar Abdulghani Abdullah[2]
[1] *Department of Software, College of Computer Sciences & Mathematics, University of Mosul,Mosul, Iraq.*
email: mohammedse44@gmail.com
[2] *Dept. of Energy Engineering, Technical College of Engineering, Duhok Polytechnic University, duhok-Iraq.*
email: batool.abdulghani@dpu.edu.krd

Abstract

Automatic person recognition using ear shape images is an active field of research within the biometric community. Similar to other biometric traits such as fingerprints, face and iris, ear also has numerous specific and unique features that aid in person identification. In this present worldwide pandemic of COVID-19 situation, most of the facial identification systems has almost failed due to the mask wearing scenario. The human ear is a perfect source of data for passive person identification as it does not involve the cooperativeness of the individual whom recognition is being attempted for and it is more easily captured at a distance. A human ear image acquisition is also easy as the ear is apparent even the users' wearing masks. In an automatic human recognition system, an ear biometric system can be integrated as a supplement to other biometric systems and offer identity cues when other system information is unreliable or even unavailable. In this paper, a comprehensive review of growing research field of feature extraction techniques and the classification technique of deep learning using convolutional neural network (CNN) was conducted to exhibit the rate of accuracy for person recognition systems using ear shape images.

KEYWORDS: *Ear recognition, CNN, Biometric, Recognition Rate*

مراجعة شاملة للتعرف على الإنسان القائم على CNN باستخدام صور شكل الأذن

محمد عبدالستار عبدالغني عبدالله1
بتول عبدالستار عبدالغني عبدالله2
1قسم البرمجيات، كلية علوم الحاسوب والرياضيات، جامعة الموصل،الموصل، العراق.
البريد الإلكتروني: mohammedse44@gmail.com
2قسم هندسة الطاقة، الكلية التقنية الهندسية، جامعة دهوك التقنية، دهوك- العراق.
البريد الالكتروني: batool.abdulghani@dpu.edu.krd

خلاصة

يعد التعرف التلقائي على الأشخاص باستخدام صور شكل الأذن مجالًا نشطًا للبحث داخل مجتمع القياسات الحيوية. على غرار السمات الحيوية الأخرى مثل بصمات الأصابع والوجه وقزحية العين، تتمتع الأذن أيضًا

بالعديد من الميزات المحددة والفريدة التي تساعد في التعرف على هوية الشخص. في ظل الجائحة العالمية الحالية لكوفيد-19، فشلت معظم أنظمة التعرف على الوجه تقريبًا بسبب سيناريو ارتداء القناع. تعد الأذن البشرية مصدرًا مثاليًا للبيانات لتحديد هوية الشخص السلبي لأنها لا تنطوي على تعاون الفرد الذي تتم محاولة التعرف عليه ويمكن التقاطها بسهولة أكبر عن بعد. يعد الحصول على صورة للأذن البشرية أمرًا سهلاً أيضًا حيث تكون الأذن واضحة حتى لدى المستخدمين الذين يرتدون الأقنعة. في نظام التعرف الآلي على الإنسان، يمكن دمج نظام القياسات الحيوية للأذن كمكمل لأنظمة القياسات الحيوية الأخرى ويقدم إشارات الهوية عندما تكون معلومات النظام الأخرى غير موثوقة أو حتى غير متوفرة. في هذا البحث، تم إجراء مراجعة شاملة لمجال البحث المتنامي لتقنيات استخراج الميزات وتقنية تصنيف التعلم العميق باستخدام الشبكة العصبية التلافيفية (CNN) لإظهار معدل الدقة لأنظمة التعرف على الأشخاص باستخدام صور شكل الأذن.

الكلمات المفتاحية: التعرف على الأذن، CNN، القياسات الحيوية، معدل التعرف

1. Introduction

In the present world, biometric systems are always implemented to increase access security levels (Sharma and Sethi, 2015). The characteristics of biometrics are permanent, distinct, universal, and quantifiable. The word "universal" refers to the requirement that each individual possess the biometric trait, which should hardly ever be lost due to illness or accident. The concept of uniqueness suggests that no two people should exhibit the same behavior. On the other hand, the permanent word indicates that a biometric property shouldn't alter over time. It shouldn't be significantly altered due to illness or aging.

The capacity to obtain and digitize a biometric feature using appropriate digital equipment or sensors without causing the subject any inconvenience is the ultimate definition of measurability (Prakash and Gupta, 2015).

The word "biometrics" refers to the process of identifying or authenticating a person based on specific physiological or behavioral traits. It is characterized as a quantitative attribute of an individual that can be utilized to identify them automatically. Behavioral traits are based on "what a person does" and are reliant on that person's conduct, whereas physiological traits are based on "what a person has" and depend on the structural information of the human body (Prakash and Gupta, 2015). Basic details on a few common physiological biometrics—the face, fingerprint, DNA, ear, iris, and palm print—are provided by the physiological characteristics. The behavioral features give the fundamental knowledge about a few common behavioral biometric traits, such as voice, gait, signature, and keystroke dynamics. Table 1 shows a summary of biometric characteristics if it is distinctive, permanence, collectability, and acceptability (Booysens and Viriri, 2022).

Table 1: A summary of biometric characteristics.

Biometric Identifier	Biometric Type	Distinctiveness	Permanence	Collectability	Performance	Acceptability
DNA	physiological	high	high	low	high	low
Ear	physiological	medium	high	medium	medium	high
Face	physiological	low	medium	high	low	high
Facial	physiological	high	low	high	medium	high
Fingerprint	physiological	high	high	medium	high	medium
Gait	behavioral	low	low	high	low	high
Hand Geometry	physiological	medium	medium	high	medium	medium
Hand Vein	physiological	medium	medium	medium	medium	medium
Iris	physiological	high	high	medium	high	low
Keystroke	behavioral	low	low	medium	low	medium
Odor	physiological	high	high	low	low	medium
Palm Print	physiological	high	high	medium	high	medium
Retina	physiological	high	medium	low	high	low
Signature	behavioral	low	low	low	low	high
voice	combination of physiological and behavioral	low	low	medium	low	high

The ear has been widely recognized as a reliable biometric for human recognition among different physiological biometric qualities. (Bhanu and Chen, 2008). The ear biometric framework can be trusted because of its constancy, uniform tone, and fixed position in the center side of the face. The size of a person's ear is more significant than a distinct fingerprint and makes it easier to capture an image of the subject without needing to gain information from the subject (Bhanu and Chen, 2008). Accurately measuring the features of the ear is very challenging. These include hiding an ear through clothing, hair, or jewelry. Another explanation might be that the photo was taken from a different angle, hiding important details of the ear's anatomy. Due to these challenges, ear recognition has become less important in methods and systems that are frequently utilized for verification and identification.

2. Ear Image Anatomy

The human ear, which is essential to hearing, fully develops in the first trimester of pregnancy. According to the medical study, the age range between 4 months and 8 years old as well as older than 70 years old is when ear variations are most obvious throughout time. From four months to eight years of age, ear growth is roughly linear.

After that, it remains constant until the age of seventy, at which point it begins to increase once more. The ear lobe is the primary target of the nonlinear stretch rate caused by gravity because of its stability and predictable fluctuations. Research is being done on ear recognition as a possible biometric (Kumar and Srinivasan, 2014).

Figure 1.1 depicts the anatomical structure of the human ear, which includes the lobe, the concha (a hollow portion of the ear), the tragus (a tiny bulge of cartilage over the meatus), and the outer rim (helix) and ridges (anti-helix) parallel to the helix. The crux is there at the start of the helix (ABDULGHANI and AL-SULAIFANIE, 2017).

The use of ears can be useful in the medical field when a patient's face cannot be recognized, particularly in cases of accidents and death. It is easy to take ear photos from a distance without the subject's knowledge. As a result, ear biometrics are appropriate for applications including monitoring, access control, surveillance, and security (Boodoo-Jahangeer and Baichoo, 2013).

Many biometrics are the subject of extensive research to ascertain their applicability for certain uses. The use of auditory biometrics for identification is the main topic of this thesis.

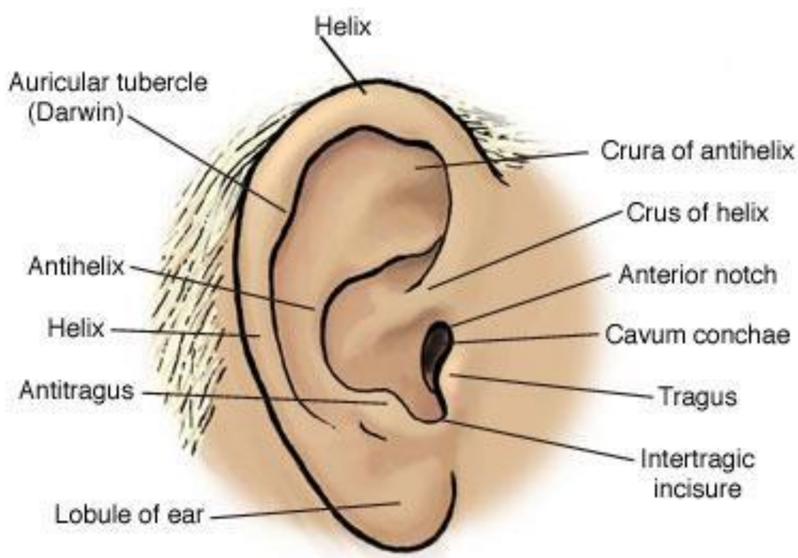

Fig. 1: External ear anatomy (D'Alessandro, 2012)

3. Convolutional Neural Network (CNN)

A convolutional neural network is consisting of numerous layers, such as convolution layers, pooling layers, and fully connected layers, and it uses a backpropagation algorithm to learn spatial hierarchies of data automatically and adaptively. The layers in a convolutional neural network are depicted in Figure 2.

2024 آذار العدد 12 A المجلة العراقية للبحوث الانسانية والاجتماعية والعلمية
No.12A March 2024 Iraqi Journal of Humanitarian, Social and Scientific Research
Print ISSN 2710-0952 Electronic ISSN 2790-1254

Deep learning methods called convolutional neural networks (CNNs) convolve input images using filters or kernels in order to extract features. When a fxf filter is used to convolve a NxN image, the convolution process learns the same feature over the whole image (Zeiler and Fergus, 2014). Following each operation, the window slides, and the feature maps learn the features. The feature maps include shared weights and biases and capture the image's local receptive field (Chauhan et al., 2018). Equation (1) shows the output matrix's size without any padding. and Equation (2) illustrates how convolution works. Padding is used to maintain the size of the input image. The output picture size is the same as the input image size when there is "SAME" padding, and there is no padding when there is "VALID" padding. Equation (3) shows the size of the output matrix with padding.

$$NXN * fXf = N - F + 1 \quad (1)$$

$$O = \sigma(b + \sum_{i=0}^{2} \sum_{j=0}^{2} w_{i,j}\, h_{a+i,b+j}) \quad (2)$$

$$NXN * f * f + (N + 2p - f) / (s+1) \quad (3)$$

Here, P is the padding, O is the output, s is the stride, b is the bias, σ is the sigmoidal activation function, $h_{x,y}$ is the input activation at position x, y and w is a 3x3 weight matrix of shared weights. CNN has many applications in fields of large scale image recognition (Krizhevsky et al.,2012), Speech emotion recognition (Huang et al., 2014) facial expression recognition (Alizadeh and Fazel, 2017), genomics, biometric systems, and many others.

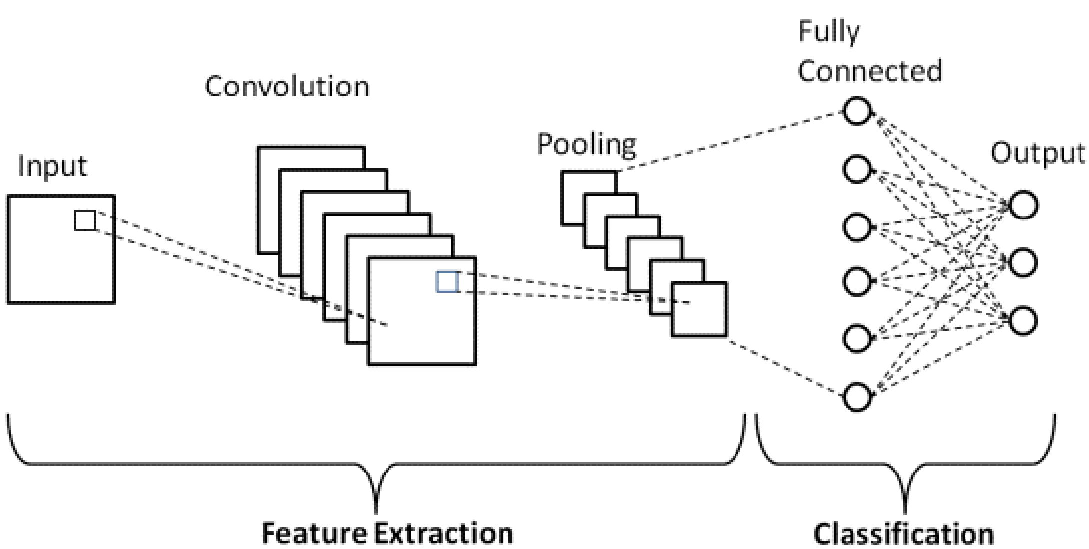

Fig. 2: Convolutional Neural Network Layers

4. Related Works

An artificial intelligence model called "deep learning" makes use of several layers to progressively understand the underlying facts. Exploring the use of deep feature extraction to ear shape photos has garnered a lot of attention lately. Notably, deep learning techniques have been developing quickly and are now widely used for data-driven learning in a variety of computer vision applications. These techniques are used in conjunction with traditional procedures to construct an end-to-end model through feature extraction and classification. One well-known example of a deep learning model for feature learning is Convolutional Neural Nets (CNNs). Furthermore, a variety of methods have been suggested to improve deep learning even more; these are described below:

Omara et al. (2018) presented a novel method for identifying human ears. It involves combining hierarchical deep characteristics. Initially, the work extracts hierarchical deep features from ear pictures using a CNN that has been pre-trained on a large dataset. Combining deep features from several layers using discriminant correlation analysis (DCA) enhances feature representation and reduces high dimensionality. Because each participant has a limited number of ear photos, the authors create paired samples and solve the ear identification problem using a pairwise (SVM) to convert it to binary classification. Four publicly accessible databases are used to assess the proposed method: USTB1, USTB2, IIT Delhi1, and IIT Delhi2. IIT Delhi I has the highest accuracy rate, which is 99.5%.

Jamil et al. (2018) introduced a new approach to capture ear photos in different uniform illumination levels, expressed in lumens or lux, which vary from 2 lux to 10,700 lux. Natural lighting was used to capture 1,100 pictures of the left and right ears of 55 individuals. The ear images were rotated at 5-degree intervals to produce 25,300 images, which was enough to provide a large enough amount of data for (CNN) training. A split of each participant's dataset was made, with 50 images set aside for testing and validation and the remaining images being used for training. After starting from scratch, the proposed CNN model was trained, and testing and validation results revealed that the recognition accuracy was 97%.

A technique for identifying unconstrained ears was presented by Dodge et al. (2018) using deep neural networks (DNNs) and transfer learning. The method consists of employing pre-existing DNNs as feature extractors to produce efficient features for ear recognition, which are subsequently used by a shallow classifier. Small image alterations are added to the training dataset to enhance performance. The study compares the performance of modified networks with feature-extraction models and addresses the issue of over-fitting resulting from small sample sizes. An average ensemble based on deep learning is proposed to mitigate the effect of over-fitting.

The suggested technique is evaluated on three datasets: AWE, CVLE, and combined AWE + CVLE. Consequently, the three datasets had an accuracy of 85.50%, 99.69%, and 93.48%, respectively.

According to Abd Almisreb et al. (2018), transfer learning was employed with the popular Alex Net (CNN) to identify people based just on pictures of their ears. In order to solve the particular problem domain, the Alex Net CNN was modified and improved. This included replacing the last completely connected layer to allow for the recognition of 10 classes rather than the original 1000 classes. To enhance the network's non-linear problem-solving abilities, a second layer of Rectified Linear Units (ReLU) was added. A dataset of 250 ear photos from 10 people was utilized to train the improved network; 50 ear images were used for testing and validation. With a validation accuracy of 100%, the results show that the suggested revised network performed remarkably well in the target application.

A unique deep learning-based method for ear localization and recognition was presented by Sinha et al. (2019). CNNs are used for ear classification, SVMs are used for ear recognition, and Histograms of Oriented Gradients (HOG) are used for ear localization in the proposed system. By merging feature extraction and recognition tasks into a single network, the proposed method addresses challenges such as variations in illumination, contrast, rotation, scale, and position. CNNs technology delivered an average recognition accuracy of 97.9% on the USTB III database. These findings suggest that the suggested method has potential for use in biometric identification applications.

In 2019, El-Naggar and Bourlai presented a human recognition system based on ear picture analysis. explicitly searched for a convolutional neural network architecture that would function well for the ear recognition challenge. The performance of four distinct network architectures—Alex Net, Squeeze Net, Google Net, and MobileNetV2—was examined in order to achieve this. Given the lack of training data, learning is enhanced by the use of transfer learning, data augmentation, and domain adaption methodologies. All things considered, this work provides a substantial insight into the suitability of different convolutional neural network topologies for ear detection and highlights the challenges associated with identifying ears in different orientations. The results demonstrate that Mobile NetV2 has a significantly higher accuracy rate of 95.67%.

A method in geometric deep learning (GDL) that extends CNNs to non-Euclidean domains was proposed by Tomczyk and Szczepaniak (2019). Using the Gaussian mixture model (GMM) in continuous space, the suggested method provides convolutional filters that produce systems with rotation equivariance qualities and facilitate simple rotation without the need for extra interpolation. An crucial task in biometric human identification, ear detection, serves as a demonstration of this approach's usefulness. Compared to traditional CNNs, the study presents a very straightforward approach with less information on visual content. Thus, 98.22% is the best accuracy that was attained.

A study on ear recognition with deep convolutional neural networks is presented by Alshazly et al. (2020). Using the EarVN1.0 dataset, eight distinct CNN architectures—AlexNet, VGGNet, InceptionV3, ResNet, and ResNeXt—are assessed using various transfer learning techniques. Techniques for ensemble learning, fine-tuning, and feature extraction are investigated. Using ResNeXt101, fine-tuning produced the greatest results with a rank-1 accuracy of 93.45%. Ensembles of optimized models achieved an additional 95.85% improvement in performance. Discriminative features were learned, as evidenced by feature visualizations.

A novel six-layer deep convolutional neural network architecture for ear recognition was proposed by Ahila Priyadharshini et al. in 2021. The potential efficacy of the proposed deep network is evaluated using the IITD-II and AMI ear datasets. The deep network model achieves 97.36% and 96.99% recognition rates with these datasets, respectively. Furthermore, the resilience of the proposed system is assessed in an uncontrolled environment using the AMI Ear dataset. When combined with a sufficient monitoring system, this method can yield significant real-world benefits for identifying certain individuals throughout large crowds.

Faster Region-based Convolutional Neural Network (Faster R-CNN), a deep learning object detection framework, was proposed by Alkababji and Mohammed (2021) for the detection of ears. With the use of a convolutional neural network (CNN), feature extraction was completed. Following that, PCA was utilized to achieve feature reduction, and a genetic algorithm was employed to achieve feature selection. As a matcher, a fully connected (ANN) was employed at the end. Based on the testing phase results, a 97.8% recognition rate was achieved.

For the classification of ear pictures, Sharkas (2022) suggested two different methods. First, a classifier was trained using a conventional machine-learning algorithm employing features extracted from the discrete curvelet transform. Prior to feature extraction, picture segmentation and enhancement were accomplished by image preprocessing techniques. To be more specific, when the ear region was selected from the backdrop, the segmented ear images were wrapped using the curvelet transform. We looked into different curvelet transform levels and divided the coarse image into blocks. Following the determination of the mean, variance, and entropy of each block, a feature vector was constructed by concatenating the statistical parameters that were similar across the sub-images at different levels. The only classifier that could produce competitive results after obtaining the feature vector for ear recognition was the ensemble classifier. Deep learning methods were applied in the second strategy to classify images of ears. Features from different end-to-end networks were extracted and given to a shallow classifier in order to categorize the ears. A 99.4% recognition percentage was attained.

Ramos-Cooper et al. (2022) enhanced the VGGFace dataset to create a new dataset. Deep learning models that were pre-trained on this dataset were then refined, and their sensitivity to several factors in the data was examined. investigated the effectiveness of a score-level fusion method in improving these models' overall recognition performance. In order to do this, the Unconstrained Environments for Recognition Challenge (UERC) dataset and the suggested dataset were used in both open-set and closed-set tests. When pre-trained face models are used instead of general image recognition models, the analysis's findings show a statistically significant improvement in recognition performance of about 9%. Additionally, a 4% boost in recognition accuracy was achieved, reaching 98.10%, by the combination of scores from both types of models.

Using CNN technology, Ebanesar et al. (2022) presented an innovative and effective approach to ear-based recognition. This non-invasive method makes use of the distinctive characteristics of the ear, one of the most widely used biometric markers for personal identification. The proposed system consists of two components: the authentication and identity components of the ear and the recognition component. This study uses the AMI dataset, which was taken using a Nikon D100 camera. The two-dimensional images are first converted into a one-dimensional space during the pre-processing phase. Then, the (LBP) approach is used for detection, recognition, and classification. To extract more discriminating factors, histogram analysis is employed. This technology uses a CNN-based model to increase the ear biometric system's security level. The findings provide a 98.99% total accuracy rate, indicating the potential for the suggested methodology to advance biometric security.

A technique for evaluating the effects of occlusion and the division of left and right ear pictures on recognition accuracy in the AWE dataset was put out by Resmi et al. in 2023. It is usual for a person's left and right ear images to be asymmetric, so a systematic assessment of each ear independently is necessary. To assess the accuracy of the left and right ear image recognition, the study used a pre-trained ResNet50-based model. The results show a significant increase in accuracy when processing the pictures of the left and right ears separately. Additionally, a novel occlusion-based data augmentation method is developed and tested using a ResNet50-based model. The findings show that the left ear has a 77.33% accuracy rate and the right ear has an 80% accuracy rate. These results demonstrate how well left and right ear pictures can be distinguished from one another and how useful occlusion is for improving recognition accuracy in biometric identification systems. 63.00% is the accuracy rate for the entire dataset.

5. Discussion

Ear biometrics has the potential to be a dependable and precise method of identifying people, as demonstrated by the evaluation of human recognition using ear features. For the field to be widely adopted, a number of important challenges still need to be resolved. Privacy concerns are the primary issues that need to be addressed. The protection of people's privacy is becoming more and more important as biometric technologies are used. Subsequent research ought to concentrate on creating ear recognition systems that protect privacy by not storing biometric or personal data in centralized databases. Strong security measures should also be included in these systems to guard against misuse or illegal access to the data. The lack of diversity in the datasets utilized for algorithm evaluation and training for ear identification is another cause for concern. The majority of the datasets that are currently accessible are not very diverse in terms of gender, age, or ethnicity. To make sure that ear recognition algorithms are inclusive and do not result in bias or discrimination, future research should concentrate on creating more varied datasets.

6. Conclusion

In conclusion, ear biometrics offers a lot of promise as a trustworthy and precise method of identifying people. However, in order to guarantee its widespread adoption, a number of important difficulties must be resolved. Standardization, privacy, and diversity should be the main topics of future research, along with possible uses outside of security and law enforcement. By addressing these problems, ear biometrics will be used in many fields in a more ethical and inclusive manner. More specifically, even though there are a number of well-established techniques for ear image recognition, the current study offers an extensive overview of several ways for identifying people based on their ear characteristics and points out areas for future development. It is therefore possible to identify persons based on ear features by employing deep learning techniques as CNN, Alex Net CNN, DNN, VGG-16, SVM, Google Net, Squeeze Net, Mobile NetV2, etc.

References

Abd Almisreb, A., Jamil, N., & Din, N. M. (2018, March). Utilizing AlexNet deep transfer learning for ear recognition. In 2018 fourth international conference on information retrieval and knowledge management (CAMP) (pp. 1-5). IEEE.

ABDULGHANI, B. A., & AL-SULAIFANIE, A. K. (2017). EAR RECOGNITION USING LOCAL BINARY PATTERN. Journal of Duhok University, 120-128.

Ahila Priyadharshini, R., Arivazhagan, S., & Arun, M. (2021). A deep learning approach for person identification using ear biometrics. Applied intelligence, 51, 2161-2172.

Alizadeh, S., & Fazel, A. (2017). Convolutional neural networks for facial expression recognition. arXiv preprint arXiv:1704.06756.

Alkababji, A. M., & Mohammed, O. H. (2021). Real time ear recognition using deep learning. Telkomnika (Telecommunication Computing Electronics and Control), 19(2), 523-530.

Alshazly, H., Linse, C., Barth, E., & Martinetz, T. (2020). Deep convolutional neural networks for unconstrained ear recognition. IEEE Access, 8, 170295-170310.

Bhanu, B., & Chen, H. (2008). Human ear recognition by computer. Springer Science & Business Media.

Boodoo-Jahangeer, N., & Baichoo, S. (2013). LBP-based ear recognition. In Proceedings of the 2013 IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 1-4). IEEE.

Booysens, A., & Viriri, S. (2022). Ear biometrics using deep learning: a survey. Applied Computational Intelligence and Soft Computing, 2022.

Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018, December). Convolutional neural network (CNN) for image detection and recognition. In 2018 first international conference on secure cyber computing and communication (ICSCCC) (pp. 278-282). IEEE.

D'Alessandro, M. P. (2012). A digital library of anatomy information. Retrieved 9 19, 2016, from http://www.anatomyatlases.org/firstaid/Otoscopy.shtml

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

Ebanesar, T., Bibin, A. D., & Jalaja, J. (2022). HUMAN EAR RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK. Journal of Positive School Psychology, 8182-8190.

El-Naggar, S., & Bourlai, T. (2019, November). Evaluation of deep learning models for ear recognition against image distortions. In 2019 European Intelligence and Security Informatics Conference (EISIC) (pp. 85-93). IEEE.

Ganapathi, I. I., Prakash, S., Dave, I. R., & Bakshi, S. (2020). Unconstrained ear detection using ensemble-based convolutional neural network model. Concurrency and Computation: Practice and Experience, 32(1), e5197.

Huang, Z., Dong, M., Mao, Q., & Zhan, Y. (2014, November). Speech emotion recognition using CNN. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 801-804).

Jamil, N., Almisreb, A., Ariffin, S. M. Z. S. Z., Din, N. M., & Hamzah, R. (2018). Can convolution neural network (CNN) triumph in ear recognition of uniform illumination invariant?. Indonesian Journal of Electrical Engineering and Computer Science, 11(2), 558-66.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

Kumar, V. N., & Srinivasan, B. (2014). Automated human identification scheme using ear biometrics technology. International Journal of Image, Graphics and Signal Processing, 6(3), 58-65.

Mewada, H. K., Patel, A. V., Chaudhari, J., Mahant, K., & Vala, A. (2020). Wavelet features embedded convolutional neural network for multiscale ear recognition. Journal of Electronic Imaging, 29(4), 043029-043029.

Omara, I., Wu, X., Zhang, H., Du, Y., & Zuo, W. (2018). Learning pairwise SVM on hierarchical deep features for ear recognition. *IET Biometrics*, *7*(6), 557-566.

Prakash, S., & Gupta, P. (2015). Ear Biometrics in 2D and 3D: Localization and Recognition (Vol. 10). Springer.

Ramos-Cooper, S., Arequipa, P., & Camara-Chavez, G. (2022). Domain Adaptation for Unconstrained Ear Recognition with Convolutional Neural Networks. CLEI electronic journal, 25(2), 8-1.

Resmi, K. R., Raju, G., Padmanabha, V., & Mani, J. (2023, January). Person Identification by Models Trained Using Left and Right Ear Images Independently. In 1st International Conference on Innovation in Information Technology and Business (ICIITB 2022) (pp. 281-288). Atlantis Press.

Sharkas, M. (2022). Ear recognition with ensemble classifiers; A deep learning approach. Multimedia Tools and Applications, 81(30), 43919-43945.

Sharma, C., & Sethi, N. S. (2015, June). An ICA Based Feature Level Fusion of Iris and Ear Biometrics. International Journal of Advanced Trends in Computer Applications (IJATCA), 2(1), 72-76.

Sinha, H., Manekar, R., Sinha, Y., & Ajmera, P. K. (2019). Convolutional neural network-based human identification using outer ear images. In Soft Computing for Problem Solving: SocProS 2017, Volume 2 (pp. 707-719). Springer Singapore.

Tomczyk, A., & Szczepaniak, P. S. (2019). Ear detection using convolutional neural network on graphs with filter rotation. Sensors, 19(24), 5510.

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13 (pp. 818-833). Springer International Publishing.