International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 8 ǁ August 2022 ǁ PP. 71-78

ResNeXt Model for Ear Recognition and Classification Based on Deep Learning

Manvendra singh, Dr Dushyant singh
[1]*Department of Electronics Science and Engineering of FET RBS, Bichpuri, Agra*
[2]*Department of Electronics Science and Engineering of FET RBS, Bichpuri, Agra*

Abstract
The identification of an individual's identity from an image of the ear is an important subject of research within the biometrics field. Ear images may be collected from a distance, making ear recognition technology an interesting option for security and surveillance applications as well as other application sectors. In contrast to other biometric modalities, the human ear is not influenced by facial emotions and does not need intimate contact as fingerprints do. In the current research, an automated method for the identification and classification of the ears is presented in this paper. An EarVN1.0: A new large-scale ear images dataset in the wild from an open repository is built using a modified ResNeXt CNN (Convolution Neural Network) model. Identification and classification of the ears using a modified ResNeXt CNN topology are based on the ear's arrangement and morphological features, such as its width and shape. Experimental classification accuracy of 95.70% indicates the effectiveness of the proposed strategy. The current study may pave the way for the future creation of a time- and cost-efficient, highly sensitive computer-based system for the identification and classification of various ears.
Keywords: *Ear Detection, Digital Image Processing, CNN, Deep Learning*

Date of Submission: 22-07-2022 Date of acceptance: 04-08-2022

I. INTRODUCTION

The process of automated human recognition based on the physical characteristics of the ears is known as ear recognition. It has a wide range of applications, including forensics, surveillance, identity verification, and unlocking user devices. Ear photographs can provide a rich supply of biometric information for creating successful recognition systems due to the distinctive structure of the ear shape. Furthermore, the human ear has several desirable qualities, including simplicity of capture from a distance, stability over time, capacity to distinguish identical twins [1], and insensitivity to emotions and facial expressions [2], [3]. We can construct and develop reliable identification systems on a variety of devices in a non-intrusive and non-distracting manner using these appealing qualities [4][5]. However, when ear pictures are obtained in unconstrained contexts, where numerous appearance variations and illumination changes must be addressed, accurate recognition can be a difficult challenge [6].

Early ear recognition research has shown significant performance improvements, particularly for ear photos recorded under controlled conditions [7]. To define the main aspects of ear pictures, most of these strategies used manual feature engineering (a.k.a. handmade) methods. The collected characteristics were then used to train a standard classifier to discover certain patterns in the extracted features that might be used to distinguish individuals. The robustness of the feature extraction method and the effectiveness of the used classier have a significant impact on the success of these ear identification techniques. In essence, these approaches have two major drawbacks. On the one hand, manually extracting useful features from photos necessitates personnel with a thorough understanding of the topic, and it is a time-consuming operation. On the other hand, if the level of appearance diversity in the given photos increases, the performance of these approaches decreases. As a result, if these restrictions are not addressed, performance suffers, particularly when detecting ear pictures under uncontrolled imaging conditions.

Deep learning techniques, particularly deep Convolutional Neural Networks (CNNs), have led to advancements in a variety of application domains in recent years, including picture classification [8][9], object detection [10][11], and biometric recognition [12]. The availability of massive volumes of labelled data, powerful hardware (i.e. GPUs) for accelerating computations, well-designed deep network designs, effective optimization approaches, and technical advancements in deep network training have all contributed to these gains. Deep CNNs, in addition to being scalable supervised learning algorithms, accomplish feature extraction and classification by training the entire system from beginning to end, eliminating the need for human feature extraction. Deep CNN training, on the other hand, necessitates the optimization of a huge number of trainable parameters (millions) as well as large-scale labelled datasets. Furthermore, collecting such large volumes of data may be prohibitively expensive in some real-world applications, limiting the deep model's high potential.

Transfer learning [13], [14] is an effective way to address the restrictions described above. It's an approach in which a deep CNN's knowledge gained on a specific task and dataset is transferred or used to train additional deep CNNs on similar but unrelated tasks and datasets. Transfer learning has become the most promising solution for tackling difficult visual recognition tasks in recent years. In this paper, we use transfer learning with deep CNNs to solve the challenge of recognizing ear photos recorded under unconstrained conditions, and we provide the results of the first ear recognition experiments on the EarVN1.0 dataset. Several significant contributions have resulted from the research efforts and outcomes, which are outlined below:

- On the unconstrained EarVN1.0 dataset, we report the first experimental study of ear identification. To do this, we use state-of-the-art deep CNN architectures of various depths and compare and analyze their recognition performance.
- For the first time, two seminal deep CNN architectures (InceptionV3 and ResNeXt) are tested in-ear recognition studies. Extensive experiments are described, as well as a thorough study of their performance and computing complexity. To train the networks, we use LAMB [15], a recently introduced layer-wise adaptive large batch optimization approach. As stated in [16], [17], the LAMB optimizer has shown success in training deep networks, beating adaptive optimization approaches such as the Adam optimizer [18].
- For CNN architectures with more than one completely linked layer, we suggest a two-step one-tuning technique. To preserve the aspect ratio of ear images, we experiment with training the networks with fixed input sizes and bespoke input sizes defined for each network. With a rank-1 accuracy of 93.45 per cent, our networks tuned with specific size inputs achieve state-of-the-art results, demonstrating the success of our proposed technique.
- To improve overall recognition accuracy, we investigate the usefulness of deep ensembles of separately one-tuned CNNs. When compared to utilizing a single network, each of the analyzed networks achieves a relative improvement in accuracy of more than 2%.
- Under each learning approach, we present visualizations of the learned features by the deep models. Our models have learned more discriminative characteristics, as evidenced by the visualizations. As a result, the acquired results are easier to understand.

II. INTRODUCTION

The literature on existing ear identification systems focuses on two aspects: feature extraction and the ear categorization approach.

- **FEATURE EXTRACTION-BASED APPROACHES**

Different types of discriminant features, such as geometrical, appearance-based, and deep features, have been proposed in the literature for feature extraction. Mu et al. used a combination of outer ear shape and inner ear curves to describe ear pictures in [19]. Shailaja and Gupta find intersection locations between the outer helix and the normal lines of maximal ear high lines in [20]. (EHLs). The EHLand horizontal reference lines were used by Rahman et al. to detect angles as features in [21], which was driven by the merger of maximal and minimum ear EHLs in [22]. However, several minor occlusions, rotations, and illuminations on-ear pictures had an impact on system performance for the majority of the prior geometrical features. To circumvent some of the aforementioned constraints, various other studies have offered holistic, local [23], and fusion aspects [24]. For example, [25] introduces eigen-ear, and [26] introduces independent component analysis (ICA). In addition, [27, 28] proposes 1D and 2D Gabor filters, and [23] illustrates most of the local features for ear recognition.

For ear recognition, Nosrati et al. [24] used 2Dwavelet and PCA, while Nanni and Lumini [29] presented a multi-matcher approach for ear verification. They separated each ear image into sub-windows, retrieved features from each sub-window using Gabor filters, and then used the sequential forward floating selection (SFFS) algorithm to choose the best match. To represent ear pictures, the authors combined the Haar wavelet transform and uniform local binary patterns (ULBPs) in [30]. The texture features of ear sub-images processed by Haar wavelet are then provided by ULBPs. The SIFT features were provided to the RGB colour channels by Zhou et al. [31], and the features were merged for ear recognition. Guo and Xu [33] and Benzaoui and Hadid [32] employed closest neighbour and SVM for classification, respectively, and combined versions of the local binary pattern for encoding ear pictures. Pflug et al. [34] recently integrated two surface and local texture descriptors (shape context, LBP, LPQ, HOG, and BSIF) to represent ear images, whereas Jacob and Raju [35] used grey-level co-occurrence matrix (GLCM), local binary pattern (LBP), and Gabor filter for ear detection.

Many ear recognition systems have recently begun to acquire deep features that can attain great accuracy and performance. Pedro et al. [36] demonstrated a bespoke neural network for ear detection and gave a quick overview of deep learning for ear recognition. Omara et al. [37], on the other hand, employed a pre-trained model to extract deep ear characteristics, merged them with CNN features, and presented a pairwise SVM classifier for ear classification. To address the low amount of labelled training data, Emersic et al. [38] offered several CNN ear recognition models with limited training data based on the augmentation procedure of the training images. Dodge et al. [39] used a shallow classifier for ear recognition and presented several deep CNNs as feature extractors. All of the above deep aspects have shown to be effective and can help the ear recognition system operate better. As a result, the suggested approach in this study uses the merging of deep CNN features to better depict ear pictures.

- **EAR RECOGNITION METHODS**

The approach of determining the similarity of two ear images to the same person is known as ear classification or recognition, for which discriminative features for each ear image should be retrieved first. These characteristics, on the other hand, may be related to the same issue in different ways. Previous famous and usual ear identification approaches [22, 39, 23, 32], and [40] focus on the feature extraction process and use basic distances or standard classifier methods to solve the ear recognition problem. Multi-view learning and metric learning algorithms have been more popular in the recent decade for pattern recognition and picture categorization. Multi-view learning techniques, and are designed to combine features from many perspectives to identify collective information or employ complementary characteristics in a specific function to aid learning tasks such as picture categorization. Multi-view learning aims to train with labelled examples from all viewpoints. Moreover, in several pattern classifications, such as image classification, speech recognition, breast cancer, and gesture recognition, many SVMs and extreme machine learning approaches have recently been published [41, 42].

A Mahalanobis distance metric, on the other hand, uses the correlation of several features as the off-diagonal elements and is scaled invariant. As a result, Mahalanobis metric distance can take advantage of both observation properties and their weights. There are two types of metric learning methods now available: pairwise and triplet constraints-based metric learning methods, as described in [43, 44], and [45]. The paired constraints cause comparable pairs' distances to be lower than a specific threshold, while dissimilar pairs' distances to be longer. Because of the triplet limitations, the distance between comparable samples is shorter than the distance between dissimilar samples. For example, LMNN learns the distance metric using a convex problem with triplet constraints. Metric learning based on triplet constraints gives more information and details for distance metric learning. It differentiates each sample's similar and dissimilar neighbours by a significant margin. The Metric Boost and FrobMetric [46, 47] approaches are proposed by Shen et al. They learn the combination parameters based on triplet constraints and parameterize the distance metric as a linear combination of rank-1 matrices. In addition, in [48], Mei et al. inspired the ITML pairwise constraints to dynamic triple constraints. They looked into the LogDet divergence-based LDMLT [48], which uses LogDet divergence to address the metric learning problem with triplet constraints (i.e., Bregman projection method ITML). When compared to ITML and other metric learning approaches, LDMLT outperformed them in terms of system accuracy and training time. As a result, the goal of this research is to look at the LDMLT approach for ear recognition.

III. PROPOSED APPROACH

In this paper, Ears are classified using a modified ResNeXt CNN model derived from the residual network, as described by Pant et al. [49]. The presented model includes ReLu activation, residual blocks, grouped convolution, and the softmax optimizer in the final dense layer. Figure 1 depicts the architecture of the proposed ResNeXt CNN model. By duplicating the filters inside a module, grouped convolutions enable the building of a new network in a deep learning model. By extracting feature maps using kernel filters, this method minimizes computation [50]. The Leaky ReLu activation function was utilized to eliminate gradient descent issues and prevent model saturation [51]. In the proposed model, batch normalization is followed by Leaky ReLu activation.

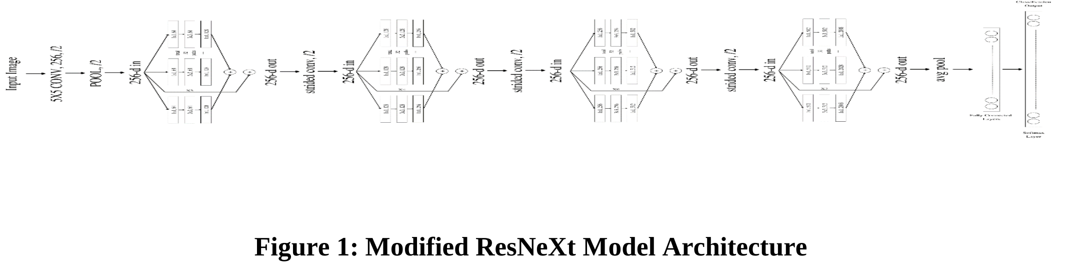

Figure 1: Modified ResNeXt Model Architecture

The inner product is a form of aggregating transformation that may be explored.

(1)

Where i.e. n channel input vector to the neuron.

The filter weight of the ith neurons is given by ω_i. The elementary transformation ($\omega_i \chi_i$) has been replaced by a function that is more general and can also be a network. As:

(2)

Where is an arbitrary function? should, similarly to a basic neuron, project into an (optionally low-dimensional) embedding and then change it. The residual block of the proposed model is defined as follows:

(3)

where is the output of the proposed model, c is the model's cardinality, and is the input vectors for neurons. The last dense layer comprises a softmax optimizer that transforms logits into the probability of transitioning to a certain ears class. This probability value may be estimated depending on the weight and bias provided as inputs. The value is finally transformed into the class value of the ears. Equations (4) and (5) are used to get the value of the SoftMax optimizer.

(4)

(5)

Where ϕ is the input vector, $\omega_0 \chi_0$ = the bias of the ith class? The network was trained using image dimensions of 256X256X3 and the cardinality hyperparameter ζ set to 32. Adjusting the last dense layer to 10 class classification. The training of the network to 20 epochs takes around 14 hours on an Nvidia GeForce GTX TITAN X GPU.

IV. RESULT & DISCUSSION

- **Dataset**

We have conducted the ear recognition experiments on the EarVN1.0 ear dataset. EarVN1.0 ear dataset is divided into three parts train, validation, and test set. Eighty per cent of the data set has been used for training, 10% has been used for validation, and the remaining 10% has been employed for testing.

- **Data Augmentation**

In the experiments, for deep CNN model training, images have been resized to 256 × 256 pixels resolution. These 256 × 256 sized images are cropped into five different images during the training phase and a single crop is taken from the centre of the image during the test phase. The crop image size for GoogLeNet [52] and VGG-16 [53] models is 224 × 224, while for AlexNet [54], it is 227 × 227.

- **Training and Validation**

Based on the Ear characteristics training and validation, the input size was ear images of size 256× 256 pixels. The labelling of each class was performed in both the training and the validation images. The network was compiled with the Adam optimizer.

The performance of the system was evaluated by the parameters accuracy (ACC), precision (Pre), recall (Re), and F1-score from the confusion matrix which can be defined as:

(6)

(7)

(8)

(9)

where, Tp (true-positive), Fp (false- positive), Fn (false-negative) and Tn (true-negative).

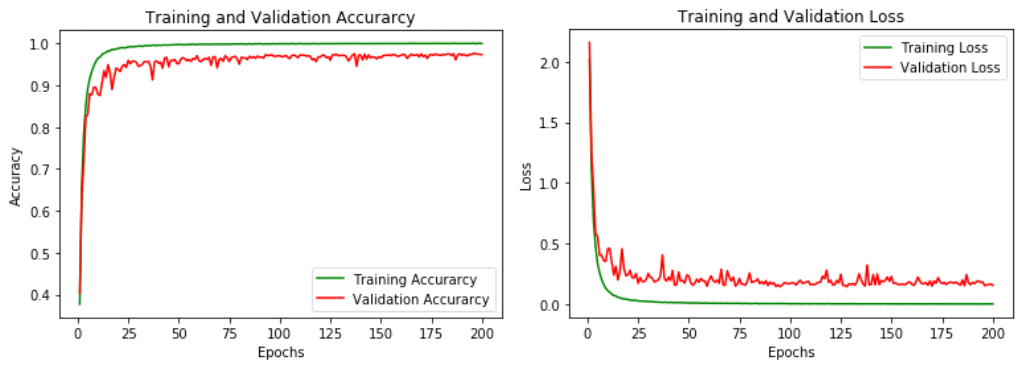

Figure 2: Training and validation of the proposed model with accuracy and loss.

We trained the model for a total of 162 people (Over 287 images) to achieve a training and validation accuracy were 100 and 95.97%, respectively (In figure 2), while training and testing loss was found to be 0.04 and 0.06%. Results show that the model performed remarkably well in the current study for ear recognition. However, by running the model for longer periods, loss in training and testing can be further minimized. The validation error of the suggested model is a key consideration when determining the number of periods. Overfitting could be indicated, for instance, if the validation error starts to rise. In the scenario, the model's number of periods should be set as large as possible, and training should be stopped based on error rates. The model's accuracy begins to saturate after a few iterations. For the suggested model, the minimum number of periods that saturates the accuracy should be fixed. After 20 periods in the current investigation, the model reaches saturation. We also ran the system through another experiment for 25 periods. In which the classification and training accuracy remained unaltered but the validation loss decreased to 0.002. As a result, 20 periods were chosen as the minimum.

- **The Error Matrix**

In the current work, an error matrix was plotted for the suggested model's sensitivity analysis according to Pant et al[49] .'s description. Predicted classes and their instances are shown in the rows and columns, accordingly. Similar physical traits in several ear shapes may affect how sensitive automated models are. Due to the same characteristics and shape of different ear types, various studies have documented a decrease in accuracy during automated recognition. A similar finding was made by Pant et al. [49] when occlusion occurs in-ear images were automatically classified using a modified ResNeXt CNN topology, and the error matrix analysis revealed 30 erroneous positive and negative results in a dataset of 42,000 pictures. Through the training of a model with less noise image and a big training dataset, the misclassification of ears with comparable occlusion ears images can be decreased. To decrease misclassification, a very high deep learning model—for instance, one with 150 layers—can be utilized [55]. The proposed model's classification accuracy was both ideal and respectable. Additionally, You can see in figure 3 that is confusion matrix of the proposed model.

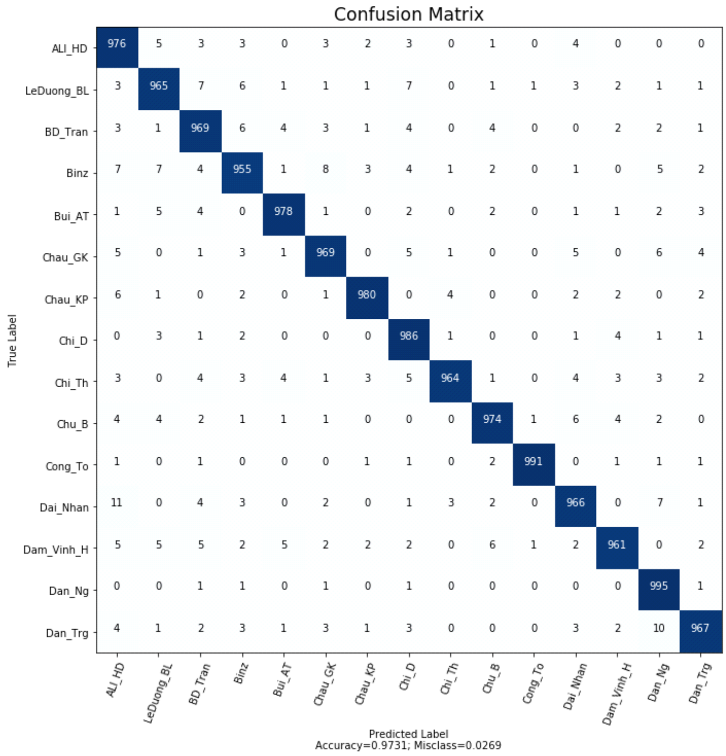

Figure 3: Confusion matrix of the proposed model

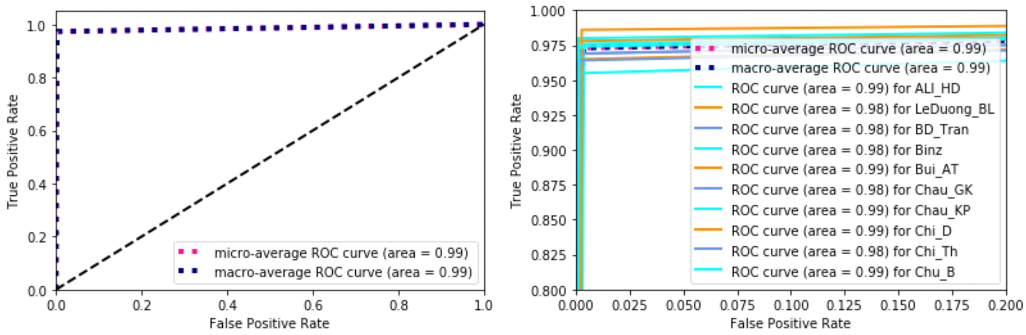

Figure 4: The ROC curve of the proposed model and class-wise ROC curve area.

- **F1-score and Receiver Operating Characteristic (ROC) curve**

Precision, recall, and F1-score were used to evaluate the model's effectiveness. The ear generates class-wise performance was evaluated [56]. Ears were recognized and categorized by the modified ResNeXt CNN model in the earlier study by Pant et al. [49]. However, because occlusion ears share a lot of characteristics and similarities, it was shown that false-positive values increased while real positive values decreased, which decreases precision.

The proposed automated model's accuracy is also assessed using Receiver Operating Characteristic (ROC) analysis. According to Figure 4, the ROC curves for the 10-class categorization of algae were over 99 per cent. Some classes' ROC curves are clsose to 100%. It is noteworthy how well the suggested modified deep ResNeXt CNN topology performs when evaluated using various parameters.

V. CONCLUSION

To reduce time and reliance on an expert for algae identification, a deep learning-based approach using a modified ResNeXt CNN topology has been created in the current study. For a supplemented data set of 80,000 pictures of 16 different person’s ear images, the proposed model confirms high accuracy and tolerable loss during training and validation. High F1-score, training accuracy, and classification accuracy, i.e. 95.9%, 100%, and 95.97%, respectively, establish the efficacy of the suggested model in conjunction with very minimal affinity in the confusion matrix and 95% ROC curve area. Compared to cutting-edge techniques, this precision is substantially superior. We will add more algal classes to the taxonomy in the upcoming investigation. To further assess the system's performance, the proposed models must also be tested on the other dataset.

REFERENCES

[1]. Nejati, Hossein, et al. "Wonder ears: Identification of identical twins from ear images." Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE, 2012.
[2]. Unar, J. A., Woo Chaw Seng, and Almas Abbasi. "A review of biometric technology along with trends and prospects." Pattern recognition 47.8 (2014): 2673-2688.
[3]. Wang, Zhaobin, Jing Yang, and Ying Zhu. "Review of ear biometrics." Archives of Computational Methods in Engineering 28.1 (2021): 149-180.
[4]. Oravec, Miloš, et al. "Mobile ear recognition application." 2016 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, 2016.
[5]. Bargal, Sarah Adel, et al. "Image-based ear biometric smartphone app for patient identification in field settings." VISAPP (3). 2015.
[6]. Barra, Silvio, et al. "Unconstrained ear processing: What is possible and what must be done." Signal and image processing for biometrics. Springer, Berlin, Heidelberg, 2014. 129-190.
[7]. Emeršič, Žiga, Vitomir Štruc, and Peter Peer. "Ear recognition: More than a survey." Neurocomputing 255 (2017): 26-39.
[8]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).
[9]. Iqbal, Muhammad Ather, et al. "Automatic fish species classification using deep convolutional neural networks." Wireless Personal Communications 116.2 (2021): 1043-1053.
[10]. Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).
[11]. Fu, Keren, et al. "Light field salient object detection: A review and benchmark." Computational Visual Media (2022): 1-26.
[12]. Minaee, Shervin, et al. "Biometrics recognition using deep learning: A survey." arXiv preprint arXiv:1912.00271 (2019).
[13]. Pan, Sinno Jialin, and Qiang Yang. "A survey on transfer learning." IEEE Transactions on knowledge and data engineering 22.10 (2009): 1345-1359.
[14]. Srivastav, Devansh, Akansha Bajpai, and Prakash Srivastava. "Improved classification for pneumonia detection using transfer learning with gan based synthetic image augmentation." 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2021.
[15]. You, Yang, et al. "Large batch optimization for deep learning: Training bert in 76 minutes." arXiv preprint arXiv:1904.00962 (2019).

[16]. Honkala, Mikko, Dani Korpi, and Janne MJ Huttunen. "DeepRx: Fully convolutional deep learning receiver." IEEE Transactions on Wireless Communications 20.6 (2021): 3925-3940.
[17]. Jahrens, Marius, and Thomas Martinetz. "Solving raven's progressive matrices with multi-layer relation networks." 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020.
[18]. Kingma, Diederik P. "&Ba J.(2014). Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2015).
[19]. Mu, Zhichun, et al. "Shape and structural feature based ear recognition." Chinese Conference on Biometric Recognition. Springer, Berlin, Heidelberg, 2004.
[20]. Shailaja, Dasari, and Phalguni Gupta. "A simple geometric approach for ear recognition." 9th International Conference on Information Technology (ICIT'06). IEEE, 2006.
[21]. Rahman, Mostafizur, Muhammad Sheikh Sadi, and Md Rezwanul Islam. "Human ear recognition using geometric features." 2013 international conference on Electrical information and communication technology (EICT). IEEE, 2014.
[22]. Omara, Ibrahim, et al. "A novel geometric feature extraction method for ear recognition." Expert Systems with Applications 65 (2016): 127-135.
[23]. Emeršič, Žiga, Vitomir Štruc, and Peter Peer. "Ear recognition: More than a survey." Neurocomputing 255 (2017): 26-39.
[24]. Nosrati, Masoud S., Karim Faez, and Farhad Faradji. "Using 2D wavelet and principal component analysis for personal identification based on 2D ear structure." 2007 International Conference on Intelligent and Advanced Systems. IEEE, 2007.
[25]. Chang, Kyong, et al. "Comparison and combination of ear and face images in appearance-based biometrics." IEEE Transactions on pattern analysis and machine intelligence 25.9 (2003): 1160-1165.
[26]. Zhang, Hai-Jun, et al. "A novel approach for ear recognition based on ICA and RBF network." 2005 international conference on machine learning and cybernetics. Vol. 7. IEEE, 2005.
[27]. Nanni, Loris, and Alessandra Lumini. "Fusion of color spaces for ear authentication." Pattern Recognition 42.9 (2009): 1906-1913.
[28]. Kumar, Ajay, and Chenye Wu. "Automated human identification using ear imaging." Pattern Recognition 45.3 (2012): 956-968.
[29]. Doghmane, Hakim, Abdelhani Boukrouche, and Larbi Boubchir. "A novel discriminant multiscale representation for ear recognition." International Journal of Biometrics 11.1 (2019): 50-66.
[30]. Sinha, Harsh, et al. "Convolutional neural network-based human identification using outer ear images." Soft computing for problem solving. Springer, Singapore, 2019. 707-719.
[31]. Zhou, Jindan, Steven Cadavid, and Mohamed Abdel-Mottaleb. "Exploiting color sift features for 2d ear recognition." 2011 18th IEEE international conference on image processing. IEEE, 2011.
[32]. Benzaoui, Amir, Abdenour Hadid, and Abdelhani Boukrouche. "Ear biometric recognition using local texture descriptors." Journal of electronic imaging 23.5 (2014): 053008.
[33]. Guo, Yimo, and Zhengguang Xu. "Ear recognition using a new local matching approach." 2008 15th IEEE International Conference on Image Processing. IEEE, 2008.
[34]. Pflug, Anika, Pascal Nicklas Paul, and Christoph Busch. "A comparative study on texture and surface descriptors for ear biometrics." 2014 International carnahan conference on security technology (ICCST). IEEE, 2014.
[35]. Jacob, Lija, and G. Raju. "Ear recognition using texture features-a novel approach." Advances in Signal Processing and Intelligent Recognition Systems. Springer, Cham, 2014. 1-12.
[36]. Galdámez, Pedro Luis, William Raveane, and Angélica González Arrieta. "A brief review of the ear recognition process using deep neural networks." Journal of Applied Logic 24 (2017): 62-70.
[37]. Omara, Ibrahim, et al. "Learning pairwise SVM on deep features for ear recognition." 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS). IEEE, 2017.
[38]. Emeršič, Žiga, et al. "Training convolutional neural networks with limited training data for ear recognition in the wild." arXiv preprint arXiv:1711.09952 (2017).
[39]. Dodge, Samuel, Jinane Mounsef, and Lina Karam. "Unconstrained ear recognition using deep neural networks." IET Biometrics 7.3 (2018): 207-214.
[40]. Omara, Ibrahim, et al. "Ear verification based on a novel local feature extraction." Proceedings of the 2017 International Conference on Biometrics Engineering and Application. 2017.
[41]. Güner, Ahmet, Ömer Faruk Alçin, and Abdulkadir Şengür. "Automatic digital modulation classification using extreme learning machine with local binary pattern histogram features." Measurement 145 (2019): 214-225.
[42]. Hu, Lin, and Jian Cui. "Digital image recognition based on Fractional-order-PCA-SVM coupling algorithm." Measurement 145 (2019): 150-159.
[43]. Davis, Jason V., et al. "Information-theoretic metric learning." Proceedings of the 24th international conference on Machine learning. 2007.
[44]. Bellet, Aurélien, Amaury Habrard, and Marc Sebban. "A survey on metric learning for feature vectors and structured data." arXiv preprint arXiv:1306.6709 (2013).
[45]. Wang, Faqiang, et al. "A kernel classification framework for metric learning." IEEE transactions on neural networks and learning systems 26.9 (2014): 1950-1962.
[46]. Shen, Chunhua, et al. "Positive semidefinite metric learning using boosting-like algorithms." (2012).
[47]. Shen, Chunhua, Junae Kim, and Lei Wang. "A scalable dual approach to semidefinite metric learning." CVPR 2011. IEEE, 2011.
[48]. Mei, Jiangyuan, et al. "Logdet divergence based metric learning using triplet labels." Proceedings of the Workshop on Divergences and Divergence Learning (ICML'13). Citeseer, 2013.
[49]. Pant, Gaurav, D. P. Yadav, and Ashish Gaur. "ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum." Algal research 48 (2020): 101932.
[50]. Zhao, Gangming, et al. "Rethinking ReLU to train better CNNs." 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018.
[51]. Zhao, Ruizhe, and Wayne Luk. "Learning grouped convolution for efficient domain adaptation." arXiv preprint arXiv:1811.09341 (2018).

[52]. Çınar, Ahmet, and Seda Arslan Tuncer. "Classification of lymphocytes, monocytes, eosinophils, and neutrophils on white blood cells using hybrid Alexnet-GoogleNet-SVM." SN Applied Sciences 3.4 (2021): 1-11.
[53]. Alippi, Cesare, Simone Disabato, and Manuel Roveri. "Moving convolutional neural networks to embedded systems: the alexnet and VGG-16 case." 2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). IEEE, 2018.
[54]. Iandola, Forrest N., et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size." arXiv preprint arXiv:1602.07360 (2016).
[55]. Durant, Thomas JS, et al. "Very deep convolutional neural networks for morphologic classification of erythrocytes." Clinical Chemistry 63.12 (2017): 1847-1855.
[56]. Rizk, Yara, et al. "Deep belief networks and cortical algorithms: A comparative study for supervised classification." Applied computing and informatics 15.2 (2019): 81-93.