Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Comparative Study On Iris Recognition Using Deep Learning

D. MEENAKSHI[1], Dr. M. SIVAJOTHI[2], Dr. M. MOHAMED SATHIK[3]

[1]Research Scholar (12108), Department of Computer Science, Sadakathullah Appa College, Affiliated to Manonmaniam Sundaranar University, Tirunelveli – 627 011, Tamilnadu, India.

[2]Associate Professor, Department of Computer Science, Sri Parasakthi College for Women, Courtallam – 627 802, An Autonomous College of the Manonmaniam Sundaranar University, Tamilnadu, India.

[3]Former Principal, Department of Computer Science, Sadakathullah Appa College, Affiliated to Manonmaniam Sundaranar University, Tirunelveli – 627 011, Tamilnadu, India.

ABSTRACT

Iris Recognition has been one of the most robust means of biometric recognition. In recent years, there are so many researches in iris recognition using deep learning. In this paper, three recent researches on iris recognition are compared and analyzed for the in-depth study on iris recognition. All the three researches use deep learning for its method which are VGGNet, Residual Convolutional Neural Network (Residual CNN) and Dense Fully Convolutional Neural Network (DFCN). Experiments results are compared on CASIA 1000 and IITD databases for all the three models. The DFCN architecture achieves good accuracy when compared to other methods.

Keywords: Deep Learning, Convolution Neural Network, Biometric Recognition

1. INTRODUCTION

From smartphone activation to cash withdrawals from ATMs to local store shopping [1], biometric technology has been quickly integrated into our daily lives. Various biometric modalities such as face, iris, retina, speech, fingerprints, palm prints, palm geometry are used in a wide variety of applications including law enforcement, border crossings and consumer applications [2, 3]. The human eye iris (the annular region between the pupil and the sclera) is of particular interest as iris is a highly distinguishable, robust and functioning biometric modality [4].

A number of key stages can roughly sum up the historical evolution of Iris recognition systems, each posing a new set of different challenges over the implementation of earlier technology:

In 1936, Burch planned to use the iris as a biometric [5] and in 1990s, the underlying technology was developed by Daughman [6] to automate iris recognition for practical deployment. Such systems have acquired iris pattern through a dedicated imaging system which restricts the target eye and uses Near-Infrared (NIR) imagery.

During the 2000s, systems were built to enable the acquisition of iris patterns from mobile persons under unconstrained acquisition conditions [7]. This system has been developed for implementation in public spaces such as airports, and allows people to walk along a designated path where multi-camera systems acquire multiple successive iris images under controlled lighting conditions.

Recently, the Iris recognition was developed and deployed on handheld devices such as smartphones [2]. Such image processing is completely unmonitored and unconstrained. It adds new artifacts that are not present in earlier conditions of acquisition which include unwanted reflections, occlusions, non-frontal iris images, poor contrast and partially blurred images. With such an acquisition environment, it has proved more difficult to distinguish iris regions properly and requires improvements to the authentication.

Most of the current iris recognition system follows these conventional steps: (i) image acquisition: a camera is used to capture an eye image (ii) iris segmentation: this image shows the region of the eye / iris accompanied by extraction of the area containing the iris. (iii) Feature extraction: specific features reflecting the uniqueness of the iris pattern are extracted from the iris area and (iv) pattern matching techniques determine the similarity of the two iris images.

The rest of the paper is organized as follows: Section II describes some related works that are made in recent years. Section III elaborates three recent methods on deep learning. Section IV analyzes those three methods with some experiments followed by conclusion in Section

II. Related Works

This section briefly discusses some innovative iris-recognition work. Most methods of recognition utilize methodologies of deep learning. Upon studying these methods, it is easy to build our own iris recognition architecture.

The iris code in [8] is determined from the optimization perspective. The traditional iris code is the solution to an optimization problem that minimizes the distance between the characteristic values and the iris codes. In addition, it is possible to obtain more efficient iris codes by adding terms to the objective function of this optimization problem. In addition, it examines two objective terms. The first objective term exploits the bits' spatial relations in various positions of an iris code. The second objective concept in the iris codes mitigates the impact of less accurate bits. The two objective concepts can be applied individually or in a combined scheme to the problem of optimization.

Four feasible network schemes are developed in [9], and the best Fully Dilated Convolution network model combining U-Net (FD-UNet) is obtained through training and testing on the same datasets. The FD-UNet uses dilated convolution to extract more global features instead of original convolution, so that image details can be processed better.

By combining Dense network with U-Net network, Iris is segmented by exploiting dense U-Net, which is smaller and has fewer parameters, and exploiting U-Net in semantic segmentation [10]. Dense connected path is derived from the dense connected network (Dense U-Net), in which enhanced information and parameters are helpful in reducing deep network training difficulty.

Deep Convolution Neural Network (DCNN) is trained to extract iris features on the basis of a large number of iris samples [11]. More specifically, the Tight Center (T - Center) Loss optimized center loss function is used to address the question of inadequate tolerance caused by the conventional loss function of Softmax layer.

The impact of iris segmentation on deep learning model performance is analyzed using a simple two-stage pipeline consisting of segmentation and recognition [12]. They evaluated how segmentation accuracy influences recognition performance and also examined whether segmentation is important at all.

Periocular information can be dynamically improved for more precise identification of iris by integrating the variations in the effective area of available iris regions. A periodic assisted dynamic framework is developed in [13] for more accurate, less-constrained identification of the iris.

An effective framework for the recognition of iris [14] is designed, consisting of three steps. Firstly, a reconstruction loss guided unsupervised pre-training phase is developed followed by supervised refinement to address the lack of labeled iris data. That drives the weights of the network to focus on patterns of discriminative iris texture. First, some texture-conscious improvisations are modified to improve the use of iris textures within a CNN. Finally, an efficient framework is built by systematic training and architecture with up to 100 fewer parameters than contemporary deep learning methodologies while achieving better recognition performance.

A phase-based iris recognition algorithm is developed in 2006 [15]. The concept of 2D Fourier Phase Code (2D FPC) for representing iris information is implemented to minimize the size of registered iris data and to prevent the exposure of iris images. 2D FPC corresponds to the quantized phase-spectrum version of the iris image, important for phase-based iris recognition. By using 2D FPC, while maintaining an adequate level of performance, the size of iris data can be reduced to below one-quarter when compared to using iris image directly as the record data. 2D FPC is particularly useful for using state-of-the-art DSP (Digital Signal Processing) technology to implement compact iris recognition devices.

III. Comparison of Deep Learning Methods on Iris Recognition

Because of tremendous success of deep learning in computer vision problems, there was a lot of interest in applying features learned by CNN on general image recognition to other tasks such as segmentation, face recognition, and object detection. In this section three methodologies of deep learning are discussed.

- VGG-Net - The first method has extracted deep features on iris recognition [16]. They used VGG-Net for the feature extraction.
- Residual CNN - The second method used residual convolutional neural network for feature extraction and recognition [17].
- DCNN - The third method used dense fully convolutional neural network and some popular optimization methods [18].

A. VGG Net Model

The application of profound features extracted from VGG-Net for iris recognition tasks is explored in this method. The VGG16 architecture is shown in Fig. 1. The trained model is viewed as a feature extraction engine and used to extract features from iris images without any fine-tuning to see if the general features are appropriate for biometric recognition. Features are extracted from various layers of this network and their iris recognition efficiency is evaluated. Then the Principal Component Analysis (PCA) is applied to minimize the dimensionality of the function, and then the multi-class Support Vector Machine (SVM) is used for recognition purposes.

It is worth noting that the segmentation step is skipped in this process to see the robustness of these features to intra-class variability, and given the presence of various variations in the CASIA-1000 dataset, this algorithm has achieved a very high precision rate. Although VGG-Net is trained to identify objects from different categories, it is very important that the CNN feature from this network works fairly well for iris recognition, which is to identify iris images of different subjects.

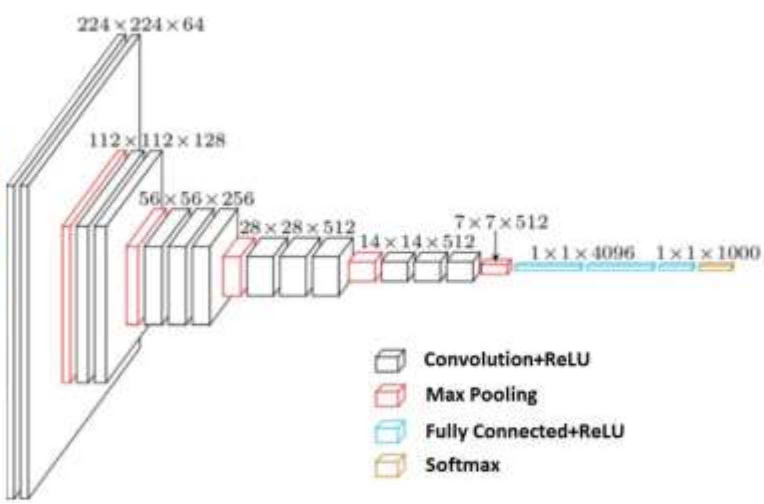

Fig. 1 VGG16 Architecture

B. Residual CNN Model

For iris recognition, an end-to-end deep learning system is built based on Residual CNN which can jointly learn the feature extraction and perform recognition. A dataset with a large number of subjects but a small number of images per subject is chosen in this process. A transfer learning algorithm is used to perform recognition using a deep convolution residual network. A ResNet50 [19] pre-trained model trained on the ImageNet dataset is used and fine-tuned on the images of the training. ResNet is common CNN software which was the winner of the ImageNet 2015 competition for visual recognition. ResNet50 's architecture is shown at Fig. 2. For more efficient training it generates easier gradient flow.

This would help the network provide a direct path to the network's very early layers, simplifying the gradient updates for those layers.

The most significant regions are visualized using a simple method when conducting iris recognition using CNN [20]. It starts from the top-left corner of an image and each time zeroes out a square region of size N x N within the image, and predicts using the model trained on the occluded image. If it makes the model mislabel the iris image by occluding the region, that area will be considered as an important area, when doing iris recognition.

On the other hand, if the elimination of that region does not impact the prediction of the model, that region is assumed as unimportant. This process is repeated for specific N x N sliding windows, a saliency map is obtained for the most significant regions in iris recognition each time they are moved with a phase of S. Most regions within the iris area are referred to as essential in this model when performing iris recognition.

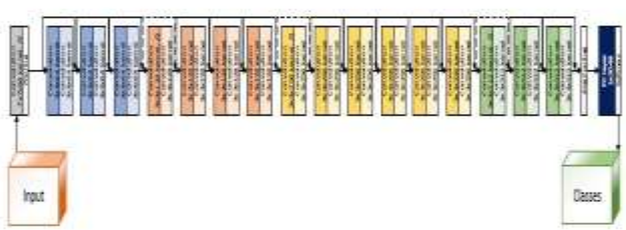

Fig. 2 Architecture of ResNet50

C. DFCN Model

The workflow of the model is as follows: First, CNN architecture is built which is then combined with dense iris segmentation blocks, referred to as a Dense-Fully Convolutional Network (DFCN), and adopts some common optimizer methods, such as Batch Normalization (BN) and dropout. Second, since the public ground-truth masks of IITD iris database do not include the designated eyelash regions, these regions are labelled using the Labelme software package as occluding the iris regions.

The architecture of DFCN is shown at Fig. 3. The DFCN encoder consists of dense blocks, and the decoder obtains the output prediction masks by transpose convolution. The cross entropy function is implemented as cost function at the end of the Dense-decoder. The Adam mini-batch algorithm [21] is used for optimization to reduce cost function.

The iris images consist mostly of non-iris regions and a few iris regions, leading to the adoption of the standard Stochastic Gradient Descent (SGD) algorithm, which obtains many non-iris features and ignores the iris features. That is, the iris features are sparse, and the non-iris features are general. The Adam algorithm, however, raises the learning rate of sparse data and decreases it for common data, and it rapidly updates for sparse features and slowly for common features. Notably, the Dense-decoder does not follow pretrained versions. It is necessary to remember that the DFCN is trained from scratch.

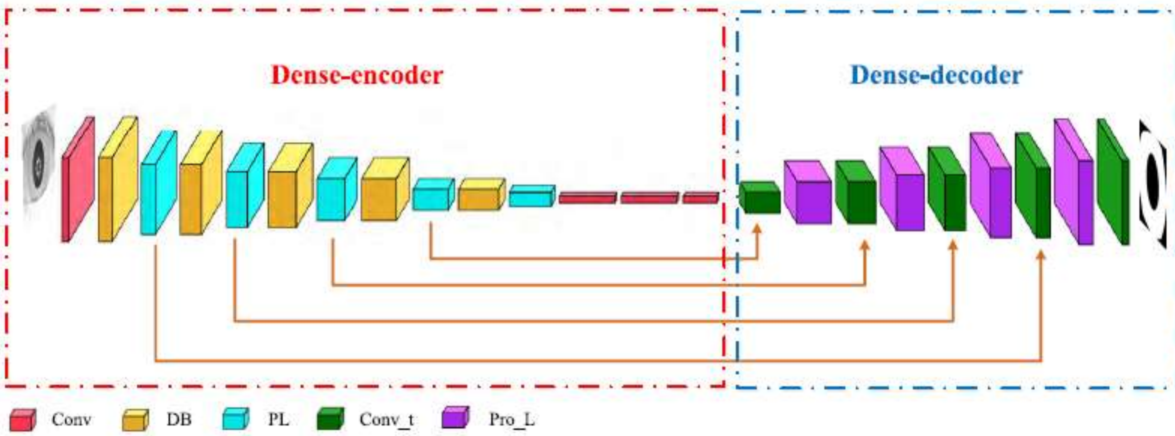

Fig. 3 Architecture of DFCN

D. Comparison of Three Models

All the three models have their own pros and cons. It is summarized in Table 1.

Table 1 Comparison of three Models

Methods	Strengths	Weaknesses
VGGNet	There is no pre-processing and architecture optimization.	The results could be further improved by training a deep network specifically for iris recognition
Residual CNN	(1) With very few training images per class, good accuracy is obtained. (2) A visualization technique is presented for detecting the most important regions while doing iris recognition.	The results are not promising.

DFCN	(1) No data augmentation methods and preprocessing methods are adopted. (2) Promising results are obtained due to dense connection	Dense connections are more time consuming. Comparing to other conventional iris segmentation methods, this method requires more training parameters.

IV. Experimental Results

All three models are tested on database of IIT Delhi (IITD). The VGGNet Model is also tested on CASIA-Iris-1000 besides the IITD database and the DFCN model is also tested on CASIA-Interval-v4 and UBIRIS.V2. In this paper CASIA Iris 1000 and IITD are used for comparison. The CASIA-Iris-1000 contains 20,000 iris images collected from 1,000 subjects using an IKEMB-100 camera [22]. In CASIA-Iris-1000 the key causes of intra-class differences are eyeglasses and specular reflections. The database of IIT Delhi includes 2240 pictures of iris taken from 224 different individuals. Such images are solved in 320x240 pixels [23]. All images are resized to 224x224 in VGGNet model to match VGG-Net inputs. Table 2 displays the features of each dataset. Sample images from CASIA 1000 and IITD dataset are shown in Fig. 4 and 5 respectively.

Table 2 Properties of Datasets

Property	CASIA 1000	IITD
Number of Classes	1000	224
Number of Images	20000	2240
Image Size	640 x 480	320 x 240
Image Format	JPG	BMP
Illumination Environment	NIR	NIR

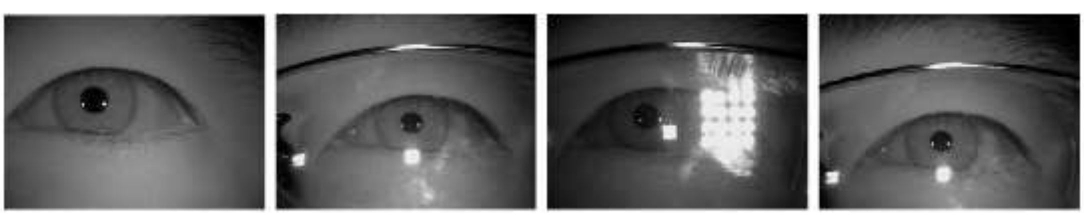

Fig. 4 Sample images from CASIA 1000 dataset

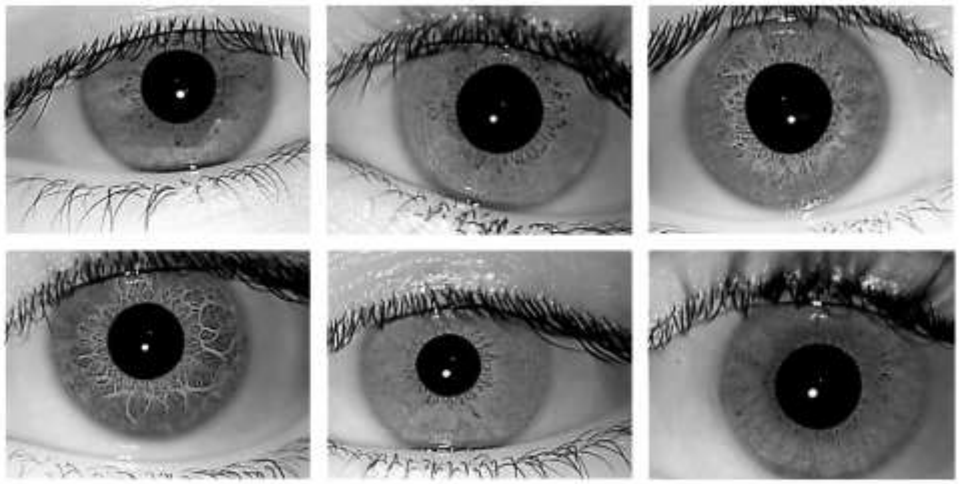

Fig. 5 Sample images from IITD Dataset

The data collection has to be broken up for training and testing to validate the models. Table 3 includes both the data set details after breaking in each model for training and testing.

In the VGGNet model, the number of training samples ranges from 1 to 5 (out of 10 samples) for each individual and the accuracy of the recognition was found. From the results, the accuracy of recognition by using 3 samples out of 10 is found to have a great benefit over using 1 or 2 samples, and it remains fairly constant by increasing the number of training samples.

Table 3 The characteristics of the adopted iris image databases after splitting.

Model	CASIA 1000		IITD	
	Number of training sets	**Number of test sets**	**Number of training sets**	**Number of test sets**
VGGNet	10000	10000	672	1568
Residual CNN	18000	2000	1344	896
DFCN	14000	6000	1792	448

The hyper-parameters used in the training process and the experimental setting for the execution of each model is presented prior to reporting the outcome of each test. Table 4 summarizes the hyper-parameters and the execution environment adopted for each model.

In the DFCN model, the learning rate decreases every 10000 steps by a factor of 10. The weights of all convolutional kernels are initialized as a truncated normal distribution with a standard deviation of 0.01, and the bias of all convolutional layers is 0. The values of $\boldsymbol{\mu}$ and v of the mini-batch Adam algorithm are set to 0.9 and 0.99, respectively

Table 4 Experimental Setup of System and Hyper-parameters of Neural Network Models

Model	**Execution Environment**	**Hyper-Parameters**
VGGNet	Core i5 CPU running at 2.2GHz.	-
Residual CNN	Nvidia Tesla GPU.	• Epochs - 100 • Batch size - 24 • Optimization - Adam optimizer is used to optimize the loss function • Learning rate - 0.0002
DFCN	Inter Core i9-7900x CPU with 32 GB memory and an NVIDIA 1080ti GPU with 11 GB memory	• Learning rate - 0.001. • Iterative step – 30000 • Batch size - 2 • Optimization - Adam optimizer

The accuracy represents the correct segmentation pixels and is calculated as follows:

$$\mathrm{acc_r} = \frac{\mathrm{T_p+T_n}}{\mathrm{T_p+T_n+F_p+F_n}} \times 100 \qquad (1)$$

where $\mathrm{acc_r}$ is the accuracy rate, $\mathrm{T_p, T_n, F_p, F_n}$ are the True Positive, True Negative, False Positive and False Negative rates. Table 5 shows the experimental results of all the three models. Apart from accuracy, the DFCN model also calculates Nice score, Precision, Recall and F1 score. For comparison purpose, only accuracy is considered in this paper.

The accuracy of recognition of the features is evaluated from the fc6 layer in the VGGNet model. First PCA is applied to all features, and the accuracy of recognition is evaluated for various number of PCA features. Surprisingly, a very high precision rate is obtained even when using few PCA features. It is observed that using 100 PCA features, an accuracy rate of 98% is obtained for IIT database, which only increases around 1% by using more PCA features.

In another experiment of VGGNet model, deep-function output is derived from different layers of VGG-Net. The number of features from each layer is limited to 256 to make a fair comparison (by taking first 256 PCA, if that layer has more than 256 output filters). An accuracy rate of over 98% can be achieved by removing features from every layer after the 7th. The accuracy of identification hits its height by using 10th layer, and then decreases. One potential explanation may be that they start collecting more abstract and high-level details by moving into the higher layers of the deep network, which doesn't distinguish much between different iris patterns, while the mid-level features in the previous layers have more discriminating power for similar recognition.

Table 5 Accuracy Comparison of three Models

	Accuracy (%)	
Method/Database	**CASIA 1000**	**IITD**
VGGNet Model	90	98.84
Residual CNN	92.4	95.5
DFCN	97.6	99.4

Among the three models, the DFCN Model achieves a higher accuracy than other models. But this model consumes more time than other methods.

References:

[1] A. Pando, "Beyond Security: Biometrics Integration into Everyday Life," forbes. [Online].
Available: https://www.forbes.com/sites/forbestechcouncil/2017/08/04/beyond-securitybiometrics- integration-into-everyday-life/#ce400be431fb.

[2] S. Thavalengal and P. Corcoran, "User Authentication on Smartphones: Focusing on iris biometrics.," IEEE Consum. Electron. Mag., vol. 5, no. 2, pp. 87–93, 2016.

[3] P. Corcoran and C. Costache, "Smartphones, Biometrics, and a Brave New World," IEEE Technol. Soc. Mag., vol. 35, no. 3, pp. 59–66, 2016.

[4] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: security and privacy concerns," IEEE Secur. Priv. Mag., vol. 1, no. 2, pp. 33–42, 2003.

[5] K. Irsch and D. L. Johns H. M. Guyton, "Anatomy of Eyes," Encycl. Biometrics, no. April, pp. 1212–1217, 2009.

[6] J. G. Daugman, "Biometric personal identification system based on iris analysis." Google Patents, Mar-1994.

[7] J. R. Matey et al., "Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments," Proc. IEEE, vol. 94, 2006.

[8] Hu, Y., Sirlantzis, K. and Howells, G., 2016. Optimal generation of iris codes for iris recognition. IEEE Transactions on Information Forensics and Security, 12(1), pp.157-171.

[9] Zhang, W., Lu, X., Gu, Y., Liu, Y., Meng, X. and Li, J., 2019. A Robust Iris Segmentation Scheme Based on Improved U-Net. IEEE Access, 7, pp.85082-85089.

[10] Wu, X. and Zhao, L., 2019. Study on iris segmentation algorithm based on dense U-Net. IEEE Access, 7, pp.123959-123968.

[11] Chen, Y., Wu, C. and Wang, Y., 2020. T-Center: A Novel Feature Extraction Approach Towards Large-Scale Iris Recognition. IEEE Access, 8, pp.32365-32375.
[12] Lozej, J., Štepec, D., Štruc, V. and Peer, P., 2019. Influence of segmentation on deep iris recognition performance. arXiv preprint arXiv:1901.10431.
[13] Wang, K. and Kumar, A., 2019. Segmentation-Aware and Adaptive Iris Recognition. arXiv preprint arXiv:2001.00989.
[14] Chakraborty, M., Roy, M., Biswas, P.K. and Mitra, P., 2020. Unsupervised Pre-trained, Texture Aware and Lightweight Model for Deep Learning-Based Iris Recognition Under Limited Annotated Data. arXiv preprint arXiv:2002.09048.
[15] Miyazawa, K., Ito, K., Aoki, T., Kobayashi, K. and Nakajima, H., 2006, December. An implementation-oriented iris recognition algorithm using phase-based image matching. In 2006 International Symposium on Intelligent Signal Processing and Communications (pp. 231-234). IEEE.
[16] Minaee, S., Abdolrashidiy, A. and Wang, Y., 2016, December. An experimental study of deep convolutional features for iris recognition. In 2016 IEEE signal processing in medicine and biology symposium (SPMB) (pp. 1-6). IEEE.
[17] Minaee, S. and Abdolrashidi, A., 2019. Deepiris: Iris recognition using a deep learning approach. arXiv preprint arXiv:1907.09380.
[18] Chen, Y., Wang, W., Zeng, Z. and Wang, Y., 2019. An adaptive CNNs technology for robust iris segmentation. IEEE Access, 7, pp.64517-64532.
[19] He, Kaiming, et al. “Deep residual learning for image recognition.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[20] M Zeiler, R Fergus. “Visualizing and understanding convolutional networks.” European conference on computer vision, springer, Cham, 2014.
[21] D. P. Kingma, J. Ba, “Adam: A method for stochastic optimization," Computer Science, 2014.
[22] http://biometrics.idealtest.org/
[23] A Kumar and A Passi,” Comparison and combination of iris matchers for reliable personal authentication,” Pattern Recognition, pp. 1016-1026, 2010.