International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

MULTIMODAL BIOMETRIC IDENTIFICATION SYSTEM USING THE FUSION OF FINGERPRINT AND IRIS RECOGNITION WITH CNN APPROACH

Prof. Ramya K M
Department of Computer Science and Engineering
Dayananda Sagar College of Engineering, Bengaluru

Pavan H
Department of Computer Science and Engineering
Dayananda Sagar College of Engineering, Bengaluru

Darshan Gowda
Department of Computer Science and Engineering
Dayananda Sagar College of Engineering, Bengaluru

Bhagavantray Hosamani
Department of Computer Science and Engineering
Dayananda Sagar College of Engineering, Bengaluru

Jagadeva A S
Department of Computer Science and Engineering
Dayananda Sagar College of Engineering, Bengaluru

Abstract -- **Multimodal biometric systems are widely applied in many real-world applications because of its ability to accommodate variety of great limitations of unimodal biometric systems, including sensitivity to noise, population coverage, intra-class variability, non-universality, and vulnerability to spoofing. during this paper, an efficient and real-time multimodal biometric system is proposed supported building deep learning representations for images of both the correct and left irises of someone, and fusing the results obtained employing a ranking-level fusion method. The trained deep learning system proposed is named IrisConvNet whose architecture relies on a mix of Convolutional Neural Network (CNN) and Softmax classifier to extract discriminative features from the input image with none domain knowledge where the input image represents the localized iris region and so classify it into one amongst N classes. during this work, a discriminative CNN training scheme supported a mixture of back-propagation algorithm and mini-batch AdaGrad optimization method is proposed for weights updating and learning rate adaptation, respectively. additionally, other training strategies (e.g., dropout method, data augmentation) also are proposed so as to gauge different CNN architectures. The performance of the proposed system is tested on three public datasets collected under different conditions: SDUMLA-HMT, CASIA-IrisV3 Interval and IITD iris database**

Keywords- **Iris recognition, Multimodal biometricm systems, Deep learning, Convolutional Neural Network, Softmax classifier, Ada Grad method)**

I. INTRODUCTION

Biometric systems are constantly evolving and promise technologies that may be employed in automatic systems for identifying and/or authenticating a person's identity uniquely and efficiently without the necessity for the user to hold or remember anything, unlike traditional methods like passwords, IDs [1, 2]. during this regard, iris recognition has been utilized in many critical applications, like access control in restricted areas, database access, national ID cards, and financial services and is taken into account one in all the foremost reliable and accurate biometric systems [3,

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

4]. Several studies have demonstrated that the iris trait features a number of benefits over other biometric traits (e.g., face, fingerprint), which make it commonly accepted for application in high reliability and accurate biometric systems. Firstly, the iris trait represents the annular region of the attention lying between the black pupil and therefore the white sclera; this makes it completely shielded from varied environmental conditions. during this paper, two discriminative learning techniques are proposed supported the mixture of a Convolutional Neural Network (CNN) and therefore the Softmax classifier as a multinomial logistic regression classifier. CNNs are efficient and powerful Deep Neural Networks (DNNs) which are widely applied in image processing and pattern recognition with the power to automatically extract distinctive features from input images even without a preprocessing step

II. OVERVIEW OF THE PROPOSED APPROACHES

In this section, a short description of the proposed deep learning approach is given, which includes two discriminative learning techniques: a CNN and a Softmax classifier. the most aim here is to examine their internal structures and identify their strengths and weaknesses to enable the proposal of an iris recognition system that integrates the strengths of those two techniques

2.1 Convolutional Neural Network

The CNN architecture, as illustrated in Fig. 1, comprises several distinct layers including sets of locally connected convolutional layers (with a particular number of various learnable kernels in each layer), subsampling layers named pooling layers, and one or more fully connected layers. the interior structure of the CNN combines three architectural concepts, which make the CNN successful in several fields, like image processing and pattern recognition, speech recognition, and NLP.

Convolutional layer during this layer, the parameters (weights) incorporate a group of learnable kernels that are randomly generated and learned by the back-propagation algorithm. The results of each kernel convolved across the full input image is named the activation (or feature) map, and also the number of the activation maps is up to the quantity of applied kernels in this layer. Figure 1 shows a primary convolution layer consisting of 6 activation maps stacked together and produced from 6 kernels independently convolved across the full input image.

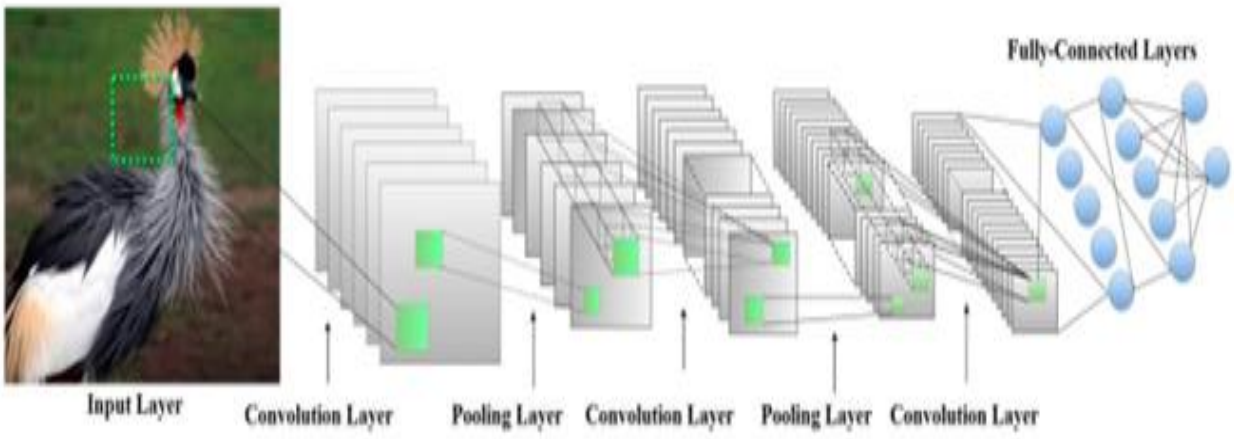

Fig. 1 An illustration of the CNN architecture, where the gray and green squares refer to the activation maps and the learnable convolution kernels, respectively. The cross-lines between the last two layers refer to the fully connected neurons (color figure online)

$$Y^{j(r)} = \max(0, b^{j(r)} + \Sigma_i k^{ij(r)} * x^{j(r)})$$

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

$$h_\theta(x_i) = \begin{bmatrix} p(y_i = 1|x_i;\theta) \\ p(y_i = 2|x_i;\theta) \\ . \\ . \\ . \\ p(y_i = K|x_i;\theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{K} e^{\theta_j^T x_i}} \begin{bmatrix} e^{\theta_1^T x_i} \\ e^{\theta_2^T x_i} \\ . \\ . \\ . \\ e^{\theta_K^T x_i} \end{bmatrix}$$

Here, $x^{i(r)}$ and $y^{j(r)}$ are the ith input and therefore the jth output activation map, respectively. $b^{j(r)}$ is that the bias of the jth output map and * denotes convolution. $k^{ij(r)}$ is that the kernel between the ith input map and the jth output map. The ReLU activation function (y = max (0,x)) is used here to add non-linearity to the network, as will be explained later on.

- Pooling layer Its main function is to scale back the spatial size of the convolutional layers output representations, and it produces a limited kind of the translational invariance. Once a selected feature has been detected by the convolutional layer, only its approximate location relative to other features is kept. As shown in Fig. 1, each depth slice of the input volume (convolutional layer's output) is split into non-overlapping regions, and for every subregion the most value is taken.

$$y^{(l)}(j) = f^{(l)}\left(\sum_{i=1}^{N^{(l-1)}} y^{(l-1)}(i).w^{(l)}(i,j) + b^{(l)}(j)\right) (3)$$

- Fully connected layers the output of the last convolutional or max-pooling layer is fed to a 1 or more fully connected layers as in a very traditional neural network. In those layers, the outputs of all neurons in layer (l − 1) are fully connected to each neuron in layer l. The output y(l) (j) of neuron j in a very fully connected layer l.
- where N(l−1) is that the number of neurons within the previous layer (l-1), w(l) (i, j) is that the weight for the connection from neuron j in layer (l − 1) to neuron j in layer l, and b(l) (j) is that the bias of neuron j in layer l. As for the opposite two layers, f(l) represents the activation function of layer l.

2.2 Softmax regression classifier

Is the Softmax regression classifier, which could be a generalized kind of binary logistic regression classifier intended to handle multi-class classification tasks. Suppose that there are K classes and n labeled training samples, where xi ∈ Rm is that the ith training example and yi∈ is that the class label of xi .

Then, for a given test input xi , the Softmax classifier will produce a K-dimensional vector (whose elements sum to 1), where each element within the output vector refers to the

estimated probability of each class label conditioned on this input feature.
Here, (θ1, θ2, …, θK) are the parameters to be randomly generated and learned by the back-propagation algorithm. The cost function used for the Softmax classifier is named as cross-entropy loss function and can be defined as follows:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{K} 1\{y_i = j\} log \frac{e_{j^{x_i}}^T}{\sum_{l=1}^{K} e_{j^{x_i}}^T}\right] + \frac{\lambda}{2}\sum_{i=1}^{K}\sum_{j=0}^{n}\theta_{ij}^2$$

III. THE PROPOSED SYSTEM

An overview of the proposed iris recognition system is shown in Fig. 2. Firstly, a preprocessing procedure is implemented supported employing an efficient and automatic iris localization to carefully detect the iris region from the rear ground and every one extraneous features, like pupil, sclera, eyelids, eyelashes, and specular reflections. during this work, the most reason for outlining the iris area because the input to CNN rather than the entire eye image is to scale back the computational complexity of the CNN. one more reason is to avoid the performance degradation of the matching and have extraction processes resulting from the looks of eyelids and eyelashes. After detection, the iris region is transformed into a normalized form with fixed dimensions so as to permit direct comparison between two iris images with initially different sizes.

The normalized iris image is further accustomed provide robust and distinctive iris features by employing the CNN as an automatic feature extractor. Then, the matching score is obtained using the generated feature vectors from the last fully connected layer because the input to the Softmax classifier. Finally, the matching voluminous both the proper and left iris images are fused to ascertain the identity of the person whose iris images are under investigation. During the training phase, different CNN configurations are trained on the training set and tested on the validation set to get the simplest one with the tiniest error that we call IrisConvNet.

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

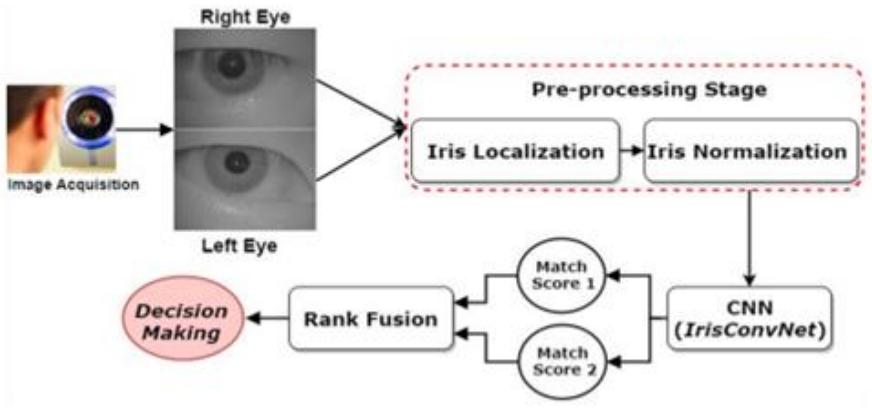

Fig. 2 An overview of the proposed multi-biometric iris recognition system

3.1 Iris localization

Precise localization of the iris region plays a very important role in improving the accuracy and reliability of an iris recognition system, because the performance of the subsequent stages of the system directly depends on the standard of the detected iris region. The iris localization procedure aims to detect the 2 iris region boundaries: the inner (pupil–iris) boundary and therefore the outer (iris–sclera) boundary. However, the task becomes tougher, when parts of the iris are covered by eyelids and eyelashes. additionally, changes within the lighting conditions during the acquisition process can affect the standard of the extracted iris region then affect the iris localization and also the recognition outcome. during this section, a short description of our iris localization procedure [39] is given where an efficient and automatic algorithm is proposed for detecting the inner and outer iris boundaries. As depicted in Fig. 3, firstly, a refection mask is calculated after the detection of all the specular refection spots within the eye image, to assist their removal. Then, these detected spots are painted employing a pre-defined refection mask and a Roifill MATLAB function. Next, the inner and outer boundaries are detected figures and tables after they're cited within the text. Use the abbreviation "Fig.1", even at the start of a sentence.

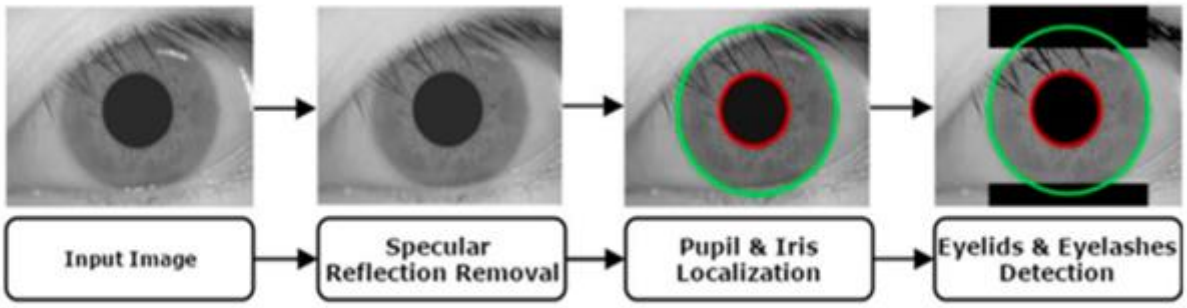

Fig. 3 Overall stages of the proposed iris localization procedure

Operations so as to scale back the computational complexity of the Circular Hough Transform (CHT), smooth the attention image and to boost the contrast between the iris and sclera region. this is often followed by applying a coherent CHT to get the middle coordinates and radius of the pupil and iris circles. Finally, the upper and lower eyelids boundaries are detected employing a fast and accurate eyelid detection algorithm, which employs an anisotropic diffusion filter with Radon transform to ft them as straight lines. For further details on the iris localization procedure.

3.2 Iris normalization

Once, the iris boundaries are detected, iris normalization is implemented to supply a set dimension feature vector that

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

permits comparison between two different iris images. the most advantage of the iris normalization process is to get rid of the dimensional inconsistencies which will occur thanks to stretching of the iris region caused by pupil dilation with varying levels of illumination. Other causes of dimensional inconsistencies include, changing imaging distance, elastic distortion within the iris texture which will affect the iris matching outcome, rotation of the camera or eye so forth.
The iris matching outcome, rotation of the camera or eye and then forth. to deal with of these mentioned issues the iris normalization process is applied using Doughman's rubber sheet mapping to remodel the iris image from Cartesian coordinates to polar coordinates, as shown in Fig. 4. Doughman's mapping takes each point (x, y) within the iris region to a pair of normalized non-concentric polar coordinates (r, θ) where r is on the interval [0, 1] and θ is that the angle on the interval [0, 2]. This mapping of the iris region will be defined mathematically as follows:

$$I(x(r,\theta), y(r,\theta)) \rightarrow I(r,\theta)$$
$$x(r,\theta) = (1-r)x_p(\theta) r x_l(\theta)$$
$$y(r,\theta) = (1-r)y_p(\theta) r y_l(\theta)$$

Here I(x, y) is the intensity value at (x, y) in the iris region image. The parameters xp, xl , yp, and yl are the coordinates of the pupil and iris boundaries along the Q direction.

3.3 Deep learning for iris recognition

Once a normalized iris image is obtained, feature extraction and classification is performed employing a deep learning approach that mixes a CNN and a Softmax classifier. during this work, the structure of the proposed CNN involves a mix of convolutional layers and subsampling maxpooling. the highest layers within the proposed CNN are two fully connected layers for the classification task. Then, the output of the last fully connected layer is fed into the Softmax classifier, which produces a probability distribution over the N class labels. Finally, a cross-entropy loss function, an acceptable loss function for the classification task, is employed to quantify the agreement between the anticipated class scores and therefore the target labels and calculate the price value for various configurations of CNN. during this section, the proposed methodology for locating the simplest CNN configuration to be used for the iris recognition task is explained. supported domain knowledge from the literature, there are three main aspects that have an excellent inference on the performance of a CNN, which need to be investigated. These include: (1) training methodology, (2) network configuration or architecture (3) input image size. The performance of some carefully proposed training strategies, including the dropout method, AdaGrad method, and data augmentation, is investigated as a part of this work. These training strategies have a big role in preventing the overfitting problem during the training process and increasing the generalization ability of the neural network. These three aspects are described in additional detail within the next section.

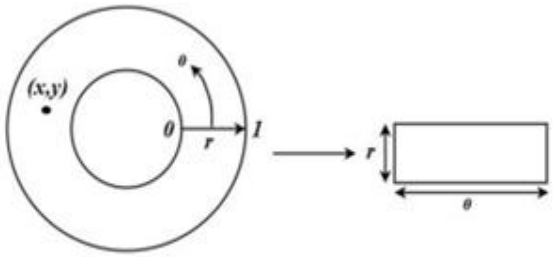

Fig. 4 Daugman's rubber sheet model to transfer the iris region from the Cartesian coordinates to the polar coordinates

3.4 Training methodology

In this work, all of the experiments were disbursed, given a selected set of sample data, using 60% randomly selected samples for training and also the remaining 40% for testing. The training methodology as in [40, 41], starts training a selected CNN configuration by dividing the training set into four sets after the information augmentation procedure is implemented: three sets are accustomed train the CNN and also the last one is employed as a validation set for testing the generalization ability of the network during the training process and storing the weights configuration that performs best thereon with minimum validation error, as shown in Fig. 5. during this work, the training procedure is performed using the back-propagation algorithm with the mini-batch AdaGrad optimization method introduced in [42], where each set of the three-training data is split into mini-batches and therefore the training errors are calculated upon each mini-batch within the Softmax layer and acquire back-propagated to the lower layers.

After each epoch (passing through the complete training samples), the validation set is employed to live the accuracy of the present configuration by calculating the price value and therefore the Top-1 validation error rate. Then, in line with the AdaGrad optimization method, the training rate is scaled by an element adequate the root of the sum of squares of the previous gradients as shown in Eq. 8. An initial learning rate must be selected; hence, two of the foremost common used learning rate values are analyzed herein, as shown in (Sect. 5.2.1). To avoid the overfitting problem, the training procedure is stopped as soon because the cost value and therefore the error on the validation set start to rise again, which suggests that the network starts to overfit the training set. This process is one in all the regularization methods called the first stop ping procedure. during this work, different numbers of epochs are investigated as explained in (Sect. 5.2.1). Finally, after the

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

training procedure is finished, the testing set is employed to live the efficiency of the ultimate configuration obtained in predicting the unseen samples by calculating the identification rate at Rank-1 as an optimization objective, which is maximized during the training process. Then, the Cumulative Match Characteristic (CMC) curve is employed to visualize the performance of the simplest configuration obtained because the iris identification system. the most steps of the proposed training methodology are summarized as follows:

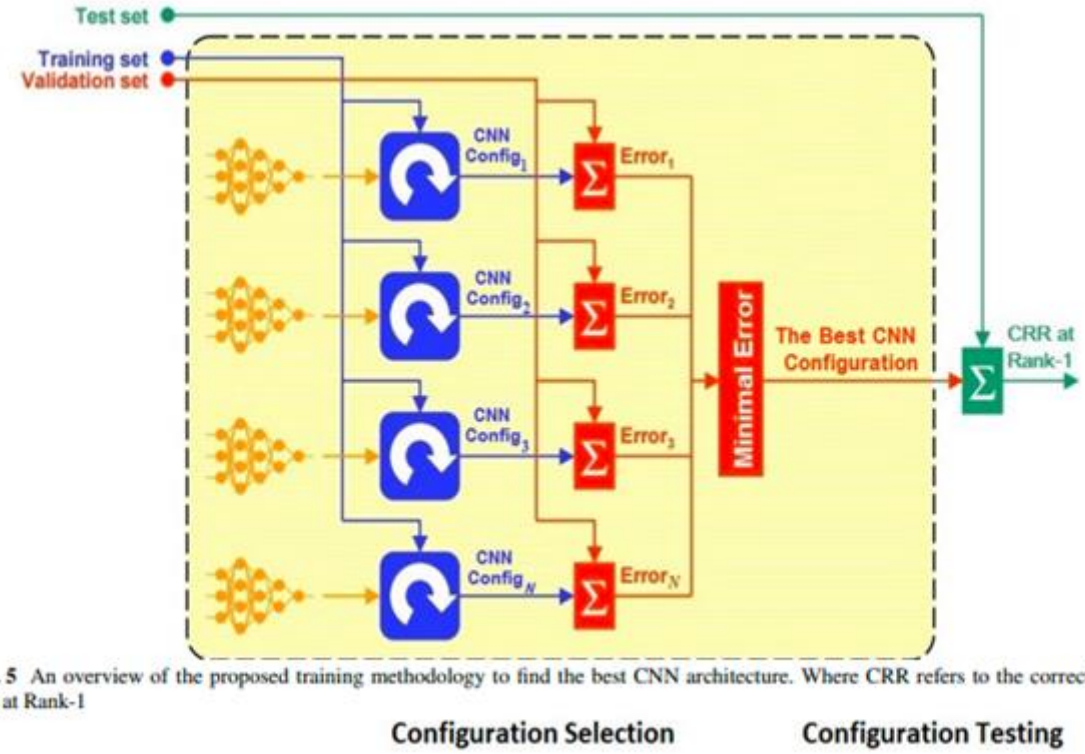

Fig. 5 An overview of the proposed training methodology to find the best CNN architecture. Where CRR refers to the correction recognition rate at Rank-1

1. Split the dataset into three sets: Training, Validation and Test set.
2. Select a CNN architecture and a collection of coaching parameters.
3. Train each CNN configuration using the training set.
4. Evaluate each CNN configuration using the validation set.
5. Repeat steps 3 through 4 using N epochs.
6. Select the most effective CNN configuration with minimal error on the validation set.
7. Evaluate the most effective CNN configuration using the test set.

3.5 Network architecture

Once the parameters of the training methodology are determined (e.g., learning rate, number of epochs, etc.), it's accustomed identify the simplest specification. From the literature, it appears that choosing the specification continues to be an open problem and is application dependent. the most concern find the most effective CNN architecture is that the number of the layers to use transforming from the input image to a high-level feature representation, together with the quantity of convolution filters in each layer.

3.6 Input image size

The input image size is one in every of the hyper-parameters within the CNN that features a significant influence within the speed and also the accuracy of the neural network. during this work, the influence of input image size is investigated using the sizes (64 × 64) pixels and (128 × 128) pixels (generated from original images of larger size as described within the Data Augmentation section below), on condition that for lower values than the previous, the iris.

IV. EXPERIMENTAL RESULTS

During this database, all images were captured using an intelligent iris capture device with the gap from the device to the attention between 6 cm and 32 cm. To the most effective of our knowledge, this can be the primary work

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 6, Issue 8, ISSN No. 2455-2143, Pages 213-220
Published Online December 2021 in IJEAST (http://www.ijeast.com)

that uses all the themes during this database for the identification task. The CASIA-Iris-V3 Interval database comprises 2566 images from 249 subjects, which were captured with a self-developed close-up iris camera. during this database, the amount of images of every subject differs and 129 subjects have but 14 iris images. These weren't employed in the experiments.

Iris localization accuracy :

$$Accurcy\ Rate = \frac{Correctly\ Localized\ Iris\ Images}{Total\ Number} \times 100$$

4.1 Experimental results Snaps

Table 1 The characteristics of the adopted iris image databases

Property	SDUMLA-HMT	CASIA-Iris-V3	IITD
Number of classes	106	120	224
Samples per subject	5 right and 5 left	7 right and 7 left	5 right and 5 left
Number of images	1060 images	1680 images	2240 images
Image size	(768 × 576) pixels	(320 × 280) pixels	(320 × 240) pixels
Image format	BMP	JPEG	BMP

Table 2 Comparison of the proposed iris localization model with previous approaches

Approach	CASIA-Iris-V3		IITD	
	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)
Jan et al. [52]	99.50	7.75	99.40	8.52
Wang et al. [53]	96.95	165.4	96.07	145.4
Mahmoud and Ali [54]	99.18	–	–	–
Uhl et al. [55]	74.00	**0.21**	–	–
Ugbaga et al. [56]	98.90	–	–	–
Umer et al. [57]	95.87	0.89	98.48	0.77
Wild et al. [58]	98.13	–	97.60	–
Aydi et al. [59]	96.51	9.049	–	–
Pawar et al. [60]	96.88	–	–	–
Mehrotra et al. [61]	99.55	0.396	–	–
Proposed iris localization	**99.82**	0.62	**99.87**	**0.51**

Bold values indicate the highest obtained recognition rates

V. REFERENCES

[1]. Hajari K (2015) Improving iris recognition performance using local binary pattern and combined RBFNN. Int J Eng Adv Technol 4(4):108–112
[2]. Lim S, Lee K, Byeon O, Kim T (2001) Efcient iris recognition through improvement of feature vector and classifer. ETRI J 23(2):61–70
[3]. Tan T, Sun Z (2009) Ordinal measures for iris recognition. IEEE Trans Pattern Anal Mach Intell 31(12):2211–2226
[4]. Abiyev RH, Kilic KI (2011) Robust feature extraction and iris recognition for biometric personal identifcation. Biometric Syst Des Appl, InTech
[5]. Hentati R, Hentati M, Abid M (2012) Development a new algorithm for iris biometric recognition. Int J Comput Commun Eng 1(3):283–286
[6]. Das A, Parekh R (2012) Iris recognition using a scalar based template in eigen-space. Int J Comput Sci Telecommun 3(5):3–8
[7]. AlMahafzah H, Zaid AlRwashdeh M (2012) A survey of multibiometric systems. Int J Comput Appl 43(15):36–43
[8]. Gad R, EL-SAYED A, Zorkany M, El-fshawy N (2015) Multibiometric systems: a state of the art survey and research directions. Int J Adv Comput Sci Appl 6(6):128–138
[9]. Ross A, Nandakumar K, Anil JK (2006) Handbook of multibiometric. J Chem Inf Model 53(9):1689–1699
[10]. Fernandez FA (2008) Biometric sample quality and its application to multimodal authentication systems. PhD Thesis, Universidad Polit´ ecnica de Madrid (UPM)
[11]. Deng L, Yu D (2013) Deep learning methods and applications. Signal Process 28(3):198–387

[12]. Pellegrini T (2015) Comparing SVM, Softmax, and shallow neural networks for eating condition classifcation. In: Sixteenth annual conference of the international speech communication association. pp 899–903
[13]. Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overftting. J Mach Learn Res 15:1929–1958
[14]. Krizhevsky A, Sulskever I, Hinton GE (2012) ImageNet classifcation with deep convolutional neural networks. In: Advances in neural information processing system. pp 1–9
[15]. Menotti D, Chiachia G, Pinto A, Schwartz WR, Pedrini H, Falcao AX, Rocha A (2015) Deep representations for iris, face, and fngerprint spoofng detection. IEEE Trans Inf Forensics Secur 10(4):864–879
[16]. Silva P, Luz E, Baeta R, Pedrini H, Falcao AX, Menotti D (2015) An Approach to iris contact lens detection based on deep image representations. In: 2015 28th SIBGRAPI conference on graphics, patterns images. pp 157–164
[17]. Daugman JG (1993) High confdence visual recognition of persons by a test of statistical independence. IEEE Trans Pattern Anal Mach Intell 15(11):1148–1161