TRANSFER LEARNING WITH CONVOLUTIONAL NEURAL NETWORKS FOR IRIS RECOGNITION

Maram.G Alaslni[1] and Lamiaa A. Elrefaei[1, 2]

[1]Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

[2]Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt

ABSTRACT

Iris is one of the common biometrics used for identity authentication. It has the potential to recognize persons with a high degree of assurance. Extracting effective features is the most important stage in the iris recognition system. Different features have been used to perform iris recognition system. A lot of them are based on hand-crafted features designed by biometrics experts. According to the achievement of deep learning in object recognition problems, the features learned by the Convolutional Neural Network (CNN) have gained great attention to be used in the iris recognition system. In this paper, we proposed an effective iris recognition system by using transfer learning with Convolutional Neural Networks. The proposed system is implemented by fine-tuning a pre-trained convolutional neural network (VGG-16) for features extracting and classification. The performance of the iris recognition system is tested on four public databases IITD, iris databases CASIA-Iris-V1, CASIA-Iris-thousand and, CASIA-Iris-Interval. The results show that the proposed system is achieved a very high accuracy rate.

KEYWORDS

biometrics, iris, recognition, deep learning, convolutional neural network (CNN), transfer learning.

1. INTRODUCTION

Today, the concept of personal identity becomes critical, and the biometrics is a popular way for authentication, which has been considered as the most secure and hardest way for authentication purpose[1]. Biometric systems are developing technologies that can be used in automatic systems for identifying individuals uniquely and effectively and it becomes a good alternative to the traditional methods such as passwords. According to a study published in [1], users choose to use smartphone biometrics as an alternate for passwords because it provides additional safety for new technologies like Apple Pay [1].

Biometrics is an automated technique for authenticating an individual depends on a physical or behavioral characteristic. Physical characteristics like fingerprints, voice, face, and iris. Behavior characteristics are traits which can be learned or acquired such as, speaker verification, keystroke dynamics and dynamic signature verification [2].

Biometric methods based on the spatial pattern of the iris are believed to allow very high accuracy, and there has been an explosion of interest in iris biometrics in recent years [3]. The iris is a visible but protected structure. Also, iris does not significantly change over time, so it is perfect and reliable for identity authentication [4, 5].

Iris recognition system has main stages. First, segment the required region of the iris, after that it is normalized to be a fixed pattern in polar coordinates. Then, the features extracted from that pattern to be recognized at the last stage [6].

Extracting effective features is the major important stage in many object recognition and computer vision tasks. Therefore, several researchers have focused on designing robust features for a variety of image classification tasks [7].

Nowadays, much attention is given to feature learning algorithms and Convolutional Neural Networks (CCN). In this algorithm, the image is fed directly to the convolutional neural networks, then the algorithm extracts the best features of this image [7, 8].

The results in [9] strongly suggest that features extracted from deep learning with CNN have to be considered in most visual recognition tasks.

In this paper, we propose transfer learning with the pre-trained convolutional neural network (VGG-16) model for iris recognition system. In transfer learning, we pass learned "knowledge" from a CCN model trained on one task to do another task [10]. Used the transfer learning help to avoid overfitting a large network if there is not sufficient data to train it from scratch [11]. Also, transfer learning protects computational resources, because training from scratch possibly will take from days to weeks [12].

The proposed system performance is investigated when the features extracting directly from the segmented iris image. The experimental study is tested on four public datasets collected under different conditions IITD iris databases [13], CASIA-Iris-V1 [14], CASIA-Iris-thousand[15], and CASIA-Iris-Interval [16].

The rest of this paper is organized as follows: Section II provides a brief background on the basic idea of Convolutional Neural Networks. In section III, some related works are presented. Section IV introduces a description of the proposed iris recognition system. The experimental results and analysis are presented in section V. Section VI provide a comparison between the performance of the proposed iris recognition scheme and the other works. Finally, the Conclusion is given in section VII.

2. CONVOLUTIONAL NEURAL NETWORKS PRELIMINARIES

The Convolutional Neural Networks (CNN) is a network architecture for deep learning, and it can be to learn image feature representations automatically. Also, it outperformed many conventional hand-crafted feature techniques. CNN belongs to a specific category of Neural network methods [17].

The Neural network (NN) is a machine learning model that has a hierarchical representation of data. NN accepts the image as a single vector then transforms the image over a sequence of hidden layers. Every hidden layer is consisting of a set of neurons, and every neuron is completely linked to very neurons in the preceding layer, and every neuron in a hidden layer fully independently also, do not share any connections. The output layer is the last fully-connected layer [18].

Convolutional Neural Network (CNN) consists of three types of layers. The layer can be either Convolutional, Pooling or fully connected [19]. For more details, the description of the layers is presented in Table 1.

CNN is usually structured in two parts. First part, named feature extraction, which using groupings of convolutional and pooling layers. Second part named classification which using fully connected layers[20].

Table 1. The description of CNN layers

Layer Name	Layer Description
Convolutional layer	This layer performed the convolution operation on the input image by using a set of filters, which is named kernel. The feature map is the output of the convolution operation[21].
Pooling layer	This layer used to decrease the size of the convolutional layer output while saving the most significant information contained in the input layer. The pooling operation can be (max or average), selecting of maximal value is the most common type of pooling operation[19].
Fully connected layers	These layers used the extracted features in the preceding layers to do the classification task [21].

Convolutional Neural Networks architecture has three properties that make it strong and effective in recognition tasks: local receptive fields, weights sharing, and downsampling operations [21]. In CNN only a small region of input layer neurons is connected to neurons in a hidden layer this region named the local receptive field. The local receptive has the same size of the convolution filter. The local receptive field is translated across an image to create a feature map. The local receptive field is used in convolutional and downsampling layers[21].

In NN each neuron has its weight, but in CNN the weight is the same in all neurons in a given layer. In CNN the weights sharing is applied to the convolutional layer to control the capacity and to decrease the complexity of the model. Finally, the nonlinear downsampling which used in the downsampling layers to decrease the spatial size of the image beside decrease the number of the free parameters of the model[21].

3. RELATED WORK

Several algorithms have been used for the iris recognition system, and we will review the most important algorithms in each iris recognition stage.

The main phases of the iris recognition system are image acquisition, image pre-processing, feature extraction, and classification [22]. The most important algorithms in every iris recognition phase will be discussed in this section
.

3.1. Image Acquisition

Image acquisition is the first phase that takes the iris images by using particularly developed cameras. Iris and pupil region must be clearly visible in these images. There are, different databases be on the internet for free and we can take the iris image from it, for the testing usage such as IITD iris databases [13], CASIA-Iris-V1 [14], CASIA-Iris-thousand [15], and CASIA-Iris-Interval [16].

3.2. Image preprocessing

The preprocessing stage is required for the iris image to get a useful iris region. Image preprocessing is included two phases: iris segmentation and iris normalization [22].

3.2.1. Iris segmentation

It is essential to separate the iris part from the other parts of the eye image also to find outer and inner boundaries of the iris[23]. Iris segmentation is significant because accurate iris part is required to create the templates for correct classification [22]. Several algorithms have been used for iris segmentation and some of them are presented here:

Daugman [24, 25] which proposed the first model of an iris recognition system and now many iris recognition systems used the Daugman's model. This model used an integro-differential operator for finding the circular iris and pupil part, in addition to the upper and lower eyelids. This model calculates the partial derivative of the average intensity of circle points, on increasing the radius. However, this algorithm is perfect since it searches over the image domain for the global maximum. Also, because it uses the first derivative information, it can compute faster. However, this model has some errors if there is a noise in the iris image, likes reflections also; it works only on a local scale.

Hough transform which used in [26] the inner and outer boundaries of an iris can be demonstrated as circles, circular Hough transform is used to detect the iris. Initially, an edge detector is performed to a grayscale iris image to produce the edge map. The edge map is produced by computing the first derivative of intensity values and thresholding the results. The Gaussian filter is performed to smooth the image to choose the proper scale of edge analysis. The Hough transform algorithm is not appropriate for real-time applications because it is computationally intensive. Also, it needs a threshold value to produce the edge map. The carefully chosen threshold value could eliminate some important edge points that lead to false circle detection.

Canny edge detection method which used in [27] it used to detect the iris and pupil boundaries of the iris image. Which offers the effective edges of the iris image. Therefore, it helps to find the correct pupil edge to detect the image. The algorithm contains five stages: They are Smoothing, Finding, Gradients, Non-maximum suppression, Double thresholding, Edge tracking by hysteresis.

In [2, 28-30] the authors proposed iris recognition system that involves the use of both a *Canny Edge Detection and a Circular Hough Transform*, to detect the iris and pupil boundaries in the iris image. Moreover, according to [23] the performance of the Canny edge detector with Hough Transform is work very well compared to the different iris segmentation methods.

3.2.2 Iris Normalization

Eye image could be of different size because it captured from a different person and different environment. Therefore, for reaching more accurate recognition irises images need to normalize from different size to the same size. This process is done by remapping the iris region into a rectangular region [31].

Daugman's Rubber Sheet model is commonly used for normalization [24, 25, 32]. The rubber sheet algorithm remaps every pixel in the segmented iris area from the cartesian coordinates to polar coordinates. The non-concentric polar representation is normalized to the rectangular block with a fixed size. This algorithm accounts for pupil dilation, non-concentric pupil displacement, and imaging distance [22]. The image may need to enhancement before the feature extraction stage when the normalized pattern has non-uniform illumination or low contrast [22].

3.2.3 Feature Extraction

Feature extraction is the process of extracting highly discriminative, yet compact features from the images and it uses a texture analysis technique to extract features from the iris image [33]. These features extracted from the iris image for correct authentication purpose [23]. The iris features can be extracted through the two types of algorithms, hand-crafted algorithms, and feature learning algorithms and will be discussed some of them in this section.

Hand-Crafted Algorithms

There are many hand-crafted algorithms used for iris recognition and some of them are presented here:

2D Gabor filters which used to extract iris features in [24, 25, 33] Gabor filter's impulse response is defined by a harmonic function multiplied by a Gaussian function. Excellent localization is offered by this algorithm in both spatial and frequency domains. Using quadrature 2D Gabor wavelets every sample is demodulated to extract its phase information. The information of the phase is quantized into four quadrants in the complex plane. Every quadrant is denoted with two bits phase information. Consequently, every pixel in the normalized image is demodulated into two bits code in the template. The phase information is extracted since it offers important information inside the iris image. It does not base on extraneous factors, for example, illumination, imaging contrast, and camera gain.

Wavelet transform which used to extract iris features in [34, 35] and has a different type like Haar wavelet that used in [28]. In the wavelet transform, the iris part is decomposed into components have diverse resolutions. The wavelet transforms better than Fourier transforms because it has both space resolution and frequency resolution. The features are localized in both space and frequency domains with different window sizes.

Laplacian of Gaussian filters which used in [26] it decomposes the iris region to encode the feature. The filtered image is realized as a Laplacian pyramid. A cascade of Gaussian-like filters is performed to the image. The Laplacian pyramid is made with four levels to produce a compact biometric template. This method compresses information to get important information. The compressed information can be stored and processed successfully.

Radon Transform Thresholding (RTT) and Gradient based Isolation (GI) used in[36] the iris image is separated into blocks of pre-defined size, and for each of these blocks, Radon transform is applied. The statistical features of the thresholded peak lines are calculated for every block and kept in the feature gallery. When combine the enhanced Gradient image with the original image produces an edge image involving the finer edges existent in the Iris texture.

Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) used in[37]. The salient iris features extracted by using both DWT and DCT. The pre-processed image is fed to the feature extraction too. The DWT is divided data into deferent frequency component and studied every component with a resolution matched to its scale.

The intersecting cortical model which used in [38] in this technique after preprocessing, the iris image after enhancement put into ICM network, and the third output is used as the iris codes. The ICM network is a basic model of Pulse-Coupled Neural Network (PCNN) model that has good performance for image segmentation, and this will help to fast the process of coding. After this processing, different features of iris like ridges, crypts, a zigzag collarette, and furrows are segmented out completely.

Circular sector DCT which used in [39] the feature extraction done by using diverse geometrical shapes. They proposed using quadrant of a circle or a sector of a circle as a geometrical shape for extraction the feature. This geometry helps to decrease the number of features compared to the conventional square technique of extraction. Decreased amount of features help to decrease the computation which leads to a decrease in the time of training and testing.

Scattering transform-based features and textural features which used in [7] in this technique extracted the scattering features globally, whereas it extracted the textural features locally. Scattering features are locally invariant and have a great deal of high-frequency information. The high-frequency information offers more distinguished power for iris recognition.

Feature Learning Algorithms

The feature learning algorithm is based on using the pre-trained model CNN to extract the feature and then make the classification by the same pre-trained model or using the machine learning algorithm like SVM to make the classification task.

The pre-trained convolutional neural network (Inception V3) [12]. In this work, the authors proposed a system that discovers the problem of diabetic retinopathy in eye fluorescein angiography photographs. This system used the pre-trained convolutional neural network named Inception V3. The Inception V3 network trained on ImageNet dataset. The experiment dataset has been delivered by eyePacs as color fundus images on the scale of 0-5. The Inception V3 network for 5-class severity classification achieved very well with an accuracy of 48.2% that is more than the human evaluation.

The pertained convolutional neural network (VGG-Net) [6]. In this work, the authors proposed an iris recognition system where they used the pertained VGG-Net to extract the deep features. Then they used a multi-class SVM algorithm for classification. They tested their system on two iris databases, IIT iris dataset, and CASIA 1000 Iris dataset. Their experiment achieved high results with the accuracy recognition of 99.4%.

The pertained convolutional neural network (LeNet-5) [19]. In this work, the authors proposed a system for the diagnosis of iris nevus depending on a CNN in addition to deep belief network. Iris nevus is defined as a pigmented growth located around the pupil or in the front of the eye. They used a pertained LeNet-5 architecture for CNN. Their accuracy rates are 93.35% for CNN and 93.67% for the deep belief network.

The convolutional neural network (SRCNN and VDSR) [40] In this work the authors proposed to explore the use of texture transfer learning for super-resolution applied to low-resolution images. They evaluated this method in a subset of CASIA Iris Database representing the authentication images to also validate if the transfer learning from the registration image subset is appropriate for this application. The result shows how the features from completely different nature can be transferred in the feature domain, improving the recognition performance if applied to bigger reduction factors comparing to the classical interpolation approaches.

The pertained convolutional neural network (Alex-Net) [41]. In this work, the authors extracted the learned features by using the pre-trained CNN (Alex-Net) then used the multi-class SVM algorithm for classification. The iris is segmented using canny edge detection and circular Hough transform and normalized using the rubber sheet model. The segmented and normalized image is fed as an input to the CNN (Alex-Net). The results showed that the recognition accuracy when extracting features from the segmented image is higher than when extracting features from the normalized image. The proposed system is tested on public datasets (IITD iris

databases, CASIA-Iris-V1, CASIA-Iris-thousand, and CASIA-Iris- Interval), and a high accuracy rate is achieved

3.3. Classification

Once the feature extraction from the iris image the classification is done. There are two approaches for classification:

1. The static formula for matching
2. Machine learning algorithms

3.3.1. The Static Formula for Matching

The matching is done using the matching metric to compare the user template with database templates. The matching metric measured the similarity among two iris templates. Also, it provides a range of values once matching templates from the same iris, and a different range of values once matching templates from dissimilar irises [22].

Hamming distance which used in [2, 24, 25, 28, 32, 42] it is defined as the fractional measure of the difference among two binary templates. The perfect match would represent a value of zero. Hamming distance while is near to 0.5 when two templates are completely independent. A threshold is used to determine the two templates belong to the same subject or diverse subjects. Hamming distance has some advantages such as the templates are in a binary format which helps to make the matching process is quick. The execution time for the exclusive-OR comparison of two templates is about 10μs. So, this method is appropriate for comparisons of millions of templates in the big dataset.

Weighted Euclidean distance which used in [43] it is used to compare between two templates to recognize an iris. The templates are consisting of integer values. To measure the similarity between two templates it used the Weighted Euclidean Distance. Which computing using the Pythagorean Theorem to get the distance between two points. It compared an iris template with each template in the database. The minimum Weighted Euclidean Distance means the two templates are matched.

The normalized correlation which used in [26] it calculated for the goodness of match between two representations. It is defined as the normalized similarity of matching points in the iris area. The correlations are done above minor blocks of pixels in four diverse spatial frequency bands. Normalized correlation accounts for local variations in image intensity. This technique is not computationally efficient because it used images for comparisons.\

3.3.2. Machine Learning Algorithms.

Machine learning algorithms learn from data and develop from experience, without human intervention. The tasks of learning might contain learning the function that transmits the input to the output [44]. There are different types of machine learning algorithms and will be discussed some of them in this section.

Support Vector Machine which used in [6, 19, 45] this algorithm is a common machine learning algorithm that used for classification. SVM classifier makes classification through discovery the hyperplane that maximizes the margin among the two classes. The support vectors are the vectors (cases) that determine the hyperplane.

The k-nearest Neighbor which used in [35] in this method the whole dataset used for the training set so, it not divided the dataset into a training set and test set. Once the new data

instance needs the result, the k-Nearest Neighbor algorithm search over the whole dataset to get the k-nearest instances to the new instance, or the k number of instances that almost like the new record. After that, the outputs the mean of the outcomes (for a regression problem) or the mode (most frequent class) for a classification problem. The user specified the value of k. The comparison among instances is computing with algorithms like Hamming distance and Euclidean distance.

Decision tree classifier which used in [46] this algorithm defined as a repetitive partition of the certain data space. This algorithm comprises of a rooted tree that directed with a node named root, the root node is the main and the other nodes are leaves. Decision tree builds a decision tree for the certain dataset automatically by making the error is minimal. It attempts to enhance the cost function to discovery a decision tree T with a certain set of L labeled samples.

4. THE PROPOSED IRIS RECOGNITION SYSTEM

The system model of the proposed iris recognition system using transfer learning with the pre-trained Convolutional Neural Network (VGG-16) is shown in Figure 1. The development of the proposed iris recognition system is discussed in two parts:

1. The pre-processing stage.

2. The feature extraction and a classification stage.
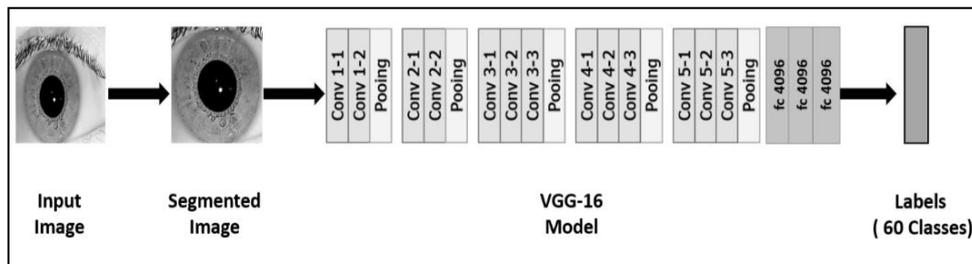
Figure 1. The system model of the iris recognition system

4.1. The Pre-Processing stage

In this stage, iris segmentation is performed. The primary goal of selecting the iris region as an input to pre-trained Convolutional Neural Network model as an alternative to the entire eye image, as proposed in [21], is to decrease the computational complexity of the model. An additional goal is to avoid the decline of matching performance.

For the iris segmentation process, the circular Hough transform is performed to detect the iris and pupil boundaries of which involve first perform Canny Edge detection to generate an edge map[23]. The main steps for iris segmentation are summarized as follows:

1. Define the range of pupil and iris radius, manually set the range of radius, according to the database used

2. Find the iris boundaries
   a. Perform canny edge detection to generate an edge map.

     i. Apply Gaussian filter
     ii. Apply gamma function
     iii. Apply non-maximum suppression
     iv. Apply hysteresis thresholding

   b. Apply circular Hough transform
   c. Draw circles of different radius for all edge point.
   d. Find the maximum in the Hough space, and it will be the parameters of the circle

3. Find pupil boundary using the same steps but just using the region within the previously detected iris boundary

4. Return the center and radius of iris and pupil circle

4.2. The feature extraction and classification stage

In this stage, we use the transfer learning with the pre-trained model of convolutional neural networks named as (VGG-16). The pre-trained CNN VGG-16 model has been trained on about 1.2 million images from the ImageNet Dataset and can classify images into 1000 object categories [47].

The VGG-16 is a deeper network which contains thirteen convolutional layers. followed by rectified linear units (ReLU), five max polling layers, and three fully-connected layers [48]

The convolutional layers in VGG-16 are separated into five groups, and every group is followed by a max pooling layer. The first group of the convolutional layer has 64 filters then after every max-pooling layer the number of filters increases by a factor of 2 till it reaches 512 in the last group. The size of filters of all convolutional layers is fixed to 3 x 3. Also, the stride and padding of all convolutional layers are fixed to 1 pixel. Max-pooling is carried out over a 2 x 2 window with stride 2[48]. The detailed explanation of VGG-16 layers is shown in Table 2.

Table 2. The detailed explanation of VGG-16 layers

Type of Layer		No. of Filter	Kernel Size	No. of Stride	No. of Padding
Image input layer					
Group1	(1st convolutional layer) Relu-1	64	3x3	1x1	1x1
	(2nd convolutional layer) Relu-2 Max pooling1	64 1	3x3 2x2	1x1 2x2	1x1 0x0
Group2	(3rd convolutional layer) Relu-3	128	3x3	1x1	1x1
	(4rd convolutional layer) Relu-4 Max pooling2	128 1	3x3 2x2	1x1 2x2	1x1 0x0
Group3	(5th convolutional layer) Relu-5	256	3x3	1x1	1x1

	(6th convolutional layer) Relu-6	256	3x3	1x1	1x1
	(7th convolutional layer) Relu-7 Max pooling3	256 1	3x3 2x2	1x1 2x2	1x1 0x0
Group4	(8th convolutional layer) Relu-8	512	3x3	1x1	1x1
	(9th convolutional layer) Relu-9	512	3x3	1x1	1x1
	(10th convolutional layer) Relu-10 Max pooling4	512 1	3x3 2x2	1x1 2x2	1x1 0x0
Group5	(11th convolutional layer) Relu-11	512	3x3	1x1	1x1
	(12th convolutional layer) Relu-12	512	3x3	1x1	1x1
	(13th convolutional layer) Relu-12 Max pooling5	512 1	3x3 2x2	1x1 2x2	1x1 0x0
Fully connected layer-6 (FC6) Relu-6					
Fully connected layer-7 (FC7) Relu-7					
Fully connected layer-8 (FC8) Softmax layer Output layer					

The pre-trained convolutional neural network (VGG-16) is fine-tuning for features extracting and classification. In the transfer learning, we use the VGG-16 model that already been trained on ImageNet Dataset, and now we have our dataset that we want the VGG-16 model to give a prediction for it. The basic logic behind transfer learning is that every layer on the pre-trained CNN model determines specific shapes particular to the worldly objects like edges, circles, etc. This information is helpful when the model gives a prediction of the new classes [12].

As stated earlier, the Convolutional Neural Network model comprises convolutional layers, pooling layer, and three fully connected layers (FC6, FC7, and FC8). Every convolution layer is activated by a ReLU function. The output of (FC6 and FC7) layers is a 4096-dimensional vector. FC8 that is the last fully connected layer is adjusted to give a prediction for 1000-class object recognition task on ImageNet dataset. The softmax which is a very last layer has a probabilistic output of 1000 for every object category[49].

After the preprocessing stage, the segmented iris image is fed to the pre-trained convolutional neural network (VGG-16) model to extract the features. Then used the feature vectors from the last fully connected layer to obtained the matching score by the Softmax classifier.

4.2.1. Softmax Regression Classifier

The Softmax classifier is the very last layer in architectures of pre-trained CNN (VGG-16). The Softmax regression classifier is a general form of binary logistic regression classifier proposed to use for multi-class classification tasks. Assume that there are K classes and n labeled training samples $((x_1, y_1),…, (x_n, y_n))$, where x_i ($i \in (1,2,…n)$) denotes the training data, and y_j ($j \in (1,2,…n)$) denotes the class label.

Then, if we have an input feature input x_i, the Softmax classifier will generate a K-dimensional vector (whose elements sum to 1), and every element in the output vector denotes to the predicted probability of every class label conditioned on this input feature [21].

All the required steps for using the pre-trained convolutional neural network model for features extracting and classification (transfer learning with Convolutional Neural Networks) are summarized as follows:

1. Load input images and its labels
2. Split the sets of the images into training and testing data.
3. Load pre-trained CNN
4. Modify the Pre-trained model: change the last fully connected layer based on the number of classes
5. Set the option of Transfer Learning
6. Pre-process images for the model
7. Train the model using a training set
8. Test the model performance using the test set
9. Compute the mean accuracy.
10. Display the mean accuracy.

5. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed system is tested using four popular iris datasets: IITD iris databases [13], CASIA-Iris-V1[14], CASIA-Iris-thousand [15], and CASIA-Iris-Interval [16]. The iris images are captured in these databases under different situations of pupil dilation, eyelids/eyelashes occlusion, a slight shadow of eyelids, specular reflection, etc.

The proposed system is tested on 60 subjects in each dataset. The specifications of these datasets are summarized in Table 3.

Table 3. The specifications of the used datasets

Dataset	IITD	CASIA-Iris-V1	CASIA-Iris-Thousand	CASIA-Iris-Interval
Number of subjects	60	60	60	60
Samples per subject	10	7	10	10
Number of images	600	420	600	600
Image size (pixels)	(320×240)	(320×280)	(640×480)	(320×280)
Image format	BMP	BMP	JPEG	JPEG

The system and its stages are implemented using MATLAB 2017 on a laptop with Core i7 CPU running at 2.8GHz. We use a different toolbox which are:

• Neural Network Toolbox
• Neural Network Toolbox Model for VGG-16 Network

The input image is resized to 224-by-224-by-3 as required for the VGG-Net model [47].

The proposed system performance is evaluated according to the recognition accuracy rate. The recognition accuracy is the fraction of labels that the classifier predicted correctly.

Also, the performance of the proposed iris recognition system investigated by using the segmented iris images. We selected to use the segmented iris image to fed as input for the CNN model because according to the study published in [41] the recognition accuracy when extracting features from the segmented iris image is higher than when extracting features from the normalized iris image.

The sets of the segmented images are split randomly into training and test data. For each person, 80% of images are used for the training, and 20% of images are used for the testing.

Once the input image is fed to the pre-trained CNN model, feature extraction and classification is performed. The pre-trained VGG-16 were trained to recognize 1000 classes[47], and we modified it to recognize just 60 classes because we tested the system on 60 subjects for each dataset.

The pre-trained VGG-16 is fun-tuning by changing the number of outputs in last fully connected layer (FC8) based on a number of our classes which is 60 classes so that the network adjusts its output to the new classification task.

In this work, we investigated the pre-trained VGG-16 using different numbers of epochs (10 and 20). An epoch is defined as the complete pass of the training algorithm over the whole training set [21]. The initial learning rate is set to 0.001. The mini-batch size is set to 22. A mini-batch is defined as a subset of the training set that is used to evaluate the gradient of the loss function and update the weights. The gradient descent algorithm updates the network parameters (weights and biases)[50].

The recognition accuracy of the VGG-16 model for the segmented iris image for each database is shown in Figure 2.
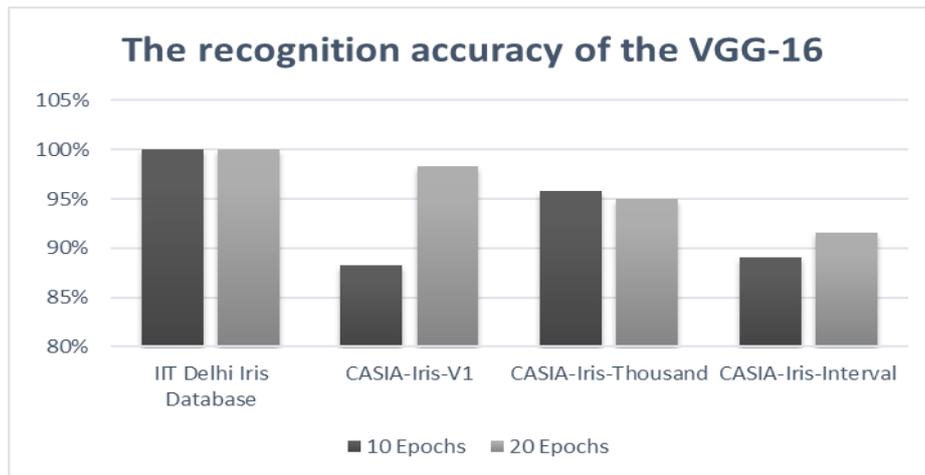
Figure 2. The recognition accuracy of the VGG-16 model for different numbers of epochs

We can see from Figure 2 the recognition accuracy of VGG-16 give better result in 20 epochs than 10 epochs**.**

6. COMPARISON BETWEEN THE PERFORMANCE OF THE PROPOSED IRIS RECOGNITION SYSTEM AND THE OTHER WORKS

In this section, we presented a comparison with other works. We compare the performances of the systems according to the recognition accuracy as illustrated in Table IV.

The work in [41] used the segmented iris image as input for pre-trained Alex-Net model for feature extraction and used multi-class SVM for the classification task.

The experiment tested on four public databases IITD, iris databases CASIA-Iris-V1, CASIA-Iris-thousand and, CASIA-Iris-Interval. The sets of the segmented iris images are split randomly into training and test data 80% of images are used for the training, and 20% of images are used for the testing. The Alex-Net model has 25 layers. Alex-Net consists of a series of five convolutional layers, followed by rectified linear units (ReLU) and, three max-pooling layers that used after the first, second and fifth convolutional layers. Next to these layers, three fully-connected layers [41].

As can be seen from Table IV, both models (Alex-Net and VGG-16) give high accuracy rate for IIT Delhi Iris database and CASIA-Iris-V1. But in CASIA-Iris-Thousand the Alex-Net with SVM gives a higher result than VGG-16. On another hand, the VGG-16 gives higher result in CASIA-Iris-Interval than the Alex-Net with SVM.

The work in [17] used the normalized iris image as the input to the different pre-trained CNN module (AlexNet, VGG, Google Inception, ResNet, and DenseNet) to extracted the features and multi-class SVM for the classification task. Their experiments tested on two iris datasets LG2200 dataset and CASIA-Iris-Thousand. The sets of the normalized images are split randomly into training and test data 70% of images are used for the training, and 30% of images are used for the testing.

The result of the work [17] in CASIA-Iris-Thousand gives the peak recognition accuracy with AlexNet that because the number of layers considered in its architecture may not fully capture the discriminative visual features in iris images. As well as it achieves the highest recognition accuracy of 98.8%,98.5%, and 98.3%, for DenseNet, ResNet and Inception respectively. VGG, with its simple architecture, only achieves 93.1% recognition accuracy.

In addition, we provide in Table 4, the comparison between the performance of the proposed iris recognition scheme and the other algorithms.

The proposed iris recognition system has overall, outperformed than other feature extraction algorithms which include, Intersecting Cortical Model (ICM) network [36], circular sector and triangular DCT [37], discrete wavelet transformation (DWT) [33], Radon transform and gradient-based isolation [34], Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT) [35], scattering transform and textural features [6], and the feature extraction using pre-trained VGG-Net [5].

Table 4. Comparison between the performance of the proposed iris recognition scheme and the other algorithms

Reference# (Year)	Feature Extraction	Classification	Database	Recognition Accuracy	Evaluation protocol
[38] (2008)	Intersecting Cortical Model (ICM) network	Hamming Distance	CASIA-Iris-V1	97.74%	4:3 (training: testing)
[39] (2012)	circular sector and triangular DCT	Euclidean distance	IIT Delhi Iris Database	97.12%	3:2 (training: testing)
[35] (2013)	discrete wavelet transformation (DWT)	k-nearest neighbor(k-NN).	IIT Delhi Iris Database	99.5	cross-validation
[36] (2014)	Radon transform and gradient-based isolation	Euclidean distance	IIT Delhi Iris Database	95.93%	4:1 (training: testing)
			CASIA-Iris-Interval	84.17%	
[37] (2015)	Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT)	Euclidean distance	IIT Delhi Iris Database	97.81%	9:1 (training: testing)
[7] (2015)	Texture+ Scattering Features	Minimum distance	IIT Delhi Iris Database	99.2%	5:5 (training: testing)
[6] (2016)	pre-trained VGG-Net	multi-class SVM	IIT Delhi Iris Database	99%	5:5 (training: testing)
			CASIA-Iris-Thousand	90%	
[17] (2017)	pre-trained AlexNet	multi-class SVM	CASIA-Iris-Thousand	Not-specified	7:3 (training: testing)
	pre-trained VGG			93.1%	
	pre-trained Inception			98.3%	
	pre-trained ResNet			98.5%	
	pre-trained DenseNet			98.8%	
[41] (2018)	pre-trained Alex-Net	multi-class SVM	IIT Delhi Iris	100%	8:2 (training: testing)
			CASIA-Iris-V1	98%	
			CASIA-Iris-Thousand	98%	
			CASIA-Iris-Interval	89%	
The Proposed scheme (2019)	**pre-trained VGG-16**	**The Softmax classifier in pre-trained VGG-16**	**IIT Delhi Iris**	**100%**	**8:2 (training: testing)**
			CASIA-Iris-V1	**98.3%**	
			CASIA-Iris-Thousand	**95%**	
			CASIA-Iris-Interval	**91.6%**	

7. CONCLUSIONS

This paper proposed an effective iris recognition system by using the transfer learning with a pre-trained CNN (VGG-16). The iris image is segmented using a circular Hough transform. The segmented image is fed as an input to the pre-trained CNN (VGG-16). The proposed system is tested on public datasets (IITD iris databases, CASIA-Iris-V1, CASIA-Iris-thousand, and CASIA-Iris- Interval), and a high accuracy rate is achieved.

In addition, the performance of the proposed system is compared against another system that used the pre-trained convolutional neural network (Alex-Net) as features extracting followed by a multi-class Support Vector Machine (SVM) algorithm to perform the classification task.

In the future, we will evaluate the performance of the proposed system using the different pretrained models in more iris datasets with other biometric recognition problems.

ACKNOWLEDGEMENTS

REFERENCES

[1] Haghighat, M., S. Zonouz, and M. Abdel-Mottaleb, CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. Expert Systems with Applications, 2015. 42(21): p. 7905-7916.

[2] Kesavaraja, D., D. Sasireka, and D. Jeyabharathi, Cloud software as a service with iris authentication. Journal of Global Research in Computer Science, 2010. 1(2): p. 16-22.

[3] Bowyer, K.W., K. Hollingsworth, and P.J. Flynn, Image understanding for iris biometrics: A survey. Computer vision and image understanding, 2008. 110(2): p. 281-307.

[4] Dehkordi, A.B. and S.A. Abu-Bakar, A review of iris recognition system. Jurnal Teknologi, 2015. 77(1).

[5] Shah, N. and P. Shrinath, Iris Recognition System–A Review. International Journal of Computer and Information Technology, 2014. 3(02).

[6] Minaee, S., A. Abdolrashidiy, and Y. Wang. An experimental study of deep convolutional features for iris recognition. in Signal Processing in Medicine and Biology Symposium (SPMB), 2016 IEEE. 2016. IEEE.

[7] Minaee, S., A. Abdolrashidi, and Y. Wang. Iris recognition using scattering transform and textural features. in Signal Processing and Signal Processing Education Workshop (SP/SPE), 2015 IEEE. 2015. IEEE.

[8] Minaee, S., A. Abdolrashidi, and Y. Wang, Face Recognition Using Scattering Convolutional Network. arXiv preprint arXiv:1608.00059, 2016.

[9] Sharif Razavian, A., et al. CNN features off-the-shelf: an astounding baseline for recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2014.

[10] Lucena, O., et al. Transfer Learning Using Convolutional Neural Networks for Face Anti-spoofing. in International Conference Image Analysis and Recognition. 2017. Springer.

[11] Torrey, L. and J. Shavlik, Transfer learning. Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, 2009. 1: p. 242.

[12] Masood, S., et al. Identification of diabetic retinopathy in eye images using transfer learning. in Computing, Communication and Automation (ICCCA), 2017 International Conference on. 2017. IEEE.

[13] IIT Delhi Database. Available from: http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm.

[14] Casia Iris Database. Available from: http://www.idealtest.org/findDownloadDbByMode.do?mode=Iris.

[15] CASIA Iris Image Database Version 4.0 (CASIA-Iris-Thousand). Available from: http://biometrics.idealtest.org/dbDetailForUser.do?id=4.

[16] CASIA Iris Image Database Version 3.0 (CASIA-Iris-Interval). Available from: http://biometrics.idealtest.org/dbDetailForUser.do?id=3.

[17] Nguyen, K., et al., Iris Recognition with Off-the-Shelf CNN Features: A Deep Learning Perspective. IEEE Access, 2017.

[18] Romero, A., C. Gatta, and G. Camps-Valls, Unsupervised deep feature extraction for remote sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, 2016. 54(3): p. 1349-1362.

[19] Oyedotun, O. and A. Khashman, Iris nevus diagnosis: convolutional neural network and deep belief network. Turkish Journal of Electrical Engineering & Computer Sciences, 2017. 25(2): p. 1106-1115.

[20] Nagi, J., et al. Max-pooling convolutional neural networks for vision-based hand gesture recognition. in Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on. 2011. IEEE.

[21] Al-Waisy, A.S., et al., A multi-biometric iris recognition system based on a deep learning approach. Pattern Analysis and Applications, 2017: p. 1-20.

[22] Ng, R.Y.F., N.Y.H. Tay, and K.M. Mok. A review of iris recognition algorithms. in Information Technology, 2008. ITSim 2008. International Symposium on. 2008. IEEE.

[23] Sangwan, S. and R. Rani, A Review on: Iris Recognition. (IJCSIT) International Journal of Computer Scienceand Information Technologies, 2015. 6(4): p. 3871-3873.

[24] Daugman, J.G., High confidence visual recognition of persons by a test of statistical independence. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1993. 15(11): p. 1148-1161.

[25] Daugman, J., How iris recognition works. Circuits and Systems for Video Technology, IEEE Transactions on, 2004. 14(1): p. 21-30.

[26] Wildes, R.P., Iris recognition: an emerging biometric technology. Proceedings of the IEEE, 1997. 85(9): p. 1348-1363.

[27] Jayachandra, C. and H.V. Reddy, Iris Recognition based on Pupil using Canny edge detection and K-Means Algorithm. Int. J. Eng. Comput. Sci., 2013. 2: p. 221-225.

[28] Daouk, C., et al. Iris recognition. in IEEE ISSPIT. 2002.

[29] Raghava, N. Iris recognition on hadoop: A biometrics system implementation on cloud computing. in Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on. 2011. IEEE.

[30] Darve, N.R. and D.P. Theng, Biometric User Authentication for Mobile Cloud Computing (MCC) Servises Through Eye Images. 2015.

[31] Elrefaei, L.A., et al., Developing Iris Recognition System for Smartphone Security. Multimedia Tools and Applications, 2017: p. 1-25.

[32] Saini, R.G.H., Generation of Iris Template for recognition of Iris in Efficient Manner. International Journal of Computer Science and Information Technologies (IJCSIT), 2011. 2(4): p. 1753-1755.

[33] Nguyen, K., et al., Long range iris recognition: A survey. Pattern Recognition, 2017. 72: p. 123-143.

[34] Pirale, D., M. Nirgude, and S. Gengje, Iris Recognition using Wavelet Transform and Neural Networks. International Journal of Science and Research (IJSR), 2016. 5(5): p. 1055-1060.

[35] Elgamal, M. and N. Al-Biqami, An efficient feature extraction method for iris recognition based on wavelet transformation. Int. J. Comput. Inf. Technol, 2013. 2(03): p. 521-527.

[36] Bharath, B., et al. Iris recognition using radon transform thresholding based feature extraction with Gradient-based Isolation as a pre-processing technique. in Industrial and Information Systems (ICIIS), 2014 9th International Conference on. 2014. IEEE.

[37] Dhage, S.S., et al., DWT-based feature extraction and radon transform based contrast enhancement for improved iris recognition. Procedia Computer Science, 2015. 45: p. 256-265
.

[38] Xu, G., Z. Zhang, and Y. Ma, A novel method for iris feature extraction based on intersecting cortical model network. Journal of Applied Mathematics and Computing, 2008. 26(1-2): p. 341-352.

[39] Abhiram, M., et al. Novel DCT based feature extraction for enhanced iris recognition. in Communication, Information & Computing Technology (ICCICT), 2012 International Conference on. 2012. IEEE.

[40] Ribeiro, E. and A. Uhl. Exploring texture transfer learning via convolutional neural networks for iris super resolution. in 2017 International Conference of the Biometrics Special Interest Group (BIOSIG). 2017. IEEE.

[41] Maram G. Alaslani, L.a.A.E., Convolutional Neural Network based Feature Extraction for Iris Recognition. International Journal of Computer Science & Information Technology (IJCSIT), 2018. 10(April): p. 65-87.

[42] Hsieh, L., W.-S. Chen, and T.-H. Li. Personal Authentication Using Human Iris Recognition Based on Embedded Zerotree Wavelet Coding. in Computing in the Global Information Technology (ICCGI), 2010 Fifth International Multi-Conference on. 2010. IEEE.

[43] Zhu, Y., T. Tan, and Y. Wang. Biometric personal identification based on iris patterns. in icpr. 2000. IEEE.

[44] Kotsiantis, S.B., I. Zaharakis, and P. Pintelas, Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 2007. 160: p. 3-24.

[45] Roy, K., P. Bhattacharya, and C.Y. Suen, Iris recognition using shape-guided approach and game theory. Pattern Analysis and Applications, 2011. 14(4): p. 329-348.

[46] Farzam, S. and H. Mirvaziri, IRIS RECOGNITION BASED ON WAVELET TRANSFORM USING SUPPORT VECTOR MACHINE AND DECISION TREE COMBINATION CLASSIFIER. Indian Journal of Fundamental and Applied Life Sciences, 2016. 6: p. 1-9.

[47] Krizhevsky, A., I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. in Advances in neural information processing systems. 2012.

[48] Yu, W., et al. Visualizing and comparing AlexNet and VGG using deconvolutional layers. in Proceedings of the 33 rd International Conference on Machine Learning. 2016.

[49] Ozbulak, G., Y. Aytar, and H.K. Ekenel. How transferable are CNN-based features for age and gender classification? in Biometrics Special Interest Group (BIOSIG), 2016 International Conference of the. 2016. IEEE.

[50] Syafeeza, A., et al., Convolutional neural networks with fused layers applied to face recognition. International Journal of Computational Intelligence and Applications, 2015. 14(03): p. 1550014.

AUTHORS

Maram G. Alaslani, female, received her B.Sc. degree in Computer Science with honors from King Abdulaziz University in 2010. her M.Sc. in 2018 at King Abdulaziz University, Jeddah, Saudi Arabia. She works as Teaching Assistant from 2011 to date at Faculty of Computers and Information Technology at King Abdulaziz University, Rabigh, Saudi Arabia. She has a research interest in image processing, pattern recognition, and neural network.

Lamiaa A. Elrefaei, female, received her B.Sc. degree with honors in Electrical Engineering (Electronics and Telecommunications) in 1997, her M.Sc. in 2003 and Ph.D. in 2008 in Electrical Engineering (Electronics) from faculty of Engineering at Shoubra, Benha University, Egypt. She held a number of faculty positions at Benha University, as Teaching Assistant from 1998 to 2003, as an Assistant Lecturer from 2003 to 2008, and has been a lecturer from 2008 to date. She is currently an Associate Professor at the faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. Her research interests include computational intelligence, biometrics, multimedia security, wireless networks, and Nano networks. She is a senior member of IEEE.