

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363521543>

# Cancelable face and iris recognition system based on deep learning

Article in *Optical and Quantum Electronics* · September 2022

DOI: 10.1007/s11082-022-03770-0

CITATIONS

19

READS

520

10 authors, including:



**Essam Abdellatef**

Faculty of engineering; Sinai University; Egypt

21 PUBLICATIONS 235 CITATIONS

[SEE PROFILE](#)

**Randa Soliman**

Menoufia University

17 PUBLICATIONS 246 CITATIONS

[SEE PROFILE](#)



**Eman M Omran**

Menoufia University

6 PUBLICATIONS 51 CITATIONS

[SEE PROFILE](#)



**Nabil A. Ismail**

Faculty of Electronic Engineering Menoufia University

209 PUBLICATIONS 767 CITATIONS

[SEE PROFILE](#)



# Cancelable face and iris recognition system based on deep learning

Essam Abdellatef<sup>1</sup> · Randa F. Soliman<sup>2</sup> · Eman M. Omran<sup>3</sup> · Nabil A. Ismail<sup>4</sup> · Salah E. S. Abd Elrahman<sup>4</sup> · Khalid N. Ismail<sup>5,6</sup> · Mohamed Rihan<sup>7,8</sup> · Mohamed Amin<sup>9</sup> · Ayman A. Eisa<sup>3</sup> · Fathi E. Abd El-Samie<sup>7,10</sup>

Received: 3 January 2022 / Accepted: 22 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Biometric recognition is an automated technique of recognising persons based on their traits. Because of their exceptional texture, the biometric features' ostensibly random nature makes them good candidates for recognition. These features are unique for each individual even for identical twins authentication. The latest developments in Deep Learning (DL) and computer vision has proved that Convolutional Neural Networks (CNNs) can extract generic descriptors that can represent complex image features. How to protect the biometric data and ensure user's privacy is a main concern, nowadays. Hence, several cancelable biometric scenarios have been proposed. In this paper, we propose a novel cancelable biometric recognition system based on a CNN model with bio-convolution. The performance metrics are estimated on different face and iris datasets. In contrary to most conventional secure biometric recognition systems, the proposed system achieves superior accuracy results, while keeping the ability to cancel the biometric traits if compromised. The experimental findings on each database are shown and compared to those of the state-of-the-art systems that have been tested on the same database. Furthermore, the recognition rates reach 99.15%, 98.35%, 97.89, and 95.48% with the LFW, FERET, IITD, and CASIA-IrisV3 databases, respectively.

**Keywords** Deep learning · Cancelable biometrics · Security

## 1 Introduction

Identity authentication is essential for several important applications. Conventional techniques for authentication based on passwords are unsuitable to today's requirements as these passwords may be lost, stolen, or forgotten, easily. On the other hand, physiological or behavioral biometric authentication techniques compared to conventional techniques have higher privacy and do not suffer from the above-mentioned problems. One of the most attractive solutions for feature representation is the utilization of Deep Learning (DL),

---

✉ Randa F. Soliman  
randa\_soliman@ai.menofia.edu.eg; randafouad2010@yahoo.com

Extended author information available on the last page of the article

which is basically a data-driven approach. Hierarchical multi-layer networks are utilized in the DL scenarios to extract optimized feature maps (Bengio et al. 2013; LeCun et al. 2015). As a result, the problem of biometric recognition can be solved by using the DL methods based on CNNs for designing alternative feature descriptors. Our present research is motivated by the inquiry “Can we introduce cancelable face and iris recognition techniques based on DL that can achieve better performance than those of the state-of-the-art techniques and at the same time preserve the biometric data from hackers?”.

In this paper, we propose an efficient cancelable biometric recognition system that can recognize face and iris images using CNNs. Believing in the concept that using deeper CNNs can enhance the system accuracy, we propose a DL model that contains eight max-pooling, batch normalization, twenty two convolutional, fully-connected, and feature normalization layers. The proposed architecture also incorporates depth concatenation and a residual learning framework.

Face and iris are chosen as the biometric traits to investigate in this paper for the following reasons. In comparison to other biometrics, the human face is believed to be one of the most useful biometrics. Moreover, face is distinguished by its low cost, lack of contact, and great acceptability during acquisition compared to other biometrics. Moreover, the iris recognition techniques achieve high matching rates with low false matching rates on large datasets. The sophisticated iris stroma texture varies from person to person. The iris genetic limited penetration and the long-term impression of its uniqueness traits are all factors that contribute to success of recognition based on iris (Muron and Pospisil 2000; Daugman 2016). The National Institute of Science and Technology (NIST) has obtained superior recognition rates in applicable iris recognition systems (NIST 2012). In addition, there are more than one billion people around the world, whose iris images are registered electronically in different datasets (Daugman 2014). These datasets are mainly based on the national ID program in Indonesia, Unique Identification Authority of India (UIDAI) program, and lastly the program of the US Department of Homeland Security. So, the iris seems to be the most dominant pattern in authentication systems around the world.

In order to preserve the biometric data from hackers, encryption techniques such as hash functions can be used. Hash functions mainly depend on computing a digest to protect biometric templates. Unfortunately, lighting changes have a significant impact on biometric templates. As hash functions cannot be inverted, different digests are generated, when the input changes slightly. Another disadvantage of integrating encryption techniques with biometrics is the necessity for data decryption, which might be exploited by hackers to get the original templates. Cancelable biometrics have recently gained a lot of interest (Ratha et al. 2001, 2007; Bolle et al. 2002; Teoh et al. 2006). This trend depends on one-way functions that can be used to transform biometric templates instead of storing the original ones. This allows revocability through the ability to re-enroll compromised patterns. In addition, a variety of transformation approaches can be employed. Moreover, privacy is preserved, because the original biometric templates cannot be easily reconstructed from the changed ones. Furthermore, because the statistical characteristics of features after transformation are roughly conserved, the recognition system accuracy is not deteriorated. The main objective of this work is to present a novel cancelable biometric system based on a CNN model with bio-convolution. This model is tested on four different datasets, namely LFW, FERET, IITD, and CASIA-IrisV3.

This remaining parts of this paper are organized as follows. The state-of-the-art face and iris recognition techniques are presented in Sect. 2. Section 3 presents the proposed cancelable biometric recognition system. The obtained experimental results are presented and discussed in Sect. 4, while the paper concluding remarks are given in the last section.

## 2 Related work

Several techniques that are used for face and iris recognition depend on applying DL technology for feature representation. DeepFace was one of the first CNN-based projects to extract features for face verification after training it on thousands of face images (Taigman et al. 2014). When tested on the LFW dataset, this network achieved a precision of 97.35%. To enhance DeepFace performance, semantic bootstrapping was utilized to find the most efficient training set from large datasets (Taigman et al. 2015).

Classification and verification losses are incorporated to increase the inter-class distance between templates, while minimizing the intra-class distance as proposed in (Sun et al. 2015). This ensemble technique achieved an accuracy of 99.47% with the LFW dataset. FaceNet has been used for face recognition, and it was trained on nearly 100 to 200 M face images combined with a total of 8 M face identities (Schroff et al. 2015). As the triplet pair selection is very essential to obtain a reasonable classification accuracy, FaceNet introduces an online triplet mining method for training-based CNNs, and it attains good accuracy. A VGG network (Simonyan and Zisserman 2014) was trained on 2622 identities of nearly 2.6 M samples gathered from the Internet, and then the model was fine-tuned by the triplet FaceNet model (Parkhi et al. 2015). This led to a high accuracy of 98.95% on LFW dataset. Moreover, domain-specific data augmentation was introduced to enhance the trained data and achieve a high accuracy on the LFW dataset (Masi et al. 2016).

To improve the performance of iris recognition, few deep networks have been introduced. DeepIris (Liu et al., 2016a, b) network that consists of 9 layers achieved high recognition rates on Q-FIRE (Johnson et al., 2010) and CASIA (2017) databases. After that, more advanced layers were used to generate a couple of DeepIrisNets that can be utilized for iris recognition (Gangwar and Joshi, 2016). The first suggested network was DeepIrisNet-A, which had eight convolutional layers, four pooling layers, three fully-connected layers, and two drop-out layers. Increasing the modeling ability was achieved by adding two inception layers in the second network, DeepIrisNet-B (Gangwar and Joshi, 2016). By applying these two networks on ND-IRIS-0405 and ND-CrossSensor-Iris-2013 (Phillips et al. 2010) databases, a superior performance was attained. Moreover, the CNNs could be also utilized with iris biometrics for spoof detection (Menotti et al. 2015), iris segmentation [Liu et al. 2016a, b], and gender classification (Tapia and Aravena 2017). Superior results were achieved by utilizing self-designed CNNs such as DeepIrisNet (Johnson et al. 2010) and DeepIrisNet (Gangwar et al. 2016), but some limitations exist in the network design as the number of training samples puts restrictions on the number of layers.

Some state-of-the-art techniques that depend on different CNN architectures have been analyzed in detail, demonstrating their favorable features. Microsoft introduced the notion of residual connection in 2015. This technique feeds successive convolutional layer outputs and bypasses the input to the adjacent layers. Moreover, the residual connection enhances the network gradient flow, which permits it to be very deep with 152 layers. As a result, ResNet (He et al. 2016) succeeded in the ILSVRC 2015 challenge with a 3.57% error rate. DeepVisage (Hasnat et al. 2017) takes the advantage of the residual learning framework. It has a deep CNN model, which is characterized by its simplicity and straightforward use for training. Moreover, DeepVisage utilizes the normalized features in computing the loss, while the pre-processing stage is included in the face detection and facial landmarks, which have been utilized for creating normalized images. Cosine similarity is computed between the pairs of face features as the verification score. A 29-layer CNN (Wu et al. 2018) was used to learn a compact embedding based on larger-scale data and massive noise labels.

In this design, the Max-Feature-Map (MFM) was first presented as a max-out activation function for every CNN convolutional layer. Moreover, max-out activation utilizes various feature maps for linear approximation of an arbitrary convex activation function, while the MFM can do that using a competitive relationship. A careful design of three networks was presented to obtain a good performance by decreasing both the number of parameters and the computational cost. Finally, the method of semantic bootstrapping was presented to let the network perdition be better and consistent with the noise labels. The results of the experiments revealed that this framework can feed larger-scale noisy data to a light model, which makes the storage and computation cost less. A 64-layer CNN (Liu et al. 2017) depends on the angular softmax loss to learn discriminative face features using Sphere-Face model with an angular margin. Better geometric interpretation can be obtained by A-Softmax loss as the learned features become discriminative on a hypersphere manifold. This intrinsically matches the prior faces, which also lay on a non-linear manifold. This connection makes A-Softmax very efficient for face representation. Comparative results on several popular face benchmarks reveal the success of this approach.

### 3 The proposed recognition system

In this paper, we employ four scenarios to provide a cancelable biometric recognition system based on DL technology. These scenarios employed for generating cancelable biometrics are summarized in Fig. 1.

For robustness verification of our scenarios, we applied them on two types of biometric traits (face and iris) with different datasets. In the first scenario, we have a biometric recognition system, where images are fed to a CNN model. Features are extracted and then classified using an SVM. For the 2nd, 3rd, and 4th scenarios, the cancelable biometric algorithm is applied at the image level, feature level, and both image and feature levels, respectively. For the iris images, we adopt the Masek method [Masek 2003] for extracting the pupil and iris circles in the segmentation step. The obtained iris region is then normalized into a rectangular unwrapped image. In the following sections, the proposed system pipeline will be explained in detail.

#### 3.1 A. The proposed CNN architecture

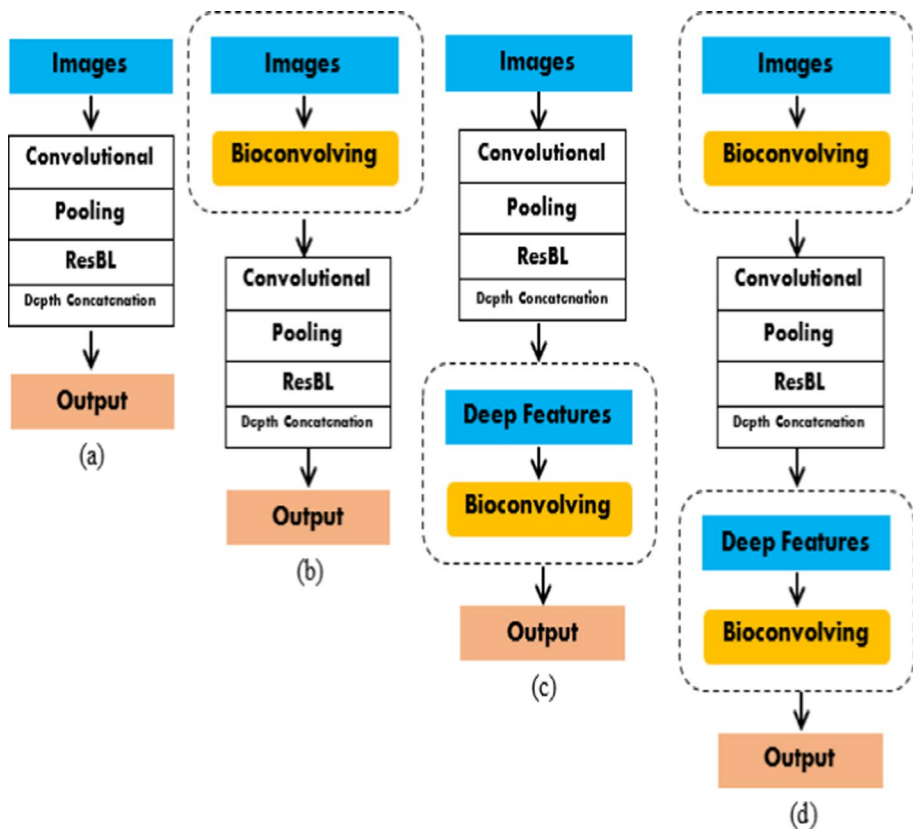
The proposed CNN architecture consists of the following layers as illustrated in Table 1;

(I) "22" convolutional layers (Ko et al. 2017; Yuan et al. 2017): These layers contain a number of filters that can be used to compute dot products between the filters and the input image by sliding them spatially over the image. In order to obtain the output feature map  $f_{x,y,k}^{C,l}$  for a particular layer  $l$  and an input  $f_{x,y}^{O_{p,l-1}}$ , Eq. (1) can be used;

$$f_{x,y,k}^{C,l} = w_k^T f_{x,y}^{O_{p,l-1}} + b_k^l \quad (1)$$

where  $w_k^l$  represents the shared weights,  $b_k^l$  represents the bias, and  $C$  is the convolution.  $O_p$  is the input image, at  $l=1$ , while for  $l>1$ , it is the convolution, pooling or activation layer.

(II) "8" pooling layers. The pooling job computes the greatest or average value in a local spatial neighbourhood, and then reduces the spatial resolution (Yuan et al. 2017).



**Fig. 1** The proposed cancelable biometric scenarios; **a** scenario\_1, **b** scenario\_2, **c** scenario\_3 and **d** scenario\_4

Max Pooling (MP) that can be implemented with equation (2), is applied in the proposed CNN architecture.

$$f_{x,y,k}^{P,l} = \max_{(m,n) \in N_{x,y}} f_{m,n,k}^{Op,l-1} \quad (2)$$

where  $P$  denotes the pooling process and  $N_{x,y}$  denotes the local spatial neighbourhood of the  $(x, y)$  coordinates. An activation function is used to ensure non-linearity of the network. For the current architecture, Rectified Linear Unit (ReLU) (LeCun et al. 2010) is applied as the activation function.

**(III)** Batch normalization layer (Ioffe and Szegedy 2015): The idea behind using this layer is to normalize the activations of the previous layer and maintain the activation standard deviation close to 1 and the mean activation close to 0. In addition, it helps in training the network faster, and making weights easier to initialize.

**(IV)** "2" Residual learning framework layers "ResBL" (He et al. 2016): These layers are used to optimize the loss of the CNN, easily. Equation (3) illustrates the output of a residual block  $R$ ;

**Table 1** The proposed CNN design

Name	Filters	Size of filter	Size of stride	Size of padding
Convolutional-1	64	$7 \times 7 \times 3$	$2 \times 2$	$3 \times 3$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$2 \times 2$	$1 \times 1$
Batch normalization layer	Batch normalization			
Convolutional-2	64	$1 \times 1 \times 64$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-3	128	$3 \times 3 \times 64$	$1 \times 1$	$1 \times 1$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$2 \times 2$	$1 \times 1$
Residual learning	1	$3 \times 3 \times 64$	$1 \times 1$	$1 \times 1$
Convolutional-4	192	$3 \times 3 \times 64$	$1 \times 1$	$1 \times 1$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$1 \times 1$	$1 \times 1$
Residual learning	1	$3 \times 3 \times 64$	$1 \times 1$	$1 \times 1$
Convolutional-5	64	$1 \times 1 \times 192$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-6	96	$1 \times 1 \times 192$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-7	128	$3 \times 3 \times 96$	$1 \times 1$	$1 \times 1$
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-8	16	$1 \times 1 \times 192$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-9	32	$5 \times 5 \times 16$	$1 \times 1$	$2 \times 2$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$1 \times 1$	$1 \times 1$
Convolutional-10	32	$1 \times 1 \times 192$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Depth Concatenation	Depth concatenation of 4 inputs			
Convolutional-11	128	$1 \times 1 \times 256$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-12	128	$1 \times 1 \times 256$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-13	192	$3 \times 3 \times 128$	$1 \times 1$	$1 \times 1$
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-14	32	$1 \times 1 \times 256$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-15	96	$5 \times 5 \times 32$	$1 \times 1$	$2 \times 2$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$1 \times 1$	$1 \times 1$
Convolutional-16	64	$1 \times 1 \times 256$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Depth concatenation	Depth concatenation of 4 inputs			
Convolutional-17	192	$1 \times 1 \times 480$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a

**Table 1** (continued)

Name	Filters	Size of filter	Size of stride	Size of padding
convolutional-18	96	$1 \times 1 \times 480$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-19	208	$3 \times 3 \times 96$	$1 \times 1$	$1 \times 1$
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-20	16	$1 \times 1 \times 480$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Convolutional-21	48	$5 \times 5 \times 16$	$1 \times 1$	$2 \times 2$
Rectified linear unit	n/a	n/a	n/a	n/a
MP layer	1	$3 \times 3$	$1 \times 1$	$1 \times 1$
Convolutional-22	64	$1 \times 1 \times 480$	$1 \times 1$	zero
Rectified linear unit	n/a	n/a	n/a	n/a
Depth concatenation	Depth concatenation of 4 inputs			
MP layer	1	$3 \times 3$	$2 \times 2$	zero
Dropout	40% dropout			
Fully-connected layer	1000 fully-connected layer			
Feature normalization	Feature normalization			
Softmax	n/a	n/a	n/a	n/a

$$f_{x,y,k}^{R,l} = f_{x,y}^{Op,l-q} + F\left(f_{x,y}^{Op,l-q}, \{W_k\}\right) \quad (3)$$

where the input feature map is represented by  $f_{x,y}^{Op,l-q}$ , the residual mapping to be learned is denoted by  $F(\cdot)$ , the parameters of the  $k$ th residual block is represented by  $W_k$  and  $q$  is the total number of stacked layers. In the proposed CNN framework, the residual blocks consist of two convolution layers, each of which is followed by a ReLU activation.

(V) "3" Depth concatenation layers (Szegedy et al. 2015): Depth concatenation means that the banks of output filters of the layers are concatenated into a single output vector, which forms the input of the next stage.

(VI) Fully-connected layer (Yakopcic et al. 2017): This layer is mainly used for classification purposes. Let layer  $(l-1)$  be a fully-connected layer. Layer  $l$  expects  $m_1^{(l-1)}$  feature maps having a size of  $m_2^{(l-1)} \times m_3^{(l-1)}$  as input. Equation (4) can be used to compute the  $i$ th unit in layer  $l$ ;

$$Y_i^{(l)} = f\left(Z_i^{(l)}\right) \text{ with } Z_i^{(l)} = \sum_{j=1}^{m_1^{(l-1)}} \sum_{r=1}^{m_2^{(l-1)}} \sum_{s=1}^{m_3^{(l-1)}} W_{i,j,r,s}^{(l)} \left(Y_j^{(l-1)}\right)_{r,s} \quad (4)$$

where  $W_{i,j,r,s}^{(l)}$  is the weight used to match the unit at position  $(r; s)$  in the  $i$ th unit in layer  $l$  and the  $j$ th feature map of layer  $(l-1)$ .

(VII) Feature normalization layer (Theodoridis and Koutroumbas 2008): This layer is used to ensure that features are treated with the same cost function. This cannot be guaranteed when the input image pixels are only normalized, because there is a series of operations at different layers, and the scale of features may change. To handle this problem, normalized features  $f_i^{N_r}$  to the softmax loss are included as  $f_i^{N_r} = \frac{f_i^{Op} - \mu}{\sqrt{\sigma^2}}$ , where  $\mu$  denotes the mean and  $\sigma^2$  represents the variance.



**(VIII)** Softmax layer (Hasnat et al. 2017): The loss is computed using this layer. The method for calculating softmax loss is shown in Eq. (5);

$$L_{\text{softmax}} = - \sum_{i=1}^N \log \frac{e^{w_j^T f_i + b_j}}{\sum_{j=1}^K e^{w_j^T f_i + b_j}} \quad (5)$$

where  $f_i$  denotes the features and  $y_i$  denotes the true class label of the  $i$ th image. The weights and bias of the  $j$ th class are denoted by  $w_j$  and  $b_j$ , respectively. The number of training samples is represented by  $N$  and the number of classes is represented by  $K$ .

**(IX)** SVM classification: The CNN extracted feature vector is fed into the classification module. Furthermore, a simple Support Vector Machine (SVM) is utilized because of its popularity for efficient image classification. The SVM is used as a one-against-all approach for  $N$  classes. This is the same as merging  $N$  binary SVM classifiers, each of which discriminates between a single class and all others. After that, the test sample is assigned to the class with the largest margin of error (Scholkopf and Smola 2001).

### 3.2 B. Cancelable biometric features

For generating cancelable biometric features, we apply bio-convolving (Maiorana et al. 2010). In bio-convolving, each original sequence,  $r(n), n = 1, \dots, F$ , has the transformed version  $f(n), n = 1, \dots, F$ , which is an encrypted version of the original biometric template. We have  $\mathbf{d} = [d_0, \dots, d_w]^T$ . It is obvious that the vector  $\mathbf{d}$  is the transformation key. The original  $r(n)$  sequence is combined with  $d(n)$ .

$$f(n) = r(n) * d(n) \quad (6)$$

The CNNs succeed in the recognition process, when we keep the features without significant changes. So, we need to make the transformed sequences have a zero mean. The transformed sequences have a zero mean and the standard deviation is unity, when a normalization step is applied. So, by just altering the size or values of  $\mathbf{d}$  applied on the original biometric templates, we can construct new templates.

## 4 Experimental results

Our experiments consist of training of the CNN models based on the four scenarios explained before. In order to verify the effectiveness, we experimented on different biometric traits with several datasets; namely FERET (Feret-database 2019) and LFW (Huang et al. 2007) datasets for face images, and IITD (ITTD Iris Database, 2019) and CASIA-IrisV3-Interval (CASIA-IrisV3 Database 2018) datasets for iris images. In the experimental results, performance metrics including accuracy, specificity, precision, recall, and  $F_{\text{Score}}$  have been used to evaluate the traditional, state-of-the-art, and proposed face and iris recognition systems. These metrics are computed according to Eqs. 7, 8, 9, 10, and 11 (Soliman et al. 2020, Soliman et al. 2018, and. Abd El-Samie et al. 2021);

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

**Table 2** Workstation specifications

System	Specifications
Model	64-bit Windows 10
CPU	Intel Xeon 5670, 12 cores
Graphics Card	NVIDIA GeForce GTX 1070
RAM	48-G memory

**Table 3** Training and testing parts of datasets

Dataset	No. of training images	No. of testing images
FERET	6975	2989
LFW	7084	3036
IITD	448	672
CASIA-IrisV3	852	213

$$Specificity = \frac{TN}{FP + TN} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F_{Score} = \frac{2 * Recall * Precision}{Recall + Precision} \quad (11)$$

TP stands for True Positive, FN stands for False Negative, FP stands for False Positive, and TN is for True Negative.

All experiments have been performed using a platform with the specifications given in Table 2.

#### 4.1 A. CNN training

The datasets were split into two sections, one for training and the other for testing of CNN models as illustrated in Table 3. The momentum was set to 0.9, and we used the stochastic gradient descent approach to train our CNN. Furthermore, with the weight decay set at  $5 \times 10^4$ , we used l2-norm regularization. With a 0.1 learning rate, we started the CNN training and halted after 5 epochs.

#### 4.2 B. Results and evaluation

Now, we evaluate our proposed model performance by applying the four scenarios described in Sect. 7.3 on different biometric traits with different datasets, and compare it with those of different existing models. The performance of the proposed system is

assessed using face and iris images to verify its effectiveness. LFW and FERET datasets include face images with different poses and expressions, while the IITD and CASIA-IrisV3-Interval include iris images.

**(I)** Labeled Faces in the Wild (LFW) dataset. It is considered as a challenging dataset for evaluation. Table 4 contains the results of the proposed model compared with the other state-of-the-art ones. It is obvious that the current model achieves a significant accuracy of 98.85% for recognition (1st scenario). For cancelable face recognition, the accuracy is 98.97%, and 97.94% for the 2nd, 3rd scenarios, respectively, and we achieved the highest recognition accuracy of 99.15% for the 4th scenario. In addition, we evaluated the proposed model by computing specificity, precision, recall, and  $F_{Score}$ . The proposed model has the best results. Table 4 confirms the performance of the proposed model on the LFW database.

**(II)** FERET dataset. The proposed model achieved a significant accuracy of 97.92% for recognition (1st scenario) as shown in Table 5. For cancelable face recognition, the accuracy is 98.11%, and 97.14% for the 2nd, and 3rd scenarios, respectively, and we achieved the highest recognition accuracy of 98.35% for the 4th scenario.

**(III)** CASIA-IrisV3-Interval dataset. To assess the suggested model, the dataset of CASIA-IrisV3-Interval (CASIA-V3) is used instead of the uniform pupil circle in CASIA-V1. The pupil area has been left uncovered, posing the actual difficulty of specular reflections and non-uniform intensity. The dataset is divided into 213 people, each of whom possesses samples from his right, left, or both eyes. From Table 6, it can be seen that the proposed model achieves the highest accuracy (87.48%) for recognition (1st scenario). For cancelable iris recognition, the values of accuracy are 85.23%, 86.9%, and 85.76% for the 2nd, 3rd, and 4th scenarios, respectively.

**(IV)** IITD dataset. Experiments have been carried out using the IITD iris database version 1.01, which comprises 1,120 NIR images from 224 different subjects. Table 7 shows that the proposed model achieves a significant accuracy of 97.34% for recognition (1st scenario). For cancelable iris recognition, the accuracy is 97.55%, and 97.49% for the 2nd, and 3rd scenarios, respectively, and we achieved the highest recognition accuracy of 97.89% for the 4th scenario. Figure 2 illustrates the visual activation results for the first five CNN convolutional layers for an image from each dataset.

It is clear that applying bio-convolving on both image and feature levels gives good cancelable biometric results. In addition, Figs. (3) and (4) show a comparison between recognition accuracies for the different scenarios of the proposed model for face and iris datasets.

Figs. (5, 6, 7 and 8) illustrate the ROC curves of the proposed CNN architecture compared with the state-of-the-art DeepVisage on LFW, FERET, CASIA-IrisV3-Interval, and IITD datasets, respectively.

## 5 Conclusion

The current work introduces a unique cancelable biometric recognition system based on DL. The main aim of this work is to present different scenarios to guarantee the ability to generate cancelable biometric templates based on a bio-convolving method in both image and feature levels. Classification is performed with CNNs, and four datasets were trained to examine the proposed system robustness. The results showed that the proposed CNN model can successfully classify different cancelable biometric traits with high performance

**Table 4** Performance comparison on LFW dataset

CNN	Scenario	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	$F_{Score}$ (%)
Proposed CNN	1	98.85	98.92	96.63	98.18	97.39
	2	98.97	99.13	95.27	96.82	96.03
	3	97.94	98.08	94.55	95.63	95.08
	4	99.15	99.27	96.91	97.95	97.42
ResNet-50 He et al. (2016)	1	97.81	97.86	95.12	96.24	95.67
	2	97.93	98.05	94.79	95.46	95.12
	3	96.96	97.14	93.31	94.36	93.83
	4	98.24	98.41	95.63	96.09	95.85
29-Layer CNN Wu et al. (2018)	1	96.53	96.58	93.99	95.88	94.92
	2	96.77	96.85	93.24	94.95	94.08
	3	95.45	95.58	92.52	93.47	92.99
	4	96.89	96.96	93.46	94.98	94.21
Deep Visage Hasnat et al. (2017)	1	98.15	98.22	95.49	97.59	96.52
	2	98.32	98.44	95.67	96.73	96.19
	3	97.33	97.45	94.27	95.48	94.87
	4	98.86	98.97	95.86	96.87	96.36
64-Layer CNN Liu et al. (2017)	1	97.41	97.47	94.82	96.79	95.79
	2	97.67	97.73	94.96	95.86	95.4
	3	96.12	96.26	93.38	94.56	93.96
	4	97.73	97.79	94.74	95.47	95.1

**Table 5** Performance comparison on FERET dataset

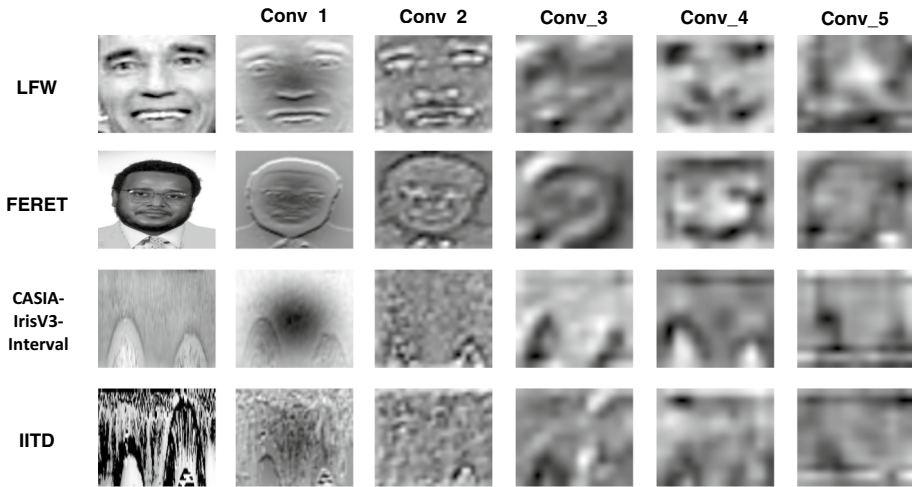
CNN	Scenario	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	$F_{score} (%)$
Proposed CNN	1	97.92	97.96	95.42	96.66	96
	2	98.11	98.25	95.54	96.59	96.06
	3	97.14	96.26	93.84	94.77	94.3
	4	98.35	98.48	95.25	96.63	95.93
ResNet-50 He et al. (2016)	1	95.33	95.39	91.48	94.88	93.14
	2	95.63	95.78	91.35	93.35	92.33
	3	94.58	94.35	91.23	92.16	91.69
	4	95.79	95.99	92.66	93.57	93.11
29-Layer CNN Wu et al. (2018)	1	93.53	93.57	90.41	92.27	91.33
	2	93.84	93.93	90.75	91.44	91.09
	3	92.62	92.58	89.33	90.36	89.84
	4	94.17	94.31	91.27	92.29	91.77
Deep Visage Hasnat et al. (2017)	1	96.93	96.99	96.08	95.44	95.75
	2	97.26	97.39	94.54	95.14	94.83
	3	96.27	95.39	92.85	93.46	93.15
	4	97.53	97.62	94.92	95.74	95.32
64-Layer CNN Liu et al. (2017)	1	94.61	94.66	90.58	93.69	92.1
	2	94.86	94.95	91.67	92.92	92.29
	3	93.88	93.45	90.22	91.83	91.01
	4	95.77	95.88	92.69	93.61	93.14

**Table 6** Performance comparison on CASIA-IrisV3 dataset

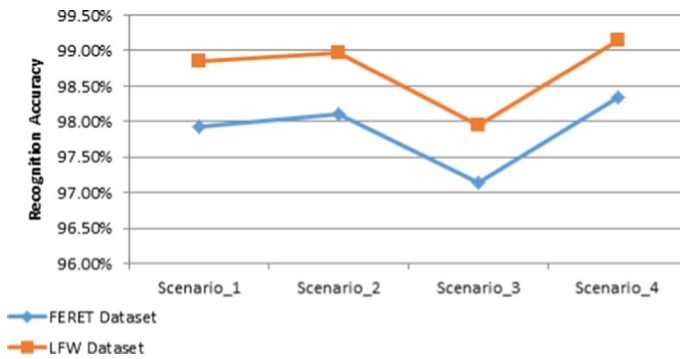
CNN	Scenario (%)	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	F <sub>score</sub> (%)
Proposed CNN	1	93.48	95.58	93.44	94.26	93.84
	2	93.23	93.29	91.13	91.95	91.53
	3	94.9	95.06	92.98	93.87	93.42
	4	93.76	93.82	91.74	92.55	92.14
ResNet-50 He et al. (2016)	1	93.64	93.68	91.54	92.37	91.95
	2	91.45	91.53	89.44	90.23	89.83
	3	93.13	93.25	91.15	91.95	91.54
	4	91.95	92.08	89.94	90.77	90.35
29-Layer CNN Wu et al. (2018)	1	89.5	89.57	87.45	88.26	87.85
	2	87	87.12	85.06	85.93	85.49
	3	89.19	89.27	87.18	88.06	87.61
	4	88.09	88.23	86.14	86.94	86.53
Deep Visage Hasnat et al. (2017)	1	94.36	94.46	92.33	93.18	92.75
	2	92.15	92.28	90.16	91	90.57
	3	93.81	93.89	91.76	92.5	92.12
	4	92.67	92.77	90.68	91.57	91.12
64-Layer CNN Liu et al. (2017)	1	91.84	91.94	89.85	90.65	90.24
	2	89.62	89.67	87.56	88.4	87.97
	3	91.31	91.37	89.24	90.05	89.64
	4	90.11	90.18	88.08	88.97	88.52

**Table 7** Performance comparison on IITD dataset

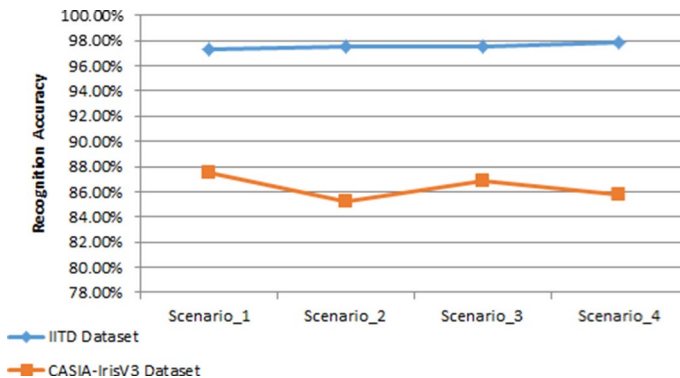
CNN	Scenario (%)	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	$F_{Score}$ (%)
Proposed CNN	1	97.34	97.42	95.27	96.12	95.69
	2	97.55	97.69	95.59	96.79	96.18
	3	97.49	97.55	95.44	96.38	95.9
	4	97.89	97.94	95.83	96.83	96.32
ResNet-50 He et al. (2016)	1	96.11	96.25	94.51	95.42	94.96
	2	96.22	96.34	94.86	95.46	95.15
	3	96.38	96.47	94.64	95.56	95.09
	4	96.44	96.49	95.03	95.93	95.47
29-Layer CNN Wu et al. (2018)	1	94.2	94.28	92.11	93.07	92.58
	2	94.49	94.57	92.62	93.82	93.21
	3	94.35	94.41	92.48	93.33	92.9
	4	94.72	94.78	92.77	93.61	93.18
Deep Visage Hasnat et al. (2017)	1	96.54	96.66	94.24	95.24	94.73
	2	96.98	96.84	94.58	95.48	95.02
	3	96.59	96.71	94.39	95.39	94.88
	4	97.18	97.23	95.22	96.12	95.66
64-Layer CNN Liu et al. (2017)	1	94.85	95.57	93.61	94.55	94.07
	2	95.21	95.89	93.88	94.75	94.31
	3	95.02	95.67	93.75	94.68	94.21
	4	95.44	95.47	94.17	95.17	94.66



**Fig. 2** Visual activation results of the first five convolutional layers of the presented system on different datasets

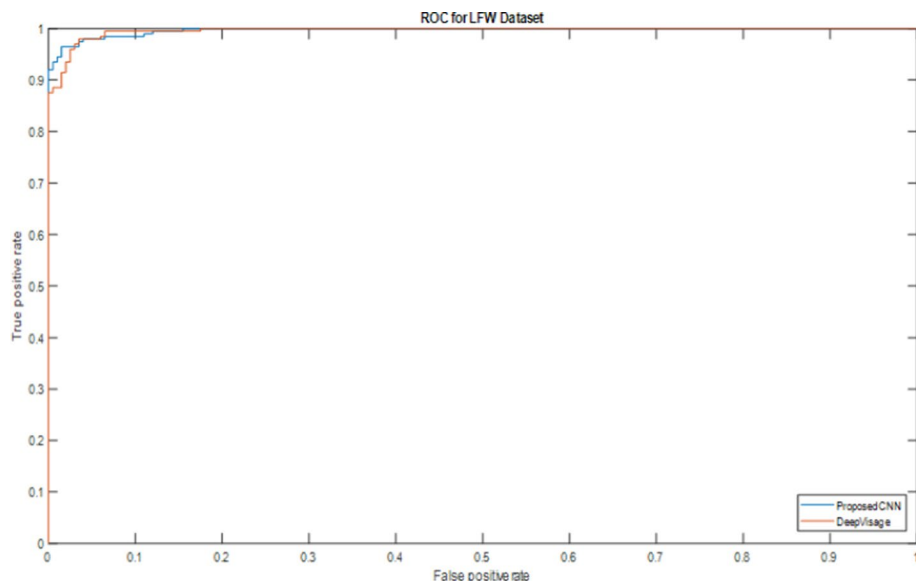


**Fig. 3** Recognition accuracy of different scenarios on FERET and LFW databases

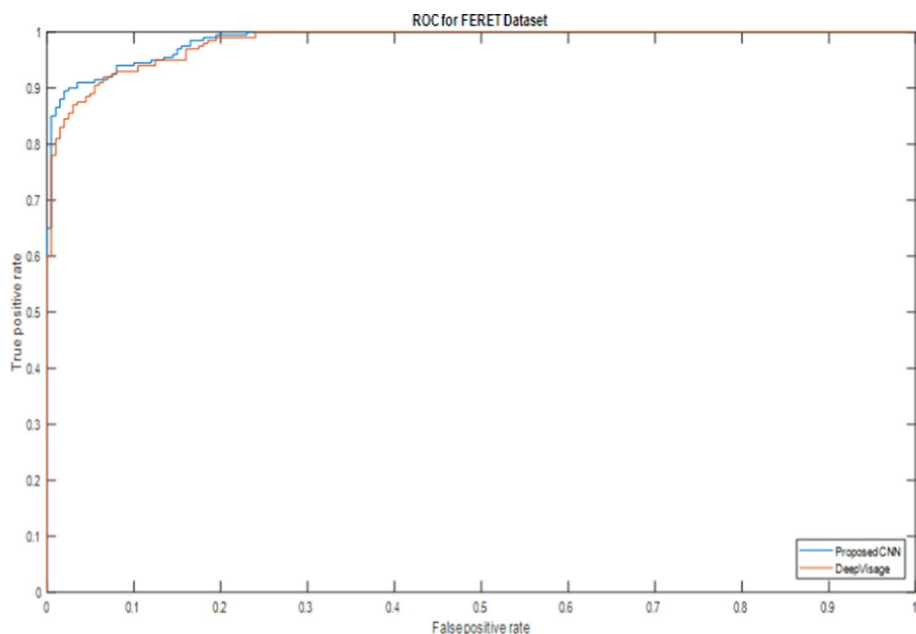


**Fig. 4** Recognition accuracy of different scenarios on IITD and CASIA-IrisV3 databases



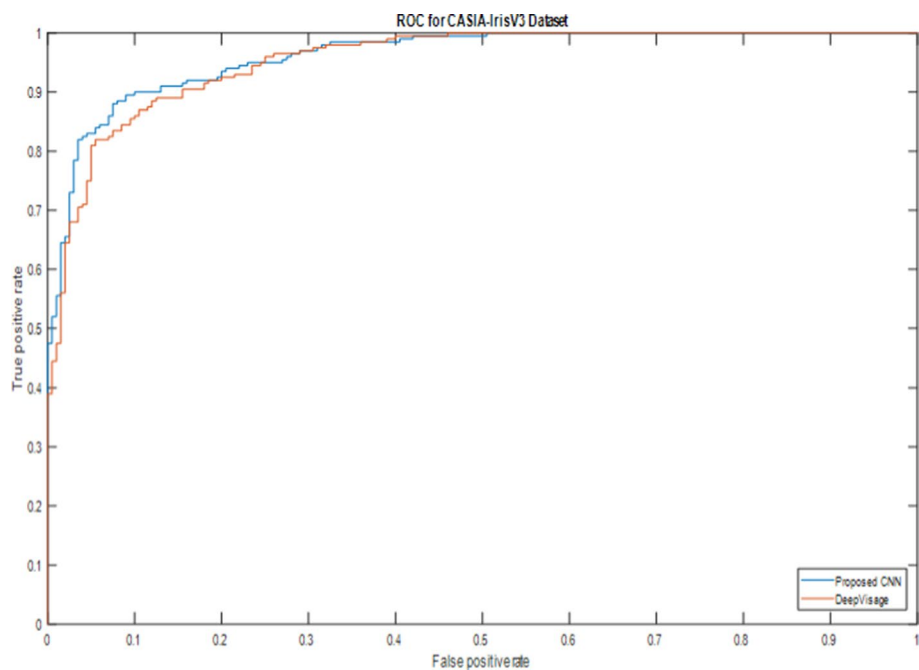


**Fig. 5** ROC curve on the LFW database

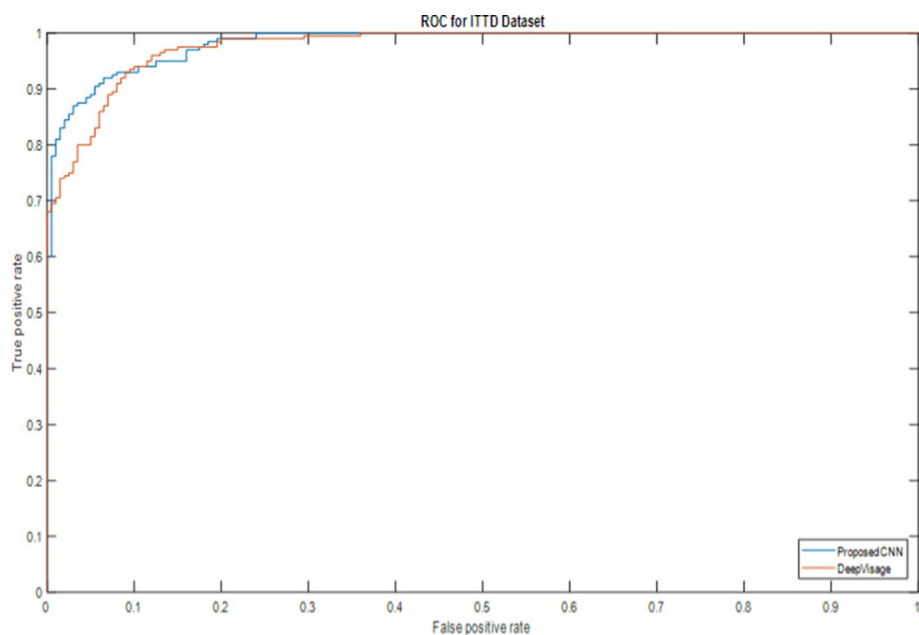


**Fig. 6** ROC curve on the FERET database

levels compared to the state-of-the-art CNNs, which proves the merit of CNNs in the field of cancelable biometric recognition. It is also clear that applying the bio-convolution on the image and feature levels achieves promising accuracy levels. To ensure robustness and



**Fig. 7** ROC curve on the CASIA-IrisV3-Interval dataset



**Fig. 8** ROC curve on the IITD dataset

effectiveness of the proposed model, performance metrics have been evaluated for different traits with different datasets: CASIA-IrisV3, IITD, FERET, and LFW. The obtained results show that the proposed CNN can extract special features from different biometric traits accurately, and solve the biometric recognition privacy problem.

**Funding** The authors have not disclosed any funding.

## Declarations

**Conflict of interest** There are no conflicts of interest declared by the authors.

## References


- Abd El-Samie, F.E., Nassar, R.M., Safan, M., Abdelhamed, M.A., Khalaf, A.A.M., Albanby, G., Zahran, O., El-Rabaie, S.M., Mohamed, A., El-Dokany, I.M., El Din Hussein, H., El-Khamy, S., Ramadan, N., Soliman, R.F., El-Shafai, W.: Cancellable multi-biometric system based on optical scanning holography. *Appl. Opt.* **60**(13), 3659–3667 (2021)
- Bengio, Y., Courville, A., Vincent, P.: Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(8), 1798–1828 (2013)
- Bolle, R.M., Connel, J.H., Ratha, N.K.: Biometrics perils and patches. *Pattern Recogn.* **35**(12), 2727–2738 (2002)
- CASIA-IrisV3 Database. <http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp>. Accessed May 2018
- Chinese Academy of Sciences Institute of Automation. CASIA iris image database. <http://biometrics.idealtest.org/>. (2017)
- Daugman, J.: Major International Deployments of the Iris Recognition Algorithms: A Billion Persons. University of Cambridge, Cambridge (2014)
- Daugman, J.: Information theory and the iriscode. *IEEE Trans. Inf. Forensics Secur.* **11**(2), 400–409 (2016)
- Feret-database. <https://www.nist.gov/itl/iad/image-group/color-feret-database>. Accessed Dec 2019
- Gangwar, A. and Joshi, A.: Deepirisnet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition. In: 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, pp. 2301–2305 (2016)
- Hasnat, A., Bohné, J., Milgram, J., Gentric, S., and Chen, L.: DeepVisage: Making face recognition simple yet with powerful generalization Skills. In: IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, pp. 1–12 (2017)
- He, K., Zhang, X., Ren, S., and Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770–778 (2016)
- Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E.: Labeled faces in the wild (LFW): A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07–49, University of Massachusetts, Amherst.
- Ioffe, S. and Szegedy, Ch.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv: 1502.03167v3 [cs.LG], pp. 1–11 (2015)
- IITD Iris atabase. <http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database> Iris.htm. Accessed Dec 2019
- Johnson, P. A., Lopez-Meyer, P., Sazonova, N., Hua, F., and Schuckers, S.: Quality in face and iris research ensemble (q-fire). In: Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), Washington, DC, USA, pp. 1–6 (2010)
- Ko, J., Mudassar, B., Na, T. and Mukhopadhyay, S.: Design of an energy-efficient accelerator for training of convolutional neural networks using frequency-domain computation. In: Proceedings of the 54th ACM/EDAC/IEEE Design Automation Conference (DAC), Austin, TX, USA, pp. 1–6 (2017)
- LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
- LeCun, Y., Kavukcuoglu, K., and Farabet, C.: Convolutional networks and applications in vision. In: Proceedings of the IEEE International Symposium, Circuits and Systems (ISCAS), Paris, France, pp. 253–256 (2010)
- Liu, N., Zhang, M., Li, H., Sun, Z., Tan, T.: Deepiris: learning pairwise filter bank for heterogeneous iris verification. *Pattern Recogn. Lett.* **82**, 154–161 (2016b)

- Liu, N., Li, H., Zhang, M., Liu, J., Sun, Z., and Tan, T.: Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In: International Conference on Biometrics (ICB), Halmstad, Sweden, pp. 1–8 (2017)
- Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., and Song, L.: SphereFace: Deep Hypersphere Embedding for Face Recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 6738–6746 (2017)
- Maiorana, E., Campisi, P., Fierrez, J., Ortega-Garcia, J., Neri, A.: Cancelable templates for sequence-based biometrics with application to on-line signature recognition. *IEEE Trans. Syst. Man Cybern. A* **40**(3), 525–538 (2010)
- Masek, L.: Recognition of human iris patterns for biometric identification. MSc. Thesis, the University of Western Australia (2003)
- Masi, I., Trần, A.T., Hassner, T., Leksut, J.T., Medioni, G.: Do we really need to collect millions of faces for effective face recognition? In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, PROCEEDINGS, Part V*, pp. 579–596. Springer International Publishing, Cham (2016). [https://doi.org/10.1007/978-3-319-46454-1\\_35](https://doi.org/10.1007/978-3-319-46454-1_35)
- Menotti, D., Chiachia, G., Pinto, A., Schwartz, W.R., Pedrini, H., Falco, A.X., Rocha, A.: Deep representations for iris, face, and fingerprint spoofing detection. *IEEE Trans. Inf. Forensics Secur.* **10**(4), 864–879 (2015)
- Muron, A., Pospisil, J.: The human iris structure and its usages. *Acta Univ Plalcki Physica* **39**, 87–95 (2000)
- NIST: IREX III-performance of iris identification algorithms. National Institute of Science and Technology, USA, Tech. Rep. NIST Interagency Report 7836 (2012)
- Parkhi, O. M., Vedaldi, A., and Zisserman, A.: Deep face recognition. In: *Proceeding of The British Machine Vision Conference*, Swansea, UK, 41.1–41.12 (2015)
- Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: Frvt 2006 and ice 2006 large scale experimental results. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 831–846 (2010)
- Ratha, N.K., Connel, J.H., Bolle, R.: Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.* **40**(3), 614–634 (2001)
- Ratha, N., Chikkerur, S., Connell, J., Bolle, R.: Generating cancelable fingerprint templates. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 561–572 (2007)
- Scholkopf, B., Smola, A.J.: *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and BEYOND*. MIT Press, Cambridge (2001)
- Schroff, F., Kalenichenko D., and Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: *IEEE Conference on Computer Vision and Pattern*, Boston, MA, USA, pp. 815–823 (2015)
- Simonyan, K., and Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *CoRR*, vol. abs/1409.1556 (2014)
- Soliman, R.F., El Banby, G.M., Algarni, A.D., Elsheikh, M., Soliman, N.F., Amin, M., Abd El-Samie, F.E.: Double random phase encoding for cancelable face and iris recognition. *Appl. Opt.* **57**(35), 10305–10316 (2018)
- Soliman, R.F., Amin, M., El-Samie, A., Fathi, E.: Cancelable iris recognition system based on comb filter. *Multimed. Tools Appl.* **79**, 2521–254 (2020)
- Sun, Y., Wang, X. and Tang, X.: Deeply learned face representations are sparse, selective, and robust. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 2892–2900 (2015)
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A.: Going deeper with convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1–9 (2015)
- Taigman, Y., Yang, M., Ranzato, M. and Wolf, L.: Deepface: Closing the gap to human-level performance in face verification. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, pp. 1701–1708 (2014)
- Taigman, Y., Yang, M., Ranzato, M. A., and Wolf, L. (2015) Web-scale training for face identification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 2746–2754 (2015)
- Tapia, J., Aravena, C.: *Gender Classification from NIR Iris Images Using Deep Learning*, pp. 219–239. Springer International Publishing, Cham (2017)
- Teoh, A., Goh, A., Ngo, D.: Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 1892–1901 (2006)
- Theodoridis, S., Koutroumbas, K.: *Pattern Recognition*, 4th edn. Academic Press, Cambridge (2008)

- Wu, X., He, R., Sun, Z., Tan, T.: A light CNN for deep face representation with noisy labels. *IEEE Trans. Inf. Forensics Secur.* **13**(11), 2884–2896 (2018)
- Yakopcic, C., Alom, M. and Taha, T.: Extremely parallel memristor crossbar architecture for convolutional neural network implementation. In: *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, pp. 1696–1703 (2017)
- Yuan, Z., Liu, Y., Yue, J., Li, J., and Yang, H.: CORAL: Coarse-grained reconfigurable architecture for Convolutional Neural Networks. In: *Proceedings of the IEEE/ACM, International Symposium on Low Power Electronics and Design (ISLPED)*, Taipei, Taiwan, pp. 1–6 (2017)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

**Essam Abdellatef<sup>1</sup> · Randa F. Soliman<sup>2</sup>  · Eman M. Omran<sup>3</sup> · Nabil A. Ismail<sup>4</sup> · Salah E. S. Abd Elrahman<sup>4</sup> · Khalid N. Ismail<sup>5,6</sup> · Mohamed Rihan<sup>7,8</sup> · Mohamed Amin<sup>9</sup> · Ayman A. Eisa<sup>3</sup> · Fathi E. Abd El-Samie<sup>7,10</sup>**

Essam Abdellatef  
essam\_abdellatef@yahoo.com

Eman M. Omran  
omran\_eman91@yahoo.com

Nabil A. Ismail  
nabil.ismail@el-eng.menofia.edu.eg

Salah E. S. Abd Elrahman  
salaheldeem@el-eng.menofia.edu.eg

Khalid N. Ismail  
khalid.n.ismail@gmail.com

Mohamed Rihan  
mohamed.elmelegy@el-eng.menofia.edu.eg

Mohamed Amin  
mohamed\_amin110@yahoo.com

Ayman A. Eisa  
Ayman\_Eisa@yahoo.com

Fathi E. Abd El-Samie  
fathi\_sayed@yahoo.com

<sup>1</sup> Electronics and Communication Department, Delta Higher Institute for Engineering and Technology (DHJET), Mansoura, Egypt

<sup>2</sup> Department of Machine Intelligence, Faculty of Artificial Intelligence, Menoufia University, Shebin El-Koom 32511, Egypt

<sup>3</sup> Department of Nuclear Safety and Radiological Emergencies, NCRRT, Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt

<sup>4</sup> Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

<sup>5</sup> School of Computing and Digital Technology, Birmingham City University, Birmingham, UK

<sup>6</sup> Information Technology Department, Faculty of Computers and Information, Menoufia University, Menoufia, Egypt

<sup>7</sup> Department of Electronics and Electrical Communications Engineering, Faculty of Electronic

Engineering, Menoufia University, Menoufia 32952, Egypt

<sup>8</sup> Department of Electrical and Information Engineering (DIEI), University of Cassino and Southern Lazio, Cassino 03043, Italy

<sup>9</sup> Faculty of Science, Mathematics and Computer Science Department, Menoufia University, Shebin El-Koom 32511, Egypt

<sup>10</sup> Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia