International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V70I1P224

Volume 70 Issue 1, 204-211, January, 2022

Implementation of Fusion of Sclera and Periocular As A Biometric Authentication System Using Deep Learning

V. Sandhya[1], Nagaratna P Hegde[2]

Department of Computer Science and Engineering, GITAM School of Technology, Bengaluru. India.
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, India

[1]svinayak@gitam.edu, [2]nagaratnaph@staff.vce.ac.in

Abstract *- In today's digitized world, individual authentication has become a customary procedure for all organizations to provide access to their stakeholders. The type of authentication has been changing from uni-modal such as a fingerprint, to multi-modal as face and fingerprint, for identifying an individual. A multi-model authentication system differs with respect to the number of times the input is captured from the user, user cooperation, willingness, and accuracy. In this proposed system, input is captured only once. The sclera and periocular features are extracted as patches from the image. The model is trained using image patches of different sizes using a deep neural network system. The resultant accuracy of the proposed system is efficient when compared with the existing multi-modal systems.*

Keywords - *Bio-Metric, Sclera, Periocular, Multi-Model, Image Patches, CNN, deep learning.*

I. INTRODUCTION

Biometric authentication is the process of identifying an individual using physical or biological traits such as fingerprint, Iris, Speech, etc. In the traditional approach, only one trait is used to authenticate a person. If one or more traits are used for identifying an individual, it is known as **multimodal Biometrics** such as the combination of fingerprint and Iris, fingerprint and face, fingerprint and palm print. These combinations require user acceptance and the number of times these inputs are captured. Different types of Bio-metric types can be combined, known as Bio-Metric ***fusion.*** If the combination is based on the inputs received from different sensors, then it is called **"Sensor-Level-Fusion",** features acquired from different sources combined into a single feature set is known as **"Feature level fusion",** the output scores from different Bio-Metric are combined to known as **"Score level fusion".**

The two new biometrics that is combined for identification are sclera and periocular biometrics. The sclera is the white part of the eye consisting of patterns of blood vessels as in Figure 1(a). The Periocular is the part of the skin that surrounds the eye, as in Figure 1(b). This part of the skin doesn't change over time and age. The new approach is to combine sclera and periocular as a multimodal biometric authentication system.

Figure1(a)Sclera

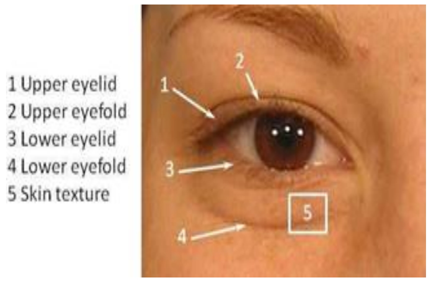


Figure 1.b) Periocular

In the existing multimodal biometric systems such as Face, Iris, and Palm Print, the input is captured multiple times from the user. Once for face, the second time for Iris, and the third time for Palm print. As user cooperation and willingness for biometric acquisition is also an important characteristic of the authentication system, capturing input three times from the user may not be an effective system. Table 1 lists out the different multimodal bio-metric system accuracy and limitations. In the proposed system, the input image is captured only once. The input image is segmented into patches, and these patches are used for identification purpose

Table 1: Existing multimodal Bio-metric Systems.

Type of Multi modal Biometrics	Proposed Method	Accuracy	Drawback/Limitation
Face,Iris and PalmPrint	Two Layer CNN	98.73%	Three times input is captured
Periocular and Iris	Feature level fusion and	92.24%	LBP
Iris and Palm Print	Sensor level	94.42	Two times input is captured
Sclera and Finger Print	Score level and GA	99.3%	Two different Input acquisition
Face and Iris	Score level and Feature level fusion	95.0	Two different Input acquisition
Fingerprint and Heartbeat	Fuzzy Based Adaptive Method.	98.82	Input accepted two times, once for Finger print and once for Heart Beat.

II. RELATED WORK

In a multimodal biometric system, input is captured from different sources such as eye, face, fingerprint, palmprint, etc. To extract the features from a multimodal system, segmentation of interest features from the biometric system is a complex task. Due to ocular occlusions and low resolution, the system may not exhibit the expected performance. To overcome these issues, a densely connected convolution encoder-decoder network is used as in [1].To identify the features, three DCCN[2] using a plain CNN and two MSF(multi-scale fusion) architectures area used to improve the performance of arrhythmia classification. In[3], the author proposed a minimal recognizable patch to identify the image by training a specialized deep network. The different type of image segmentation methods, such as semantic and instance-level segmentation by applying convolution pixel training fully, is covered in [4]. In [5], the authors propose a dual-stream CNN to identify periocular region using RGB-OCLBCP by fusing the two layers. Due to the increasing limitations of the unimodal system, [6] the author proposed a multimodal CNN by applying feature fusion strategies. The most widely used biometric authentication system is Iris. In [7], the author proposed a new deep learning method known as capsule network for Iris identification. The migration method was introduced to train the network even for small data sets. To classify images instead of training the images, [8] the author divided the images into patches using the two methods OPOD and APOD for two histopathological classes. In image processing, deep learning systems perform well when compared to machine learning methods. The author in [9] provides insights into different pre-trained models such as LeNet, AlexNet, GoogleNet, and VGG16.[10] a new CNN-based skip connection model was introduced to identify the different species of birds using the features extracted from the data set. A new approach of periocular biometrics was introduced [11] based on the chain-based CNN to find the region of interest without discarding remaining eye components such as iris and sclera. A real-time eye detector was proposed [12] to find out the eye in the facial components using CNN to classify the region as left eye or right eye. A deep Segnet architecture [13] was built to perform multi-class eye segmentation with a limited number of eye samples for an ocular biometric system. Sclera, the white part of the eye with a vein pattern, is also used for biometric identification. In [14], a new sclera segmentation method was proposed to extract the sclera part from the eye without losing any important features as an update to the OSTU method. A novel deep CNN was introduced [15] to extract features by applying two CNNs. The CNN was trained to identify the features extracted from the face and Iris for identification purposes. In [16], the author presented a study of various periocular biometrics for identification in case of low resolution, partial face occlusions, and limitations of Iris texture acquisition from a distance. To regulate the outputs of CNN,[17] conditional random fields are applied to extract the features automatically. To identify the minutiae sets in fingerprints [18] proposed an automatic minutiae extractor using CNNS. In an unconstrained environment, periocular biometrics perform well when compared to other systems. In [19] proposed a local descriptor known as LBP for feature extraction. Sclera recognition has some challenges, such as iris gaze directions, capturing the eye from a distance. The author [20] proposed a segmentation algorithm for blood vessel enhancement and feature extraction. The multimodal system of Face and Iris involves the score level and feature level fusion [21] using LBP and PCA to extract local and global features for face and iris recognition. The similarity between images is an ideal task based on the overlap of the region of interest. The author [22] proposed a new method to find the corresponding patch correspondence between images known as patch-based evaluation of Image segmentation.

III. IMAGE PATCHES

One of the important steps in Image Processing involves Image segmentation, which identifies boundaries of an Image such as edges, lines, and curves. Using the boundaries, the Image is divided into sub-parts based on the pixel values. These subparts can be used for the classification or detection of an object in an Image. Some of the limitations of segmentation are irregular objects, illumination, and overlapping of segments. Also, Image recognition algorithms fail due to incomplete occlusions. An alternative approach for segmentation is Image Patching.

Figure 2(a) Eye Image

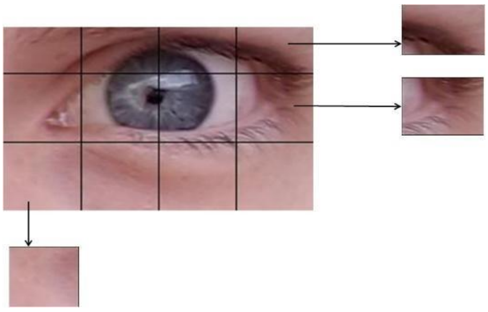


Figure 2(b) Image patches

Given an Image, I can be divided into smaller parts or regions. In Image processing, the user may consider a specific part of an image when the user tries to extract a subpart of an image that represents local features, known as a patch. Patches can be local or global [1]. A Patch P can be represented by height h and width w. The size of a Patch of an Image I is h x w pixels. Generally, the patch is chosen as a square shape. It can be rectangular or circle also. The similarity between Images patches can be calculated by comparing the pixel values of two patches for a given region. As shown in Figures 2(a) and 2(b), Image I is of size 24 x 24, and twelve patches are generated for I.

A. Image Patch generation

In this proposed system, the network is trained by considering the three different total numbers of images in the data sets. The performance of a deep learning model depends on the number of images in the training set. To generate three data sets, a data set with patches of different sizes are included. Model is trained with three datasets of images. The three data sets are named **Dataset with** 2,199 patches, **Image set** with 7,379 patches, and Images with 31,964 Patches. Figure 3(a) is the given image; the image is divided into patches using a software tool. From the generated patches, Algorithm 1 is to select only periocular patches 3b and sclera patches Figure 3(c).

Figure 3: Patch Generation

B. Algorithm 1: Selecting sclera and periocular patches

Given N number of images I_i { i=1,2,3……… N}

1. For each Image I_i :
2. Upload the Image I of size h x w('h' height and 'w' width)
3. Choose patch size as Ph x Pw (Ph ->patch height, Pw->patch width)
4. Generate non-overlapping patches $\{P_j\}_{j=1}^{M}$ of I as

   $y=f_{\Theta}(P_j)$, for j=1 ,2, ………… M,

   f_{Θ} is a function used to select patches of the image.
5. If patch P_j of Image I_i mean pixel intensity is zero, then

   remove the patch P_j from the data set
6. Repeat steps 1 to 5 for all images in the data set.

IV. CLASSIFIERS

A. Traditional Classifier

In a traditional classifier, Image classification is a method of classifying an Image by accepting an input image, process the image, and display as belonging to a particular class. The first step is to extract the features of the image such as textures, shape known as pre-processing of the image. Using the features, group the images into classes and classify Image. Figure 4(a) Image is fed to a Feature extractor. The output of the extractor is taken as input to the classifier.

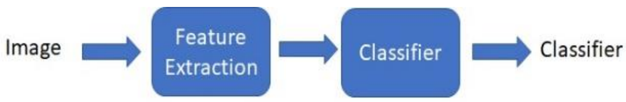


Figure 4(a) Traditional Classifier

B. Deep Learning classifier

Deep Learning Classifier is a class of Machine Learning techniques used to classify Images using N layers of the neural network. CNN is a type of feed-forward Deep neural Network that accepts raw pixels of an Image and trains to extract features of the image. Figure 4(b) input Image is passed across N layers to extract the features based on a number of layers and classify the image as belonging to a particular class.

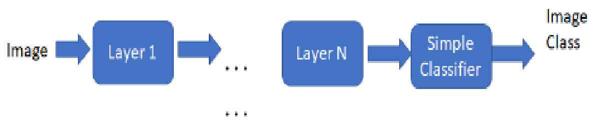


Figure 4(b)Deep Learning Classifier

C. CNN for fusion of Sclera and Periocular

In the proposed fusion of the sclera and periocular biometric system, the acquired eye image is segmented into patches. Patches corresponding to the sclera and periocular are fussed into a training data set. As shown in Figure 5, deep learning classifier CNN is used for the training and validation purpose of the biometric fusion. A test image is used to test the model by generating patches for the test Image against the trained model. The output of the model identifies as a match or a no match for the existing data set.

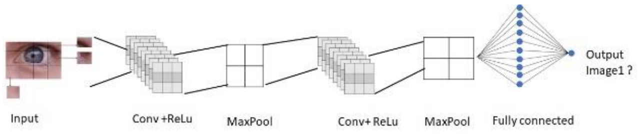


Figure 5: CNN for sclera and Periocular Fusion.

a) CNN(Convolution Neural network)

CNN has three layers Convolution layer, Pooling layer, and fully connected layer.

1) Convolution Layer.

The important building block is the convolution layer that accepts the input image, which is of three dimensions height x depth x width made up of RGB pixels. To extract the features of the image, a feature detector known as kernel or filters are moved across the image known as **"Convolution".**

The kernel size is typically 3x3 applied to the image starting from the top left corner. A dot product is calculated between image pixels and filter. The resultant dot product is fed to the next layer. The kernel is moved to the next 3 x3 areas of the image, and the process is repeated for the entire Image. The number of filters chosen determines the features extracted from the given Image. A nonlinear layer is added after each convolution operation using an activation function.

The given image is of size 6x6. Choose a patch of the image as 3 x3 and perform a dot product with a 3 x 3 filter or kernel. The result is 31. The same steps are repeated for the entire image as shown in Figure 6. a)

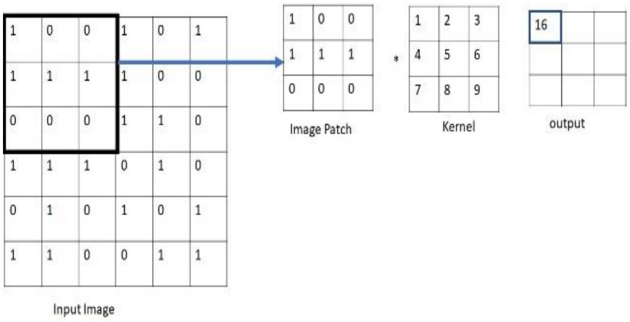


Figure 6.a)Convolution operation

For given 2D image I and a 2D filter or Kernel K

Convolution operation X can be written as

$$X=I * K$$

$$X[i,j]=\sum_{p=-k}^{k}\sum_{q=-k}^{k} I[p,q]K[i-p,j-q]$$

2) Pooling Layer.

In this layer, downsampling of the Image is used to reduce the parameters in the input. In this layer, also filter is moved over the image but without weights assigned. The kernel applies a function for the Image patch. There are two types of Pooling. Pooling is used to reduce overfitting.

i) MaxPooling

The maximum value pixel is selected from the filter of the given size. An image of size 4x4 with a filter of size 2x2. Figure 6. b)

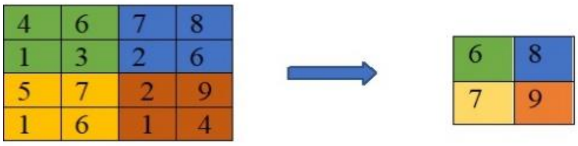


Figure 6.b) Max pooling

ii) Average Pooling

Calculates the average value of the selected filter figure 6. c)

Figure 6. c) Average pooling

3) Fully Connected Layer

This layer is used to classify the input Images based on the features extracted from the previous layer. In this layer, the SoftMax activation function is used to classify the images. Given an image of size 3x3 is converted into a vector as shown in figure 6.d)

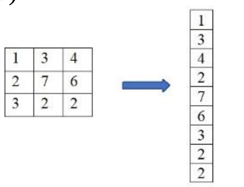


Figure 6.d) Fully connected Layer.

d) SoftMax

This function is used to compute the probabilities for all class labels, is used as the output layer for multi-class classification problems. In this proposed system to identify the images of ten classes, SoftMax is used as an activation function. The probability values of each class lie between 0 and 1.

D. Image Patch Classification Model

Let P_j ,(j=1,2,3…..M) denote M patches .The patch j is extracted from Image I, where M is the number of patches. y_{Pi} denote the true class of the image I. $y_{Pi}^{\wedge}$ = C_k, where k ϵ { 1,2,3,4,5,6,7,8,9,10} classes labels for ten Image classes. Figure 7 shows the proposed model. Image patches are assigned with corresponding Image class labels. In the next step, the models are trained using pre-trained CNN models such as VGG16 and MobileNet. To test the Image classifier, patches of the Imageis passed as input to the trained model. The proposed model predicts the score using the SoftMax classifier based on conditional probability distribution. $y_{Pi}^{\wedge}$ denotes the predicted output of the Image patch classifier.

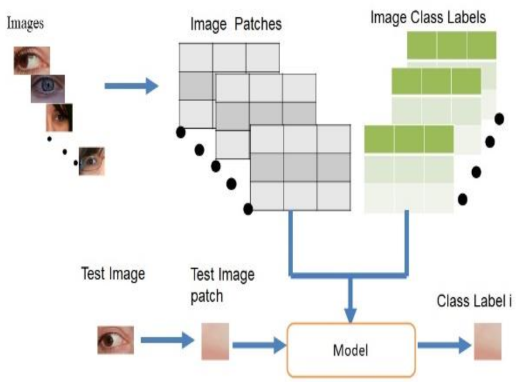
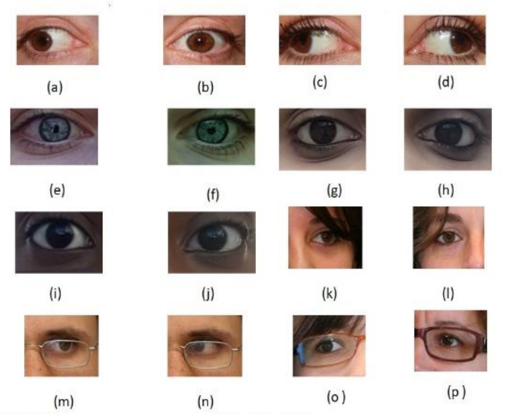


Figure 7: Sclera-Periocular Patch-based classification

V. CHALLENGING DATA SET



(a),(b),(c) ,(d) from SBVPI data set (e),(f),(g),(h),(i),(j) from VSIOB samsung_day_light_short
(k),(l),(m),(n),(o),(p) from UBIPr data set

Figure 8: challenge Data set

To train and validate the model, images are collected from three different Data sets as VISOB2 .0, UBRIS2 .0, and SBVPI. As shown in figure 8 (a),(b) are normal eye images, 8(c),(d) are images here the upper eyelid is colored. All these images are from the SBVPI data set. Figure 8(e), (f) images are with different brightness,8(g)(h) images have lower eyelids colored as black. Images from figure 8(e) to (j) are considered from the VSIOB Samsung daylight Data set. Figures8 (k) to (p) are from the UBris data set and have different images with occlusions in the form of eye spectacles and hairs covering the eye.

Visible **l**ight Mobile **O**cular **B**iometrics(VISOB) is a public data set consisting of eye Images captured from 550 healthy volunteers using smartphones iPhone 5s, Samsung Note 4, and Oppo N1. UBRIS 2.0 is a publicly available data set, has 11,000 images of Sclera Blood Vessel, Periocular, and Iris data set(SBBPI) publicly available data set and has 1858 images from 55 subjects. Table 2 list the important features of the three data sets.

Table 2 Features of three Data sets

	VISOB 2.0 T	UBRIS 2.0	SBVPI
Number of Volunteers	550	261	55
Device	SmartPhone (Samsung iPhone 5s,Samsung Note 4 and Oppo N1.)	Canon EOS 5D	(DSLR) (Canon EOS 60D
Manually cropped images	240 x160 pixels	400 x 300 pixels	3000×1700 pixel
Image Capturing	Selfie Images using front camera with 8 to 12 inches from Face	Looking at marks to rotate head and eyes with a distance of 8 and 4 meters.	20–40 centimetres
Sessions	Two visits	Two visits	Single Session
Lighting	Office Light, Day Light, Dim Light	Natural and Artificial Light sources	
Resolution	iPhone 720p, Samsung and Oppo 1080p	72 dpi	High end

VI. MODEL TRAINING

The data set has two folders, one for Training and one for testing. Each of these folders has ten folders (1,2,3…10), one for each image.

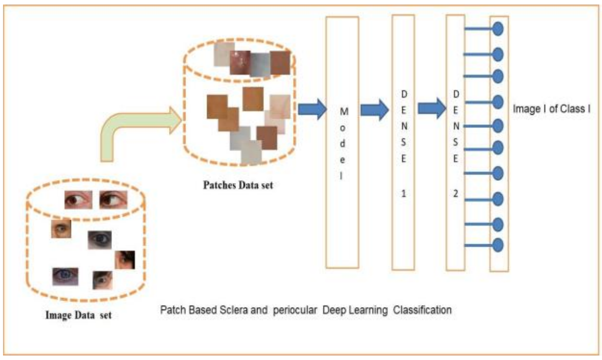


Figure 9: Block Diagram

Algorithm 1 is applied to the images in the challenge data set. From Figure 9, The resultant data set has images of the sclera and periocular patches only. On the Patch data set images, a predefined model such as VGG16 and MobileNet is applied to classify the 10 images based on patches to the resultant classes. Two dense layers are used to connect neurons to all the neurons of the previous layer. The output is flattened by using a Fully connected layer to convert the output into a one-dimensional vector.

Figure 10 shows the trained model and the layers used for training purposes. Adams optimizer is used to update network weights iteratively in the training process. The loss function used is categorical cross-entropy to classify the ten images. The image labels are assigned as 0 or 1 using one-hot encoding.

$$\text{Loss}_{(1)} = -\sum_{i=1}^{outputsize} y_i \,.\log y_i^{\wedge}$$

y_i is the target value, $y_i^{\wedge}$ is the ith value in the model output, and the output size is the number of scalar values in the model output.

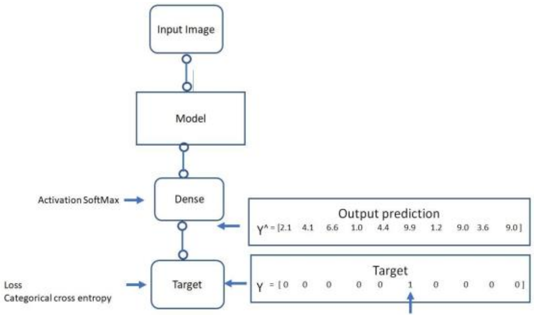


Figure 10: Model

Calculations of the trainable parameters of the model using the probability distribution function for the given image set are as follows.

$$P\left(C_i / I_j\right) = \frac{P\left(I_j / C_i\right) . P(C_i)}{P(I_j)} \quad \text{, where } I_j \text{ is assigned to class } C_i \tag{2}$$

C_i represents the Image classes as i= 1,2,..... N

I_j represents the Patches of Image I as j=1,2,....... M

$P(I_j)$ is the multivariate probabity density function of Patch j of image I

$P\left(C_i / I_j\right)$ is the conditional probability of Image class C_i given the image patch I_j

$P(C_i)$ is the prior probabilty of Image Patch in the class i

$$P\left(C_i / I_j\right) = \max P\left(C_i / I_j\right) \quad i = 1, 2, \ldots.. N \text{ classes} \tag{3}$$

$P(I_j)$ can be discarded as $P\left(C_i / I_j\right)$ is a max function

$$P\left(I_j / C_i\right) \sim \frac{I_i}{M_j}$$

$I_i$represents the image patch of class I and M_iIs the number of Training Image patches of Class i.

Equation (1) determines the classification probabilities for each class of the image data set. As the model classifies more than two or more images, the probabilities of all the classes for a given patch are determined. The output is passed to a SoftMax classifier that determines the correct Image –patch classifier. From the output, the maximum probability classifier is selected as the output and discard remaining probabilities using equations (2) and (1).

A. Adam Optimizer

To update weights in a network, iteratively, SGD (Stochastic Gradient Descent) is used to model training data. This optimizer occupies less memory and can be applied to very large data sets with more parameters. In SGD, the learning rate is maintained the same for all updates of weights. Adam is a combination of AdaGrad (Adaptive Gradient Algorithm) and RMSProp (RMS Propagation). AdaGrad updates weights for each learning rate parameter, whereas RMSProp updates average values of the gradient for each learning rate parameter. Adam optimizer utilizes weights averaged in an exponential form. In this model, Adam optimizer is applied to train the network parameters with a learning rate of **0.0001.**

B. Categorical Cross Entropy Loss

To train the model over multiple classes, we use categorical cross-entropy. Equation (4)

$$Loss\ (I) = -\sum_{i=1}^{N} y_{Pj,Ci}\ .log\left(\hat{y}_{Pj,C_i}\right) \tag{4}$$

$y_{pj,\,Ci}$ refers to the true patch Pj of the given image class Ci.

$\hat{y}_{Pj,C_i}$ Refers to the predicted patch Pj of class Ci.

VII. EXPERIMENTAL RESULTS AND ANALYSIS

The pre-trained deep learning model VGG16 and the Mobile net are used to train the patches. The training accuracy and validation accuracy of each model is shown in the table. The first data set has few image patches, the training accuracy is 99.28%, but the validation accuracy is very poor as 41% for VGG16. This indicates that the model is not able to identify the patches on the testing set. In the next training phase, the number of patches is increased to 7,379 but the same patch size as 100x100. The training accuracy is 82.5%, and validation accuracy is 77.33%, which is better when compared to the first VGG model. Now the patch size is reduced to 50x50 with 31,964 patches. The same thing is repeated by using another pretrained deep learning model MobileNet for the three data sets. From the table, we can identify that Mobile net with 7,379 patches exhibits a very good performance than all other data sets, with training accuracy as 98.87% and validation accuracy as 90%.

Table 3-Patch based sclera-Periocular authentication using VGG16, MobileNet, and Proposed CNN

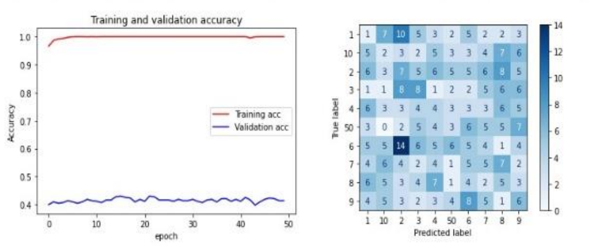
Model	Patch Size	Number of Patches of the Image	Training Image Patches	Testing Image Patches	Training Accuracy	Validation Accuracy (epochs=50
CNN	100x100x3	7,379	5958	1421	96.88% (epochs=50)	89.13% (epochs=50)
VGG16	100x100x3	2,199	1767	432	99.28% (epochs=50)	41% (epochs=50)
VGG16	100X100x3	7,379	5958	1421	82.5% (epochs=50)	77.33% (epochs=50)
MobileNet	100x100x3	2,199	1767	432	99.34% (epochs=50	49.0% (epochs=50)
MobileNet	100X100x3	7,379	5958	1421	98.87% (epochs=50)	90% (epoche=50)
MobileNet	50x50x3	31,964	26,246	5718	89% (epochs=20)	79 % (epochs=20)

VIII. PERFORMANCE OF VGG16 AND MOBILE NET ON SCLERA AND PERIOCULAR PATCHES

A. Method I (Applying VGG16)

The pretrained VGG16 is a CNN with 16 layers. This model has 3x3 Convolution layers, pooling layers of 2x2 filter size, padding as 'same', two fully connected layers, followed by a SoftMax layer. The two dense layers have 4096 elements and activation function as 'ReLu'. Upon training this model Figure 11(a) on the 2,199 patches, we achieved a training accuracy of 99.28%, validation accuracy of 41%. Since the performance of the model on the testing set is poor, the number of patches has been increased to 7,739. In the second VGG model, Figure 11(b), the training accuracy is 82.5%, validation accuracy has been improved to 77.3%. If the model is training for 100 epochs, we can achieve accuracy nearer to 99%. From the confusion matrix, we can observe that the number of misclassifications is reduced, and the efficiency of the model has been improved.

a)Training , Validation Accuracy for 50 epochs and Confusion Matrix for 2,199 patches



b) Training ,Validation Accuracy for 50 epochs and Confusion Matrix for 7,739 patches

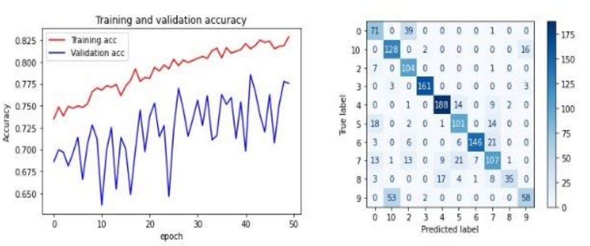
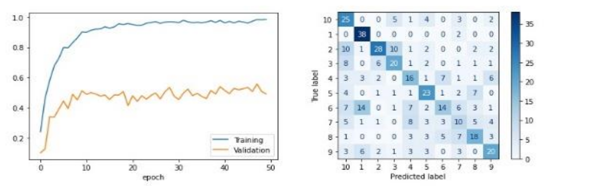


Figure 11: VGG16 performance on patch-based Sclera-periocular biometrics.

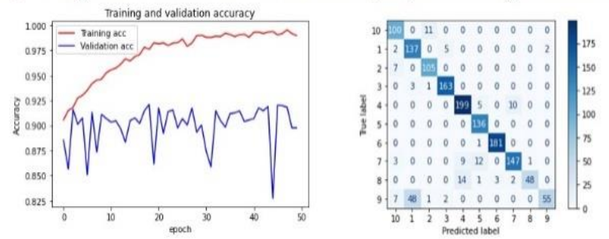
A. Method II (Applying MobileNet)

MobileNet is a depth-wise CNN with 33 layers. Each convolution layer has 2 strides. The number of filters is doubled in each layer. The three datasets in this approach are trained with the different numbers of patches, from figure 11(c) a Training accuracy of 99.34%, validation accuracy of 49% similar to that of VGG16. As the first data set size is very small, the model resulted in overfitting. We tried to increase the number of patches in the second phase of training the performance has been drastically improved to 98.87% training accuracy, Figure 11(d) 90% validation accuracy. From the confusion matrix, the classification of the image patches can be identified. In the third phase of the MobileNet training model, the patch size is reduced to 50 x 50, and the number of patches has been increased to 31,964. Using this model training accuracy Figure 11(e) of 89%, validation accuracy of 79% is achieved within 20 epochs. As depicted in the confusion matrix, the number of the sclera and periocular patches classification has been increased.

c) Training, Validation and Confusion matrix for 2,199 patches using MobileNet



d) Training, Validation and Confusion matrix for 7,379 patches using MobileNet



e) Training, Validation and Confusion matrix for 31,964 patches using MobileNet

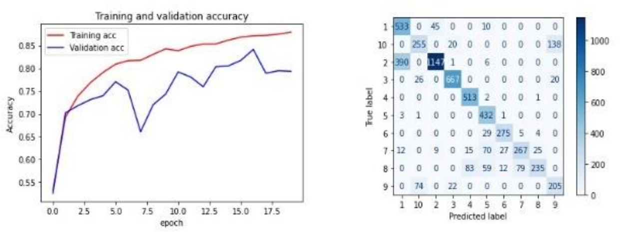


Figure 11: MobileNet performance on patch-based Sclera-periocular biometrics.

B. Method III (Applying CNN)

CNN (convolution Neural Networks) performs very well in image classification. In a traditional image classifier, features are extracted manually or by applying LBP. CNN can extract the features level by level by applying filters of specified size. In this paper, CNN has 15layers. As shown in Figure 11. f). The first three layers, each layer has 8 filters of size 3x3 that can extract the features of the sclera and periocular. In the second three layers, each layer has 16 filters of size 3x3 that can extract features from the first layer. In the third layer, three Conv2D with 32 filters at each layer of size 3x3.In the last layer also three conv2D layers are applied with 64 filters of size 3x3. At each layer Maxpool layer is applied after three convolution layers of filters of size 2x2. In all these 12 layers, Relu is used as an activation function. In order to learn the network independently of each layer, we applied batch normalization, which can normalize the output values at each layer. To update the weights at each layer w three dense layers of sizes 512 and 256 are applied. The last dense layer is 10 to classify the ten images of the data set.

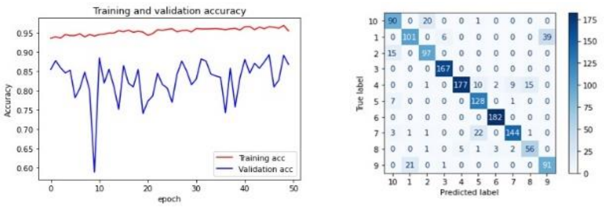


Figure 11.f) CNN performance

IX. CONCLUSION

In the proposed system, the model is trained using Image patches and classified into ten classes from a challenging data set. These patches belong to the sclera and Periocular regions of the image. The proposed model can overcome the limitation of the existing system, such as eye occlusion because of spectacle, eyelids coloring, and gaze direction. The trained model was able to identify the image of the corresponding class with improved accuracy. The testing accuracy performance has been better for a data set with a large number of patches for a given image than with a few image patches.

ACKNOWLEDGMENTS

We like to express gratitude to the University of Missouri-Kansas City for providing access to **the VISOB** data set and the Department of Computer Science, the University of Beira Interior, for providing access to the **UBIPr**and SBVPI Data set.

REFERENCES

[1] Bilal Hassan1, Ramsha Ahmed, Taimur Hassan, NaoufelWerghi, SIP-SegNet: A Deep Convolutional Encoder-Decoder Network for Joint Semantic Segmentation and Extraction of Sclera, Iris and Pupil based on Periocular Region Suppression, cs.CV, (2020).

[2] Hao Dang, Yaru Yue, DanqunXiong, Xiaoguang Zhou, Xiangdong Xu, And Xingxiang Tao, A Deep Biometric Recognition And Diagnosis Network With Residual Learning For Arrhythmia Screening Using Electrocardiogram Recordings, IEEE Access, September (2020).

[3] Mark Fonaryov, Michael Lindenbaum, On the Minimal Recognizable Image Patch, cs.CV, (2020).

[4] Shervin Minaee, Yuri Boykov, FatihPorikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos, Image Segmentation Using Deep Learning: A Survey, cs.CV, 15 (2020).

[5] Leslie Ching Ow Tiong, Yunli Lee, Andrew BengJin Teoh Periocular Recognition in the wild: Implementation of RGB-OCLBCP Dual-Stream CNN, Applied Sciences, (2019).

[6] Hui, Miao Qi, YinghuaLu, Multimodal Biometrics based on Convolution Neural Networks by Two-Layer Fusion, IEEE (2019)

[7] Tianming Zhao, Yuanning Liu, GuangHuo, Xiaodong Zhu, A Deep Learning Iris Recognition MethodBased On Capsule Network Architecture, IEEE Access, April (2019).

[8] Kaushiki Roy, DebapriyaBanik, Debotosh Bhattacharjee, MitaNasipuri. Patch-based system for Classification of Breast Histology images using deep learning, Computerized Medical Imaging and Graphics, Elsevier, (2019).

[9] SajjaTulasi Krishna, Hemantha Kumar Kalluri, Deep learning and transfer learning approaches for image classification, International Journal of Recent Technology and Engineering (IJRTE), 7 (2019).

[10] Yo-Ping Huang, Haobijam Basanta, Bird Image Retrieval and Recognition Using a Deep Learning Platform, IEEE, Access, 4 (2019).

[11] Hugo Proenc¸Jo˜ao C. Neves, Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning Frameworks, IEEE Transactions On Information Forensics And Security, 13(4) (2018)

[12] Bin Li, Hong Fu , Real Time Eye Detector with Cascaded Convolutional Neural Networks, Hindawi Applied Computational Intelligence and Soft Computing Volume (2018).

[13] Peter Rot,ZigaEmersic,Vitomir Struc,Peter. Deep Multi-class Eye Segmentation for ocular Biometrics,IEEE, (2018).

[14] Wei Dong, Han Zhou, Dongu Xu, A New Sclera Segmentation and Vessel Extraction Method for Sclera Recognition , IEEE (2018).

[15] Supreetha Gowda HD, Mohammad Imran, Hemantha Kumar G (2018)

[16] Fernando Alonso Fernandez , Josef Biguna,' A Survey on Periocular Biometrics Research,ELSVEIR, (2018).

[17] Russel Meshab, Brendan McCane ,Steven Mills ,Conditional Random Fields Incorporate CNN for Human eye Sclera Semantic Segmentation, IEEE (2017)

[18] Dinh-Luan Nguyen, Kai Cao and Anil K. Jain, Robust Minutiae Extractor:Integrating Deep Networks and Fingerprint Domain Knowledge, cs.CV, 26 (2017).

[19] K.Kishore Kumar, Movva Pavani, LBP Based Biometric Identification using the Periocular Region, IEEE, (2017).

[20] S.Alkassar, W.L.Woo,S.S.Dlay and J.A.Chambers Enhanced Segmentation and Complex-Sclera Features for Human Recognition with Unconstrained Visible–Wavelength Imaging, (2016).

[21] Maryam Eskandari Onsen Toygar Hasan Demirel, Feature extractor selection for face-iris multimodal recognition, (2014).

[22] Christian Ledig_ Wenzhe Shi Wenjia Bai Daniel Rueckert, Patch-based Evaluation of Image Segmentation, CVR (2014).