Received: February 24, 2022. Revised: April 1, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Hybrid Deep Learning Model Based on Autoencoder and CNN for Palmprint Authentication

Firas Muneam Bachay[1]* **Mohammed Hasan Abdulameer[2]**

[1] *Department of Computer Science, Faculty of Computer Science and Mathematics, University of Kufa, Iraq*
[2] *Department of Computer Science, Faculty of Education for Women, University of Kufa, Iraq*
* Corresponding author's Email: firasm.bachay@uokufa.edu.iq

Abstract: Palmprint authentication has received a lot of attention as one of the most prevalent biometric methods. A palmprint is a portion of the palm's surface that has special characteristics that could be used for authentication. Getting the most valuable features out of a palmprint, on the other hand, is a major challenge. Another challenge is coming up with an efficient strategy to authentication that uses fewer images, especially with approaches that require a high number of images in the training phase, which is a major issue. The majority of recently developed approaches rely on primary lines, wrinkles, and creases, which in some situations are insufficient to separate two people due to resemblance. Deep learning methods have recently been viewed as a critical component for extracting deep features such as texture features in these types of tasks. We concentrated on the palmprint authentication challenge in this work by creating a hybrid model called AE+CNN, which is based on the autoencoder (AE) model and a convolutional neural network (CNN) model. The proposed model comprises three phases: pre-processing, region of interest extraction (ROI), and feature extraction and matching using hybrid AE+CNN. The experiments used the COEP palmprint database, which has a limited number of palmprint images, posing a significant challenge for deep learning models that require a large number of images for training. The F1-score and the accuracy metric were both employed in the evaluation. with a score of 97.85 % for accuracy and 96.81 % for F1-score. Gradient-weighted class activation mapping (Grad-CAM) was also applied to figure out which parts of the palmprint are the most discriminative for class classification.

Keywords: Palmprint, Biometric, Deep learning, Autoencoder, Authentication.

1. Introduction

Biometrics is a method for identifying or verifying an individual's identity based on their physical or behavioural traits. Human identity and recognition have relied on biometric traits such as the face, iris, fingerprint, hand shape, palmprint, and signature [1]. Among all of these capabilities, palmprint recognition has recently acquired a lot of interest as a viable personal identification method. In terms of principal lines, creases, and wrinkles, Palmprint contains a lot of detail. Principal lines, geometry, and texture have all been studied for discerning features [2].

Several recent studies have worked on palmprint authentication using various methods, where researchers suggest in [3] a palmprint feature extraction and recognition method based on double half-orientation with three different datasets. Though, they didn't represent the orientation property of cross points with a single dominating orientation. They did not rely on extracting other types of distinct features accessible in the palmprint, instead depending on just one. The authors in [4] proposes palmprint recognition system (PRS), which combines direction, local binary pattern (LBP) features, C5.0, and K-nearest neighbour (KNN) techniques for two datasets and shows promising results. The proposed method, on the other hand, separates the image into several little sections, resulting in the loss of many critical elements that are dependent on the entire image structure. The study [5] uses the dual-tree complex wavelet

Received: February 24, 2022. Revised: April 1, 2022.

transform to offer a palmprint detection image feature descriptor (DT-CWT). The approach makes use of spatial structural information, which divides each image into many subblocks. As a result, many crucial characteristics that depend on the entire image structure are lost. Also, database is also acquired by a scanner, which reduces user satisfaction and inconvenient situations. By integrating the discrete cosine transform (DCT) and an autoregressive (AR) signal modelling a palmprint identification method is introduced by [6]. Nevertheless, the technique is interested in the texture features of palmprints and ignores other features. The study [7] suggests a palmprint personal identification system based on the merging of local and global data. The discrete orthonormal stockwell transform is used to extract the local features of the improved palmprint. By lowering the scale of the discrete orthonormal stockwell transform to infinity, the global feature is achieved. Though, it is difficult to precisely get the local features due to the sensitivity of the palmprint picture, since the performance of the local features declines if the palmprint quality and capture area are inadequate. Furthermore, the combination with the global features took a long time to compute. Deep learning approaches have recently been recognized as effective methods for a variety of domains, one of which is biometrics. For instance, convolutional neural networks have been investigated with palmprints by [8]. The experimental results showed that this approach achieves very good accuracy on the PolyU dataset. Though, the suggested network is made up of eleven convolutional layers, which increases the model's computational complexity while also increasing network training time. Likewise, the researchers in [9] suggest an enhanced deep convolutional generative adversarial net (DCGAN) to create high-resolution palmprint images by swapping the convolutional transpose layer by linear upsampling and introducing the structure similarity (SSIM) index into the loss function. The authors employ traditional data augmentation techniques and proposed deep convolutional GAN to increase data (DCGAN). Conversely, the majority of the researches showed effective results without using any data augmentation techniques. In [10] also employs the Alexnet convolutional neural network (CNN) structure. At initial, only the ROI region of the palmprint was cut away. The ROI region is then used as the input to the convolutional neural network after it has been processed. They used a standard convolutional neural network architecture only and the database in their work is massive as well as multispectral in nature. Which is make that the training duration is excessively long and also the training images' environment is not realistic. The researchers in [11] developed a deep neural network termed a palm convolutional neural network to solve the palmprint verification challenge (PCNN). The model has a simple structure with only one convolutional layer, which may not be able to extract deep and important characteristics in the palm. In recent times, autoencoder a deep learning model has demonstrated impressive image reconstruction skills based on image discriminative properties with feature reduction skills [12, 13, 14]. In this paper, we introduce a new hybrid model based on an autoencoder and a convolutional neural network (AE+CNN) that combines the advantages of both models into a single hybrid model. The following is how the rest of the paper is ordered: Part 2 gives the problem statement; Part 3 gives the theoretical background; Part 4 the proposed palmprint authentication technique; Part 5 displays and discusses the experimental results; and Part 6 conclusion the paper.

2. Problem statement

Palmprint authentication necessitates the extraction of palmprint features prior to classification, which has an essential effect on the classification rate. The most significant features, on the other hand, have a direct impact on the ultimate authentication process. Another issue is classification of the extracted palm features, which can be a problem for any authentication method, therefore choosing the optimal classification approach is crucial too. The majority of traditional solutions rely on fundamental lines, wrinkles, and creases features that are missing to distinguish two people precisely due to likeness. Also, there are many advanced features on the palm that can be utilized to create authenticity. Furthermore, past palm systems relied on touch directly between the palm pattern and the capture system device, which could reduce user adoption. As a result, current research has concentrated on contact-free solutions, which make it more pleasant and hygienic by eliminating the need for physical contact. Furthermore, the acquisition device may be costly for capturing high-resolution photos, particularly for the palmprint biometric feature. For several works, this resulted in the use of low-resolution hand images in the acquisition module. Deep learning approaches like CNN have recently been successful in a variety of disciplines, including biometrics. However, the vast majority of deep learning models,

Received: February 24, 2022. Revised: April 1, 2022.

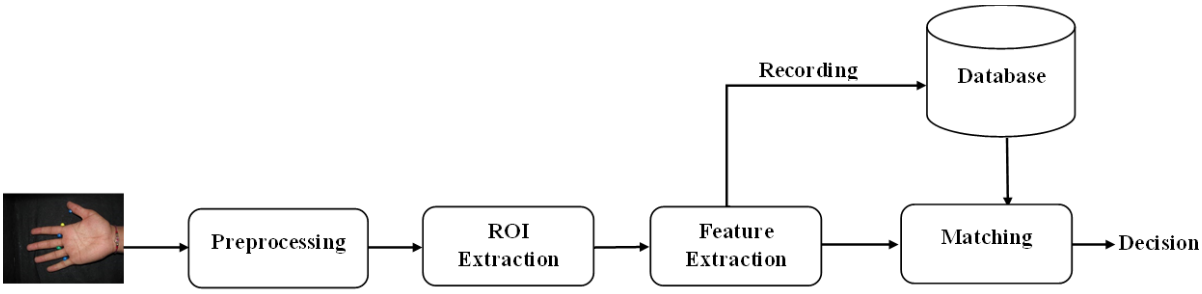


Figure. 1 The standard structure of the palmprint authentication procedure

on the other hand, necessitate a huge dataset in order to process and train the neural network. According to the over mentioned issues, our objectives are to provide a deep network-based approach that increases accuracy, lowers the cost of the biometric system, effective with little data, and enhances user acceptance by focusing on only the most important aspects of the image. Therefore, we developed a hybrid model that can benefit from both an autoencoder network and a CNN simultaneously.

3. Theoretical background

This section provides an overview of palmprint authentication before moving on to a quick explanation of deep learning networks (autoencoder and CNN).

3.1 Palmprint authentication procedure in general

Palm print collecting, pre-processing, region of interest extraction (ROI), feature extraction, matching, and decision are the five components of a typical palmprint authentication system, as depicted in Fig. 1.

3.2 Autoencoder network

An auto-encoder is a deep neural network that encodes and decodes data effectively via unsupervised feature learning. It can learn robust features from unlabelled data automatically. This approach is divided into two stages: encoding and decoding. contains one or more hidden layers with an input and output layer. The input is compressed into a lower-dimensional feature with a meaningful representation during the encoding stage. Fig. 2 depicts the simple architecture of an autoencoder [15].

3.3 Convolutional neural network

Deep learning is an artificial neural network with numerous layers. One of the most widely used deep neural networks is CNN. Fig. 3 shows CNN's

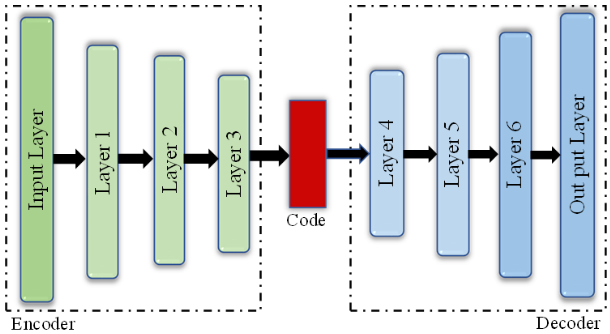


Figure. 2 The standard structure of autoencoder

numerous layers, which include a convolution layer, a pooling layer, and a fully-connected layer (FC). Convolution is the initial layer of CNN, and it consists of a set of filters. These filters are initialized by the convolutional neural network to make them more suitable for the task at hand. There is potential to add additional layers after the input layer to improve the efficiency of this method. Each layer can have its own set of filters. Stride, padding, and an activation function are among the various adjustments available. Eq. (1) explains the computations for this layer.

$$M_{r,x,y^f} = B_y^f + \sum_{i=-k_h^f}^{k_h^f} \sum_{j=-k_w^f}^{k_w^f} \sum_{y^{f-1}=1}^{y^{f-1}} W^{y^f}{}_{i+k_h^f,j,k_w^f,y}{}^{f-1}m_{r+i,x+j,y^{f-1}} \quad (1)$$

Where: $M_{r,x,y}{}^f$ is an output of the convolutional layer, (r,x) is a pixel coordinate, B_y^f is a channel bias, $W^{y^f}_{i,j,y^{f-1}}$ is the kernel weighs, k_w^f and k_h^f are respectively width and height of convolutional layer kernel, y is channel number, f is the present layer, and $f-1$ is the before layer.

Pooling is the next layer following convolution. Pooling's main purpose is to down-sample succeeding layers to minimize their complexity. It's equivalent to decreasing the resolution in the context of image processing. Pooling has no effect on the number of filters. Max-pooling is one of the most diffuse kinds of pooling strategies. It separates the image into sub-region rectangles and only returns the largest value of each. The maximum pooling computations can be demonstrated using Eq. (2).

$$q_{a^l,b^l,y} = \underset{0\le a<p_h,0\le b<p_w}{MAX} \quad O_{a^l\times ph+a,b^l\times pw+b,y} \qquad (2)$$

Where: $q_{a^l,b^l,c}$ is an output of pooling layer, $0 \le a^l < p_h^l, p_h^l$ is height of pooled channel, $0 \le a^l < p_w^l, p_w^l$ is width pooled channel, $0 \le y < y^l = y^{l-1}$, p_h is the height of the pooled window and p_w is width of the pooled window.

The last layer is the fully-connected layer where each node in it is directly related to every node in the layers above and below it. A fully connected layer also contains a large number of parameters that necessitate complex training computing. As a result, the dropout technique can be used to minimize the number of nodes and connections in a network [16]. Eq. (3) shows how this layer is calculated [11].

$$g_r = \sum_{a=1}^{m_1^{l-1}} \sum_{d=1}^{m_2^{l-1}} \sum_{e=1}^{m_3^{l-1}} W_{a,d,e,r}^l Z(e)_{a,d}, \forall 1 \le r \le m^l \qquad (3)$$

Where: g_r is an output of fully connected layer, m_1^{l-1} is width of previous channel, m_2^{l-1} is height of previous channel, m_3^{l-1} is number of previous channels, $Z(e)_{a,d}$ is vector of pooling layer outputs, $W_{a,d,e,r}^l$ is the weights between the pooling and fully connected layers, and m^l is the required number of categories.

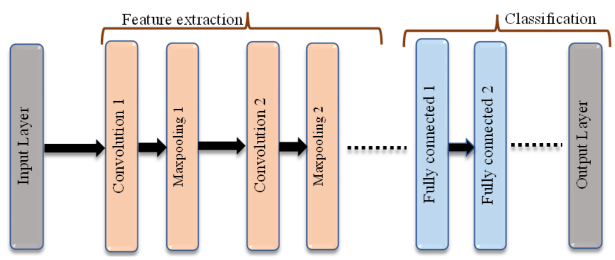


Figure. 3 The standard CNN structure

3.4 Hybrid network

A hybrid model is made up of two or more generic machine learning models that are combined to improve the model's overall performance. Typically, a single machine learning algorithm is created for a specific purpose. When two or more algorithms are integrated, the hybrid model's performance improves dramatically. CNN+GAN, CNN+AE, GAN+RL...etc. examples to hybrid models [15].

4. The proposed palmprint authentication technique

The suggested palmprint authentication technique is described in this part, and it comprises of two primary stages: region of interest (ROI) extraction and feature extraction and matching using the hybrid model. The proposed palmprint authentication technique is shown in Fig. 4.

4.1 Region of interest extraction

The ROI of palmprint images must be split before characteristics can be extracted. For the region of interest extraction approach in this paper, we used the method described by [17]. To smooth down the noise in the input palmprint, the low-pass Gaussian filter is applied initially. The palmprint border is calculated using a boundary tracking algorithm after the smoothed image is transformed to a binary image by direct thresholding. Third, reference points are located at the bottom of the gaps among the index and middle fingers, as well as the ring and little fingers. Fourth, the perpendicular bisector of the line segment among two reference landmarks is set up to locate the point of ROI. Finally, at a specific place, the subimage is cropped and scaled [17,18]. The ROI of the palmprint image in this paper is 192×192 pixels. The ROI steps can be briefed as illustrated in Fig. 5.

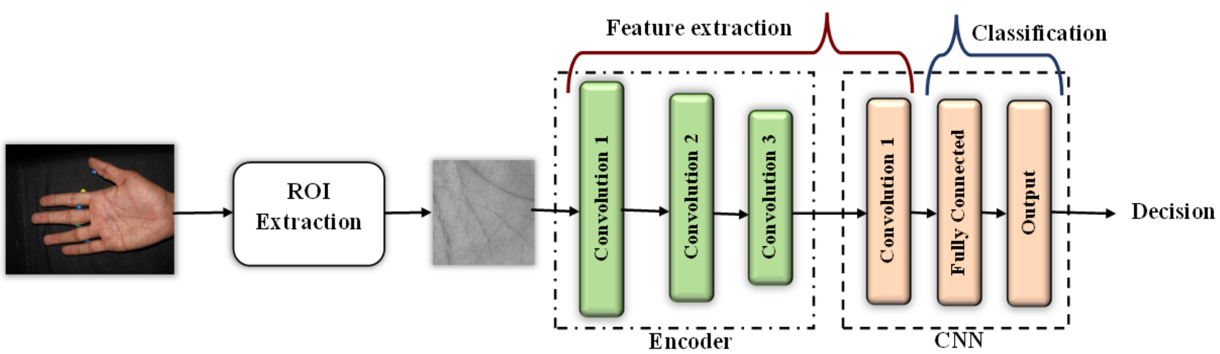


Figure. 4 The suggested palmprint authentication technique

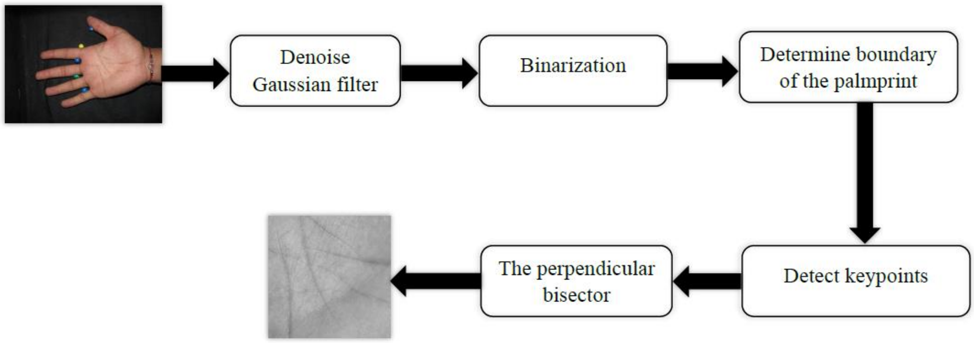


Figure. 5 The region of interest extraction steps

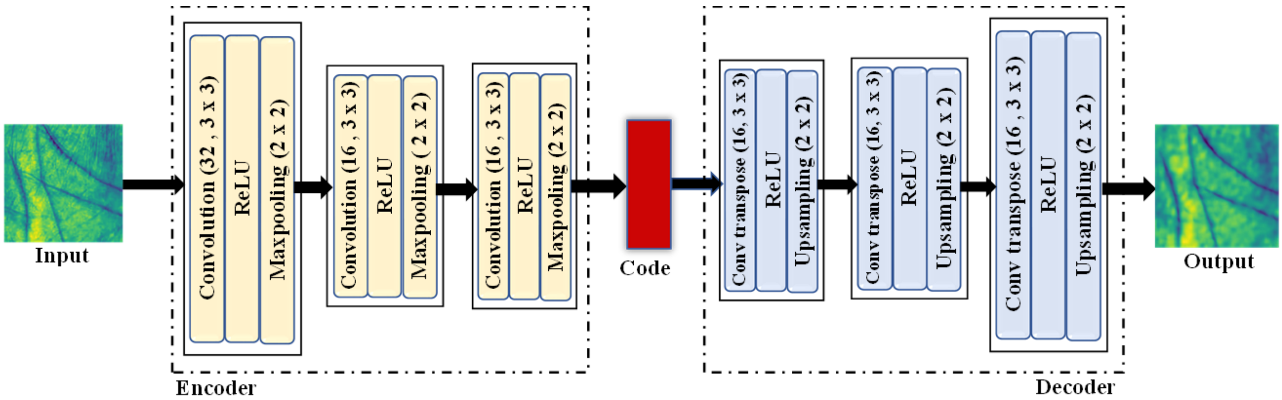


Figure. 6 The proposed autoencoder model

4.2 Autoencoder model training

The many layers that appear in the proposed autoencoder-based initial model are described in depth in this subsection. We chose the deep convolutional autoencoder because we are dealing with images. There are multiple layers to the model. The input layer of the encoder is set to receive a grayscale ROI palmprint image with a resolution of 192×192 pixels. After that, the convolution layer is used. There are three convolution layers proposed. The activation function of each convolution layer is a rectifier linear unit (ReLU). After each convolution layer, the pooling layer is applied. Windowing and maximum operations are employed. During the decoding process the encoder employs three convolutional transpose layers, which is the same number as the encoder's convolutional layers. A rectifier linear unit (ReLU) activation function is present in each convolutional transpose layer. Each convolution transpose layer is followed by an upsampling layer. The output layer is set to create a grayscale ROI palmprint image with the same dimensions as the input image. The autoencoder model is used to train the network to extract powerful features, and then the trained encoder part is combined with the CNN network to create the suggested authentication method. Table 1 summarizes the proposed autoencoder architecture. In addition, the following Fig. 6 depicts the structure of the suggested autoencoder model.

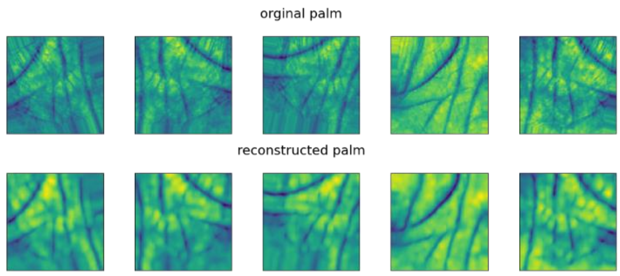


Figure. 7 The reconstructed images

Furthermore, the suggested autoencoder model's reconstructed images are shown in Fig. 7, demonstrating that the features derived from the reconstructed image using the autoencoder are expected to be robust and important.

4.3 Feature extraction and matching via hybrid model

The proposed hybrid model, which we developed and built exclusively for palmprint images, is described in depth in this subsection. The

Table1. The suggested autoencoder architecture

Layer Kind	Filter size	No. of filter	Input Shape	Output Shape	No. of parameter
Convolution 1	3x3	32	192,192,1	192,192,32	320
Max pooling 1	2x2	-	190,190,32	96,96,32	0
Convolution 2	3x3	16	96,96,32	96,96,16	4624
Max pooling 2	2x2	-	96,96,16	48,48,16	0
Convolution 3	3x3	16	48,48,16	48,48,16	2320
Max pooling 3	2x2	-	48,48,16	24,24,16	0
Conv transpose 1	3x3	16	24,24,16	24,24,16	2320
Upsampling 1	2x2	-	24,24,16	48,48,16	-
Conv transpose 2	3x3	16	48,48,16	48,48,16	2320
Upsampling 2	2x2	-	48,48,16	96,96,16	-
Conv transpose 3	3x3	32	96,96,16	96,96,32	4640
Upsampling 3	2x2	-	96,96,32	192,192,32	-
Output	3x3	1	192,192,32	192,192,1	289
Total parameters: 16,833 Trainable parameters: 16,833 Non-trainable parameters: 0					

Table2. The suggest hybrid model architecture

	Layer Kind	Filter size	Activation Function	No. of filter	FC units Layer	Input Shape	Output Shape	No. of parameter
Encoder	Convolution 1	3x3	ReLU	32	-	192,192,1	190,190,12	320
	Max pooling 1	2x2	-	-	-	190,190,32	96,96,32	0
	Convolution 2	3x3	ReLU	16	-	96,96,32	96,96,16	4624
	Max pooling 2	2x2	-	-	-	96,96,16	48,48,16	0
	Convolution 3	3x3	ReLU	16	-	48,48,16	48,48,16	2320
	Max pooling 3	2x2	-	-	-	48,48,16	24,24,16	0
CNN	Convolution 4	3x3	LeakyReLU	8	-	24,24,16	22,22,8	1160
	Max pooling 4	4x4	-	-	-	22,22,8	5,5,8	-
	Flatten	-	-	-	-	-	200	-
	Fully connected	-	Leaky ReLU	-	512	-	-	102912
	Dropout (rate = 0.5)	-	-	-	512	-	-	0
	Fully connected	-	Softmax	-	163	-	-	83619
	Total params: 194,955 Trainable params: 187,691 Non-trainable params: 7,264							

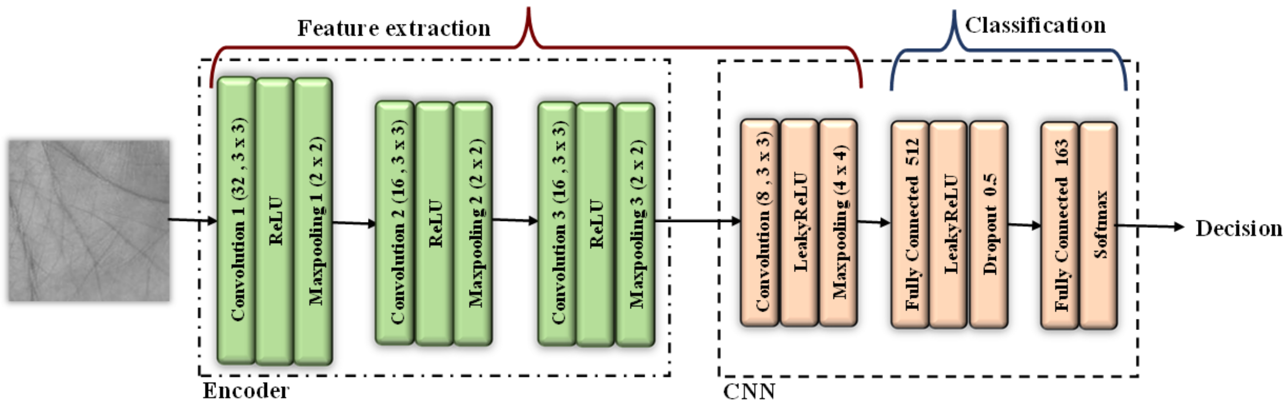


Figure. 8 The proposed hybrid model

encoder element of the proposed autoencoder model and a proposed convolutional neural network make up the model. We first used the encoder portion, which had been trained and prepared to accept a grayscale ROI palmprint image with a size of 192×192 pixels and extract the majority of the

Received: February 24, 2022. Revised: April 1, 2022.

features. Then the proposed convolutional network is added. There are numerous layers to CNN: In the beginning, the convolution layer which has a rectifier linear unit (leaky ReLU) activation function, which is based on the ReLU notion as well. There is a difference only when the input value is negative. Leaky ReLU multiplies the value by a small integer α instead of zeroing it like ReLU does (generally 0.01). As a result, the negative part gains a value, although a very small one. It's an attempt to solve the dying ReLU problem. After the convolution layer, the pooling layer is applied. Windowing and maximum operations are employed. After that, a fully connected layer is applied. This layer adjusts for the number of nodes in the before layer and the number of categories needed. After the fully connected layer, we add a dropout layer to produce a model that has better generalization and is less prone to overfitting the training data. Finally, we add a final layer that consists of softmax functions, which are widely used to solve multiple categorization problems. This layer has been modified to accept the number of categories. The appropriate size of the suggested hybrid model design is determined empirically by gradually modifying the number of convolutions and max-pooling, then the number of filters, and finally selecting the network with the highest performance. Table 2 summarizes the proposed hybrid model design. The proposed hybrid model's structure can be described by the Fig. 8.

5. Experimental results

The suggested palmprint authentication technique is built using Google Colab, which provides Python3 in a free environment, using a GPU Nvidia Tesla k80 12GB processor with 13GB of RAM. Furthermore, the region of interest is extracted using MATLAB v13.a. The performance of the recommended technique is evaluated using the COEP palmprint database, taken from the college of engineering, Pune-411005, and the proposed authentication technique as a whole is evaluated in terms of accuracy

5.1 COEP dataset

According to the file's attribute detail, the palm images in this database were captured with a Canon PowerShot SX120 IS camera with a resolution of 1600×1200×3 pixels and an image density of 180 dots per inch (DPI). The database took a year to compile, and the download collection has 1304 photos of 163 palms, each containing eight images. Fig. 9 shows palmprint samples from the COEP dataset.

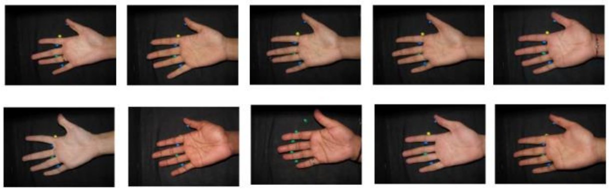


Figure. 9 Samples of palmprint images from the COEP database

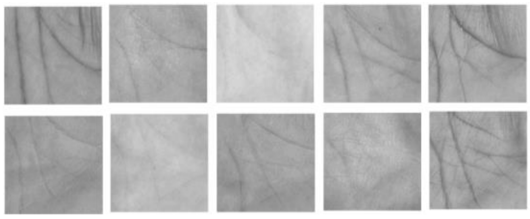


Figure. 10 Samples of region of interest

5.2 Region of interest extraction results

We utilized [17] to extract ROIs from 1304 images from the COEP dataset, with each extracted ROI image having a size of 192×192 pixels and resulting in a grayscale image. Fig. 10 shows a variety of ROI palm print images.

5.3 Evaluation metrics

The proposed technique is assessed using accuracy and F1-score metrics in this study. Eq. (4) defines the first metric, which is accuracy [19, 20].

$$ACC=\frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

Where TP (true positives) is the numeral of palmprints that are classified correctly and recognized, TN (true negatives) represents the numeral of palmprints that are negatively classified, FP (false positives) denotes the numeral of palmprints that are classified incorrectly, and FN (false negatives) indicates the numeral of palmprints that are identified as misclassified. The harmonic mean of precision (P) and recall (R) is the F1-score, which is calculated as follows Eq. (5).

$$F1\_score=\frac{2\times P\times R}{P+R} \tag{5}$$

Precision and recall are two evaluation metrics applied to assess the effectiveness of a procedure, and they are represented in Eqs. (6) and (7) [20].

$$P=\frac{TP}{TP+FP} \tag{6}$$

$$R=\frac{TP}{TP+FN} \tag{7}$$

 DOI: 10.22266/ijies2022.0630.41

Received: February 24, 2022. Revised: April 1, 2022.

Table 3. The performance results of assessing the proposed technique

No. of filters in Conv	Maxpooling	Activation function	No. of hidden layers	Accuracy	Precision	Recall	F1_score
16	4x4	Leaky ReLU	1x512	96.93%	98.43%	96.32%	96.61%
12	4x4	Leaky ReLU	1x512	95.71%	96.58%	95.40%	95.44%
8	4x4	Leaky ReLU	1x512	97.85%	97.53%	96.93%	96.81%
8	6x6	Leaky ReLU	1x512	93.25%	95.81%	91.10%	91.94%
8	8x8	Leaky ReLU	1x512	75.77%	93.68%	54.60%	57.91%
8	2x2	Leaky ReLU	1x512	95.09%	96.88%	95.09%	95.15%
8	4x4	ReLU	1x512	92.02%	95.81%	91.10%	91.88%
8	4x4	ELU	1x512	96.01%	97.20%	95.71%	95.71%
8	4x4	Swish	1x512	94.17%	95.02%	93.56%	93.60%
8	4x4	Leaky ReLU	1x256	95.40%	96.26%	94.79%	94.74%
8	4x4	Leaky ReLU	2x256	95.40%	97.16%	94.48%	94.81%

5.4 Evaluation of the hybrid model via accuracy and F1-score

The ROI was retrieved and divided into 978 training palmprint images and 326 testing palmprint images. After that, the network is trained on Keras, a prominent deep learning platform. Many tests were built and analyzed to find the proper autoencoder model parameters. After obtaining the best results, the coding part was used in the hybrid model with CNN. In hybrid model, the RMSprop algorithm was used with a learning rate of 0.001, and 100 epochs of training. Many experiments were set up and analyzed in order to determine the right CNN parameters in the hybrid model. After multiple rounds of training, the best accuracy rate was 97.85 % when the number of filters in the convolutional layer was 8, with maxpooling 4×4. The results of evaluating various CNN part parameters are shown in Table 3. The parameters of the convolutional and pooling layers are evaluated one by one in this table, by tweaking one parameter and settling the values of the other parameters. The proposed technique's assessment metrics are accuracy, precision, recall, and F1-score. Table 4 shows the changing learning rate values in the RMSprop algorithm and their effect on assessment results. Fig. 11 depicts the relationship between the proposed network's accuracy and the number of iterations used during the training phase. Also illustrated Fig. 12 the relationship between the proposed network's loss function (categorical_crossentropy) and the number of iterations used during the training phase.

From the above figure, we can notice the start of the accuracy with a dramatic rise. After that, the indicator begins to gradually increase at epoch 40, and the accuracy of access is relatively stable after epoch 80.

Table 4. The performance results of changing the learning rate values in RMSprop technique

Learning rate	Accuracy	Precision	Recall	F1_score
0.0001	39.57%	100%	0.61%	0.61%
0.0005	93.25%	95.51%	91.41%	92.45%
0.001	97.85%	97.53%	96.93%	96.81%
0.005	95.40%	95.96%	94.79%	94.76%
0.01	95.09%	95.38%	95.09%	94.67%
0.05	84.05%	84.05%	84.05%	82.73%
0.1	86.20%	86.20%	86.20%	84.14%
0.5	91.10%	91.10%	91.10%	90.43%

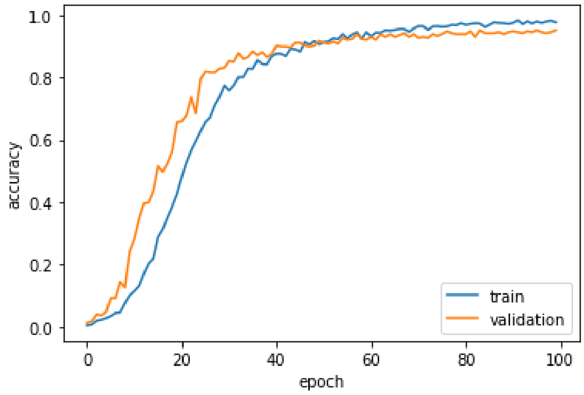


Figure. 11 The accuracy of the proposed model

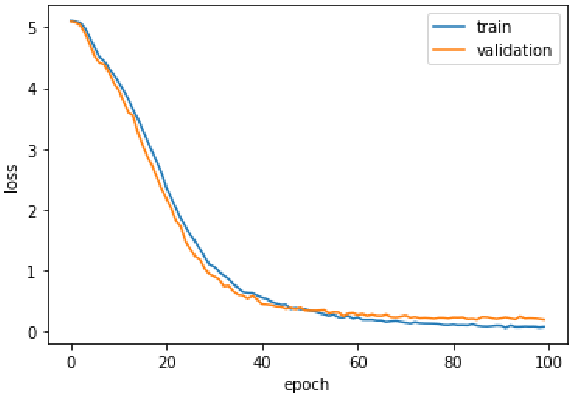


Figure. 12 The loss function of the proposed model

Received: February 24, 2022. Revised: April 1, 2022.

Table 5. comparison of the proposed model with the previous models

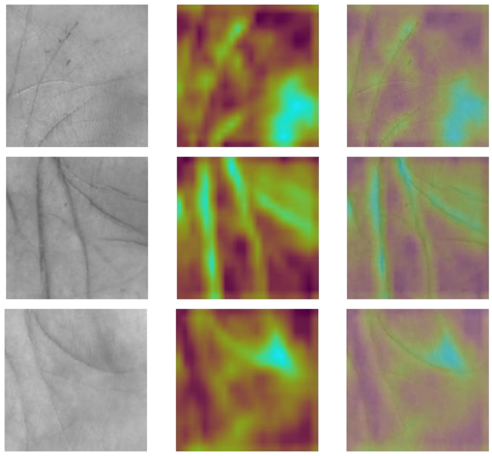
Method	No. Of con. layers	No. of FC layers	Dataset	No. of images	No. of cases	Accuracy
CNN [8]	11	2	PolyU	6000	500	99.95%
CNN [10]	5	3	PolyU Multi-Spectral	24000	500	99.99%
CNN [11]	1	3	PolyUC	1000	100	97.67%
The proposed AE+CNN	4	2	COEP	1304	163	97.85%

From the above figure, we can notice the start of the loss function with a large drop. After that, the indicator begins to steadily fall at epoch 40, and the loss function is relatively stable after epoch 80.

To evaluate the proposed hybrid model, we compared it with recent developed works, and the comparison contained two parts. The first part was a comparison with some deep learning methods in terms of the number of convolutional layers used, the number of fully connected layers, the database used, its size, and the attained accuracy of the method. The results of the proposed technique were promising despite the small number of images used for the database, which is considered an important characteristic of our proposed technique. The comparison showed that the proposed technique was well-organized in terms of building the model, as it does not use a large number of layers, which leads to the computational complexity of the model and increases the training time. At the same time, our model does not resemble a simplistic model with a single layer that may not be capable of extracting deep features. Table 5 shows the details of the comparison results. The other part of the comparison was with the traditional methods that used the COEP database used in our research, noting that the database was not used with deep learning methods, and we believe that this is due to the small number of images in it, which is one of the biggest challenges in deep learning techniques, as we showed the method used and the results obtained compared to the proposed technique in Table 6. The proposed technique also has an additional advantage, which is to ensure that the features extracted from the palm print are important features and able to be distinguished with high accuracy. This is based on the autoencoder model used in the hybrid model that uses the extracted features for the purpose of reconstructing the image with high accuracy, which can highlight the strength of the proposed model on this side as we proved in Fig. 13 which provided the gradient weighted class activation map (Grad-CAM) for the entire model.

Table 6. comparison of the proposed model with the classical methods

Method	Accuracy
Stockwell transform [7]	100%
PRS [4]	99.7%
The proposed AE+CNN	97.85



(a) (b) (c)

Figure. 13 Illustration of the GRAD-CAM from the model: (a) original image, (b) points of interest in the final convolutional layer, and (c) Softmax activation output

5.5 Evaluate by Grad-CAM result

To show regions of interest in the final convolutional layer and the softmax activations output from the model, we developed a gradient weighted class activation map (Grad-CAM). As shown in Fig. 13, the GRAD-CAM has identified the specific areas of the palmprint that CNN finds relevant for class differentiation. The network concentrates on the parts highlighted in yellow, which represent the most relevant spots in the palm image, and red, which signifies the least important points, in order to make the classification decision.

5.6 Using the feature maps technique to visualize convolution layers

The feature maps technique is used to display the applied filters in this section. The first six filters are displayed in Fig. 14. Each filter is represented by a figure with six rows of three images. Dark region symbolizes little or inhibitory weights, while light region represents large or excitatory weights.

We also use the feature maps method to evaluate and understand the model's predictions. This way helps to understand how the model learns different filters and how data is transferred through the layers. The premise is that feature maps near the input identify small or fine-grained details, while feature maps near the model's output capture more generic information. The visualized feature maps for the first convolution layer, which contains 32 filters, are shown in Fig. 15. We can see from the figure that the consequence of applying the filters in the first convolutional layer gives different sorts of features.

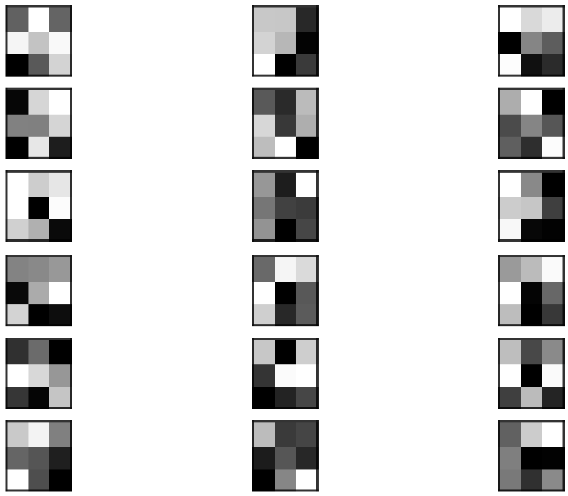


Figure. 14 Feature map technique for first six filters

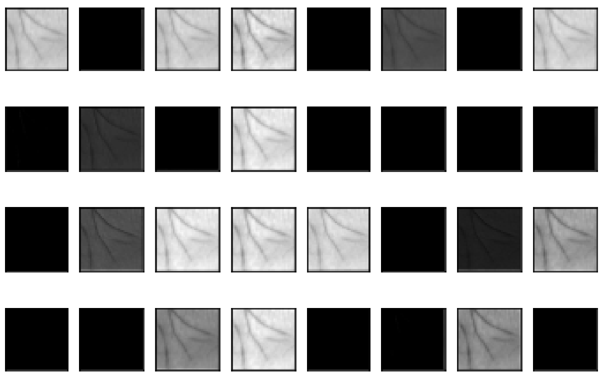


Figure. 15 An illustration of the first convolution layer, which uses the feature maps technique to interpret 32 filters.

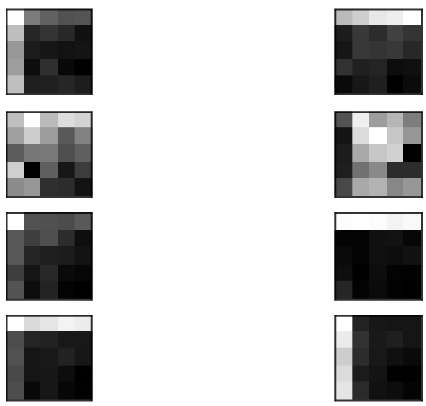


Figure. 16 An illustration of the fourth convolution layer, which incorporates interpret eight feature maps-based filters.

Fig. 16 also depicts feature maps for the fourth convolution layer, which includes eight filters. Where the feature maps display fewer information, we can tell that the model works deeply. This context is expected at this convolution level, but it has the potential to provide useful features for categorization. We lose our capacity to read these deeper feature maps in most cases, yet they are very apparent for the model.

6. Conclusion

In this work, we developed a palmprint authentication technique (AE+CNN) based on deep learning approaches, specifically a hybrid of autoencoder and CNN. With the suggested approach, we obtained deep features from the hybrid model that led to valuable authentication results with a smaller number of images, which is a major issue in any deep learning model. We used the COEP dataset, which has never been used before with palmprint authentication based on deep learning. Several experiments have been conducted to see how the parameters of various network hybrid layers evolve over time. The network has shown that it can adapt to a wide variety of palmprints, resulting in strong authentication. We compared our findings with conventional approaches and deep learning models. In compared to other methodologies, the best results obtained were 97.85 % accuracy and 96.81 % F1-score, which are excellent findings. In future work, we suggest adopting deep learning networks to be used in different stages of the palmprint authentication process, from ROI extraction to classification. We came to the conclusion that not all deep learning models require a large number of images to provide effective results.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper conceptualization, Mohammed and Firas; methodology, Firas; software, Firas; validation, Mohammed and Firas; formal analysis, Firas; investigation, Mohammed; resources, Firas; data curation, Mohammed and Firas; writing—original draft preparation, Mohammed and Firas; writing—review and editing, Mohammed; visualization, Firas; supervision, Mohammed; project administration, Mohammed.

References

[1] Z. Guo, D. Zhang, L. Zhang, and W. Zuo, "Palmprint verification using binary orientation co-occurrence vector", *Pattern Recognition Letters*, Vol. 30, No. 13, pp. 1219-1227, 2009.

[2] S. Verma and P. Mishra. "A survey paper on palm prints based biometric authentication system", *International Journal of Electrical and Electronics Engineering (IJEEE)*, Vol. 1, No. 3, pp. 20-27, 2012.

[3] L. Fei, Y. Xu, and D. Zhang. "Half-orientation extraction of palmprint features", *Pattern Recognition Letters*, Vol. 69, pp. 35-41, 2016.

[4] M. Kadhm, H. Ayad, and M. Mohammed, "Palmprint Recognition System Based on Proposed Features Extraction and (C5. 0) Decision Tree, K-Nearest Neighbour (KNN) Classification Approaches", *J. Eng. Sci. Technol*, Vol. 16, No. 1, pp. 816-831, 2021.

[5] M. Mu, Q. Ruan, and Y. Ming, "Shape parameters of Gaussian as descriptor for palmprint recognition based on dual-tree complex wavelet transform", In: *Proc. of IEEE 10th International Conf. on SIGNAL PROCESSING PROCEEDINGS*, Beijing, China, pp. 1406-1409, 2010.

[6] B. Ergen, "Scale invariant and fixed-length feature extraction by integrating discrete cosine transform and autoregressive signal modeling for palmprint identification", *Turkish Journal of Electrical Engineering & Computer Sciences*, Vol. 24, No. 3, pp. 1768-1781, 2016.

[7] N. Kumar and K. Premalatha, "Palmprint authentication system based on local and global feature fusion using DOST", *Journal of Applied Mathematics*, Vol. 2014, pp. 1-11, 2014.

[8] X. Dong, L. Mei, and J. Zhang, "Palmprint recognition based on deep convolutional neural networks", In: *Prof. of 2018 2nd International Conf. on Computer Science and Intelligent Communication (CSIC 2018)*, Leipzig, Germany, pp. 82-88, 2018.

[9] G. Wang, W. Kang, Q. Wu, Z. Wang, and J. Gao, "Generative adversarial network (GAN) based data augmentation for palmprint recognition", In: *Proc. of IEEE International Conf. 2018 Digital Image Computing: Techniques and Applications (DICTA)*, Canberra, ACT, Australia, pp. 1-7, 2018.

[10] W. Gong, X. Zhang, B. Deng, and X. Xu, "Palmprint recognition based on convolutional neural network-AlexNet", In: *Proc. of 2019 Federated Conf. On Computer Science and Information Systems (FedCSIS)***,** Leipzig, Germany, pp. 313-316, 2019.

[11] L. Albak, R. A. Nima, and A. Salih, "Palm print verification based deep learning", *Telkomnika*, Vol. 19, No. 3, pp. 851-857, 2021.

[12] J. Mehta and A. Majumdar, "Rodeo: robust de-aliasing autoencoder for real-time medical image reconstruction", *Pattern Recognition*, Vol. 63, pp. 499-510, 2017.

[13] C. Chaitanya, A. Kaplanyan, C. Schied, M. Salvi, A. Lefohn, D. Nowrouzezahrai, and T. Aila, "Interactive reconstruction of Monte Carlo image sequences using a recurrent denoising autoencoder", *ACM Transactions on Graphics (TOG)*, Vol. 36, No. 4, pp. 1-12, 2017.

[14] J. Zheng and L. Peng, "An autoencoder-based image reconstruction for electrical capacitance tomography", *IEEE Sensors Journal*, Vol. 18, No. 13, pp. 5464-5474, 2018.

[15] M. Fuad, A. Fime, D. Sikder, M. Iftee, J. Rabbi, M. A. Rakhami, A. Gumae, O. Sen, M. Fuad, and M. Islam, "Recent Advances in Deep Learning Techniques for Face Recognition", *IEEE Access*, Vol. 9, pp. 99112-99142, 2021.

[16] S. Albawi, T. Mohammed, and S. A. Zawi, "Understanding of a convolutional neural network", In: *Proc. of 2017 International Conf. On Engineering and Technology (ICET)*, Antalya, Turkey, pp. 1-6, 2017.

[17] W. Li, B. Zhang, L. Zhang, and J. Yan, "Principal line-based alignment refinement for palmprint recognition", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 42, No. 6, pp. 1491-1499, 2012.

[18] L. Fei, G. Lu, W. Jia, S. Teng, and D. Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 49, No. 2, pp. 346-363, 2018.

[19] P. Dhandapani and A. Varadarajan, "Multi-Channel Convolutional Neural Network for Prediction of Leaf Disease and Soil Properties", *International Journal of Intelligent Engineering and Systems,* Vol. 15, No. 1, pp. 318-328, 2022.
[20] S. Behdenna, B. Fatiha, and G. Belalem, "Ontology-Based Approach to Enhance Explicit Aspect Extraction in Standard Arabic Reviews", *International Journal of Computing and Digital Systems*, Vol. 11, No. 1, pp. 277-287, 2022.