IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 January 2024, accepted 22 January 2024, date of publication 24 January 2024, date of current version 7 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358209

RESEARCH ARTICLE

Improved Unsupervised Deep Boltzmann Learning Approach for Accurate Hand Vein Recognition

RANA NOUR[1], HOSSAM EL-DIN MOUSTAFA[2], (Senior Member, IEEE), EHAB H. ABDELHAY[2], (Member, IEEE), AND MOHAMED MAHER ATA[3,4]

[1]Department of Electronics and Communications Engineering, Mansoura Higher Institute for Engineering and Technology, Mansoura 35516, Egypt
[2]Department of Communications and Electronics Engineering, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
[3]School of Computational Sciences and Artificial Intelligence (CSAI), Zewail City of Science and Technology, 6th of October City, Giza 12578, Egypt
[4]Department of Communications and Electronics Engineering, Misr Higher Institute for Engineering and Technology, Mansoura 35511, Egypt

Corresponding author: Mohamed Maher Ata (mmaher844@yahoo.com; momaher@zewailcity.edu.eg)

ABSTRACT Dorsal hand vein (DHV) recognition is a burgeoning biometric technology that has recently garnered considerable attention. This article uses image processing and deep learning to present a novel DHV recognition approach. It involves detecting and identifying the unique patterns present in the DHV. The proposed system begins with the preprocessing mechanism that is applied to enhance the quality of the acquired images, including contrast enhancement and noise reduction, by using some filters such as Median and Contrast Limited Adaptive Histogram Equalization (CLAHE). Next, a deep learning model, such as a convolutional neural network (CNN), is employed to automatically abstract discriminative features from the preprocessed vein images. The empirical outcomes prove the influence and reliability of the proposed technique for vein recognition, making it a promising solution for biometric authentication systems. Compared with traditional CNN, the proposed approach shows good accuracy and classification rate results. The suggested model achieved a high recognition rate accuracy, recall, precious, and f-score of 99.7%,97%,96%, and 96%, respectively, and a recognition time of about 1283.45 s. To enrich the model's capability for feature recognition and reduce recognition time, decrease the intricacy of learning and the connectivity CNN structure, an alternative approach based on Restricted Boltzmann Machines (RBM) was assessed. This strategy exhibits superior accuracy in comparison to other contemporary algorithms. The proposed RBM achieved a high recognition rate accuracy, recall, precious, and f-score of 99.9%,99%,99%, and 99%, respectively, and a recognition time of about 137.235s.

INDEX TERMS Biometrics, image processing, DHV, CNN, deep learning, RBM.

I. INTRODUCTION

DHV is a biometric authentication technique that leverages the distinctive patterns of veins on the back of the hand to distinguish and authenticate individuals. Biometric recognition automatically recognizes an individual's properties based on anatomic/behavioral features. Several biometric methods have been proposed based on various anatomical and behavioral characteristics. These include palm print, fingerprint, finger veins, hand veins, foot veins, palm veins, iris, voice recognition, gait, palates, facial expression, heartbeat, signature, body language, DNA recognition, and face shape. Applications that use biometrics are becoming more popular as well as being used more commonly since they provide security, accuracy, and quick performance. The dorsal hand veins, which are located beneath the skin surface, have a unique pattern and shape that can be used to identify individuals accurately. We present in this work a method for DHV authentication that combines deep learning and image processing approaches, as illustrated in Figure 1.

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

The improvement of deep learning processes has established the ability of computer vision systems to recognize these patterns. Using vast datasets of hand vein images, deep learning algorithms may be trained to recognize the complex and frequently subtle patterns that differentiate one person's hand from another. These models can then be used to identify individuals in real-time based on a captured image of their DHV patterns. One major advantage of using DHV recognition over other biometric authentication methods, such as facial recognition, is that the vein patterns in the hands are more unique and less subject to changes over time. It makes hand vein recognition a reliable and accurate way of identifying individuals for security purposes. With the support of learnable weights and biases, CNN is a Deep Learning framework capable of analyzing an input image, assigning relative value to various objects or features, and distinguishing between them. When compared to other classification algorithms, CNN needs substantially less preprocessing. The convolutional layer, fully connected layer, pooling layer, and activation functions (e.g., Sigmoid, Tanh, ReLU, etc.) are all fundamental components of CNN architecture. In this study, an RBM-based deep learning scheme for vein recognition is developed. An energy-based Artificial Neural Network (ANN) known as RBM is employed in many different fields, including pattern recognition, image processing, regression, image processing, and topic modeling. RBM is an unsupervised learning and feature extraction system based on a two-layered stochastic recurrent neural network. The test/training sets are located in the visible layer of the RBM, and the feature extractor is located in the invisible layer. Every unit in the unhidden layer is related to every unit in the invisible layer; in contrast, there is no relationship between any of the units in the visual layer, and the same is true of the invisible layer's units [1].

Furthermore, in recognition tasks, RBM has shown remarkable results as both a practical first step in training deep neural network classifiers and a feature extractor for text and picture data [2].

CNN technique gives better accuracy as compared to the traditional methods. This technique has significantly improved the ability to learn biometric characteristics to offer perfect and forceful results. This paper's contributions are summed up as follows:

1. Suggesting an image-processing approach for vein patterns such as filtering techniques, contrast enhancement strategies, detecting DHV patterns, and reducing noise.
2. Proposing deep learning method working on two different datasets: Dr. Badawi and Dorsal Hand Vein Image Database.
3. Evaluating the traditional classical CNN models (ResNet, AlexNet, VGG16, VGG19, and InceptionV3) in the hand vein authentication.
4. Comparing a proposed deep learning method with the traditional CNN models. The result shows that performance is better than other conventional CNN models.
5. Suggesting the RBM technique for DHV recognition to decrease the time for authentication and achieve high accuracy.
6. Boosting accuracy, sensitivity, and specificity in accordance with the suggested CNN and RBM methods.

The following are the outstanding sections of the paper: Section II presents the related work. CNN and RBM model architectures and methodology used in this article are covered in Section III. Discuss the trial results and valuation of the suggested system in Section IV. Finally, discuss the conclusion in Section V.

II. RELATED WORK

Kumar et al. [3] offered a deep-learning model that determines recognition accuracy by individually evaluating DHV datasets from children and adults. This study predicts how useful vein patterns will be for biometric identification because children up to 12 undergo significant physical changes [1].

Alashik et al. [4] have posited the DL-GAN method to confirm biometric identity. The DL-GAN approach has the advantage of higher accuracy because it employs a backhand structure for authentication. The main limitation of this research is the unlimited complexity. It was observed that this research generated many simulated samples for training to increase the accuracy and potency of this image classification method. One limitation of DL-GAN is its insistence on a sizable dataset, which can be challenging to compile. There are disadvantages to the DL-GAN approach. First, achieving high accuracy requires a lot of training data.

Kuzu et al. [5] published a unique CNN pipeline for vein-based biometrics. Using the weights of the network from the A dataset of pre-trained ImageNet and adjusting the Densenet-161 architecture, according to the findings adjusting lasting CNN structural design would perform better than developing a network from the ground up for a specific purpose area (such as vein biometrics).

Daas et al. [6] observed unimodal and multimodal biometric tools for identification based on the FV. AlexNet, VGG16, and ResNet50 are three pre-training models used for feature extraction. SVM and Softmax classifiers are utilized in recognition performance calculations and improvements. The recognition accuracy is improved by the suggested ResNet50-Softmax with weighted sum fusion (score level fusion).

Qin et al. [7] have documented an iterative deep-learning theory to use a deep article demonstration to estimate the likelihood of pixels being veins in the background by using a deep feature performance to estimate the likelihood of pixels being veins in the background. In order to make vein segmentation easier, a Deep Belief Network (DBN) is recommended for the extraction of venous features. The DBN avoids laborious and error-prone automatic labeling and performs well since it is iteratively taught to rectify the wrong labels. The trial consequences reduce the verification error rate and reach cutting-edge performance.

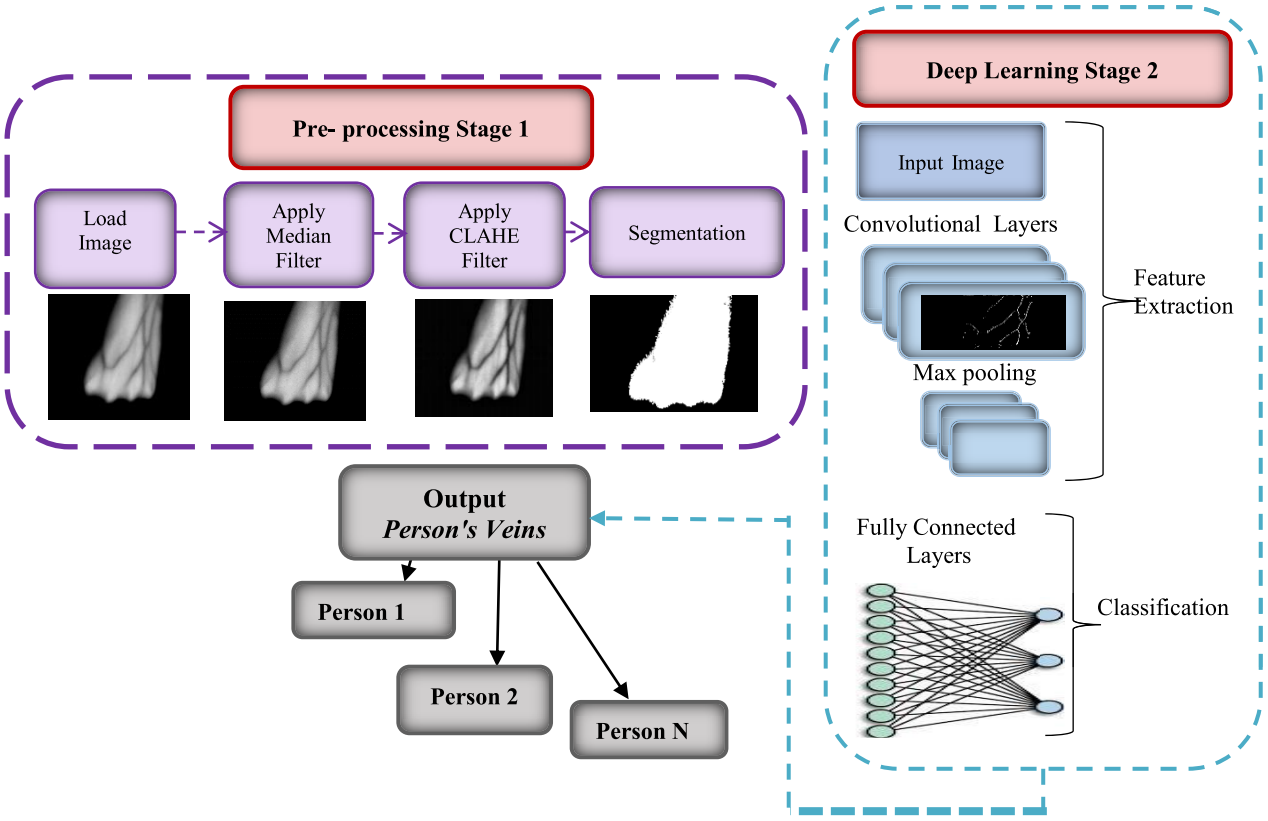


FIGURE 1. Block Diagram for Combination of Image Processing and Deep Learning Techniques for Dorsal Hand Vein .Authentication.

Song et al. [8] have shown a way to identify finger veins based on a deep Dense Net. This system reached high recognition accuracy. Additionally, the composite image was found to be more robust against noise than the difference image when tested with a noisy image.

Obayya et al. [9] have used trained models to examine the authentication issue by palm veins and developed a finely tailored deep CNN-based model by applying the Bayesian technique for optimization. The developed CNN model gained a high level of identification accuracy.

Zhou et al. [10] have a significantly lightweight CNN and its methods of instruction. The lightweight CNN is utilized from images of finger veins to extract more condensed and discriminatory features. One advantage of the presented training approach is that lightweight CNN performs nearly as well as pertained weights-based CNN while using substantially fewer resources. Additionally, the suggested approach can be used in other fields lacking training data, such as palm prints and palmar veins.

Zeng et al. [11] have gained a new CNN to fulfill the pixel-wise task involving the segmentation of finger veins. The experimental results demonstrate favorable segmentation outcomes compared to other comparable models. Also, performance improvements can be attributed to the incorporated conditional and random fields.

Trabelsi et al. [12] proposed efficient identification systems for quick palm print recognition, including unimodal and multimodal approaches. The recommended way is the simplified PalmNet-Gabor algorithm. Experiment results substantiated the proposed method's robustness and performance by realizing high recognition accuracy with considerably fewer features.

Kuzu et al. [13] developed a unique extracting features methodology for vein-based biometric recognition in terms of templates built to utilize a CNN alone. The suggested model has documented effective recognition better to up to date on the dorsal, palm, and finger pattern of veins.

A unique model for palm vein recognition that uses decision-level fusion to connect a CNN-based architecture with a texture-based system has been thoroughly studied by Babalola et [14]. In addition, a CNN structure established on deep learning was constructed by the AlexNet architecture. However, with fewer filters, CNN's decision, and the decision to reach a final decision, five sub-regions are fused. Linked to the additional methods, the proposed model achieved the highest accuracy.

A unique CNN model based on Transfer Learning was created by Garcia-Martin et al. [15] and tested on smartphones for contactless vein biometric detection. The proposed TL model has four different configurations. VGG16 is a well-known CNN shape or architecture. ResNet50, ResNet152, and VGG19. The model under consideration demonstrates enhanced performance and yields optimal outcomes to Equal Error Rate (EER).

Shao et al. [16] designed a deep biometric hash learning (DBHL) model to accurately analyze and manage Palm print, palm vein, and DHV recognitions. In order to transform binary codes from images, implementing a comprehensive network from start to finish is executed. Palm vein and fingerprint recognition achieve optimal performance when the DHV identification is recognized with a lower EER.

Zhong et al. [17] have thoroughly investigated the PHD (Palmprint Hybridized with Dorsal hand vein) model. PHD is a novel multi-biometrics algorithm that combines palm print and DHV recognition. This model takes advantage of palm print recognition's high accuracy and DHV recognition's liveness recognition. This model can produce optimal outcomes, i.e., FAR and FRR are close to zero.

A novel multi-biometric approach for personal authentication utilizing fingernail plates and finger knuckles has been investigated by Choudhury et al. [18] inside a deep learning framework. After the two traits are optimized, extensive investigations demonstrate that a flawless 100% identification accuracy may be achieved under specific fusion settings.

Tiong et al. [19] have proposed a Multi-feature Deep Learning Network (MDLN) architecture that improves recognition performance by combining modalities from the face and regions with texture descriptors. The proposed MDLN enhanced biometric recognition performance under difficult situations, incorporating illumination variances, features, and posing misalignments.

Shahreza et al. [20] attempt to secure and improve current finger vein authentication procedures by utilizing a DNN to minimize the number of biometric characteristics and then retaining the reduced-dimension features of Bio hashing. This framework accomplishes superior performance results.

Arora et al. [21] offered a deep learning architecture for an iris recognition system to identify and detect various spoofing attacks. The suggested framework performs better than existing ones in detecting attacks with printed images and contact lenses.

Jayanthi et al. [22] have recommended an efficient DL-based incorporated model for detecting and recognizing iris from input photos. The suggested technique has produced superior iris recognition performance with maximum recognition accuracy. Outperforming other ways such as AlexNet, UniNetV2, VGGNet, Inception, ResNet, and DenseNet models.

Toygar et al. [23] have to present a multimodal CNN architecture combining three biometric features with decision-level data biometrics from the palm, wrist, and dorsum. The suggested CNN architecture produced better performance compared to all other approaches. CNN architecture showed the effectiveness of three-vein biometrics, Enhanced performance, and the ability to defend against multiple attacks.

Jhong et al. [24] created a prototype device for palm vein detection using a convolutional neural network utilizing the Raspberry Pi operating system and a cloud-calculating platform. The suggested study produced successful identification of palm veins without contact. The theory outcomes presented that the offered method has great identification accuracy in several databases.

Das et al. [25] offered a CNN approach for finger-vein identification, demonstrating robust symmetry capabilities under various ambient situations. The detected outcomes show that a maximum identification accuracy of 95% may be reached for all utilizing the proposed CNN architecture for all databases.

Xu et al. [26] conducted palm print image recognition and verification studies on two public contactless palm print databases using a residual network (ResNet) and spatial transformation networks (STN). Several state-of-the-art procedures were compared in extensive experiments, and the results showed that the method was effective.

Yin et al. [27] offered a novel approach to feature transfer for deep face identification training that examines the UR classes with an imbalance problem. Across regular, UR, and unseen classes, the proposed technique is shown to acquire superior representations, consistently improving performance. As shown in Table 1.

III. DORSAL HAND VEIN DATABASE

Two datasets are passed down in this work; the first is identified as ''Badawi hand vein dataset.'' It comprises one hundred hands, each containing five photographs per individual per hand (500 images). It's related to 50 different people for both left and right hands. People with left and right hands are linked to 50 diverse people. Photos of both right and left hands are involved in the dataset. The organization of Dr. Badawi's hand vein dataset enables its application in the context of identification. The study of changes between the right and left hands and the listing of people by the dorsal hand veins. People's left hands are depicted in the data between 1 and 50, whereas right hands are shown in the data between 51 and 100. The study does not differentiate between the right and left hands; rather, it generates one hundred distinct classes using assumptions that the left and right hands represent different people [28]. The second one is known as '' Dorsal Hand Veins Image Database 1.'' With 138 individuals and four images per individual per hand, the database contains 1,104 pictures. The time interval between the data obtained in the sessions for Database 1 is two months [29].

IV. METHODOLOGY

A. PREPROCESSING FOR DORSAL HAND VEIN

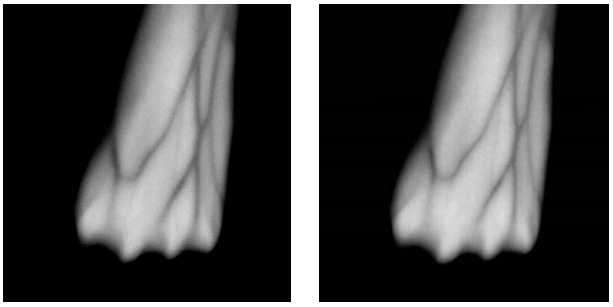
Image processing is a method utilized to enhance the content of an image by eliminating extraneous data present in different regions of the picture. This method is required to improve image visualization before identification. Preprocessing also aids in The procedure of enhancing pixel density and improving image quality. After preprocessing, image quality improves. Image preprocessing is frequently the first and most practical step in the identification process. Algorithm 1 explained the steps to execute image processing.

TABLE 1. Related work.

Author/ Ref # / Year	Goal	Advantages	System Accuracy	Challenge
Kumar et al.[3] 2021	This paper is important in the deep learning technique that affects recognition accuracy.	Because children up to the age of 12 undergo considerable physical changes, this study offers predictions regarding how functional vein patterns will be for biometric identification.	99%	The model is very complex.
Alashik et al.[4]2021	The DL-GAN approach for confirming biometric identity was created in this study.	The use of a backhand framework improves authentication precision.	98.3%	This approach is highly sophisticated. DL-GAN method needs a massive dataset.
Kuzu et al.[5] 2020	This article offered a novel CNN pipeline for vein-based biometrics.	This system was developed using transfer learning. It makes the system is not complicated.	96%	Make use of pre-trained CNN only.
Daas et al. [6] 2020	The Fisher Vector (FV) is used in this study to propose novel unimodal and multimodal detection biometric techniques.	Softmax is used as a classifier to evaluate and improve recognition ability.	99.8%	Refrain from using outside databases to assess the suggested system's effectiveness.
QIN et al. [7] 2019	This study suggested a deep learning-based method for estimating likelihood that learns from errors.	The DBN achieves superior performance by bypassing the time-consuming and error-prone automated labeling process.	98%	Requires extra processing time.
SONG et al. [8] 2019	This manuscript presented a method for identifying finger veins based on deep DenseNet.	These approaches produced a high degree of recognition accuracy.	100%	Using DenseNet model only.
OBAYY A et el. [9] 2020	This work investigated the problem of palm vein recognition using trained models and produced a finely tuned deep CNN.	These strategies were successful in achieving high recognition accuracy.	99.4%	Inability to use a dependable and effective segmentation method.
Zhou et al. [10] 2019	This paper presented a lightweight CNN along with its training methodology.	The lightweight CNN outperforms the pertained-weights-based CNN while using far fewer resources.	97%	Concentrate on the loss function while ignoring ROI extraction.
ZENG et al. [11] 2020	This paper offers a unique CNN network for pixel-by-pixel segmentation of finger veins.	The experimental results yield decent segmentation results.	90.67%	Accuracy needs to be improved.
Trabelsi et al. [12] 2022	This research demonstrates functional unimodal and multimodal identification processes for quick palm print authentication.	The suggested method's robustness and outcomes provide good recognition accuracy.	99%	Not very good with large datasets.
Kuzu et al. [13] 2020	This paper presented a unique model for vein-based biometric recognition based on character extraction using CNN.	On the dorsal, palm, and finger patterns, the proposed model exceeds current best practices for vein detection.	87.4%	Accuracy is very low.
Babalola et al. [14] 2021	A novel model for palm vein detection was presented. It fuses a CNN architecture with a texture-based approach.	The proposed model achieved the highest accuracies.	100%	The model is ineffective for biometric DHV identification and wrist veins.
Garcia- et al. [15]2021	Using TL, a novel CNN model for contactless vein biometric recognition.	For the Equal Error Rate (EER), the model works better and yields the best outcomes.	98 %	The model is built on Transfer Learning.
Shao et al. [16] 2021	This article developed a DBHL model to examine and manage palm print, palm vein, and DHV.	Palm print and vein identification performs best, with an EER lower than dorsal hand veins.	99%	Using fundamental optical principles, fingerprint detection captures images that are easily replicable.
Zhong et al. [17] 2019	PHD model was presented in this study. Palm print and DHV identification are combined.	Accomplishing recognition's high accuracy and DHV recognition's liveness recognition.	95%,	Improving the method for detecting unseen samples.
Choudhury et al. [18] 2019	This study provided a novel way for configuring personal authentication within a deep learning framework.	When the two attributes are optimized, perfect identification accuracy can be obtained.	97.19%	The model is very complicated.
Tiong et al. [19]2019	This paper describes the MDLN architecture. By combining face and region modalities with texture	MDLN Enhanced biometric recognition performance in challenging	92.38%	Facial recognition is a tough process because to the quick appearance

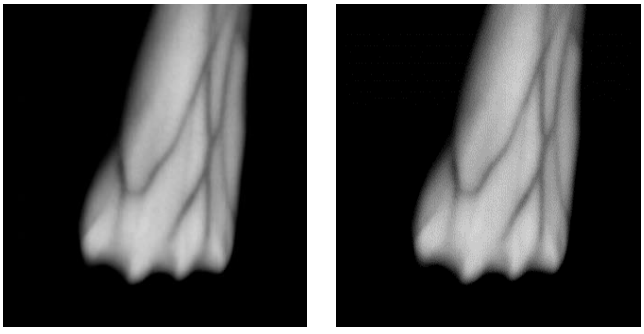
TABLE 1. *(Continued.)* **Related work.**

	descriptors.			changes.
Shahreza et al. [20] 2021	This research proposed a deep learning-based architecture to preserve and improve current finger vein detection methods.	This framework accomplishes good performance results.	98%	The preprocessing methods require more execution time.
Arora et al. [21] 2020	In this study, a deep learning architecture for an iris recognition system was proposed.	In terms of identifying assaults, the proposed approach outperforms existing ones.	95%	It cannot outperform the system's previous spoofing techniques.
Jayanthi et al. [22] 2021	This article presents an excellent DL-based model for finding and identifying iris from pictures sent in.	This study proposes an outstanding DL-based model for detecting and recognizing iris in submitted images.	99.14%	RoI features must be pixel-to-pixel, implying that they have small feature maps.
Toygar et al. [23] 2020	This study describes a multimodal CNN architecture that incorporates three biometric features: palm, wrist, and dorsum.	The recommended CNN architecture yielded the best results.	98.67%	Training computation times, with CNN frequently taking the longest.
Jhong et al. [24] 2020	The prototype device and CNN-based palm vein recognition system developed in this research were constructed.	The proposed method can accurately identify people in a large number of databases.	96.54%	Because the model has been trained 2500 times, it will take a very long time.
Das et al. [25]2018	This paper described a CNN-based system for detecting finger veins.	The observed results show a maximum identification accuracy.	95%	The model requires a picture-enhancing method to boost the feasible identification performance.
Xu et al. [26] 2021	This research investigated palm print picture recognition and verification.	The results proved how effective the strategy was.	96.88%	Many low-resolution contactless palmprint photographs show the main lines and wrinkles.
Yin et al. [27] 2019	This work demonstrates a novel method for moving features for deep face recognition training.	The suggested strategy can consistently raise performance percentage in ordinary, unseen, and UR classes.	99.5%	Cannot address the core issue of insufficient and diverse samples in UR classes.



(a) (b)

FIGURE 2. Dorsal hand image in preprocessing phase a) original image b) result obtained after applying image sharping.



(c) (d)

FIGURE 3. Dorsal hand image in preprocessing phase c) original image d) result obtained after applying a Median Filter.

The sharp masking filter is used to bring out the vascular structure even more. This filter is done to improve the contrast between different ranges of pixel intensities. Image sharpening contrasts the texture and hidden details of an image; it does not add additional detail, as shown in Figure 2.

After using the sharp masking filter, it is critical to improve the image quality. Subsequently, a median filter is applied. In addition to reducing distortion, this filter effectively detects image boundaries, an advantage over other filters. The resultant image subsequent to the execution of the median filter is depicted in Figure 3.

After applying sharping and median filters, we found that the CLAHE filters can strengthen the contrast and visibility of details in pictures of dorsal hand veins, making it easier to detect patterns and identify the vein's location. This filter achieved the best result for enhancing vein pattern contrast. The CLAHE procedure can expressed as:

$$\delta = \frac{x}{y}(1 + \frac{\sigma}{100}(C_{max} - 1)) \quad (1)$$

δ is the bound value (clip limit), x signifies the region size, y denotes the grey-level value (256), σ is the clip factor

Algorithm 1 *Pre-Processing DHV Dataset*

***Input*:** *Dorsal Hand vein Dataset I, every pixel $b_{ij,}$threshold Th_{ij}.*
Output: *Enhance image and detect vein for hand image.*
*# **Start** Procedure*
Load Dataset
Preprocessing image (Data Set):
***Resizing** I to $M \times M$.*
***Convert** I to greyscale.*
***Sharpen** the image.*
***Apply** the Median filter to smooth the image.*
***Apply** CLAHE filter as a transformation function to contrast and visibility of details in photos.*
The optimal threshold process:
For each grayscale image
  Scan grayscale image
  For every pixel b_{ij}, threshold Th_{ij}
    If $b_{ij} > Th_{ij}$–k
      $b_{ij} = 225$;
      Else
      $b_{ij} = 0$;
    End If
  End For
  ***Output** binary image.*
End For
Mathematical morphology processes:
***Apply** ← Skeletonize to reduce foreground regions in a binary image.*
***Apply** Erosion ← to remove extraneous parts.*
***Apply** Dilatation ← to expand the binary image and remove gaps.*
***Finding** contours area* by *applying an adaptive threshold to show that all noise near vein pattern.*
***Perform**an adaptive threshold to show that all noise is near.*
***Removing**noise*
***Apply**a median blur filter and canny edge filter.*
***#End** Procedure.*

expresses the sum of a histogram limit with a value ranging from 1 to 100 [26]. C_{max} is the extremely acceptable slope.

After that, the optimal threshold process separates the background's vein pattern. This process leads to the acquisition of the targeted vein picture. To detect a vein's image, more important steps are done as:

- Converting the vein pattern into white on a black background extracts the vein pattern.
- The resultant binary hand vein contains the output hand vein pattern.
- Applying mathematical morphology processes for hand vein pattern enhancement by three basic operations: Skeletonize is the process of reducing binary things to 1-pixel width representations. This can be helpful in vein extraction. Erosion is the process that analyses the image, narrows it, and removes extraneous parts, and Dilatation is the process that enlarges the binary image and eliminates gaps. It broadens the picture and removes broken edges, as shown in Figure 4.
- Finding contours area by applying an adaptive threshold to show that all noise near vein pattern appears in green points. The result is shown in Figure 4.
- Finally, to reduce the effect of these unwanted defects and remove noise, apply a median blur and canny edge filter. The canny edge filter ensures that the vein image obtained is continuous and has no breakpoints. The outcome is depicted in Figure 5.

V. DEEP LEARNING ARCHITECTURES

In recent times, deep neural networks, namely Convolutional Neural Networks (CNN), have exhibited a notable capacity to acquire proficient feature representations from input data. CNN deep learning algorithm and essential modules were used to analyze the network's overall structure. CNNs are designed for image and video recognition tasks. They

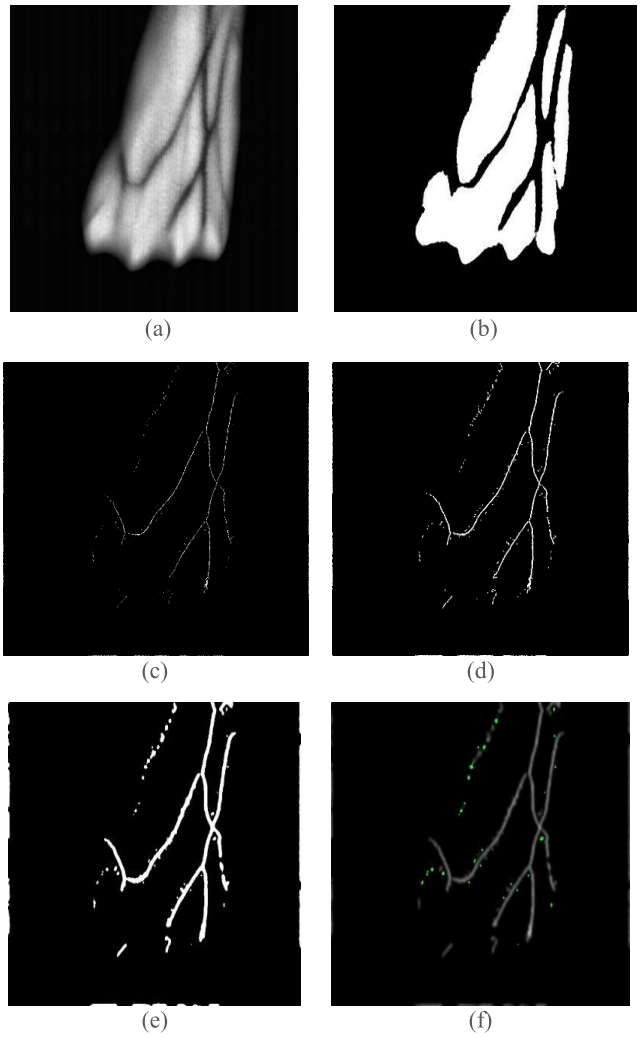


FIGURE 4. Dorsal hand image in preprocessing phase a) CLAHE filter applied on hand vein image b) result image obtained after applying threshold v = 127 c) Skelton image d) Eroded image e) Dilated image f) hand image with noise.

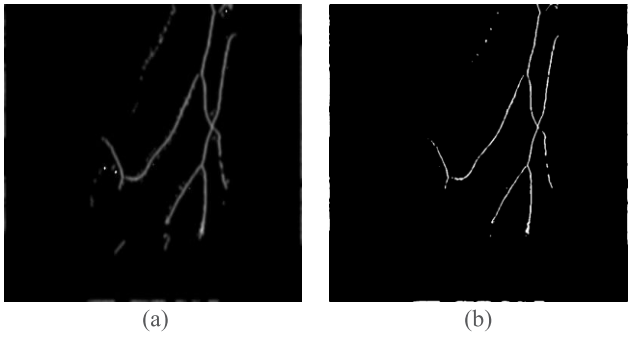


FIGURE 5. Dorsal hand image in preprocessing phase a) Image with noise b) removing noise by a canny edge.

extract features from input data using convolutional layers, followed by numerous fundamental concepts, like the Convolutional Layer, activation functions, Pooling Layer, and Fully-Connected Layer.

A. CONVOLUTIONAL LAYER

One of the primary constituents of a CNN is the convolution process. The convolutional layer is the first stage in which the diverse features are extracted from the input images. Input data, a filter, and a feature map are the few components that are needed. The filter (also called kernels) is applied to an image to obtain a specific map, resulting in multiple feature maps. The features are different depend on the filter that is used. Input pixels and filter weights are used to determine the dot product when applying a filter to an image. The filter matrix is commonly 3×3 (but the size might vary). Following the filter's n-pixel shift, or "stride," the dot product of the first pixels and the filter weights are used in the calculation. This method is also called convolution, as the equation (2) shows.

$$\mathrm{P}\left[\mathrm{x}, \mathrm{y}\right] = (\mathrm{m} * \mathrm{n})\left[\mathrm{x}, \mathrm{y}\right] = \sum_{\mathrm{a}} \sum_{\mathrm{b}} \mathrm{m}\left[\mathrm{a}, \mathrm{b}\right] \mathrm{n}\left[\mathrm{x} - \mathrm{a}, \mathrm{y} - \mathrm{b}\right] \tag{2}$$

where m represents the input image, and n signifies our kernel. The indexes representing the rows and columns of the result matrix are represented by x and y correspondingly.

B. POOLING LAYE

The purpose of this layer is to diminish the feature maps, which, when the depth of the image is held constant, are completely unrelated to the depth dimension. The pooling layer performs feature summarization within a specific region of the feature map generated by a convolutional layer. Consequently, succeeding operations are executed on summarized features rather than those produced by the convolution layer that is precisely positioned. Two main types of pooling layers are commonly used in neural networks: max Pooling and average Pooling. The formula for maximal Pooling over "Y" is as follows:

$$Y_b^{a,b} = \max_{a,b=1}^{p} (Y_{a,b}) \tag{3}$$

where $Y_b^{a,b}$ describes the output after the pooling layer. The size of the pooled kernel is denoted by p.

C. RELU ACTIVATION FUNCTION

After the convolutional layer, a nonlinear layer (or activation layer) called a rectified linear unit (ReLU) is typically used. Sigmoid, Tanh, and ReLU are the activation functions that are widely used. ReLU is used in this paper because it helps the network train more quickly. Equations (4), (5), and (6), respectively, define the functions of the ReLU, the Sigmoid, and the Tanh:

$$\mathrm{ReLU} \rightarrow \mathrm{R}\,(y) = \begin{cases} 0 & \text{"}if\text{"}\, y < 0 \\ y & \text{"}if\text{"}\, y >= 0 \end{cases} \tag{4}$$

$$\mathrm{Sigmoid} \rightarrow S(y) = \frac{1}{1 + e^{-y}} \tag{5}$$

$$\mathrm{Tanh} \rightarrow \mathrm{Tanh}\,(y) = \frac{e^{y} - e^{-y}}{e^{y} + e^{-y}} \tag{6}$$

D. FULLY CONNECTED LAYERS (FC)

The fully connected layer is the last stage of a CNN. FC is a neural network in which, via a weights matrix, each neuron applies a linear change to the incoming vector. Subsequently, the product undergoes a nonlinear transformation by utilizing a nonlinear activation function denoted as f. It is also known as a dense layer since it has a high density of connections. As the equation (7) shows

$$x_{jk}(y) = f\left(\sum_{i=1}^{n_H} w_{jk} y_i + w_{j0}\right) \tag{7}$$

where w represents the weights matrix, and y represents the input vector. The bias term, marked as w_{j0}," can be incorporated inside the nonlinear function, whereas the symbol f represents the activation function.

E. MODEL OPTIMIZATION

When mapping inputs to outputs, optimization models find the weights that make the least amount of mistakes. These optimization models, also called optimizers, have a big effect on how accurate the deep learning model is. They also affect how fast the model trains. The most popular model optimizers used in CNN are Stochastic Gradient Descent (SGD), RMS, and Adam.

1) STOCHASTIC GRADIENT DESCENT (SGD)

Rather than utilizing the complete dataset for every iteration, SGD selects data batches randomly. That is, we take only a few samples from the dataset. SGD can be uttered by:

$$w := w - \eta \nabla Q_i(w) \tag{8}$$

where w is the weight parameter, η is the learning rate, and $\nabla Q_i\,(w)$is the gradient of the weight parameter.

2) RMS PROP

RMS prop, or Root Mean Square Propagation, is a popular optimizer in the deep learning community. The algorithm keeps every weight's moving average of squared gradients, which then divides the gradient by the mean square's square root. RMS Prop can be expressed by:

$$v(w,t) := \gamma v(w,t-1) + (1-\gamma)(\nabla Q_i(w))^2 \tag{9}$$

where γ is the disregarding factor, the below formula updates weights:

$$w := w - \frac{\eta}{\sqrt{v(w,t)}} \nabla Q_i(w) \tag{10}$$

3) ADAM

Adaptive Moment Estimation (Adam), an optimization strategy that adapts the learning rate during training, is employed to enhance the proposed technique. In contrast to RMS Prop, Adam changes learning rates using the second moment of the gradients instead of merely the first. Adam's formula can be described as:

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \tag{11}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_t^t} \tag{12}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{13}$$

where the default values of β_1 is 0.9, β_2 is 0.999, and ϵ is 10^{-8}, $\hat{m}_t$ and $\hat{v}_t$ are the first moments of gradients and second moments of gradients, respectively.

F. DATA AUGMENTATION

Regularly, training the constructed CNN with an adequate number of images is recommended to ensure its best performance during testing. Unfortunately, there are times when a dataset doesn't have enough pictures. Data augmentation refers to the process of incorporating more data that is synthetically generated from pre-existing training data. Various image manipulation techniques such as resizing, flipping, rotating, cropping, padding, and other similar operations are employed. It helps fix problems like overfitting and lack of data and strengthens the model to perform better.

VI. TRADITIONAL CNN ARCHITECTURES

A common deep neural network called CNN handles input that is presented as several arrays, such as a grey image (one 2D array) and a color image (three 2D arrays). This paper uses pre-trained CNN models to recognize the vein pattern by relating the initial DHV image to the trained DHV datasets system. The most common deep learning models are Vgg-Net, ResNet50, Inception V3, Xception, DenseNet, AlexNet, Mobile Net, and Efficient Net. All pre-trained models were tested on the ImageNet dataset, and each model is described in detail below.

A. VGG-NET

Andrew Zisserman and Karen Simonyan established the Visual Geometry Group (VGG) system [30]. The following three dual convolutional layers are a single max pooling layer and a ReLU with 64,128 and 256 distinct filters of size 3×3 and stride size 1. Three convolutional layers, a ReLU with 512 separate filters of size 33 and stride 1, and a single max pooling layer comprise VGG-16's fourth and fifth layers. On the other hand, the VGG-19's fourth and fifth layers have four convolutional layers, a ReLU with 512 different filters, each of size 33 and stride size 1, and a single max pooling layer.

B. RESNET50

The concept of deep residual learning for image classifiers was introduced by He et al. [31]. The depth of Resnet50 far exceeds that of the VGG family of architectures. Respectively, Every ResNet is planned by 7×7 and 3×3 Kernel Sizes for the initial Convolution and Max-Pooling. Resnet50 Can Express by:

$$X_{m+1}^{k} = Y\left(X_{1\to m}^{k}, k_{1\to m}\right) + X_{1}^{k} m \geq 1 \tag{14}$$

$$X_{m+1}^{k} = Y_a(X_{m+1}^{k}) \tag{15}$$

$$Y_c\left(X^k_{1\to m}, k_{1\to m}\right) = X^k_{m+1} - X^k_l \tag{16}$$

where $Y_c\left(X^k_{1\to m}, k_{1\to m}\right)$ is an altered signal, and X^k_l is an input of $\left(X^k_{1\to m}, k_{1\to m}\right)$ as a result, an aggregated output X^k_{m+1}, appears in the subsequent layer once the activation function Y_a is added.

C. INCEPTION V3

Szegedy et al. [32] proposed the Inception model, a deep CNN architecture. It contains 42 hidden levels. There are 11 inception modules in the architecture of Inception V3.

D. XCEPTION

The Xception model was proposed by Chollet et al. [33]. The Xception structural design is an upgraded version of the Inception V3 architecture. There are 36 convolutional layers in the Xception network design, and they are responsible for the feature deduction. The Xception system has 14 modules.

E. ALEXNET

AlexNet is suggested by Krizhevsky et al. [34]. AlexNet had eight learned layers. Five of them were convolutional, and the other three were fully connected. A 1000-way softmax coupled to the last fc layer generates the categorization results. AlexNet can be expressed as:

$$k(A) = i(A) + j(A) \tag{17}$$

where $i(A)$ and $j(A)$ are the output goal and the total companion goal, respectively.

F. MOBILENET

MobileNet model was proposed by Google [35]. MobileNet comprises 28 layers, some named deep convolution, one-point convolution, batch normalization, ReLU, an average collecting layer, and softmax, respectively.

G. DENSENET

DenseNet was offered by Huang et al. [36]. Each layer in DenseNet has additional inputs from layers that came before it. A feature layer, numerous dense blocks, and a few transition layers are all components of DenseNet.

H. EFFICIENTNET

Efficient Net was recommended by Tan et al. [37]. The whole number of convolution layers is 18, or D = 18, and each layer has a kernel of type k (3,3) or k. (5,5). The formula for scalability of depth, width, and resolution is displayed below:

$$\text{depth} : \mathrm{D} = \varepsilon\psi, \text{width} : \mathrm{W} = \alpha\psi$$
$$\varepsilon \geq 1, \alpha \geq 1 \tag{18}$$

VII. PROPOSED CNN ARCHITECTURE

The proposed approach uses deep learning to estimate the likelihood that a given pixel is a vein. The model comprises several significant components, including convolutional layers, filters, strides, max Pooling, and fully connected layers. By identifying vein patterns, reducing data dimensions, and injecting nonlinearity, these layers' primary responsibility is to boost model performance. Figure 7 depicts the proposed CNN model's overall computational structure. The structure of the Deep Learning model proposed has four convolutional layers, three max Pooling layers, three fully connected layers, and a softmax layer. Batch normalization is implemented following each convolutional layer to mitigate overfitting and uphold the stability of the model. The model under consideration underwent training for 50 epochs with 16 batch sizes. The learning rate employed during training was 0.00000001. The learning rate controls CNN models' problem-solving agility. Whereas Adam, the optimizer for the study, uses a cross-entropy function to obtain the most precise loss measurements, the whole network uses ReLU as its activation function and L regularization. Our best results were achieved with a value of (0.0001), minimizing training loss and preventing the model from becoming overfitting. The dense layer with the "softmax" activation function is the last layer of this model. The final dense layer categorizes the

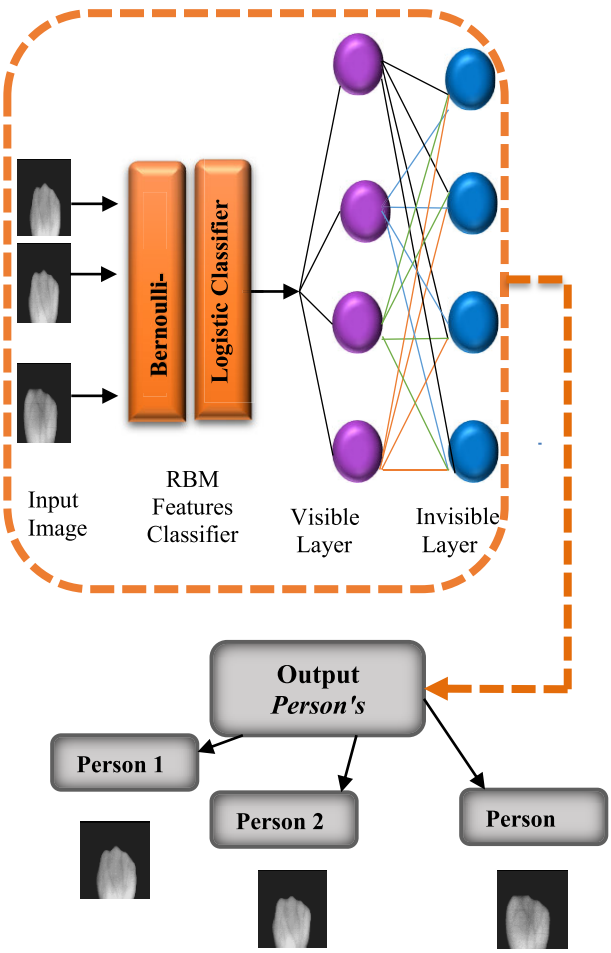


FIGURE 6. The structure of proposed RBM.

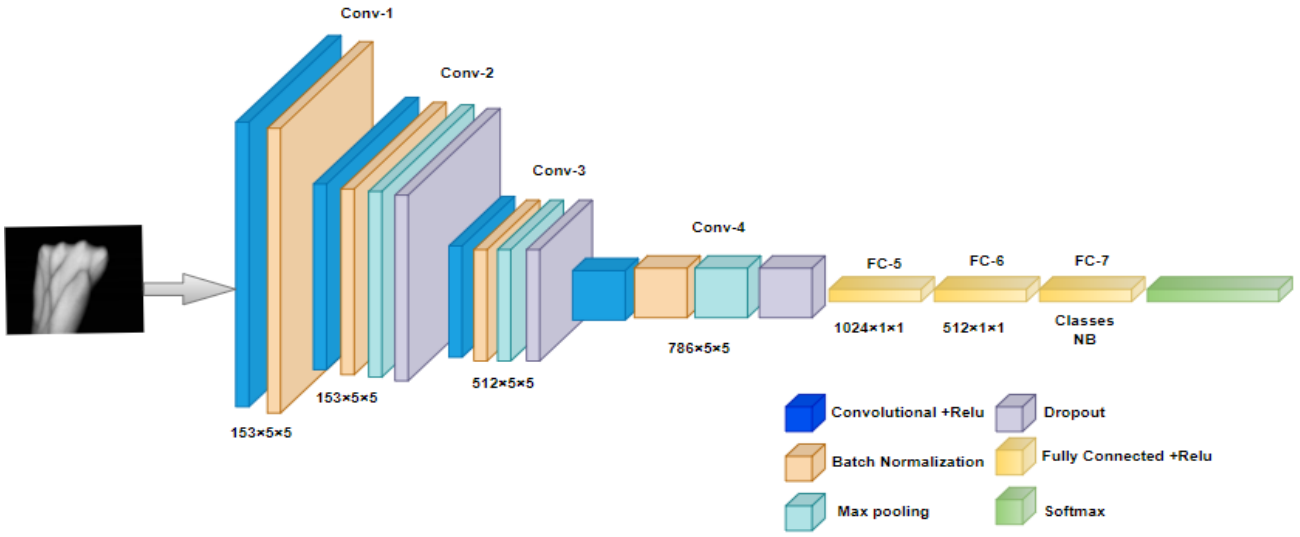


FIGURE 7. The structure of the proposed deep learning model.

50 classes when using the Dr. Badawi datasets, and the kernel size of the final dense layer grows to 138 when using the DHV Image datasets. This model works if the dorsal vein patterns of the left and right hands are very different. Setting the scale to 3 for Batch Normalization and adding a 0.3 dropout after each convolution layer keeps the model stable, and overfitting is avoided. Algorithm 2 listed the pseudo-code of CNN mode.

VIII. RESTRICTED BOLTZMANN MACHINES

Increasing the ability to recognize features of deep model transfer learning and decreasing recognition time, we present a deep learning technique for identifying images based on restricted Boltzmann machines (RBMs). This approach integrates the learning capabilities of two models., which conduct subject categorization by exacting structural higher-order statistical aspects of images. While the system transfers the taught convolutional neural networks to the target datasets, restricted Boltzmann machine layers can substitute fully connected layers. RBM, which has one visible layer and one invisible layer, operates unsupervised. The unseen layer attempts to rebuild the input as closely as possible after the information has been delivered to the visible layer. The neurons are stochastic binary units in both the visible and unseen layers. Figure 6 provides the Structure of proposed RBM. The pseudo-code of RBM is listed in Algorithm 3.

A certain joint configuration between two layers receives its energy from:

$$E(a, b) = -\sum_{i=1}^{n} x_i a_i - \sum_{j=1}^{z} y_j b_j - \sum_{i=1}^{n} \sum_{j=1}^{z} a_i b_j. \tag{19}$$

where $a_i$and $b_j$represent the units for the binary state in question, x_i and y_j are the biases, M_{ij} is the connection weight between the seen unit i and unseen unit j. According to the concept of the energy function, the probability of the visible and hidden layers existing is defined as:

$$p(a, b) = \frac{1}{Z} e^{-E(a,b)}. \tag{20}$$

where$Z = \sum_{a,b} e^{-E(a,b)}$ is the function of partitioning. As a result, the likelihood of the visible vector being assigned can be calculated by adding the probabilities of all possible hidden vectors, as shown below.

$$p(a) = \frac{1}{Z} \sum_{b} e^{-E(a,b)}. \tag{21}$$

The unit's conditional probability b_j's binary state is set to 1. Approximating a given visible vector a is possible, assuming no direct connections exist between the hidden units.:

$$p(a_j = 1 \mid b) = \sigma\left(y_j + \sum_i b_i M_{ij}\right). \tag{22}$$

where the sigmoid function is $\sigma(\cdot)$, Also b is a hidden vector. The probability that the visible unit will be 1 could be calculated as:

$$p(a_i = 1 \mid b) = \sigma\left(x_i + \sum_j b_j M_{ij}\right). \tag{23}$$

Using training examples, a generative model could be trained via a log-likelihood minimization technique. A more precise description of the log-likelihood would be as follows:

$$\mathcal{L}(C_{\text{train}}) = \sum \log p(a, b). \tag{24}$$

where the training dataset is C_{train}. can calculate the gradient of the log-likelihood concerning M correspondingly as:

$$\frac{\partial \log(p(a, b))}{\partial M_{ij}} = \left[\frac{\partial \log \mathcal{L}(C_{\text{train}})}{\partial M_{ij}}\right]_{\text{data}} - \left[\frac{\partial \log \mathcal{L}(C_{\text{train}})}{\partial M_{ij}}\right]_{\text{model}}. \tag{25}$$

Algorithm 2 *Deep Learning Proposed Method*

Input : *Dorsal Hand Dataset Size = (150,150), E: Epochs, CNN model parameters η: Learning Rate, b: Size of the batch; X_{train}: Dorsal Hand Training Dataset; X_{test}: Dorsal Hand Test Dataset*
Output : *Computed model evaluated metrics and predicted the person's veins*
*#**Start** Procedure*
Load Dataset
Preprocess *the data by applying (resizing, grayscale image, and sharping).*
Implement *a filtering process incorporating (MEDIAN filter, CLAHE filter)*
Apply *the morphological operation (e.g., skeleton, dilating, eroding, undesirable object removal).*
Split *(dataset): Prepare training, testing, and validating.*
    $X_{train}, y_{train} \leftarrow pre\ data(X_{train})$
    $X_{test}, y_{test} \leftarrow pre\ data(X_{test})$
Set *hyper-parameters.*
    $E \leftarrow 50$
    $\eta \leftarrow 10^{-7}$
    $b \leftarrow 16$
    $loss \leftarrow categorical\_crossentropy$
for local epoch e $\leftarrow$ *1 to E do*
    for $\boldsymbol{b} = (x, y) \in$ *random batch*
        Optimize the model using ADAM optimizer
    end for
end for
Initialize *the CNN model.*
    model. Sequential ()
Adjust *the model layer by adding*
    Add Conv2D
    Add Batch normalization
    Add MaxPooling2D
    Add Dropout
    Add Dense
Model.*compile (loss, Learning Rate).*
Implement *image augmentation techniques to expand the dataset.*
Training *model for vein recognition.*
    Model .fit (X_{train}, y_{train})
Load *the suggested model.*
for j=1:num test datasets
    split and evaluate (X_{test}, y_{test})
    (y_{pred}) = model predicts (X_{test})
    Acc = accuracy score (y_{test}, y_{pred})
    Loss = $(y_{true}\ log\ (y_{pred}) + (1\text{-}y_{true})\ log\ (1\text{-}y_{pred}))$
end for
Detect *Veins for Dorsal Hand for Persons.*
Compute*Accuracy.*
Plot *acc curve-loss curve.*
Prediction = *Recognition (Train CNN, Test dataset).*
Return *prediction.*
Evaluate *Metrics (Precision, Recall, F1-score).*
Train*(Vgg-Net, ResNet50, Inception V3, Xception, DenseNet, AlexNet, Mobile Net, and Efficient Net).*
Analogize *the models.*
Match *for all models.*
*#**End** Procedure.*

Algorithm 3 *Proposed RBM Methodology*

Input: Input Size = (50,50), I: Iterations
Output: Computed RBM model evaluated metrics " and predicted the person's veins.
*#**Start** Procedure*
Load Dataset
***Preprocess** the data by applying (resizing, grayscale image, and contrast).*
Notation: $x \leftarrow y$ means l is set to y. $l \sim k$ means l is sampled from k.
Require:
    A training set of N data vectors: $\{\boldsymbol{an}\}_n^N = 1$.
    Learning rate α.
    Iterations Number I.
    Components Number N.
    Parameters $\theta = \{\boldsymbol{M}, \boldsymbol{g}, \boldsymbol{p}\}$
***Set** hyper- parameters*
    $I \leftarrow 30$
    $\eta \leftarrow 0.1$
    $N \leftarrow 256$
***Load** the proposed model.*
for $q = 1$ *to* E
    for $n = 1$ *to* N
        $\boldsymbol{a}' \leftarrow \boldsymbol{an}, \boldsymbol{b}' \sim sigm\,(\boldsymbol{g} + \boldsymbol{Mva}')$
        $\boldsymbol{a}'' \leftarrow sigm\,(\boldsymbol{p} + \boldsymbol{Mb}'), \boldsymbol{b}'' \leftarrow sigm\,(\boldsymbol{p} + \boldsymbol{Ma}'')$
        $\theta \leftarrow \theta + \alpha \frac{\partial}{\partial \theta} E(\boldsymbol{a}', \boldsymbol{p}') - \frac{\partial}{\partial \theta} E(\boldsymbol{a}'', \boldsymbol{b}'')$
    end for
end for
RBM_Features_Classifier
    Bernoulli-RBM
    LogisticRegression
RBM_Features_Classifier.fit
***Detect** Veins for Dorsal Hand for Persons.*
***Prediction** = classification (RBM feature).*
***Compute**Accuracy.*
***Evaluate** Metrics (Precision, Recall, F1-score).*
***Compare** the RBM result with Deep learning result.*
*#**End** Procedure*

Allowing for Eqns. (18), (19), (23), and (24), it could be obtained by:

$$\frac{\partial \log(p(a,b))}{\partial M_{ij}} = \left[a_i b_j\right]_{\text{data}} - \left[a_i b_j\right]_{\text{model}}. \tag{26}$$

where[·]indicates the probability that the data or model will behave as expected.

IX. EVALUATION METHODS

The recommended model of deep learning is run on "Badawi hand vein dataset" with 50 classes and "Dorsal Hand Vein Image Database" with 138 classes. In this part, we compared the outcomes of the proposed deep learning model and the suggested RBM model to that of traditional models such as Vgg-Net, ResNet50, Inception V3, Xception, DenseNet, AlexNet, Mobile Net, and Efficient Net.

A. EVALUATION METRICS

Checking out how well the system is working is important. Numerous evaluation parameters are used for this reason. When the model recognizes the vein patterns, it provides four effective outcomes: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). While TNs are correctly identified as negative instances, while TP displays appropriately anticipated positive cases, FPs, and FNs are incorrectly labeled as positive and negative instances. As performance evaluation criteria in this study, recall, precision, accuracy, and F1-score are used, respectively. This study uses recall, precision, accuracy, and F1-score as evaluation

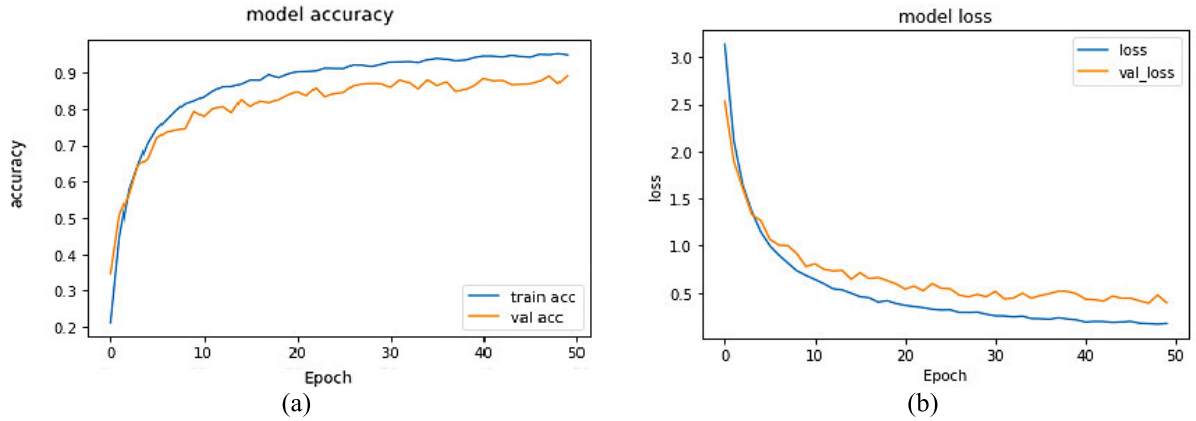


FIGURE 8. The accuracy (a) and loss curves (b) during training vgg-16 model.

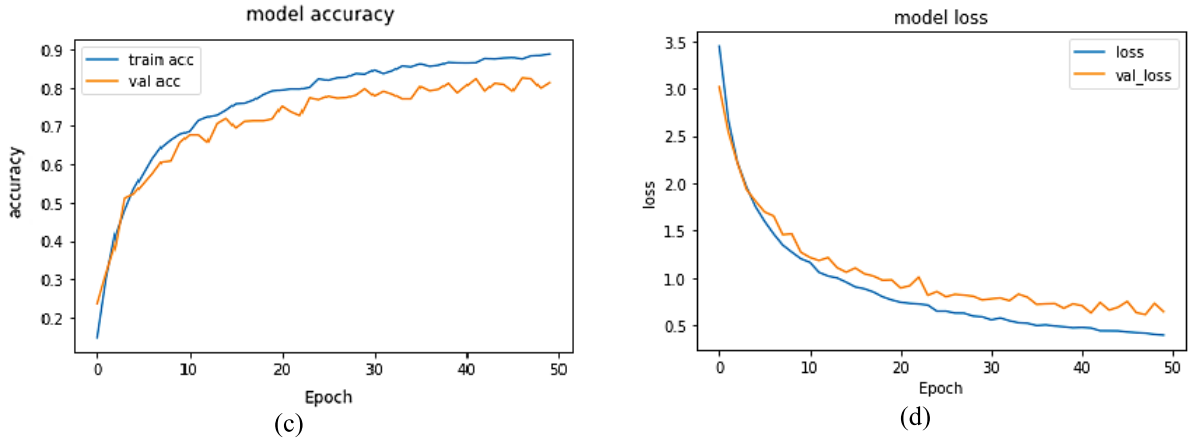


FIGURE 9. The accuracy (c) and loss curves (d) during training vgg-19 model.

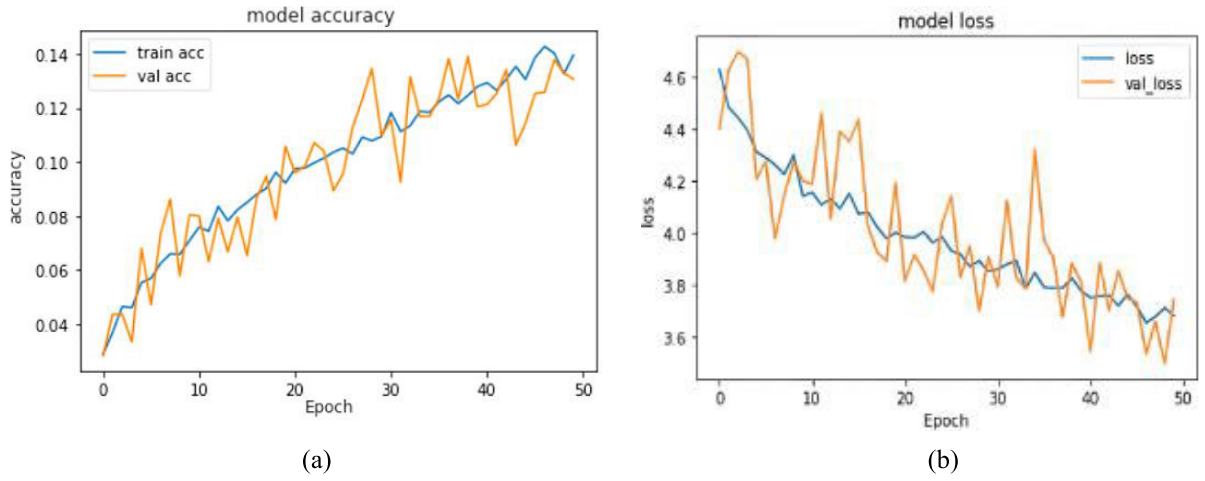


FIGURE 10. The accuracy (a) and loss curves (b) during training ResNet50.

metrics of achievement [38]. Accuracy can be defined as the fraction of vein patterns that are effectively identified compared to the entire number of vein patterns, as presented in equation (27).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{27}$$

As shown in equation (28), the ratio of accurately detected vein patterns to the whole number of vein patterns identified by the classifier is known as precision.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{28}$$

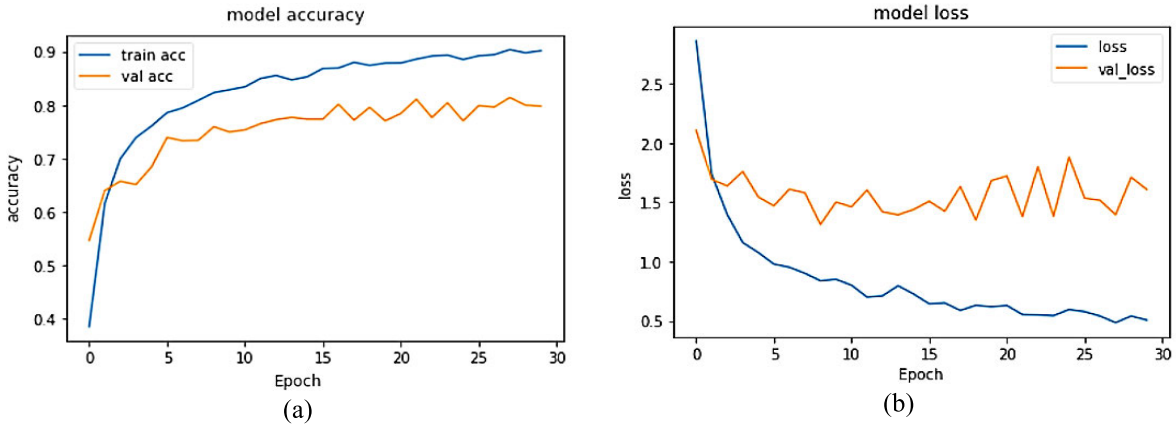


FIGURE 11. The accuracy (a) and loss curves (b) during training Inception V3 model.

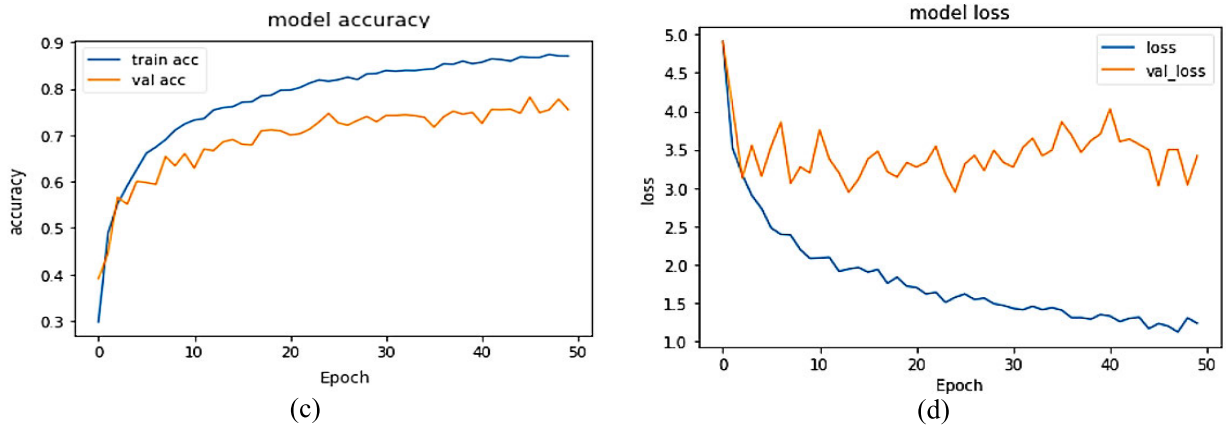


FIGURE 12. The accuracy (c) and loss curves (d) during training Xception model.

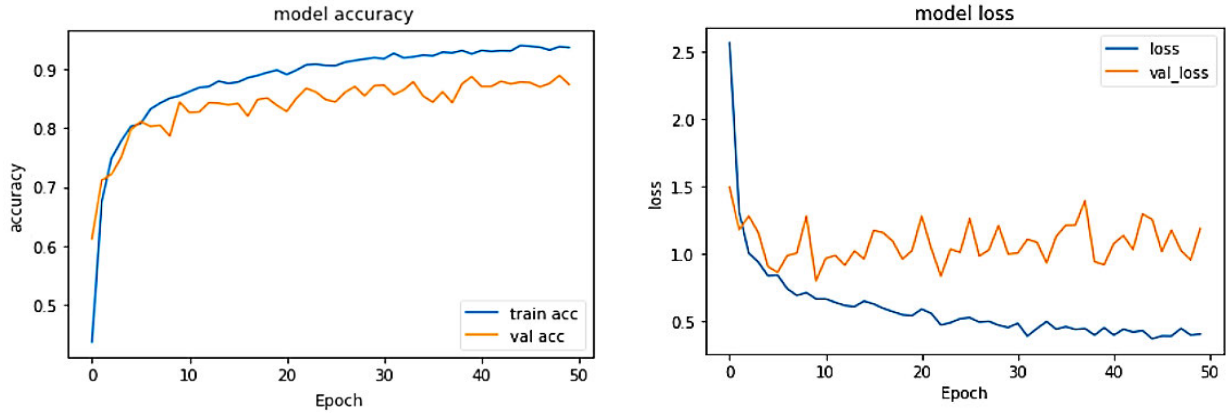


FIGURE 13. The accuracy (e) and loss curves (f) during training DenseNet model.

As shown in equation (29), the concept of recall is operationalized as the ratio between the count of accurately identified vein patterns and the overall count of vein patterns

$$\text{Recall} = \frac{TP}{TP + FN} \tag{29}$$

Model performance can be estimated using an evaluation metric called F1 score, which is an average of the model's precision and recall, as shown in equation (30).

$$F_{1\text{score}} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{30}$$

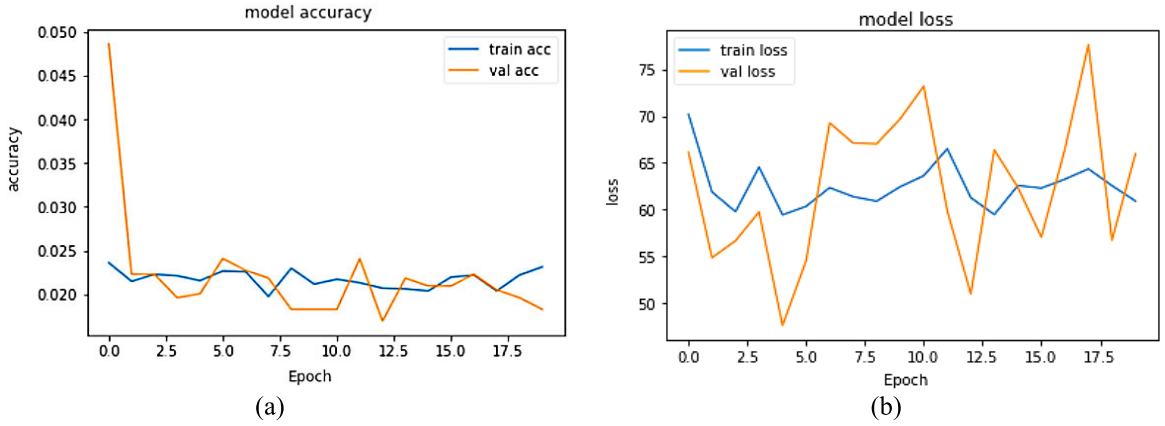


FIGURE 14. The accuracy (a) and loss curves (b) during training AlexNet model.

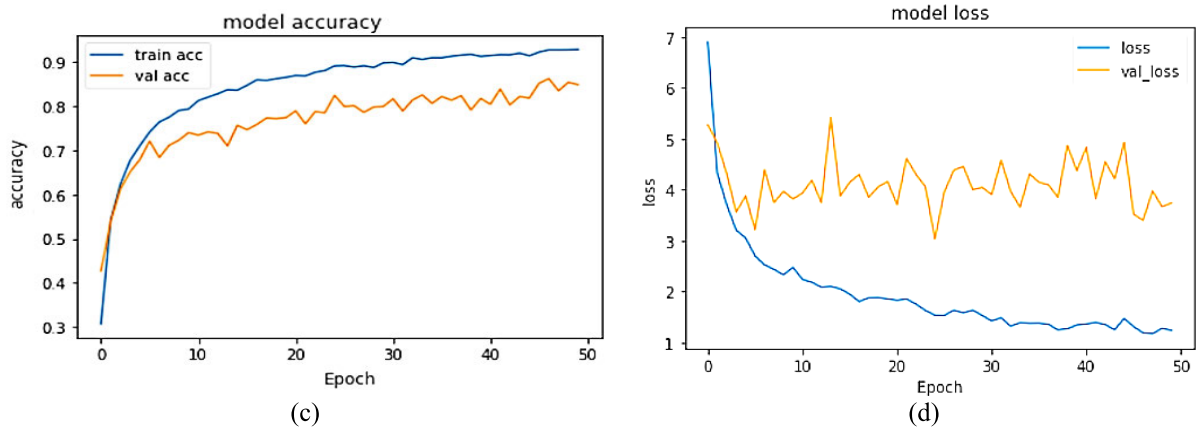


FIGURE 15. The accuracy (c) and loss curves (d) during training MobileNet model.

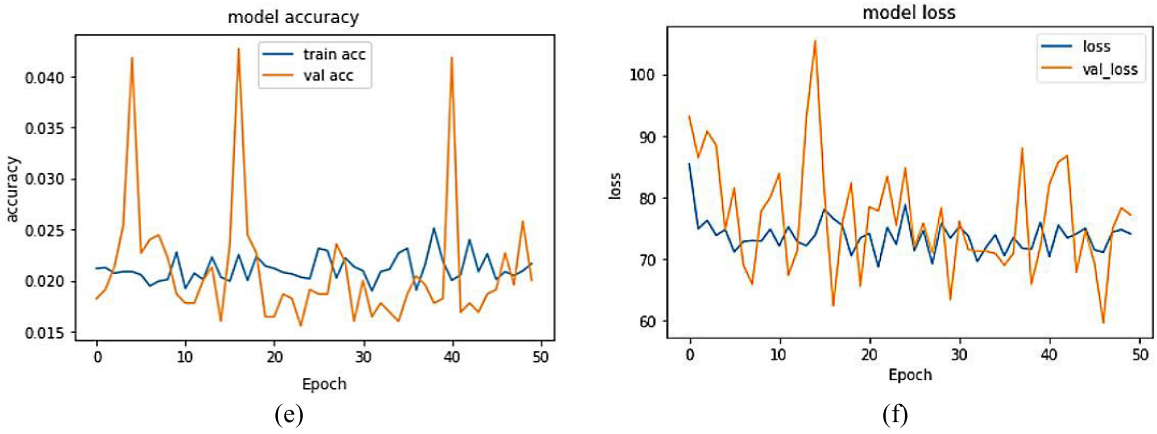


FIGURE 16. The accuracy (e) and loss curves (f) during training EfficientNet model.

X. RESULTS AND DISCUSSION

The nine common experiments used in this work utilized various deep-learning models. First, we compared the performance of the traditional CNN models Vgg-Net, ResNet50, Inception V3, Xception, DenseNet, AlexNet, MobileNet, and Efficient Net. Next, we focused on the suggested CNN model and RBM model. In the final step, the suggested CNN model and proposed RBM are compared with traditional CNN. Measures of performance such as accuracy, precision, recall, F1-score, and recall were tested in each experiment. Note that

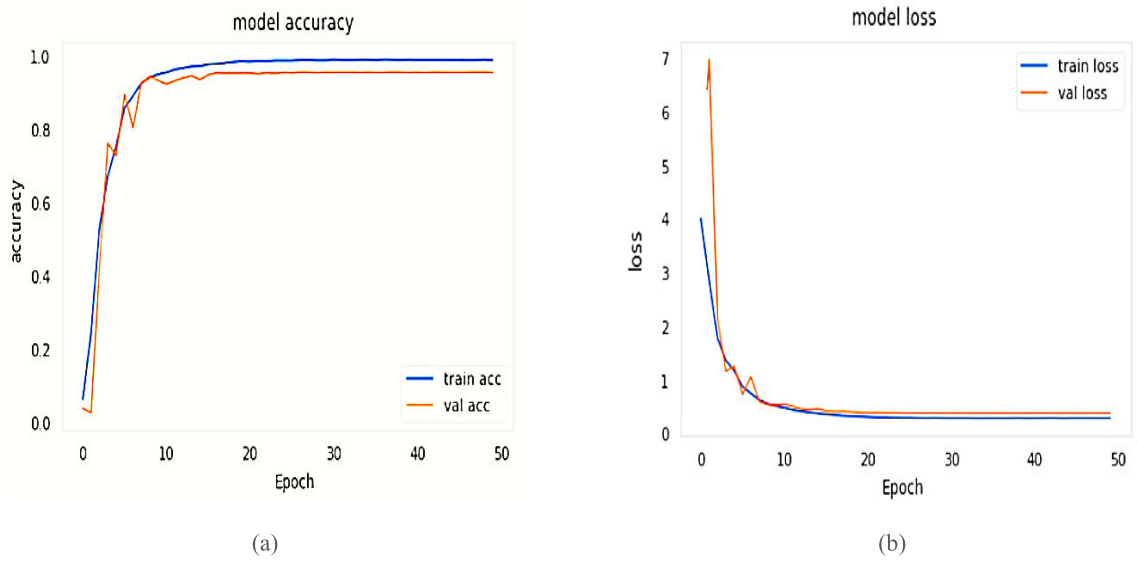


FIGURE 17. The proposed deep learning model for dorsal hand vein image datasets (a) Accuracy curve (b) Loss curve.

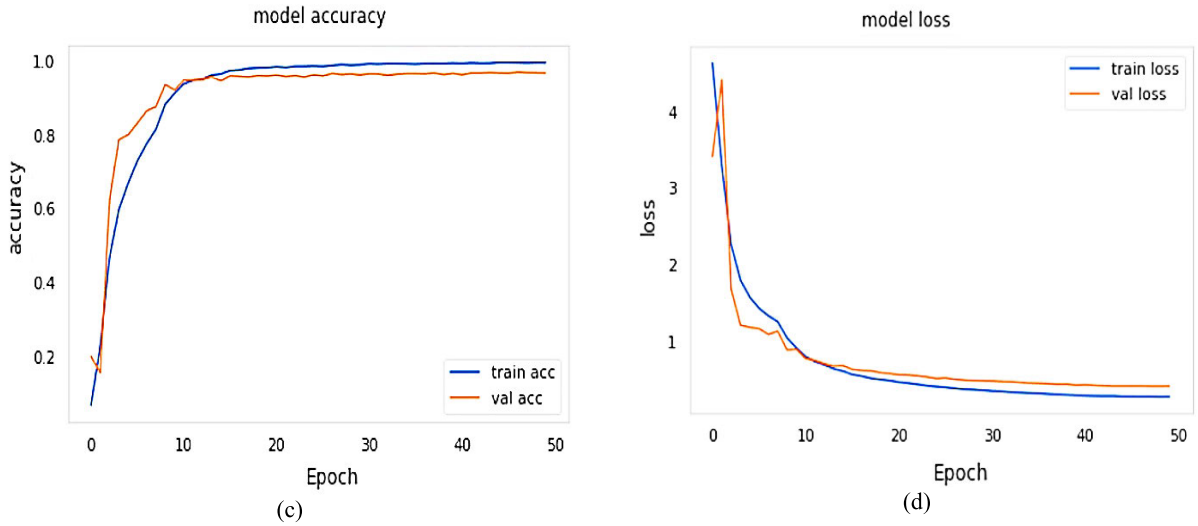


FIGURE 18. The proposed deep learning model for Badawi Hand Vein datasets (c) Accuracy curve (d) Loss curve.

the suggested models and the nine previously trained models were trained for 50 epochs. All of the trials were carried out on a PC with the following specifications: Microsoft Windows 10 operating system, a 5-core processor running at 4.0 GHz, 4 GB of RAM, and an NVidia Tesla NVIDIA Tesla P100 GPU.

A. PRE-TRAINED MODE

CNN's abilities are shown by putting nine distinct architectures, VGG-Net, ResNet50, Inception V3, Xception, DenseNet, AlexNet, MobileNet, and EfficientNet to work. The execution result of each model was evaluated in Table 2. Respectively, Figure 8,9,10,11,12,13,14,15,16 shows the traditional CNN model accuracy and loss curve. In Table 2 discusses the performance of traditional CNN models. EfficientNet achieved the lowest accuracy with 2.50%, followed by ResNet50 with 14.9%, and results started to improve in each architecture with AlexNet at 73.9, Xception at 87.3, VGG-19 at 88.8%, MobileNet at 92.8%, Inception V3 at 93.2%, and DenseNet at 93.9%, with the best accuracy being based on VGG-16.

B. PROPOSED CNN AND RBM MODELS

After evaluating pre-trained models, we proposed a deep learning model to achieve high accuracy and good performance greater than pre-trained methods. It makes hand vein recognition a reliable and accurate way of identifying individuals for security purposes. RBM has been suggested to reduce

TABLE 2. Traditional CNN architectures classification report.

Model	Accuracy(%)	Precision (%)	Recall (%)	F1-score (%)
VGG-16	95.2	89.1	88.1	88.5
VGG-19	88.8	82.5	82.1	82.2
ResNet50	14.9	13.9	12.6	13.2
Inception V3	93.2	90.3	89.2	89.7
Xception	87.3	78.1	77	77.5
DenseNet	93.9	88.8	85.2	86.9
AlexNet	73.9	70.2	70.1	70.14
MobileNet	92.8	86.2	80.2	83.0
EfficientNet	2.50	4.26	4.12	4.1

TABLE 3. Results of the proposed deep learning method.

Datasets	Methodology	Acc (%)	Val-acc (%)	Loss(%)	Val_loss(%)	Computational Time
Dorsal Hand Vein Image [29].	CNN	99.5	96.4	29.7	45.2	1145.52 s
Dr. Badawi hand vein [28].	CNN	99.7	95.8	30.2	39.4	1283.45 s
Dorsal Hand Vein Image[29].	RBM	99.9	96.7	31.1	40.6	137.235 s
Dr. Badawi hand vein [28].	RBM	99.9	95.5	30.8	43.5	142.745 s

TABLE 4. The proposed deep learning model classification report.

Datasets	Methodology	Precision (%)	Recall (%)	F1-Score (%)
Dorsal Hand Vein Image[29].	CNN	97	96	96
Dr. Badawi hand vein [28].	CNN	96	95	95
Dorsal Hand Vein Image [29].	RBM	99	99	99
Dr. Badawi hand vein [28].	RBM	99	99	99

TABLE 5. Overall performance as compared to other methods.

Reference	Dataset	Year	Accuracy (%)	Computational Time
Toygar[23]	Dr. Badawi	2020	98.67	4172.63 s
Al-johania andElrefaei[39]	Dr. Badawi	2019	99.25	
Babalol[40]	Dr. Badawi	2023	99.07	
Proposed Model	Dr.Badawi	2023	99.75	1283.45 s

computational time and increase the accuracy rate. Badawi hand vein dataset and the Dorsal Hand Vein Image Database was used to train the recommended models. The experimental outcomes of the proposed approach are presented in Table 3. The performance of the proposed deep learning model was evaluated in Table 4. Figures 17 and 18 show the curves for accuracy and loss of our experimental results.

C. THE COMPARISON BETWEEN CLASSICAL CNN AND THE PROPOSED METHOD

The proposed deep learning model's end result is compared with traditional CNN. The study aims to determine the proposed CNN model's ability to detect veins for the dorsal hand to achieve the best performance. The proposed methods achieve 99.50% recognition accuracy for Dorsal Hand Vein Image datasets and 99.76% for Dr.Badawi dataset rate in 50 epochs. In the other traditional models, Dr.Badawi dataset with VGG16 gives better accuracy than different architectures, reaching to 95.2% rate in 50 epochs. According to observations, the proposed CNN model appears to perform the best results. Table 5 lists further research projects that have been worked on in the area of DHV as a characteristic across time. The existing dorsal vein databases are shown, along with instances in which they have been utilized in the literature. The methods for recognition employed in those studies are also briefly described, along with the main findings of the experiments.

XI. CONCLUSION

This study presents ten deep-learning models, including customized CNN and RBM models, which were tested on two datasets to determine how well they identified DHV. Two datasets that are freely available to the public were used. To evaluate how various factors, act on the overall

performance of the models that are being tested, experiments were carried out by gradually increasing the training data size, the epochs number, the parameters, and the network complexity. The proposed models exhibited superior performance across the training, validation, and testing stages, with the acquired features contributing to their good performance. The proposed CNN technique outdid the other methods in terms of accuracy, with precision, recall, and F1 scores ranging from 97 to 96%, 96 to 95%, and accuracy scores from 99.2% to 99.7%, respectively. The proposed RBM technique outdid precision, recall, and F1 scores ranging from 99%, 99 to 99%, and accuracy scores from 99.9 %, respectively. This study predicts how useful vein patterns will be for biometric identification, proving that each person's hand has a different vein pattern.

REFERENCES

[1] M. S. Hosseini-Pozveh, M. Safayani, and A. Mirzaei, "Interval type-2 fuzzy restricted Boltzmann machine," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 5, pp. 1133–1142, May 2021.

[2] W. Yi, J. Park, and J.-J. Kim, "GeCo: Classification restricted Boltzmann machine hardware for on-chip semisupervised learning and Bayesian inference," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 53–65, Jan. 2020.

[3] R. Kumar, R. C. Singh, and S. Kant, "Dorsal hand vein recognition using very deep learning," in *Proc. Macromolecular Symp.*, 2021, vol. 397, no. 1, Art. no. 2000244.

[4] K. M. Alashik and R. Yildirim, "Human identity verification from biometric dorsal hand vein images using the DL-GAN method," *IEEE Access*, vol. 9, pp. 74194–74208, 2021, doi: 10.1109/ACCESS.2021.3076756.

[5] R. S. Kuzu, E. Maiorana, and P. Campisi, "Vein-based biometric verification using transfer learning," in *Proc. 43rd Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2020, pp. 403–409, doi: 10.1109/TSP49548.2020.9163491.

[6] S. Daas, A. Yahi, T. Bakir, M. Sedhane, M. Boughazi, and E. Bourennane, "Multimodal biometric recognition systems using deep learning based on the finger vein and finger knuckle print fusion," *IET Image Process.*, vol. 14, no. 15, pp. 3859–3868, Dec. 2020, doi: 10.1049/iet-ipr.2020.0491.

[7] H. Qin, M. A. El Yacoubi, J. Lin, and B. Liu, "An iterative deep neural network for hand-vein verification," *IEEE Access*, vol. 7, pp. 34823–34837, 2019, doi: 10.1109/ACCESS.2019.2901335.

[8] J. M. Song, W. Kim, and K. R. Park, "Finger-vein recognition based on deep DenseNet using composite image," *IEEE Access*, vol. 7, pp. 66845–66863, 2019, doi: 10.1109/ACCESS.2019.2918503.

[9] M. I. Obayya, M. El-Ghandour, and F. Alrowais, "Contactless palm vein authentication using deep learning with Bayesian optimization," *IEEE Access*, vol. 9, pp. 1940–1957, 2021, doi: 10.1109/ACCESS.2020.3045424.

[10] S. Tang, S. Zhou, W. Kang, Q. Wu, and F. Deng, "Finger vein verification using a Siamese CNN," *IET Biometrics*, vol. 8, no. 5, pp. 306–315, Sep. 2019, doi: 10.1049/iet-bmt.2018.5245.

[11] J. Zeng, F. Wang, J. Deng, C. Qin, Y. Zhai, J. Gan, and V. Piuri, "Finger vein verification algorithm based on fully convolutional neural network and conditional random field," *IEEE Access*, vol. 8, pp. 65402–65419, 2020, doi: 10.1109/ACCESS.2020.2984711.

[12] S. Trabelsi, D. Samai, F. Dornaika, A. Benlamoudi, K. Bensid, and A. Taleb-Ahmed, "Efficient palmprint biometric identification systems using deep learning and feature selection methods," *Neural Comput. Appl.*, vol. 34, no. 14, pp. 12119–12141, Jul. 2022, doi: 10.1007/s00521-022-07098-4.

[13] R. S. Kuzu, E. Maiorana, and P. Campisi, "Vein-based biometric verification using densely-connected convolutional autoencoder," *IEEE Signal Process. Lett.*, vol. 27, pp. 1869–1873, 2020, doi: 10.1109/LSP.2020.3030533.

[14] F. O. Babalola, Y. Bitirim, and Ö. Toygar, "Palm vein recognition through fusion of texture-based and CNN-based methods," *Signal, Image Video Process.*, vol. 15, no. 3, pp. 459–466, Apr. 2021, doi: 10.1007/s11760-020-01765-6.

[15] R. Garcia-Martin and R. Sanchez-Reillo, "Deep learning for vein biometric recognition on a smartphone," *IEEE Access*, vol. 9, pp. 98812–98832, 2021, doi: 10.1109/ACCESS.2021.3095666.

[16] H. Shao, D. Zhong, and X. Du, "A deep biometric hash learning framework for three advanced hand-based biometrics," *IET Biometrics*, vol. 10, no. 3, pp. 246–259, May 2021, doi: 10.1049/bme2.12014.

[17] D. Zhong, H. Shao, and X. Du, "A hand-based multi-biometrics via deep hashing network and biometric graph matching," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3140–3150, Dec. 2019, doi: 10.1109/TIFS.2019.2912552.

[18] S. Hom Choudhury, A. Kumar, and S. H. Laskar, "Biometric authentication through unification of finger dorsal biometric traits," *Inf. Sci.*, vol. 497, pp. 202–218, Sep. 2019, doi: 10.1016/j.ins.2019.05.045.

[19] L. C. O. Tiong, S. T. Kim, and Y. M. Ro, "Implementation of multimodal biometric recognition via multi-feature deep learning networks and feature fusion," *Multimedia Tools Appl.*, vol. 78, no. 16, pp. 22743–22772, Aug. 2019, doi: 10.1007/s11042-019-7618-0.

[20] H. O. Shahreza and S. Marcel, "Towards protecting and enhancing vascular biometric recognition methods via biohashing and deep neural networks," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 3, no. 3, pp. 394–404, Jul. 2021, doi: 10.1109/TBIOM.2021.3076444.

[21] S. Arora and M. P. S. Bhatia, "Presentation attack detection for iris recognition using deep learning," *Int. J. Syst. Assurance Eng. Manag.*, vol. 11, no. 2, pp. 232–238, Jul. 2020, doi: 10.1007/s13198-020-00948-1.

[22] J. Jayanthi, E. L. Lydia, N. Krishnaraj, T. Jayasankar, R. L. Babu, and R. A. Suji, "An effective deep learning features based integrated framework for iris detection and recognition," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 3, pp. 3271–3281, Mar. 2021.

[23] Ö. Toygar, F. O. Babalola, and Y. Bitirim, "FYO: A novel multimodal vein database with palmar, dorsal and wrist biometrics," *IEEE Access*, vol. 8, pp. 82461–82470, 2020, doi: 10.1109/ACCESS.2020.2991475.

[24] S.-Y. Jhong, P.-Y. Tseng, N. Siriphockpirom, C.-H. Hsia, M.-S. Huang, K.-L. Hua, and Y.-Y. Chen, "An automated biometric identification system using CNN-based palm vein recognition," in *Proc. Int. Conf. Adv. Robot. Intell. Syst. (ARIS)*, Aug. 2020, pp. 1–6, doi: 10.1109/ARIS50834.2020.9205778.

[25] A. Avci, M. Kocakulak, and N. Acir, "Convolutional neural network designs for finger-vein-based biometric identification," in *Proc. 11th Int. Conf. Electr. Electron. Eng. (ELECO)*, Nov. 2019, pp. 580–584, doi: 10.23919/ELECO47770.2019.8990612.

[26] N. Xu, Q. Zhu, X. Xu, and D. Zhang, "An effective recognition approach for contactless palmprint," *Vis. Comput.*, vol. 37, no. 4, pp. 695–705, Apr. 2021, doi: 10.1007/s00371-020-01962-x.

[27] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Feature transfer learning for face recognition with under-represented data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5697–5706, doi: 10.1109/CVPR.2019.00585.

[28] A. M. Badawi, "Hand vein biometric verification prototype: A testing performance and patterns similarity," in *Proc. Int. Conf. Image Process., Comput. Vis., Pattern Recognit.*, 2006, vol. 1, no. 3, pp. 3–9.

[29] F. Wilches-Bernal, B. Nuñez-Álvares, and P. Vizcaya, "A database of dorsal hand vein images," 2020, *arXiv:2012.05383*.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015, pp. 1–14.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[32] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

[33] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[35] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[36] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[37] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019 pp. 10691–10700.
[38] P. Ghosh, S. Azam, R. Quadir, A. Karim, F. M. J. M. Shamrat, S. K. Bhowmik, M. Jonkman, K. M. Hasib, and K. Ahmed, "SkinNet-16: A deep learning approach to identify benign and malignant skin lesions," *Frontiers Oncol.*, vol. 12, Aug. 2022, Art. no. 931141.
[39] N. Al-johania, L. Elrefaei, and B. University, "Dorsal hand vein recognition by convolutional neural networks: Feature learning and transfer learning approaches," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 3, pp. 178–191, Jun. 2019.
[40] F. O. Babalola, Y. Bitirim, and Ö. Toygar, "Dorsal hand vein biometrics with a novel deep learning approach for person identification," *Acta Scientiarum. Technol.*, vol. 45, Apr. 2023, Art. no. e61948.

RANA NOUR received the M.Sc. degree in communications engineering from Mansoura University, in 2019. She is currently an Assistant lecturer with the Mansoura Higher Institute for Engineering and Technology, Mansoura, Egypt. Her research interests include image and video processing, communication systems, and both machine and deep learning methodologies.

HOSSAM EL-DIN MOUSTAFA (Senior Member, IEEE) is a Professor with the Department of Electronics and Communications Engineering and the Founder and the former Executive Manager with the Biomedical Engineering Program (BME), Faculty of Engineering, Mansoura University. His research interests include biomedical imaging, image processing applications, and bioinformatics.

EHAB H. ABDELHAY (Member, IEEE) received the B.Sc. degree in communications engineering, and the M.Sc. and Ph.D. degrees from Mansoura University, Egypt, in 2005, 2010, and 2015, respectively. He is an Associate Professor with the Faculty of Engineering, Mansoura University. He is also the Programs Director of the Faculty of Engineering, Mansoura National University, Egypt, and the IEEE MNU Student Branch Counselor. He worked as a Demonstrator at the Department of communication and electronics - Faculty of Engineering, Mansoura University, from 2006, a Lecture Assistant from 2011, an Assistant Professor from 2015 to May 2022, and an Associate Professor from May 2022 till now. His research interests include 5G and beyond, WSNs, the IoT, cloud computing, AI, and cyber security.

MOHAMED MAHER ATA received the Ph.D. degree from the Electrical Communication and Electronics Department, Faculty of Engineering, Tanta University, Egypt, with the cooperation of Regina University, Canada. He is currently an Assistant Professor with the School of Computational Sciences and Artificial Intelligence (CSAI), Zewail City of Science and Technology, 6th of October City, Giza, Egypt. He is also an Assistant Professor with the Misr Higher Institute for Engineering and Technology, Mansoura, Egypt. He has published many indexed research articles (SJR-indexed and ISI-indexed) in the state-of-the-art of biomedical engineering, astrophysics, electrical communication, bioinformatics, encryption, cyphering, and intelligent transportation systems (ITS). His research interests include signal processing, image processing, multimedia, machine learning, deep learning, video processing, and computer vision.