Palm vein recognition based on convolution neural network

Ali Salam Al-jaberi [a], Ali Mohsin Al-juboori [b] *

[a] [b] University of Al-Qadisiyah, Al-Qadisiyah, Iraq, College of Computer Science and Information Technology

[a] Email: com.post02@qu.edu.iq

[b] Email: Ali.mohsin@qu.edu.iq

A R T I C L E   I N F O

A B S T R A C T

This paper presents a new validation method using a convolutional neural network for palm vein recognition. Unlike fingerprint and face. Vein patterns are endogenous biometric features that do not change over time and that make them difficult to identify and replicate in people. The proposed paper aims to provide a new way to identify people through their veins. This paper used the CASIA dataset, which consists of several wavelengths, in this research used the 850nm wavelength, which is clear in the veins, In addition, we divided the data into 3 cases. The first case is when the training and testing ratio is 50/50, the second case when it is 70/30, and the last case when it is 90/10. Obtained an accuracy of 98% in the case 90/10. In addition, to the proposed network, and used a well-known global network, the AlexNet network, where did the same work on it to compare the results of our proposed network with it. As proposed network outperformed it in terms of accuracy and speed, where the accuracy was 96% in the case 90/10.

1. Introduction

Personal validation, which is linked to individual verification, is a much-needed technique for security access systems. The correct certification of individuals is particularly important in managing the operations of many systems at the airport and in companies[1][2][3].

Traditional biometric characterization techniques are based on behavioral patterns or physiological characteristics such as fingerprints, faces, or eyes. However, these patterns present several defects in the determination of the individual[4]l [5]. Recently, the palm vein was being recognized and proved an effective effect of biometric validation.

In this paper, the system of biometrics in palm vein was applied in images of the acquisition of the generation of the palm within the hand and like other biometric patterns, the blood vascular patterns of palm are any individual. Unlike other biometric validation techniques, the receptacles are under the skin so it is difficult to falsify them.

∗Corresponding author: Ali Salam Al-jaberi

Email addresses: com.post02@qu.edu.iq

Communicated by: *Dr. Rana Jumaa Surayh aljanabi.*

In addition, the visual features of near-infrared facilitate the identification of palm vein, which seeks the thermal difference between the blood distribution in the blood vessels and the body-specific skin. In addition, palm vein[6].

The validation was discovered by a Charge Couple Device camera. The hemoglobin in a hand comfortable can absorb the Light Emitting Diode light by infrared at 700-900 nm after passing through the vein. A sensitive infrared CDD camera can then capture palm vein patterns [7][8]. A sufficient vein certification has many advantages compared to other biometric validation techniques: high precision, not susceptible to impersonation, and internal annexes so that they cannot be stolen. Moreover, the designs of palm veins do not change over the years (time), lack of blood or wounds, and are unique even for identical twins who share the same Deoxyribonucleic Acid pattern.

Deep literacy patterns, in particular convolution neural network networks, have demonstrated their capacity in diverse image recognition applications, such as visualization and verbal communication in biometrics. Deep literacy allows the automatic discovery of internal frameworks for high-dimensional training information.

This paper proposes a methodology based on a convolution neural network for biometric validation in palm veins. A convolution neural network is a multi-layer nervous network consisting of a series of cascade layers intercepted with a classification that ends with an intertwined network of layers. In one training able unit, a convolution neural network integrates the graphic sections of images, features, and components. It is designed to accept a two-dimensional, nonrecurring image that requires little preprocessing and recovery of the two-dimensional composition during the process. During training, classification is performed, ending with the final weights that act as an extracted feature of the classification of the entry sample. Convolution neural networks have been used in various case studies. For example (definition of vision, decoding, manual writing, and document identification).

The rest of the paper deals with the following sections: The second section explains the database that used in our work and how it is divided during the training process of the images. The third section explains the methods we used in the preprocessing stage to extract a clear pattern of the veins to insert into the network. The fourth section explains in detail artificial neural networks and how they work. The fifth section explains in detail the deep learning technology and the basic layers that we used in our network that we designed to identify people through the veins in the palm. The sixth section explained the structure of convolution neural networks that were used in our research. The seven sections decide the results and are discussed. The last section is the conclusion.

2.　Dataset

The database that used in our work contains 7200 images taken for 100 people, both males and females, taken with 6 waveforms and for two periods between each time interval. In addition, each person contains six images. The 850nm spectrum is used in our work because of the clarity of the veins in them due to the infrared rays penetrating the skin, which makes it easy to extract the veins through them. This rule is known as the CASIA dataset. In our work, divided this database in the training process into three cases (50-50, 70-30, and 90-10). Table (1) shows the split use in the training process[9].

TABLE 1 SPLIT OF DATABASE IMAGE

Number of images	Training	Test and validation
600	300	300
600	400	200
600	500	100

3. Preprocessing stage

This stage is considered one of the most important stages in extracting the vein pattern of the palm because the images captured in the CASIA database contain a lot of problems and noise. Since the images taken for each palm, will cause us problems in extracting the veins; we have implemented many methods on the images until we extracted the pattern of the veins and inserted it into the proposed network and AlexNet network. At first, extracted the region of interest, which is the region that contains the veins, which is free of problems such as hair and others to make it easier for us to extract the veins. extracted the region of interest by converting the image into image binary because the image is gray. When converted, some noise from the background appeared around the palm. Determine the joint points to extract the region of interest and draw the rectangle region of interest. After extract ROI, we applied some filters to remove the noise and improve the image to show us a clear vein pattern were applied the anisotropic diffusion filter, median filter, and closing morphological operation. After applying anisotropic diffusion filter, median filter, and closing morphological operation we subtract one image from another or subtract constant from image and Adjust image intensity values or color map to get a clear vein pattern until entering it in the proposed convolution neural network. Fig. 1 shows us the steps that we have applied to extract a clear vein pattern that entered into CNN.
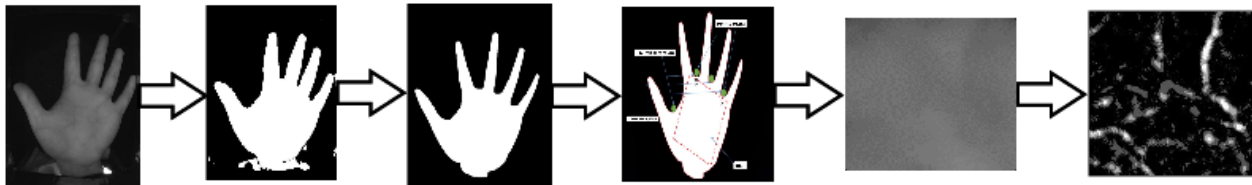

Fig. 1 - the preprocessing stapes

ALGORITHM 1 OTSU THRESHOLD

Goal: binary image
Input : rot_image   // rotate image
Output : bin_image
Begin:
Step 1:
Compute the histogram and probabilities of each intensity level
Step 2:
Set up initial ωi $(0) and$ $\mu i(0)$
Step 3:
Step through all possible thresholds $t = 1, … ….$

Maximum intensity

1. update ωi and μi

2. compute $\sigma_b^2(t)$

Step 4 :

The desired threshold corresponds to the maximum $\sigma_b^2(t)$

Step 5 :

end

ALGORITHM 2 EXTRACTS ROI

Goal: extract the region of interest

Input : Boundary extraction

W,H // the width and height for image

Output : ROI

Begin :

Step1:

 Determine thumb direction

Set c=0

   for all i **do** {where $0 \leq i \leq W$}

  for all j **do** {where $0 \leq j \leq H$}

if Img (i, j) =green & the i>point Then set point=i

c: **stop**

   end for

  end for

Step2:

Determine reference points

for all i, j **do** {where $0 \leq i \leq$ point-1, $0 \leq j \leq H$}

  if Img (i, j) = green **&** c =0 **then set** p1x = i, p1y = j, increment

c: **stop**

else If Img (i, j) =green **&** c=1 **then set** p1x = i, p1y = j, increment

 c: **stop**

else If Img (i, j) =green **&** c=2 **then set** p2x = i, p2y = j, increment

c: **stop**

else If Img (i, j) =green **&** c=3 **then set** p3x = i, p3y = j, increment

 c: **stop**

end for

Step 3:

draw line between two reference points

draw line (p1, p3)

 stop

Step4:

draw rectangle (ROI)

set h = line – Distance (green_point, End_line)/2 set w = h

draw rectangle (p1, p3, w, h)

stop

Step5:

end

ALGORITHM 3 MEDIAN FILTER

Goal: create enhancement image

Input : ROI

Output : Med_img   // median image

Begin:

---

for every pixel in the image **do**

    Sort values in the mask

    Pick the middle one in the sort list

    Replace the pixel value with median one

end

4. *The Artificial Neural Networks (ANN)*

ANN is several interconnected computational contracts. ANN is a computational approach to problems in which adequate representation is found to solve the problem; it is difficult for traditional computer programmers. Artificial neural networks are sometimes referred to as boxes black because it is impossible to understand their work (function) [10][11].

4.1 The Multi-Layer Perceptron structure (MLP)

Multi-Layer Perceptron (MLP) It is a forward feeding structure for artificial neural networks containing more hidden layers, and each layer consists of a simple, mathematically connected combination of the contract known as neuro cells as shown in Fig. 2 [12][13][14].
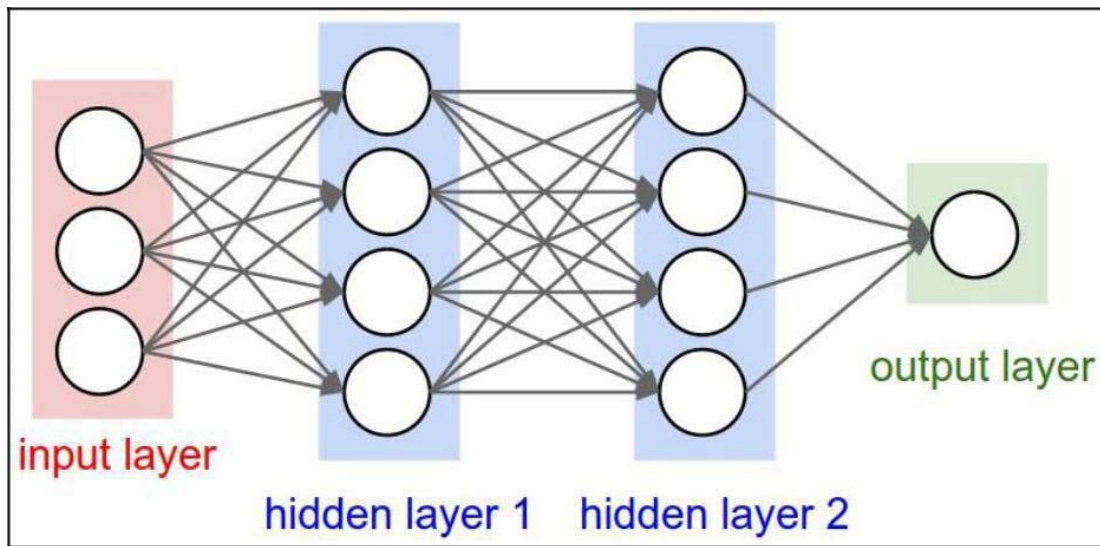

Fig. 2 - Multi-Layer Perceptron Structure (MLP) [11]

The value of the output is described for each j neural cell in layer l in equation (1) [11] [13].

$$z = \sum_i^N (Wi, j. Xi) + b \qquad (1)$$

Where N is the input number, i.e. inputs that passed from (l-1), (Wi, j) are the weight of the edge linking the neural of the presentation layer with the neural of the former layer, and (b) like bias, the equation (1) is written in simplified matrices to be as follows [11] [12] :

$$Z = W^T x + b \qquad (2)$$

4.2 Activation Functions

This function is to be determined, as the neuro cell should be activated or not activated by calculating the total weight and addition with bias. The main purpose of using the activation function is to give the irritation to extract nervous cells, for learning and carrying out complex tasks. Back Propagation makes it possible the most active function used is the corrected written unit (ReLU), and equation (3) illustrates the representation of the RLU[14].

$$O(x) = \max(x, 0) \qquad (3)$$

The output layer needs to assess the probability of individual categories. "For this reason, softmax is the most common activation function used to equate the output layer (4). Clarify this function.

$$O(x) = \frac{e^{xk}}{\sum_j^{j=1} e^{xj}} \qquad (4)$$

Where xk the corresponding production of category k and j is the total number of classes [10] [15] [13].

4.3 Loss Function

The loss function is used in the neurological network to estimate the prediction of error there are different types of loss functions such as cross-entropy and mean square error, and this method use in Neural networks for multiple categories using softmax.

$$L = \frac{1}{n} \sum_{k=1}^{n} \ln o(x^k) + (1 - z^k)\ln(1 - o(x^k)) \qquad (5)$$

Where "n" represents the number of categories, "z^k" is the required output of category k, while "o (x^k)" is the estimated probability of category k, calculated using the SoftMax activation function described in the equation (4) [11] [14] [15].

In the MLP layers, weight edited (adjusted) during training, the backpropagation algorithm is the most common algorithm used to control weight in the training process.

4.4 The Back Propagation

The backpropagation is a supervised algorithm to be used in the training of neural networks. Because of their high efficiency and simplicity. This algorithm uses the proportions technique to reduce neural system errors.

At the beginning of the training process, weights and biases are set up in the network with a random number, and there are two phases of counter-proliferation [10] [16].

I.   Propagation Forward

The Input instance is propagated through the entire network. Layer by layer beginning from input to output using equation (1) and equation (3) to produce a prediction value.

II.   Propagation Backward

Backward the second part of the forward end, which begins with the calculation of the error and the propagation of a layer after another from outputs to inputs, and the weights and biases are updated accordingly.

To do backward propagation, a job that continues with derivatives is required. There are two types of active posts based on the type of appointment required from entry to graduation.

The first is a nonlinear function such as sigmoid, and the second type of activation function is the liner function (ReLU) explained in equation (3). When these stages are. Repeat all inputs. This is called an epoch. The neural network can run. For many epochs. As it is required to find it. Solution [9].

5. The Technique Deep Learning

Deep learning is a category of automated learning algorithms that utilize multiple layers to extract benefits at a progressively higher level than initial inputs. For example, in the processing of images, the lower classes may identify the margins, while the upper layers may define personal related concepts such as numbers, letters, or faces. This process is carried out through successive layers that further complicate the steps taken. Each layer is productive. It passes as an entrance. To the next layer, it is used to learn the advantages of the higher (more complex) level, as explained. Figure (3) [9] [13]
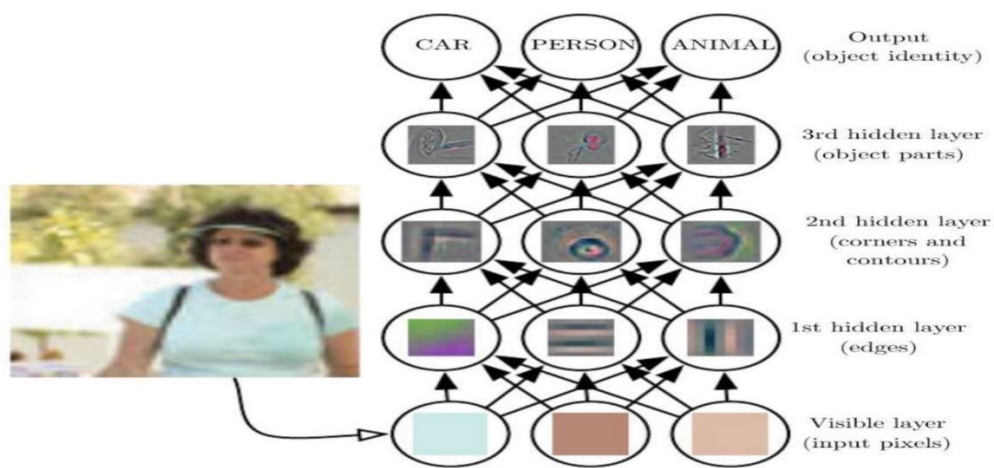

Fig. 3 - Abstract of Deep Neural Network diagram [9].

5.1 The Deep Neural Networks (DNN)

Theoretically, deep neural networks (DNN) are artificial neural networks (ANN) with many hidden layers. MLP is one of the most commonly used ANN structures for DNN. Because a neuron network consists of layers of interconnected neurons, it is almost impossible to successfully train more than a few hidden layers. Given the number of weights in the network that can easily reach thousands or even millions, DNN requires very large calculus and feeding data for training stages.

5.1.1   The structures of Convolution Neural Networks (CNN)

The convolution neural network is one of the most widely used networks in the problem of computer vision, a category of Deep Neural Networks that rely on multi-layered perception (MLP) and techniques backward propagation [8]. They differ from the traditional MLPsr by combining several locally connected layers used for extraction features, followed by several fully connected layers used for classifying [17][18].

The most important characteristic of the NN is local communication and the use of common weights, so that they can identify features local of the portray [14].

The CNN model consists of three different layers:

- The layers Convolution.
- The layers pooling (or subsamples).
- The layers are fully connected (FC).

I.   Convolution layers

When you enter the inputs to this layer, it convolutes it with a constant kernel K to produce n of the feature map, as shown in equation (6)

$$Y^l = K^l \otimes X^{l-1} \quad (6)$$

K^l Refers to the ice in the aversion class, $\otimes$ referring to the operation of convolution. Figure (4) shows one neural with a single entry map. The output of this layer is mostly through patch normalization to speed the training process and through ReLU non-line activation function as in equation (3) [10] [19].
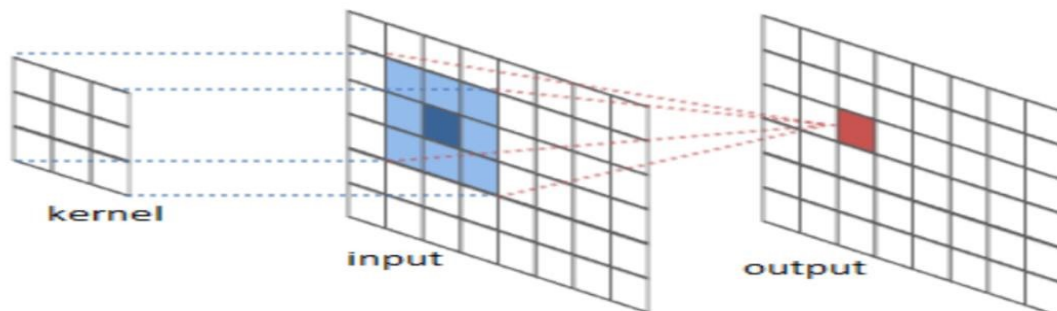

Fig. 4 - Example of convolution layer [10]

II.   Batch normalization

This layer is used to speed up the training process and eliminate sensitivity to network development. It reduces a large number of each channel. First, the activation of each channel is normalized by introducing the mean mini-batch and dividing the standard deviation of the mini-batch. Batch Normalization of its entrance xi by calculating mean $\mu\beta$ and variation $\sigma^2\beta$ through a mini-batch and across each input channel.

$$\hat{xi} = \frac{xi - \mu\beta}{\sqrt{\sigma^2\beta + \epsilon}} \qquad (7)$$

Here, ϵ (epsilon properties) improves numerical stability when the small difference between the mini patch variance is very small. To allow for the possibility that the inputs with zero mean difference and unity are not ideal for the layer that follows the batch normalization layer.

$$yi = \gamma\hat{xi} + \beta \qquad (8)$$

Here, the scale factor γ and offset β are parameters learnable to update during the training of the network. [20] [21].

III.   Pooling layers

Its main objective is to limit. The spatial size of the features maps of the production of the filtering layers, which used almost by the stride $s \in N \geq 2$, which is used to reduce the data to $\frac{1}{s^2}$ of data. Figure (5) shows the average pooling how it works.
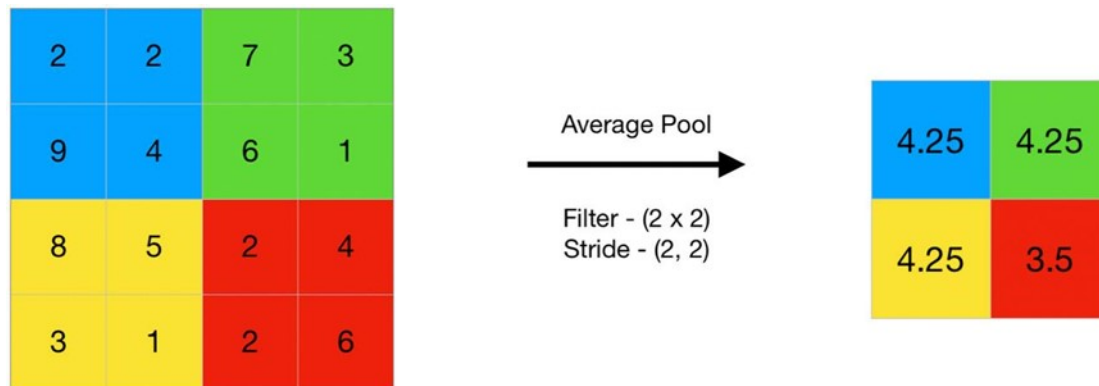

Fig. 5 - Example of average pooling

IV.   The layer is Fully connected

The fully connected layer is the traditional layer of MLP described in section 4.1 where all neural cells in the class are connected to all neural cells in the next class, which is fully related to being used for classification. The dropout technique can be applied to this layer to prevent the problem of over-fitting [10].

6.   The Architecture of Convolution neural network

This research aims to develop a new convolution neural network structure and to assess the extent to which CNN-based solutions can practice by identifying palm veins. As we mentioned earlier, the first one adopts the design of a network that has trained in the CASIA database. The second is AlexNet.

6.1 The proposed CNN

Using Convolutional neural networks CNN in which the network is based on convolution layers; in our experiment, we have baser our architecture on three convolutions, each convolution is successes by a batch normalization layer and a ReLU activation layer (as all recommendations on convolution neural network implementation). So have nine layers (conv1 + batch Nrml + ReLU) *three and have the input and output layers those are 11 layers the remain three layers are the average pooling layer + two obligatory layers which are (softmax layer and classification layer).
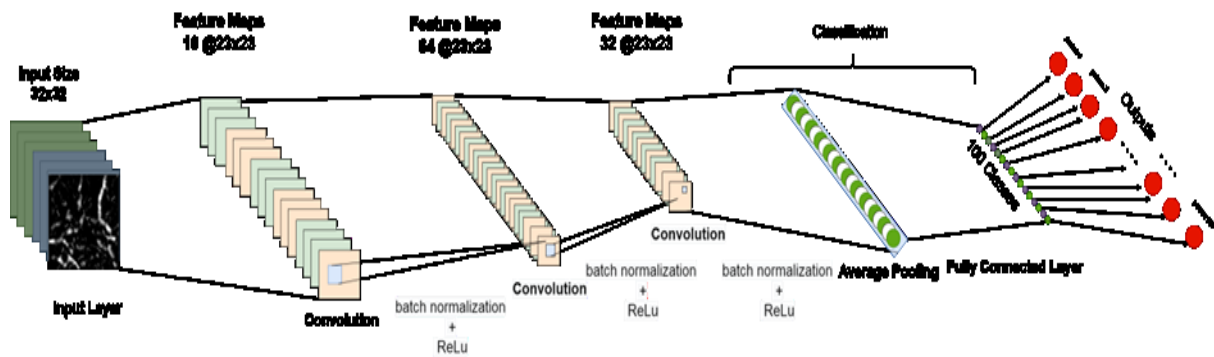

Fig. 6 - Proposed CNN Architecture for palm vein recognition

TABLE 2 THE ANALYSIS RESULT OF OUR PROPOSED CNN.

Layer	Name	Type	Activation	Learnable
1	Input	Image input	32*32*1	-
2	Conv_1	Convolution	32*32*16	Weight 23*23*1*16 Bias 1*1*16
3	BN_1	Batch normalization	32*32*16	Offset 1*1*16 Scale 1*1*16
4	Relu_1	ReLU	32*32*16	
5	Conv_2	Convolution	16*16*64	Weight 23*23*16*64 Bias 1*1*64
6	BN_2	Batch	16*16*64	Offset

		normalization		1*1*64 Scale 1*1*64
7	Relu_2	ReLU	16*16*64	
8	Conv_3	Convolution	16*16*32	Weight 23*23*64*32 Bias 1*1*32
9	BN_3	Batch normalization	16*16*32	Offset 1*1*32 Scale 1*1*32
10	Relu_3	ReLU	16*16*32	
11	Avpool	Average pooling	8*8*32	
12	Fc	Fully connected	1*1*100	Weight 100*2048 Bias 1*100
13	Softmax2	Softmax	1*1*100	-
14	Class output	Classification output	-	-

6.2 AlexNet

AlexNet has eight layers with parameter learnable. The model consists of five layers with a set of max pooling, followed by three fully connected layers, and uses rectified linear unit's activation in each of these layers, except for the output class. They discovered that the use of rectified linear units as an activation function has accelerated the training process about six times. They also used dropout layers, which prevented their model from over-processing. Furthermore, the model is being trained in the Image net dataset. The Image net dataset contains approximately 14 million images across 1,000 categories.

7. The results and discussions

The dataset CASIA palm image was implemented. It contains 100 people with six samples of each hand (left and right palm).

In addition, we divide the data set into 100 classes, and each class contains six images of training, testing, and validation. The training was conducted with 20 epoch with 120 iterations and a learning rate of 0.01 for our CNN, while AlexNet with 100 epochs with 5000 iterations and a learning rate of 0.00001. The system works on Intel(R) Core(TM) i7-8565U, RAM 16 Gigabyte, Hard 1 Terabyte with Hard 500-Gigabyte m.2, VGA 4 Gigabyte.

Table (3) shows us the results obtained after training the images of the palm in the database CASIA. In addition, we observe the superiority of our network that we designed on the Alex network with accuracy and at the time of training as the AlexNet network take a lot of time to obtain this accuracy due to the many iterations, as for our designed network takes less time in the training process and has demonstrated high efficiency.

TABLE 3 THE ACCURACY OF PROPOSED CNN AND ALEXNET.

Images split	Our Proposed CNN	AlexNet
50%-50%	Accurate =97%	Accurate =87.33%
70%-30%	Accurate =97.50%	Accurate =91%
90%-10%	Accurate =98%	Accurate =96%

8. Conclusion

In this paper, we have identified people through the veins inside the palm in a new way, which is by using the convolutional neural network. As the vein in the palm is unique and not recurring in people, even twin brothers. used the global database CASIA. Where designed network for the proposed convolutional neural and after extracting the vein pattern from the palm of the preprocessing processes and the process of extracting the region of interest. We showed a pure vein pattern and inserted it into the proposed network, where divided the data into 100 rows on the number of people and divided them into three cases. The first case is when the training and testing process is 50/50. The second case is when it is 70/30, and the last case is when it is 10/90. Also trained the vein pattern extracted on the AlexNet network to compare it with the proposed network. The proposed network showed a clear superiority in terms of accuracy and speed in training.

References

[1]   A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, 2004, doi: 10.1109/TCSVT.2003.818349.

[2]   S. C. Eastwood, V. P. Shmerko, S. N. Yanushkevich, M. Drahansky, and D. O. Gorodnichy, "Biometric-Enabled Authentication Machines: A Survey of Open-Set Real-World Applications," *IEEE Trans. Human-Machine Syst.*, vol. 46, no. 2, pp. 231–242, 2016, doi: 10.1109/THMS.2015.2412944.

[3]   A. F. Sequeira and J. S. Cardoso, "Fingerprint liveness detection in the presence of capable intruders," *Sensors (Switzerland)*, vol. 15, no. 6, pp. 14615–14638, 2015, doi: 10.3390/s150614615.

[4]   Z. M. Noh, A. R. Ramli, M. Hanafi, and M. I. Saripan, "Review on a Palm Vein Infrared Image Acquisition System," no. December, pp. 16–17, 2013.

[5]   B. Prommegger and A. Uhl, "Rotation Invariant Finger Vein Recognition," *2019 IEEE 10th Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2019*, no. 61173069, pp. 151–156, 2019, doi: 10.1109/BTAS46853.2019.9185982.

[6]   X. Zhu and D. Huang, "Hand Dorsal Vein Recognition Based on Hierarchically," no. 2, pp. 157–164, 2012.

[7]   J. M. Cross and C. L. Smith, "Thermographic imaging of the subcutaneous vascular network of the back of the hand for biometric identification," *IEEE Annu. Int. Carnahan Conf. Secur. Technol. Proc.*, no. v, pp. 20–35, 1995, doi: 10.1109/ccst.1995.524729.

[8]   L. Wang and G. Leedham, "Near- and far- infrared imaging for vein pattern biometrics," *Proc. - IEEE Int. Conf. Video Signal Based Surveill. 2006, AVSS 2006*, pp. 52–57, 2006, doi: 10.1109/AVSS.2006.80.

[9]   A. S. Al-jaberi and A. Mohsin Al-juboori, "Palm vein recognition, a review on prospects and challenges based on CASIA's dataset," pp. 169–176, 2021, doi: 10.1109/dese51703.2020.9450241.

[10]   N. G. Peppes and N. G. Peppes, "Ship Detection in Satellite Images With Deep Learning and a Pythonic Interface on a Hadoop HDFS Platform

MASTER THESIS Ship Detection in Satellite Images With Deep Learning and a Pythonic Interface on a Hadoop HDFS Platform MASTER THESIS," 2018.

[11]   Nada B.jarah, "Deep Learning in Wireless Sensor network," vol. 13, no. 1, pp. 11–17, 2019.

[12]   E. Calli, "Faster Convolutional Neural Networks," p. 11, 2017, [Online]. Available: https://theses.ubn.ru.nl/bitstream/handle/123456789/5233/Çalli%2C_E._1.pdf?sequence=1.

[13]   H. Mohsin and H. Hasan, "Convolution Neural Network with Dual Tree Complex Wavelet Transform Preprocessor for Universal Image Steganalysis," vol. 11, no. 3, pp. 43–58, 2019.

[14]   H. Mohsin Ahmed and H. Hasan Mahmoud, "Convolution Neural Networks for Blind Image Steganalysis: A Comprehensive Study," *J. Al-Qadisiyah Comput. Sci. Math.*, vol. 11, no. 2, pp. 53–64, 2019, doi: 10.29304/jqcm.2019.11.2.573.

[15]   H. I. Abdulrazzaq, "Robust Authorization Based on Multi Biometric Techniques," 2018.

[16]   T. M. Mezza, N. R. Armstrong, G. W. Ritter, J. P. Iafalice, and M. E. Kenney, "Electrochemical studies on stacked-ring silicon phthalocyanines," *J. Electroanal. Chem.*, vol. 137, no. 2, pp. 227–237, 1982, doi: 10.1016/0022-0728(82)80038-7.

[17]   L. M. Q. De Santana, R. M. Santos, L. N. Matos, and H. T. Macedo, "Deep Neural Networks for Acoustic Modeling in the Presence of Noise," *IEEE Lat. Am. Trans.*, vol. 16, no. 3, pp. 918–925, 2018, doi: 10.1109/TLA.2018.8358674.

[18]   A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. M. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Anal. Appl.*, vol. 21, no. 3, pp. 783–802, 2018, doi: 10.1007/s10044-017-0656-1.

[19]   M. Thoma, "Analysis and Optimization of Convolutional Neural Network Architectures," Karlsruhe Institute of Technology, MSC thesis, "2017.

[20]   Laurent, César, et al. "Batch normalized recurrent neural networks." 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE,"2016.

[21]   Chen, Jun-Cheng, Vishal M. Patel, and Rama Chellappa. "Unconstrained face verification using deep CNN features." 2016 IEEE winter conference on applications of computer vision (WACV). IEEE,"2016.