DOI: 10.21917/ijivp.2022.0399

PALM VEIN CLASSIFICATION FROM LARGE DATASETS USING DEEP CONVOLUTIONAL FUSION LEARNING

B. Chellapraba[1], K. Periyakaruppan[2], D. Manohari[3] and M.S. Kavitha[4]
[1]Department of Information Technology, Karpagam Institute of Technology, India
[2,4]Department of Computer Science and Engineering, SNS College of Engineering, India
[3]Department of Computer Science and Engineering, St. Joseph's Institute of Technology, India

Abstract
Biometric techniques are currently among the most widely used methods all over the world for determining a person identity. This trend is expected to continue in the near future. In this study, we focused on palm vein is used as a strategy to improve biometric authentication systems by combining a method that is based on texture with a method that is based on a convolutional neural network (CNN). The simulation is used to test the performance of the model on several different datasets. In simulations, the suggested method routinely achieves better results than the current best practice on each and every dataset.

Keywords:
Palm Vein, Classification, Convolutional Neural Network

1. INTRODUCTION

The development consistently collect the biometric features has been of great assistance to the field of human recognition, which uses these features to identify individuals [1]. Numerous studies have been conducted in this sector on the issue of unimodal and multimodal biometric features for the purpose of identifying or verifying individuals [2]. The user's face, iris, fingerprints, palmprints, stride, and voice are the many forms of biometric information that are most frequently utilized in these kinds of systems. On the other hand, in more recent times, a trend toward employing vein photographs taken from the finger, palm, and even the wrist has emerged.

In recent years, there has been a rise in interest in palm vein biometrics. This can be attributed to the fact that palm veins are located beneath the skin, which makes them nearly impossible to fake. In addition to this, there is no blockage and it does not make any noises like hair does. The use of sensors that are capable of capturing palm vein patterns at a low cost and that are becoming more generally available has led to an increase in popularity for their implementation in high-security authentication systems. People are more inclined to utilize these kinds of devices if taking pictures with them is simple and can be done in a short amount of time [7].

When we speak of precision, what we are truly referring to is the correctness of our categorisation. The proportion of accurate predictions made in comparison to the total number of samples fed into the system is one way to evaluate the performance of a classification algorithm [8]. It is easy to be fooled by the high level of accuracy that is stated for classifications. The issue arises due to the increased likelihood that specimens taken from a segment of the population may have their identities incorrectly assigned.

There has been a significant amount of investigation into this topic. For example, previous studies that were based on deep learning for palmprint recognition and palm vein detection only used simple networks and lacked in-depth study. The accuracy of CNNs' pattern identification will continue to improve over the next few years as a result of the exponential proliferation of CNNs in the market [9]. It is projected that convolutional neural networks, often known as CNNs, will become increasingly important as a crucial tool for detecting palmprints and palm veins in both 2D and 3D formats. It is necessary, as a result, to conduct extensive research on the recognition performance of conventional CNNs for the purpose of 2D and 3D palmprint identification as well as palm vein recognition. The purpose of this work is to accomplish just that, and in order to do so, it evaluates the performance of conventional CNNs in terms of reading 2D and 3D palmprints as well as reading palm veins. In specifically, seventeen well-known and established CNNs are accessed for the purpose of research and analysis.

2. RELATED WORKS

A comparable palm vein biometric mode was proposed in [10] for use as an access control mechanism in a security system with multiple layers. They found that their system could authenticate subjects with a success rate of 92.0% on average by using techniques such as template matching and PCA as palm vein verification algorithms. They found this out by using their system.

The authors in [11] developed a vast contactless palm vein dataset that was assembled by researchers at Tongji University. This dataset was introduced in 2018. A palm print and palm vein recognition system based on a deep convolutional neural network was presented in this paper. The results of the experiment involving palm print and palm vein identification were revealed.

In [6], the authors constructed a system that was based on a near-infrared camera in order to gather images of palm veins. After acquiring these images, Lee employed a 2-D Gabor filter in order to extract properties from them. Following their studies, the authors stated that their method had an accuracy of 99.18% and an EER of 1.82%.

In [4], the authors used an adaptive Gabor filter to improve upon, which required the parameters to be initialized at the beginning by selecting the optimal Gabor filter parameters at various orientations and frequencies. This was done in order to improve upon [3], which required the parameters to be initialized. It is said that the accuracy can reach as high as 99.38% at its peak.

In [5], the authors investigated the susceptibility of palm vein recognition to spoof attacks within the context of the Print Attack category. During the course of that investigation, the VERA palm vein database was presented, and experimental results were written up and published for two different study regions. The highest EER that could be achieved was determined to be 3.33 percent.

The low-cost equipment designed in [6] to acquire vein images consists of a web camera and infrared LED lighting, and it was developed expressly for this purpose. The authors of the study claimed that it had an accuracy rate of 93.54%.

3. PROPOSED METHOD

During the step of feature extraction, the features of both the training and test ROI images are retrieved, and then they are compared during the stage of matching. A comparison is made between the vector of the test image and its equivalent vector in the training set with the use of the Manhattan distance measurement. According to the Manhattan distance *d*, which is defined as the distance between two points that is the measure of their dissimilarity, the features p_1, which are located at coordinates (x_1,y_1), and p_2, which are also located at coordinates (x_2,y_2), are not very comparable to one another. This can be seen when comparing the two sets of coordinates.

$$d=|x_1-x_2|+|y_1-y_2| \quad (1)$$

3.1 SCORE-LEVEL FUSION

The process of fusion takes place on a number of distinct layers inside multimodal biometric systems. Score-level fusion is one of the most used strategies for integrating distinct features, and it is also one of the most accessible methods. In this investigation, we compared the features of each of the five subregions to one another and then summed up their respective scores.

In order to do score fusion for several attributes, it is required to normalize the scores so that they are all based on the same scale. Only then can the scores be combined. Because all of the features came from the same region of interest (ROI) image and were extracted using the same method, normalization was not required for this experiment because all of the features were on the same scale.

The match scores from each of the subregions are combined into a single match score vector through a process called score-level fusion. After that, the classification stage makes use of this vector to assign classes to examples of test photos.

Applying a k-NN classifier to the problem of determining which of the system's training images is most comparable to the test picture results in the generation of Decision. The k-nearest neighbor algorithm selects the subjects in a dataset can be easily set to k classes, and the Manhattan distance can be used to compute the distance between fused scores. k-NN was also selected because of these two facts. This approach can swiftly handle large datasets due to the ease with which it may be computed, and the fact that additional data can be supplied without causing the operation to end.

3.2 CNN CLASSIFICATION

The AlexNet model was used as the foundation for the deep learning-based CNN architecture that was used to construct Decision II. The structure is made up of five convolution layers. Although it is comparable to AlexNet, it has less filters so that it can do computations more quickly.

In order to bring each convolution layer into action, a series of filters are applied to the input image in order to map out the values. We are able to construct the feature map by using the Eq.(3), where f is the picture that is being input and h is the filter. The resulting matrix has indexes for rows and columns that are denoted by the numbers *m* and *n*. The form of the feature map function *G* is as follows:

$$G(m,n)=(f*h)(m,n)=\sum_j\sum_k h(j,k)\times f[m-j,n-k] \quad (2)$$

The activation function is a Rectified Linear Unit (ReLU), which is defined as y=max(0,x). A ReLU is linear (identity) for all positive values and zero for all negative values. We perform batch normalization on the data, using Eq.(3) as a guide, in order to turn the data into a normal distribution.

$$y_i=\frac{x_i-\mu_B}{\sqrt{\sigma_B^2+\varepsilon}} \quad (3)$$

where

μ_B - mean of training set and

σ_B^2 - variance of training set.

After each convolution layer comes a max pooling layer. This layer generates a new output matrix in which each element is the maximum of a region in the original input. It does this by taking the maximum of a 2×22×2 region as its input and producing a new output matrix. In order to complete the design, we utilize a dropout layer with Softmax serving to normalize the input values to the interval [0,1].

$$y=\frac{\exp^{x_i}}{\sum_j \exp^{x_i}} \quad (4)$$

where the exponential of each value that is inputted (represented by x_i) is divided by the exponential sum of all values that are inputted (represented by x_j to x_K).

3.3 DECISION-LEVEL FUSION

A Weighted OR Rule is used to combine the previous two decisions, known as Decision I and Decision II, in order to arrive at the final decision for the technique that has been suggested. At the end of each call, True is returned if the recognition was successful, and False is displayed if it was not (False). Every verdict that is correct is worth one, whereas every verdict that is incorrect is worth nothing at all. The eventual decision is arrived at by adding up the relative weight of each option and contrasting the total with a predetermined threshold value, which in this instance is set at 0.90.

4. RESULTS AND DISCUSSION

The study gives the standard operating procedure for the experiment, which includes an explanation of the hyperparameters and hardware settings that will be utilized by default during the study. Before palmprint and palm vein ROI images can be input into a network, they need to be up-sampled to a large enough size to satisfy the needs of the network. The requirements of different networks can vary. We also included a random flip operation (just during training) to improve the network's robustness; this means that, for a training image, the image may be flipped horizontally with some probability before being fed into the network. This was done so that the network could better recognize patterns in the data.

To begin things rolling, rather than employing a random parameter initialization, we began by making use of the parameters that were provided by the pre-trained model in either the ImageNet or CIFAR datasets. We choose to make use of the pre-trained model even when an official model is available that can be trained on the ImageNet dataset. In that case, we will use a model that has already been trained using the CIFAR dataset. In most cases, the database image containing the palmprint and palm vein ROI will simply have a single channel. The grayscale channel has been copied three times due to the fact that the model requires an RGB image as input. This produces an RGB image with three channels as the final product.

When we speak of precision, what we are truly referring to is the correctness of our categorisation. The proportion of accurate predictions made in comparison to the total number of samples fed into the system is one way to evaluate the performance of a classification algorithm. It is easy to be fooled by the high level of accuracy that is stated for classifications. The issue arises due to the increased likelihood that specimens taken from a segment of the population may have their identities incorrectly assigned.

Table.1. Accuracy of Various Datasets during Training

Image	**FYO DB**	**CASIA DB**	**Tongji DB**	**PUT DB**	**VERA DB**
10	91.24	92.44	94.44	94.12	95.82
20	84.64	88.52	94.44	93.64	94.95
30	91.85	93.48	95.08	94.12	95.82
40	83.74	91.48	94.44	93.64	95.39
50	92.14	93.48	94.92	94.12	95.61
60	95.00	91.40	94.12	87.88	90.81
70	96.04	95.88	96.04	95.08	96.04

Table.2. Accuracy of Various Datasets during Training

Image	**FYO DB**	**CASIA DB**	**Tongji DB**	**PUT DB**	**VERA DB**
10	93.34	92.13	95.03	95.35	96.75
20	89.38	85.46	94.55	95.35	95.87
30	94.39	92.74	95.03	96.00	96.75
40	92.37	84.55	94.55	95.35	96.31
50	94.39	93.04	95.03	95.84	96.54
60	92.29	95.92	88.73	95.03	91.69
70	96.81	96.97	96.00	96.97	96.97

Table.3. Precision

Image	**FYO DB**	**CASIA DB**	**Tongji DB**	**PUT DB**	**VERA DB**
10	94.25	93.03	95.97	96.29	97.70
20	90.26	86.30	95.48	96.29	96.81
30	95.31	93.64	95.97	96.94	97.70
40	93.27	85.38	95.48	96.29	97.26
50	95.31	93.95	95.97	96.78	97.48
60	93.19	96.86	89.60	95.97	92.59
70	97.76	97.92	96.94	97.92	97.92

Table.4. Recall

Image	**FYO DB**	**CASIA DB**	**Tongji DB**	**PUT DB**	**VERA DB**
10	95.11	93.87	96.84	97.16	98.59
20	91.08	87.08	96.34	97.16	97.69
30	96.18	94.50	96.84	97.83	98.59
40	94.12	86.16	96.34	97.16	98.14
50	96.18	94.80	96.84	97.66	98.37
60	94.04	97.74	90.42	96.84	93.43
70	98.65	98.81	97.83	98.81	98.81

When it comes to the recognition of 2D palmprints and palm veins, the performance of NAS approaches is, on average, comparable to that of conventional methods. On certain datasets, the recognition performance of conventional techniques is significantly higher. The recognition efficiency of NAS methods is significantly higher than that of other databases.

5. CONCLUSION

In this study, a comprehensive analysis of the recognition performance of representative NAS techniques was conducted for the purposes of 2D and 3D palmprint recognition as well as palm vein recognition. In order to validate the evaluation criteria, we tested them with both isolated and combined data.

This is the first time that a complete performance evaluation of common NAS techniques for recognizing 2D and 3D palmprints and palm veins has ever been carried out in the work that is being discussed here. The findings of the experiments show that the NAS is an emerging technique for recognizing 2D and 3D palmprints as well as palm veins. In our upcoming study, we intend to make use of NAS technology in order to devise novel approaches to the recognition of 2D and 3D palmprints as well as palm veins. Because of this, we will be able to improve the degree to which they may be recognized.

REFERENCES

[1] Y.Y. Fanjiang, "Palm Vein Recognition based on Convolutional Neural Network", *Informatica*, Vol. 32, No. 4, pp. 687-708, 2021.

[2] S. Li and B. Zhang, "Joint Discriminative Sparse Coding for Robust Hand-Based Multimodal Recognition", *IEEE Transactions on Information Forensics and Security*, Vol. 16, pp. 3186-3198, 2021.

[3] R. Hernandez-Garcia and N. Guil, "Large-Scale Palm Vein Recognition on Synthetic Datasets", *Proceedings of International Conference of the Chilean Computer Science Society*, p. 1-8, 2021.

[4] R.A. Mustafa, "Palm Print Recognition based on Harmony Search Algorithm", *International Journal of Electrical and Computer Engineering*, Vol. 11, No. 5, pp. 1-12, 2021.

[5] H.N. Bendini and D.D.M. Valeriano, "Exploring a Deep Convolutional Neural Network and Geobia for Automatic Recognition of Brazilian Palm Swamps (Veredas) using Sentinel-2 Optical Data", *Proceedings of IEEE International Conference on Geoscience and Remote Sensing*, pp. 5401-5404, 2021.

[6] A.S. Al Jaberi and A.M. Al-juboori, "Palm Vein Recognition based on Convolution Neural Network", *Journal of Al-Qadisiyah for Computer Science and Mathematics*, Vol. 13, No. 3, pp. 1-6, 2021.

[7] A.S.M. Htet and H.J. Lee, "TripletGAN VeinNet: Palm Vein Recognition Based on Generative Adversarial Network and Triplet Loss", *Proceedings of International Conference on Computer Engineering and Artificial Intelligence*, pp. 454-458, 2021.

[8] S. Athisayamani, A. Robert Singh and A. Sivanesh Kumar, "Recurrent Neural Network-Based Character Recognition System for Tamil Palm Leaf Manuscript using Stroke Zoning", *Proceedings of International Conference on Inventive Communication and Computational Technologies*, pp. 165-176, 2021.

[9] X. Liang and D. Zhang, "CompNet: Competitive Neural Network for Palmprint Recognition Using Learnable Gabor Kernels", *IEEE Signal Processing Letters*, Vol. 28, pp. 1739-1743, 2021.

[10] K. Zhang and Tao, "Class Constraint-based Discriminative Features Learning Algorithm for Palm Print and Palm Vein Fusion Recognition", *Proceedings of International Conference on Signal and Image Processing*, pp. 275-280, 2022.

[11] G. Ananthi and S. Arivazhagan, "Human Palm Vein Authentication using Curvelet Multiresolution Features and Score Level Fusion", *The Visual Computer*, Vol. 38, No. 6, pp. 1901-1914, 2022.