Fingervein Verification using Convolutional Multi-Head Attention Network

Anonymous WACV 2024 APPLICATIONS TRACK submission

Paper ID 0358

Abstract

Biometric verification systems are deployed in various security-based access-control applications that require user-friendly and reliable person verification. Among the different biometric characteristics, fingervein biometrics have been extensively studied owing to their reliable verification performance. Furthermore, fingervein patterns reside inside the skin and are not visible outside; therefore, they possess inherent resistance to presentation attacks and degradation due to external factors. In this paper, we introduce a novel fingervein verification technique using a convolutional multihead attention network called VeinAtnNet. The proposed VeinAtnNet is designed to achieve light weight with a smaller number of learnable parameters while extracting discriminant information from both normal and enhanced fingervein images. The proposed VeinAtnNet was trained on the newly constructed fingervein dataset with 300 unique fingervein patterns that were captured in multiple sessions to obtain 92 samples per unique fingervein. Extensive experiments were performed on the newly collected dataset FV-300 and the publicly available FV-USM and FV-PolyU fingervein dataset. The performance of the proposed method was compared with five state-of-the-art fingervein verification systems, indicating the efficacy of the proposed VeinAtnNet.

1. Introduction

Biometric verification systems have enabled magnitude of access control applications including border control, smartphone access, banking, and finance applications. Fingervein biometric characteristics are widely deployed in various applications, particularly in banking sector. Fingervein biometrics represent the vein structure underneath the skin of the finger, which can be captured using near-infrared sensing. The blood flow in the fingervein absorbs near-infrared light and appears dark compared to the neighborhood region, indicating the visibility of the fingervein (refer Figure 1). The fingervein structure has been shown to

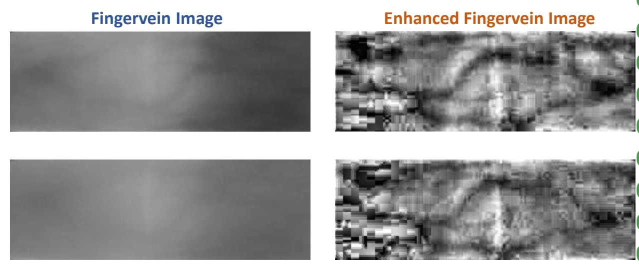


Figure 1: Example fingervein images with and without image enhancement for the same identity collected in first (top row) and second session (bottom row).

be unique [1, 34, 28] between fingers of same data subject and between the data subjects. Compared to other biometric characteristics, fingervein biometrics are known for their accuracy and usefulness, and are less vulnerable to distortion. Furthermore, fingervein biometrics provide a natural way of protecting biometric features, as they reside inside the skin and thus more challenging to spoof.

Fingervein biometrics have been widely studied in the literature, resulting in various fingervein biometric verification algorithms [33, 9]. Early works are based on extracting fingervein patterns such that the vein region is labeled as one and the background is labeled as zero. Techniques such as Maximum Curvature Points (MCP) [22], Repeated Line Tracking (RLT) [21], Wide Line Detectors (WLD) [10], Mean Curvature (MC) [37] and Randon transform [26] have been developed for reliable fingervein recognition. As these techniques can extract the structure of the fingervein pattern, the use of a simple comparator based on template matching using correction can achieve reliable performance. However, these features are sensitive to a small degree of fingervein rotation, noise, and reflection properties of the skin and NIR illuminator.

The global feature representation of fingervein patterns such as Local Binary Patterns (LBP) [15], Gabor filters [19], Local Directional Code [35], Wavelet Transform [23], Histogram of Gradients (HoG) [19] and pyramid image fea-

Table 1: State-of-the-art fingervein verification using deep learning techniques

Authors	Year	Deep Learning Technique
Huafeng Qin et al., [24]	2015	Serial CNN architecture with 3 convolution layers and 2 fully connected layer.
Itqan et al., [11]	2016	Serial CNN architecture with 3 convolution layers and 1 fully connected layer.
Syafeeza Radzi et al., [27]	2016	Serial CNN architecture with 2 convolution layers.
Huafeng Qin et al., [25]	2017	Serial CNN architecture with 4 convolution layers. Path based training of CNN.
Cihui Xie et al., [41]	2019	Siamese network with 5 conventional layers and triplet loss function.
Jong Min Song et al.,[36]	2019	Serial CNN architecture with 8 convolution layers. Composite fingervein image is generated by converting the 1-channel input image to 3-channel input image.
Rig Das et al., [5]	2018	Serial CNN architecture with 5 convolution layers.
Hyung Gil Hong et al.,[7]	2017	Serial CNN architecture with 12 convolution layers and 3 fully connected layers.
Su Tnag et al., [39]	2019	Siamese network with residual CNN architecture.
Borui Hou et al., [8]	2019	Convolutional autoencoder.
Junying Zeng et al., [43]	2020	Deformable convolution with U-NET type architecture.
Ridvan Salih Kuzu et al., [14]	2020	Serial CNN architecture with 6 convolution layers and 2 fully connected layers with LSTM for classification.
Hengyi Ren et al., [31]	2021	Feature extraction using ResNet with squeeze and excitation on the encrypted fingervein images.
Rıdvan Salih Kuzu et al., [13]	2021	Custom DenseNet 161 with additive angular penalty and large margin cosine penalty loss function.
Weili Yang et al., [42]	2022	Multi-view fingervein with individual CNNs and view pooling.
Huafeng Qin et al., [42]	2022	U-Net based architecture with attention module.
Tingting Chai et al., [4]	2022	Serial CNN architecture with 5 convolution layers and one fully connected layer.
Ismail et al., [3]	2022	Serial CNN architecture with 3 convolution layers and two fully connected layer.
Weiye Liu et al., [17]	2023	Residual Attention block with inception architecture.
Zhongxia Zhang et al., [44]	2023	Light weight CNN with spatial and channel attention module.
Chunxin Fang et al., [6]	2023	Light weight Siamese network with attention module.
Bin Wa et al., [20]	2023	Serial CNN architecture with 3 convolution layers and bilinear pooling with multiple attention module.
This work	**2024**	**Serial CNN architecture with 3 convolution layers and multi-head attention module connected in parallel with normal and enhanced fingervein.**

tures [18] are also developed for the fingervein verification. These features are often used with Support Vector Machines (SVM) or Euclidean distances as comparators. As these techniques are based on global features, they are highly sensitive to variations in finger rotation and illumination.

The representation of a fingervein image to binary codes was developed to improve template security, together with reliable verification. Binary coding techniques include Discriminative Binary Codes [16], binary hash codes [38], DoG code [28], ordinal code [28], contour Code [28] and competitive codes [28]. Because these techniques can generate binary codes for the finger vein, the Hamming distance is used as the comparator. Binary coding techniques exhibit good verification accuracy; however, these features are sensitive to variations in rotation and illumination.

Deep-learning-based fingervein verification has been extensively studied in the literature. Table 1 summarizes the deep-learning-based techniques proposed for fingervein recognition. Early works are based on the serial convolution architecture, which is inspired from existing CNN architectures that are evaluated on the ImageNet dataset. Both shallow serial CNN networks with two convolution layers and a deep CNN network with 12 convolutional layers have been studied in the literature. However, the quantitative results indicate that lightweight serial networks with a smaller number of convolution layers exhibit better performance than deep serial networks. The possible degraded performance of deep serial networks can be attributed to limited data availability. The use of a pre-trained CNN for feature extraction has also been explored in the literature, together with fine-tuning and augmented pre-trained CNN networks. The quantitative results reported indicate a performance similar to that of end-to-end trained deep CNN networks. The Siamese network for fingervein verification was studied using different CNN configurations and U-NET-based architectures. The quantitative performance is similar to that of the serial CNN architecture. Recently, attention modules with lightweight (three to four convolution layers) serial networks have been widely explored. Different types of attention modules, including spatial, channel, and multi-attention modules, were introduced. The quantitative performance of the attention networks are comparable to that of other deep learning-based techniques implemented for fingervein verification.

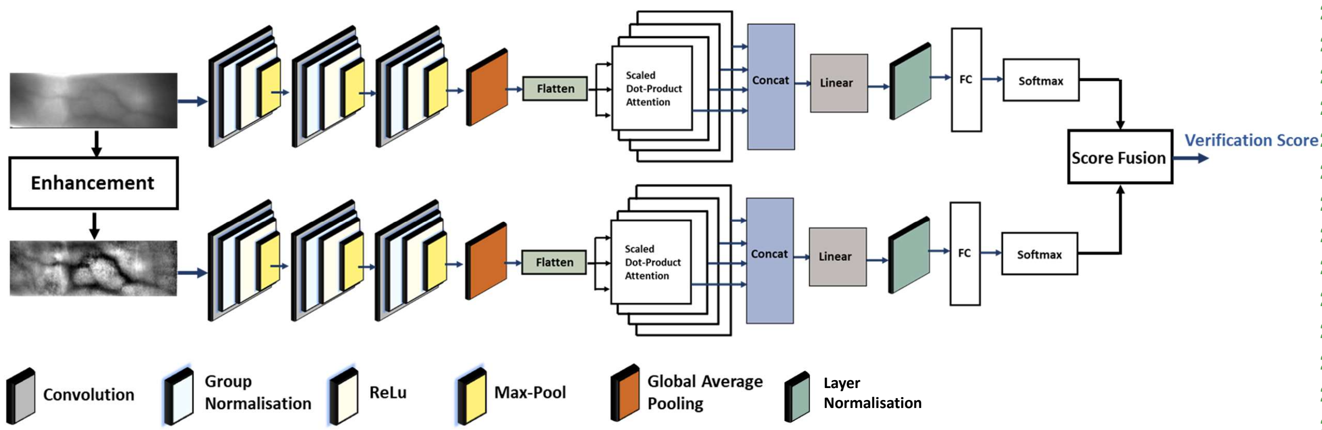


Figure 2: Block diagram of the proposed method for fingervein verification

Even though the deep learning techniques are widely studied for the reliable fingervein verification, the existing deep learning techniques indicates the following drawbacks: (a) limited data: Existing techniques are evaluated on the small-scale datasets that has 6-12 samples per data subject. This limits the effectiveness of deep learning and leads to an over fitting. (b) Lack of a consistent evaluation protocol: Even though most of the existing works have used public datasets, the evaluation protocols are not consistent across existing studies. This results in a limited comparison of existing techniques for finger vein verification. In this study, we address the above-mentioned limitation by introducing a new large-scale dataset with 75 data subjects, resulting in 300 unique identities (as we collected four fingers per data subject). For each unique fingervein, we collected 92 samples in multiple sessions, varying from 1-4 days duration. Furthermore, we propose a novel lightweight CNN architecture based on a convolutional multi-head attention module. The main contributions of this study are as follows:

- A novel fingervein verification technique based on a convolutional multi-head attention network (VeinAtnNet) is proposed.
- Introduced a new fingervein dataset with 300 unique identities captured from 75 data subjects, resulting in $300 \times 92 = 27600$ fingervein images. The dataset is available publicly for research purpose.
- Extensive experiments were performed on both the newly introduced dataset and the publicly available FV-USM and FV-PolyU datasets. The performance of the proposed method was compared with that of five state-of-the-art fingervein verification methods.

The rest of the paper is organised as follows: Section 2 discuss the proposed method for the fingervein verification, Section 3 presents the quantitative results of the proposed method with the state-of-the-art techniques and Section 4 draws the conclusion.

2. Proposed Method

Figure 2 shows a block diagram of the proposed VeinAtnNet architecture for reliable fingervein verification. The novelty of the proposed approach is that it leverages the convolutional Multi-Head Attention (MHA) framework to achieve accurate and reliable fingervein verification. The utility of MHA, together with convolutional features, leads to a discriminant feature representation that can contribute to the robust performance of the fingervein verification.

The proposed VeinAtnNet is a lightweight architecture with three Consecutive Convolution Layers (CCL) and a Multihead Self-Attention (MSA) mechanism. VeinAtnNet is connected independently with normal and enhanced fingervein images whose comparison scores from the softmax layer are fused to make the final verification decision. Given the captured fingervein image, preprocessing is performed using Contrast Limited Adaptive Histogram Equalization (CLAHE) [29] to enhance the fingervein pattern. In this work, we employed the Contrast Limited Adaptive Histogram (CLAHE) as the fingervein enhancement method by considering (a) the high quality of the fingervein enhancement achieved when compared to other enhancement techniques, as discussed in [9]. (b) Widely employed enhancement techniques in fingervein literature that have reported high verification accuracy. Both the normal (without enhancement) and enhanced fingervein images were resized to $224 \times 224 \times 3$ pixels. The CCL performs the initial feature learning of the fingervein images, which is further processed to obtain a rich feature representation using MSA. Given the fingervein image $F_v^{R \times C \times D}$, the final output features of MSA can be represented as follows:

$$F_{MSA} = MSA(F_{CCL}); where F_{CCL} = CCL(F_v); \quad (1)$$

Where, F_{MSA} denote the output features from MSA, F_{CCL} denotes the output features of CCL block. The F_{MSA} is then used with softmax classifier to make the final decision. In the following, we discuss the building blocks of the proposed VeinAtnNet.

2.1. Consecutive Convolution Layers (CCL)

The CCL has three convolution modules that are serially connected. Each convolution module has four different convolution layers (conv): a group normalization (norm), an activation function layer (ReLu), and a pooling layer (maxpool). Three convolution layers were used to extract the global features from the fingervein images. The conv-1 layer has a filter size of 7×7 the conv-2 layer has a 5×5, conv-3 has a 3×3 filter, and the number of filters in all three conv layers is set to 32. The gradual decrease in filter size ensures fine grinding (from global to local) of the fingervein features. The convolution features were normalized using group normalization, which reduced the sensitivity of the network for initialization. In particular, we employed group normalization because it outperforms batch normalization with a small size. The normalized features are then fed to the activation unit (ReLU), which can introduce sparsity and improve the network training speed. Finally, a pooling operation was performed to achieve a compact feature representation. In this study, we employed max pooling, which can capture texture information suitable for fingervein verification. The output after three convolution modules is then passed through the group-average pooling layer to obtain a compact representation of the features. Finally, the features were flattened before being fed into the MSA module.

2.2. Multihead Self-Attention (MSA)

The features from the CCL module are then fed to the MSA module to further refine the features F_{CCL} to extract discriminant features suitable for fingervein verification. In this study, we employed multihead attention [40] with four different heads and 64 channels for keys and queries. Basically, MSA runs the attention mechanism across all heads multiple times in parallel. The independent attention outputs are then concatenated and transformed linearly. MSA can be represented as follows [40]:

$$Mu - Head(Q.K, V) = [H_1, H_2, H_3, H_4]W \quad (2)$$

where W is the learnable parameter and Q, K and V represent the queries, keys, and values, respectively. In this study, we employed scaled dot-product attention across heads using Q, K and V as follows:

$$Attention(Q.K, V) = softmax(\frac{QK^T}{sqrt(d_k)})V \quad (3)$$

The outcome of the MSA module was passed through the layer normalization layer to generalize the final features. Finally, the normalized features are passed through the fully connected and softmax layers to obtain the comparison score.

2.3. Score Level Fusion

The proposed VeinAtnNet was employed independently on normal and enhanced finger vein images. Thus, given the test fingervein image, the proposed method provides two comparison scores corresponding to normal and enhanced fingerveins. We combined these two comparison scores using the sum rule to make the final verification decision. Let the comparison score from the normal fingervein image be C_n and enhanced fingervein image be C_e, then final verification score is computed as $V_s = (C_n + C_e)$.

2.4. Implementation Details

The proposed network is based on Adaptive Moment Estimation (ADAM) optimization to calculate loss. In this work, we employ the cross-entropy loss, which can be defined as $-\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K}(T_{ni}\log(Y_{ni}))+(1-T_{ni})\log(1-Y_{ni})$, where N and K denote the number of samples and classes, respectively, T_{ni} is the corresponding target value to Y_{ni}. During training, the learning rate was set to 0.0001, the mini-batch size was set to 16, and the number of epochs was set to 150. Furthermore, we performed data augmentation, which included image reflection, translation, rotation, reflection, scaling, and random noise with three different variances. This resulted in nine different images for every image used in training the proposed method. Finally, the proposed method is lightweight with only 58.2 K learnable parameters. While the existing SOTA employed in this work namely; Bin Wa et al., [20] has approximately 17.8M and Ismail et al., [3] has approximately 467.1K learnable parameters respectively.

3. Experiments and Results

In this section, we discuss the quantitative results of the proposed and existing fingervein verification algorithms. The quantitative performance is presented using the False Match Rate (FMR) and False Non-Match Rate (FNMR), together with the Equal Error Rate (EER) value computed at FMR = FNMR. The performance of the proposed method was compared with recently proposed fingervein recognition algorithms based on multiple attentions [20] and deep fusion [3] by considering their verification performance. Furthermore, we compared the performance of the proposed method with well-established fingervein verification techniques, such as MCP [22], RLT[21] and WLD [10]. In the following section, we describe the newly collected fingervein dataset, followed by the quantitative results.

3.1. FV-300 Fingervein dataset

In this study, we introduced a new fingervein dataset comprising 300 unique fingerveins corresponding to 75 unique data subjects. The fingervein images were collected using a custom camera system desired using a monochrome CMOS camera with a resolution of 744×480 pixels with two lighting sources to illuminate the finger from both the back and side. The design aspects of the finger vein capture device were inspired by [30]. The data collection was carried out under indoor conditions, and for every data subject, two fingers (index and middle) were captured from both the left and right hands, resulting in four unique fingers. For each data subject, we captured 92 fingervein images corresponding to individual fingers in multiple sessions. The duration between sessions varies from to 1-4 days. The FV-300 dataset contained 75 data subjects $\times$ 4 fingers $\times$ 92 = 27600 fingervein samples. Figure 3 shows an example of the fingervein images from the FV-300 dataset.

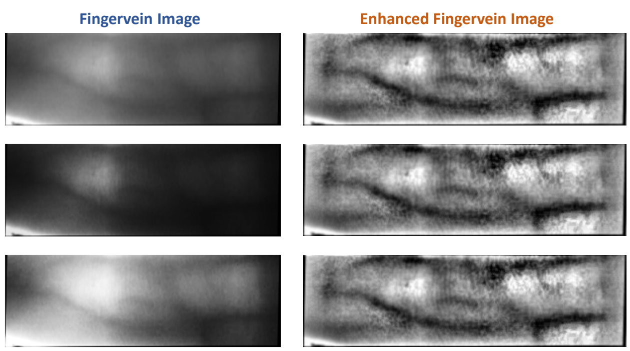


Figure 3: Example fingervein images from FV-300 dataset representing same identity captured in three different sessions.

3.2. Experimental protocol

To effectively benchmark the performance of the proposed method, we used three fingervein datasets: FV-300, FV-USM [2] and FV-PolyU [12]. To evaluate the performance of the fingervein algorithm on FV-300 dataset, the fingervein samples corresponding to each finger were divided into three independent sets such that the training set had 70 images, the validation set had 12 images, and the testing set had 10 images. This resulted in $300 \times 10 = 3000$ genuine and $300 \times 299 \times 10 = 897000$ impostor scores, respectively.

The verification performance of the FV-USM [2] dataset was evaluated by training the fingervein verification algorithms on FV-300 dataset and fine-tuning the trained networks on the FV-USM dataset. The FV-USM [2] dataset comprised 492 unique fingervein identities captured in two sessions with six samples each. Thus, the proposed method (and the existing methods employed in this work that includes multiple attention [20] and deep fusion [3]) are trained on the FV-300 dataset and fine-tuned using the first session data (from FV-USM dataset) that has 6 samples per subject. Testing was performed using the second-session data (from FV-USM dataset) with six samples per subject. However, the conventional fingervein state-of-the-art techniques (MCP [22], RLT [21] and WLD [10]) employed in this study do not require a training set for learning. Therefore, we used the first-session data from the FV-USM dataset as enrolment, and the second session data were used for testing. This resulted in $492 \times 6 = 2952$ genuine and $492 \times 491 \times 6 = 1449432$ impostor scores.

The verification performance of the fingervein algorithms (deep learning based on the proposed method) on the FV-PolyU dataset was performed using a procedure similar to that discussed for the FV-USM dataset. The fingervein algorithms trained on the FV-300 dataset were fine-tuned using the FV-PolyU dataset. The FV-PolyU dataset [12] employed in this work comprises 156 unique identities, from which the finger vein index and middle fingers are captured in two sessions with six samples each. Thus, the FV-PolyU dataset has 312 unique identities, and data from the first session are used to fine-tune both the proposed and SOTA deep learning methods, which include multiple attention [20] and deep fusion [3]) that are trained on the FV-300 dataset. Testing was performed on the second session data, which resulted in $312 \times 6 = 1872$ genuine and $312 \times 311 \times 6 = 582192$ impostor scores.

3.3. Results and discussion

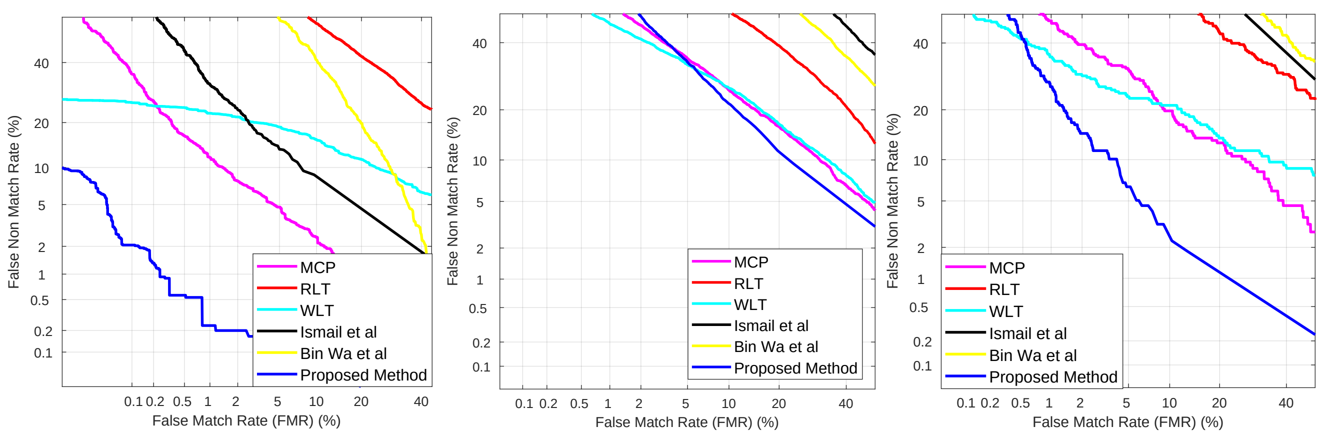
Table 2 shows the quantitative performance of the proposed and existing fingervein verification techniques on both FV-300, FV-PolyU and FV-USM datasets, and Figure 4 shows the DET curves. Existing methods were trained using enhanced fingervein images to optimise the best performance. Based on the results, the following are important observations:

- Training and testing on the same dataset will indicate the improved verification results of the deep learning based techniques. Therefore, the performance of the deep learning techniques indicated an improved performance on FV-300 compared to FV-USM and FV-PolyU dataset.

- Traditional fingervein techniques (MCP [22], RLT [21] and WLD [10]) that are based on template matching using correlation indicates the superior performance on the FV-300 dataset compared to FV-USM and FV-PolyU dataset. However, the performance of RLT [21] and WLD [10] do not indicate a significant difference

Table 2: Quantitative Performance of the proposed and state-of-the-art fingervein verification methods

Data set	Algorithms	EER(%)	TAR = (100-FNMR%) @ FMR =		
			1%	**0.1%**	**0.01%**
FV300	MCP [22]	4.74	88.93	64.41	44.35
	RLT [21]	31.30	19.35	9.52	4.68
	WLD [10]	13.55	77.47	74.11	73.15
	Ismail et al., [3]	9.14	72.92	40.95	7.34
	Bin Wa et al., [20]	19.94	21.25	5.31	0.95
	Proposed Method	**0.54**	**99.73**	**99.13**	**90.36**
FV-USM	MCP [22]	17.74	46.95	27.25	18.36
	RLT [21]	29.63	32.29	29.70	29.16
	WLD [10]	18.17	54.17	34.89	16.73
	Ismail et al., [3]	41.78	4.38	1.16	0.34
	Bin Wa et al., [20]	37.53	5.25	1.65	0.35
	Proposed Method	**15.35**	**40.45**	**14.25**	**5.11**
PolyU	MCP [22]	14.25	52.29	34.40	20.64
	RLT [21]	33.48	19.26	8.25	4.28
	WLD [10]	16.53	65.29	48.16	38.53
	Ismail et al., [3]	42.12	5.96	2.29	1.37
	Bin Wa et al., [20]	40.90	3.21	0.45	0.45
	Proposed Method	**5.52**	**74.77**	**30.27**	**19.26**



(a) FV-300 Dataset (b) FV-USM Dataset (c) FV-PolyU Dataset

Figure 4: DET Curves showing the verification performance of the proposed and state-of-the-art fingervein verification methods

in the verification performance between three different fingervein datasets employed in this work.

- Among three traditional fingervein techniques employed in this work, the MCP [22] indicated the best performance on both datasets. Furthermore, MCP [22] demonstrated improved performance compared to the state-of-the-art deep learning methods employed in this study.

- The proposed method has indicated an outstanding verification performance with EER = 0.54% and TAR = 90.36% @ FMR = 0.01% on FV-300 dataset. The proposed method also indicated the best performance with an EER of 15.35% on the FV-USM dataset. Similar performance is also noted on the FV-PolyU dataset with an EER = 5.52% . However, the verification performance degraded at a lower FMR on both FV-USM and FV-PolyU dataset.

- Based on the results, it is worth nothing that, the deep learning techniques depends on the training data and indicate the limitation to generalize on the another dataset due to the limited number of samples available for fine-tuning. However, compared with existing deep learning methods, the proposed VeinAtnNet exhibits

Table 3: Verification performance of the proposed method with normal and enhanced fingervein data

Data Type	Algorithms	EER(%)	TAR = (100-FNMR%) @ FMR =		
			1%	0.1%	0.01%
Normal Fingervein	Proposed Method	1.85	97.89	83.55	54.18
Enhanced Fingervein	Proposed Method	1.13	98.87	90.53	60.76

Table 4: Ablation study of the proposed method on FV-300 dataset

Consecutive Convolution Layers (CCL)			Multi-head	Proposed method
Conv-1	Conv-2	Conv-3	Self-Attention (MSA)	EER (%)
✓	X	X	✓	8.29
✓	✓	X	✓	2.38
✓	✓	✓	✓	0.54

superior verification performance on three fingervein datasets employed in this work.

3.4. Ablation Study of the proposed method

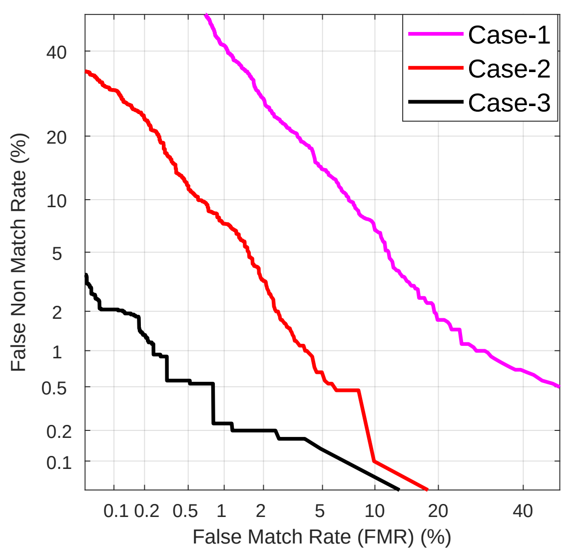


Figure 5: DET Curves indicating the performance of the proposed method with different cases of ablation study

In this section, we present an ablation study of the proposed method by using the FV-300 dataset. We considered three different cases in which Case-1 represent the performance with Conv-1 and MSA together. Case-2 shows the performance of Conv-1, Conv-2, and MSA while Case-3 indicates the performance of the proposed method with Conv-1, Conv-2, Conv-3, and MSA. Table 4 and Figure 5 show the performance of the proposed method for different ablation studies. The addition of convolutional layers with MSA can improve the overall performance of the proposed VeinAtnNet for reliable fingervein verification.

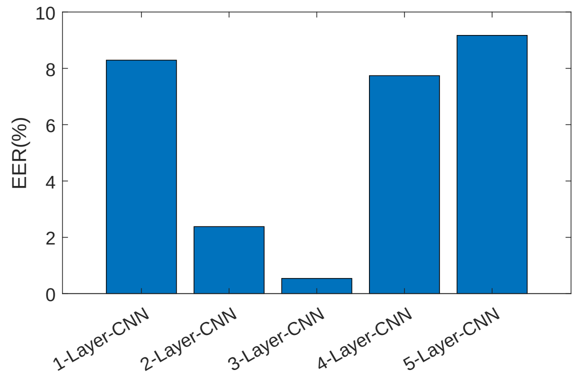


Figure 6: EER of the proposed method with different number of convolution layers.

We further investigated the role of adding additional convolution layers with MSA to improve the verification accuracy. To this extent, we start computing the verification accuracy starting with one Conv layer and increasing it to five consecutive Conv layers with MSA. Figure 6 shows the verification performance with EER for different depths of convolution layers. It should be noted that the use of three consecutive layers with MSA can achieve the best performance and further increase the depth by adding convolution layers. This further justifies the choices made in designing the proposed method that has indicated the best generalized verification performance compared to the five different SOTA.

3.5. Interpretation of the proposed method

To interpret the decision achieved by the proposed method, we employed Local interpretable model-agnostic explanations (LIME) [32] to explain the perdition's on the probe fingervein images. Because the proposed method is based on both normal and enhanced fingervein images, we present the qualitative and quantitative results for both fingervein image types. Table 3 indicates the quantitative performance of the proposed method with normal and enhanced image alone. The obtained results indicated a similar verification performance with EER and higher FMR values. However, with lower FMR values, the proposed method exhibited better performance with enhanced fingervein samples. Thus, the availability of the enhanced fingervein pattern indicates more discriminant information to

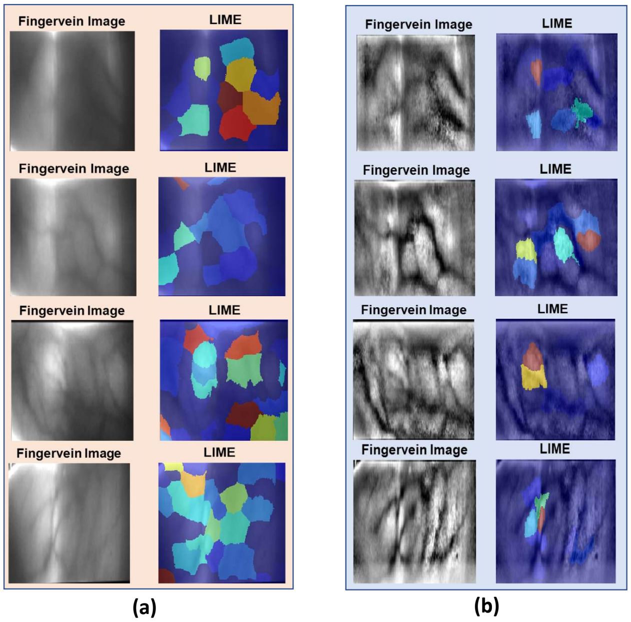


Figure 7: Illustration of LIME based explainability on the proposed method based on the (a) normal and (b) enhanced fingervein images.

improve verification accuracy at low FMR values.

Figure 7 shows the qualitative results of the LIME method for visualizing important regions in the fingervein image, which has contributed to successful verification. The LIME explanation is shown on the fingervein images from the FV-300 dataset for successful verification prediction at FAR = 0.01%. As shown in Figure 7, the proposed method utilises more image regions with normal fingervein images compared with the enhanced fingervein to make the decision. However, with enhanced fingervein images, the decision is based on a smaller number of regions associated with vein pattern and particularly on the minutiae points of the fingervein. These observations justify the improved performance of the proposed method with enhanced fingervein images compared with normal fingervein images.

4. Conclusion

Fingervein biometrics are widely employed in various secure access control applications. In this study, we proposed a novel method based on a convolutional multi-head attention module for reliable fingervein verification. The proposed VeinAtnNet is based on three consecutive convolution layers and multihead attention with four heads and 64 channels connected in parallel to the normal and enhanced fingervein samples. Finally, the decision is made using the score-level fusion of the normal and enhanced fingerveins. Extensive experiments were performed on both publicly and newly collected finger vein datasets. The quantitative performance of the proposed method was benchmarked using five state-of-the-art fingervein verification methods. The obtained results indicate the superior performance of the proposed method on both publicly available and newly collected fingervein datasets.

References

[1] Finger vein authentication. In *White Paper*, 2006. 1

[2] Mohd Shahrimie Mohd Asaari, Shahrel A Suandi, and Bakhtiar Affendi Rosdi. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Systems with Applications*, 41(7):3367–3382, 2014. 5

[3] Ismail Boucherit, Mohamed Ould Zmirli, Hamza Hentabli, and Bakhtiar Affendi Rosdi. Finger vein identification using deeply-fused convolutional neural network. *Journal of King Saud University-Computer and Information Sciences*, 34(3):646–656, 2022. 2, 4, 5, 6

[4] Tingting Chai, Jiahui Li, Shitala Prasad, Qi Lu, and Zhaoxin Zhang. Shape-driven lightweight cnn for finger-vein biometrics. *Journal of Information Security and Applications*, 67:103211, 2022. 2

[5] Rig Das, Emanuela Piciucco, Emanuele Maiorana, and Patrizio Campisi. Convolutional neural network for finger-vein-based biometric identification. *IEEE Transactions on Information Forensics and Security*, 14(2):360–373, 2018. 2

[6] Chunxin Fang, Hui Ma, and Jianian Li. A finger vein authentication method based on the lightweight siamese network with the self-attention mechanism. *Infrared Physics & Technology*, 128:104483, 2023. 2

[7] Hyung Gil Hong, Min Beom Lee, and Kang Ryoung Park. Convolutional neural network-based finger-vein recognition using nir image sensors. *Sensors*, 17(6):1297, 2017. 2

[8] Borui Hou and Ruqiang Yan. Convolutional autoencoder model for finger-vein verification. *IEEE Transactions on Instrumentation and Measurement*, 69(5):2067–2074, 2019. 2

[9] Borui Hou, Huijie Zhang, and Ruqiang Yan. Finger-vein biometric recognition: A review. *IEEE Transactions on Instrumentation and Measurement*, 71:1–26, 2022. 1, 3

[10] Beining Huang, Yanggang Dai, Rongfeng Li, Darun Tang, and Wenxin Li. Finger-vein authentication based on wide line detector and pattern normalization. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1269–1272, Aug 2010. 1, 4, 5, 6

[11] KS Itqan, AR Syafeeza, FG Gong, N Mustafa, YC Wong, and MM Ibrahim. User identification system based on finger-vein patterns using convolutional neural network. *ARPN Journal of Engineering and Applied Sciences*, 11(5):3316–3319, 2016. 2

[12] Ajay Kumar and Yingbo Zhou. Human identification using finger images. *IEEE Transactions on Image Processing*, 21(4):2228–2244, 2012. 5

[13] Rιdvan Salih Kuzu, Emanuele Maiorana, and Patrizio Campisi. Loss functions for cnn-based biometric vein recognition. In *2020 28th European Signal Processing Conference (EUSIPCO)*, pages 750–754. IEEE, 2021. 2

[14] Ridvan Salih Kuzu, Emanuela Piciucco, Emanuele Maiorana, and Patrizio Campisi. On-the-fly finger-vein-based biometric recognition using deep neural networks. *IEEE Transactions on information Forensics and Security*, 15:2641–2654, 2020. 2

[15] Eui Chul Lee, Hyunwoo Jung, and Daeyeoul Kim. New finger biometric method using near infrared imaging. *Sensors*, 11(3):2319–2333, 2011. 1

[16] Haiying Liu, Lu Yang, Gongping Yang, and Yilong Yin. Discriminative binary descriptor for finger vein recognition. *IEEE Access*, 6:5795–5804, 2017. 2

[17] Weiye Liu, Huimin Lu, Yifan Wang, Yupeng Li, Zhenshen Qu, and Yang Li. Mmran: A novel model for finger vein recognition based on a residual attention mechanism: Mmran: A novel finger vein recognition model. *Applied Intelligence*, 53(3):3273–3290, 2023. 2

[18] Yu Lu, Sook Yoon, Shiqian Wu, and Dong Sun Park. Pyramid histogram of double competitive pattern for finger vein recognition. *IEEE Access*, 6:56445–56456, 2018. 2

[19] Y. Lu, S. Yoon, S. J. Xie, J. Yang, Z. Wang, and D. S. Park. Finger vein recognition using histogram of competitive gabor responses. In *22nd International Conference on Pattern Recognition*, pages 1758–1763, Aug 2014. 1

[20] Bin Ma, Kaixuan Wang, and Yueli Hu. Finger vein recognition based on bilinear fusion of multiscale features. *Scientific Reports*, 13(1):249, 2023. 2, 4, 5, 6

[21] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications*, 15(4):194–203, 2004. 1, 4, 5, 6

[22] Naoto Miura, Akio Nagasaka, and Takafumi Miyatake. Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE TRANSACTIONS on Information and Systems*, 90(8):1185–1194, 2007. 1, 4, 5, 6

[23] Kang Ryoung Park. Finger vein recognition by combining global and local features based on svm. *Computing and Informatics*, 30(2):295–309, 2011. 1

[24] Huafeng Qin and Mounîm A El-Yacoubi. Finger-vein quality assessment by representation learning from binary images. In *Neural Information Processing: 22nd International Conference, ICONIP 2015, Istanbul, Turkey, November 9-12, 2015, Proceedings, Part I 22*, pages 421–431. Springer, 2015. 2

[25] Huafeng Qin and Mounim A El-Yacoubi. Deep representation-based feature extraction and recovering for finger-vein verification. *IEEE Transactions on Information Forensics and Security*, 12(8):1816–1829, 2017. 2

[26] Huafeng Qin, Xiping He, Xingyan Yao, and Hongbing Li. Finger-vein verification based on the curvature in radon space. *Expert Systems with Applications*, 82:151–161, 2017. 1

[27] Syafeeza Ahmad Radzi, Mohamed Khalil Hani, and Rabia Bakhteri. Finger-vein biometric identification using convolutional neural network. *Turkish Journal of Electrical Engineering and Computer Sciences*, 24(3):1863–1878, 2016. 2

[28] R. Raghavendra and C. Busch. Exploring dorsal finger vein pattern for robust person recognition. In *International Conference on Biometrics (ICB)*, pages 341–348, 2015. 1, 2

[29] R. Raghavendra, Kiran B raja, Jayachander Surbiryala, and Christoph Busch. A low-cost multimodal biometric sensor to capture finger vein and fingerprint. In *International Joint Conference on Biometrics (IJCB)*, pages 1–7, Sep 2014. 3

[30] Raghavendra Ramachandra, Kiran Bylappa Raja, Sushma Krupa Venkatesh, and Christoph Busch. Design and development of low-cost sensor to capture ventral and dorsal finger vein for biometric authentication. *IEEE Sensors Journal*, 19(15):6102–6111, 2019. 5

[31] Hengyi Ren, Lijuan Sun, Jian Guo, Chong Han, and Fan Wu. Finger vein recognition system with template protection based on convolutional neural network. *Knowledge-Based Systems*, 227:107159, 2021. 2

[32] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ” why should i trust you?” explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016. 7

[33] Kashif Shaheed, Hangang Liu, Gongping Yang, Imran Qureshi, Jie Gou, and Yilong Yin. A systematic review of finger vein recognition techniques. *Information*, 9(9), 2018. 1

[34] Mohamed Shahin, Ahmed Badawi, and Mohamed Kamel. Biometric authentication using fast correlation of near infrared hand vein patterns. *International Journal of Biological and Medical Sciences*, 2(3):141–148, 2007. 1

[35] Pragati Sikarwar et al. Finger vein recognition using local directional pattern. In *2016 International Conference on Inventive Computation Technologies (ICICT)*, volume 1, pages 1–5. IEEE, 2016. 1

[36] Jong Min Song, Wan Kim, and Kang Ryoung Park. Finger-vein recognition based on deep densenet using composite image. *Ieee Access*, 7:66845–66863, 2019. 2

[37] Wonseok Song, Taejeong Kim, Hee Chan Kim, Joon Hwan Choi, Hyoun-Joong Kong, and Seung-Rae Lee. A finger-vein verification system using mean curvature. *Pattern Recognition Letters*, 32(11):1541–1547, 2011. 1

[38] Kun Su, Gongping Yang, Lu Yang, Dunfeng Li, Peng Su, and Yilong Yin. Learning binary hash codes for finger vein image retrieval. *Pattern Recognition Letters*, 117:74–82, 2019. 2

[39] Su Tang, Shan Zhou, Wenxiong Kang, Qiuxia Wu, and Feiqi Deng. Finger vein verification using a siamese cnn. *IET biometrics*, 8(5):306–315, 2019. 2

[40] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 4

[41] Cihui Xie and Ajay Kumar. Finger vein identification using convolutional neural network and supervised discrete hashing. *Pattern Recognition Letters*, 119:148–156, 2019. 2

[42] Weili Yang, Junduan Huang, Zhuoming Chen, Junhong Zhao, and Wenxiong Kang. Multi-view finger vein recognition using attention-based mvcnn. In *Biometric Recognition: 16th Chinese Conference, CCBR 2022, Beijing,*

China, November 11–13, 2022, Proceedings, pages 82–91. Springer, 2022. 2

[43] Junying Zeng, Fan Wang, Jianxiang Deng, Chuanbo Qin, Yikui Zhai, Junying Gan, and Vincenzo Piuri. Finger vein verification algorithm based on fully convolutional neural network and conditional random field. *IEEE access*, 8:65402–65419, 2020. 2

[44] Zhongxia Zhang and Mingwen Wang. Finger vein recognition based on lightweight convolutional attention model. *IET Image Processing*, 2023. 2