Palm-vein verification based on U-Net

Peng Wang\textsuperscript{*} and Huafeng Qin\textsuperscript{2}

\textsuperscript{1}National Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing, 400067, China
\textsuperscript{2}Chongqing Engineering Laboratory of Detection Control and Integrated System, Chongqing Technology and Business University, Chongqing, 400067, China
\textsuperscript{*}Corresponding author’s e-mail: 2017658007@email.ctbu.edu.cn

Abstract: As one of the biometric characteristics, palm-vein have received more and more attention in recent years. However, in practical applications, the vein image capturing is affected by various factors, so many low quality images are produced in recognition system. Generally, the vein networks extracted from such low quality images contains many noises, which degrades the recognition accuracy. To address these problems, this paper proposes an end-to-end convolutional neural network to extract vein feature for verification. Firstly, we label the palm-vein pixel by the combination of some handcraft-based palm-vein image segmentation methods and build a training dataset. Secondly, a U-Net network is trained based on the resulting dataset and its outputs are the probability of pixels to belong to vein pattern. Thirdly, we propose a scheme to encode the outputs of U-Net to obtain the vein network patterns. The experiment results on the public CASIA palm-vein dataset implies the effectiveness of our proposed method.

1. Introduction
With the development of the economy and wide applications of internet technology, the demand for information and property security is increasing\cite{1}. As one of personal identification, biometrics characteristics have developed rapidly in past years. Biometric recognition technologies are divided into two categories: (1) Extrinsic characteristics: finger print\cite{2}, iris\cite{3}, gait\cite{4}, face\cite{5}. (2) Intrinsic characteristics: finger-vein\cite{6}, palm-vein\cite{7}. Unlike the extrinsic biometric features, vein verification technology has many advantages such as intrinsic features, living identification, uniqueness, not easy to replicate, etc. In addition, compared with the finger-vein, the palm vein has more complex lines and structures, which are beneficial to improve the verification rate of palm-vein. Based on these excellent characteristics, the palm-vein verification technology has received more and more attention from researchers and industry in recent years.

2. Related Works
Despite recent advances in palm-vein recognition, palm-vein still faces many challenges in real life. In practical application, imaging process is affected by many factors such as temperature, equipment, user habits, illumination, etc. so the palm-vein images contains ambiguous regions where the separability between vein and background is poor. In general, matching such regions degrades the recognition accuracy. To solve these problems, many palm-vein segmentation methods are proposed to extract vein network. In general, they can be divided into two categories as follows. Handcraft-based segmentation approaches. These methods mainly build mathematical models to extract vein based on the assumption that the vein pattern shows in valleys or straight-line, etc. Han\cite{7} proposed an innovative and robust...
adaptive Gabor filtering method, which encodes palm vein features into bit string representations. Based on the Scale-invariant feature transform (SIFT) method, Kang[8] proposed the ROOTSIFT method, a more stable local invariant feature extraction. Zhang[9] extracted the palm vein by multi-scale filtering. Deep learning-based segmentation methods. Convolutional neural network have outperformed the state of the art in computer vision field[10-12]. Different with handcraft-based segmentation methods, deep learning segmentation methods is an end-to-end architecture without the manual attribute distribution assumption. Some researchers have made some explorations. For example, Qin[13] proposed an iterative deep neural network for hand-vein verification. These existing methods do not always perform well in practical application. For example, handcraft-based segmentation methods are based on the distribution assumption. However, many vein patterns show more complex shape instead of valley or straight-line, so these assumption may not be always effective. Based on the good performance of U-Net[14] network in the field of image segmentation, we employ it for palm vein segmentation.

3. The Proposed Method
In this paper, we design a convolutional neural network structure to extract palm vein features. Firstly, we assign the vein pixel label by some handcrafted segmentation methods in the training datasets. Secondly, we build a U-Net and train it. Finally, we match the vein features for verification. The overview of the proposed method is shown in Figure 1.

![Figure 1. The overview of the proposed system](image)

3.1. Label vein pixel
In this Section, we will use four baselines e.g. repeated line tracking[15], mean curvature[16], maximum principle curvature[17], and Gabor filters[18] to segment vein pattern for each vein images. For each image, four binary vein images is extracted based on four approaches. Finally, we average four binary images to obtain a map which is subject to binarization by a threshold of 0.5 to compute a labeling image (vein pixel and background are labeled as 1 and 0). The label approach is detail in [13].

3.2. U-Net model
Unlike general convolutional neural network, this network consists of encoder layers and decoder layers. Encoder layers gradually reduce the spatial dimension of the layers, and decoders gradually repair the details and spatial dimension of objects. There is usually a fast connection between the encoder layers and the decoder layers, so it can help the decoder layers to repair the details of the target better. In the final layer, the outputs are the probability of each pixel which belongs to palm vein or the background. The network architecture is illustrated in Figure 2.
As shown in the Figure 2, the U-Net architecture consists of a contracting path and an expansive path. In the contracting path, down-sampling is a classical convolutional neural network, which consists of five layers. In each layer, there are two 3×3 convolution layers, followed by a rectified linear unit (ReLU) and a 2×2 max pooling layer with stride 2. And the number of convolution kernels in the next layer is twice times than the number of convolution kernels in the previous layer, and finally the number of convolution kernels in the fifth layer is 1024. In the expansive path, each layer have two 2×2 up-convolution layers, followed by two 3×3 convolution layers and a rectified linear unit (ReLU). With the number of layers increasing, we halve the number of convolution kernels. After training, the U-Net outputs a map in which the values of each pixel is its probability being to vein pattern. Then, the map is further subject to binarization for matching.

3.3. Supervised Feature Encoding

In some approach, the output map of U-Net is encoded by a threshold of 0.5 for matching. To achieve robust matching, we encode the feature map to a threshold which is related to matching scores. In our encoding scheme, a threshold is obtained by maximizing the distance between intra-class and inter-class, and employed to encode feature map. We assume that there are m samples, and each sample have n classes. Then there are m × n images. We generate intra-class scores by matching binary images from the same classes using Hamming distance in [19], and generate inter-class scores by matching binary images from the different classes. Finally, we obtain $a_1 = m \times c_n^2$ genuine matching scores $U_1 = \{d_1(T), d_2(T), ..., d_{a_1}\}$ and $a_2 = n \times m \times c_m^2$ imposter matching scores $U_2 = \{d'_1(T), d'_2(T), ..., d'_{a_2}\}$. To get a robust threshold, we formulate the following optimization objective function:

$$max J(T) = \frac{|u_1(T) - u_2(T)|}{D_1(T) + D_2(T)} \tag{1}$$

Where $T$ denote the threshold, $u_1(T)$ and $u_2(T)$ are the means of the scores in the sets $U_1$ and $U_2$, and $D_1(T)$ and $D_2(T)$ are the variance of the scores in the sets $U_1$ and $U_2$. The parameter T is assigned from 0 to 255, so 256 different values $J(T)$ are computed using Eq. (1).
3.4. Enhanced hamming distance
Hamming distance [19] is used to calculate the difference between two vectors. The vein network are stored in binary image, so we employ Hamming distance to compute the similarity between two palm-vein images. Because there are translations among palm-vein images, we propose an enhanced hamming distance to match them. We assume $E$ and $F$ are two binary feature images with size of $w \times h$. The length and width of image $E$ are extended to $w + 2i$, $h + 2j$. Its expanded images $\bar{E}$ are expressed as:

$$
\bar{E}(w, h) = \begin{cases} 
E(w-i, h-j) & \text{if } 1 + i \leq w \leq g + i, \\
-1 & \text{otherwise} \end{cases} 
$$

The parameters $w$ and $h$ that control the translation distance in horizontal and vertical directions are set to 100 and 100, and the matching distance between $E$ and $F$ is obtained by

$$
d(E, F) = \min_{0 \leq m \leq 2l, 0 \leq n \leq 2j} \frac{\sum_{w=1}^{x} \sum_{h=1}^{y} \text{hamdis}(\bar{E}(w+m, h+n), F(w, h))}{\sum_{w=1}^{x} \sum_{h=1}^{y} \theta(\bar{E}(w+m, h+n), -1)}
$$

hamdis denotes the hamming distance between two binary images and

$$
\theta(U, V) = \begin{cases} 
1 & \text{if } U \neq V \\
0 & \text{otherwise} \end{cases}
$$

4. Experiment and results

4.1. CASIA Datasets
There are 1200 images in the CASIA Multi-Spectral Palm-print Image Database. In our experiments, a pre-processing method[13] is employed to extract the region of interest (ROI) image. we split the database into two datasets. There are 100 hands associated with 600 images in the training dataset, and 100 hands associated with 600 images in the test dataset.

4.2. Verification results
In this experiment, we visually assess the performance of various approach. We employ the proposed to extract the vein network and segmentation results is shown in Figure 3. In addition, the vein patterns extracted by four baselines are illustrated in Figure 3 for comparison. From Figure 3, we can see that the proposed model can extract more continuous and smoothness vein patterns compared to four handcrafted approaches. This can be explained by the following fact. the proposed model can directly learn rich prior knowledge acquired by training the U-Nets on a huge train set from different images without depending any assumptions.
In this section, the experiments are carried out to estimate the verification performance of our approach. Based on Eq (1), we encode the output of U-Nets and obtain the binary images. In test dataset, there are 600 (100 x 6) genuine scores are generated by matching samples from same hand and (100x99x6/2) impostor scores are generated by matching samples from different hands. The FAR and FRR are computed using genuine sores and imposter scores. The EER is the error rate when FAR is equal to FRR. The receiver operating characteristics of various approaches is depicted in Figure.4 and the corresponding EER is listed in Table 1. The experimental results from Figure.4 and Table 1 implies that the proposed approach outperform handcrafted approaches and achieve lowest EER, e.g. 0.47%.

| Methods                     | EER(%) |
|-----------------------------|--------|
| Gabor filters               | 0.93   |
| Maximum principle curvature | 1.85   |
| Mean curvature              | 0.93   |
| Repeated line tracking      | 3.59   |
| Proposed method             | 0.47   |
5. Conclusion
In this paper, we present a U-Net based palm segmentation approach to extract palm vein pattern. Firstly, we label the vein pixel by the combination of several handcraft-based segmentation methods. Secondly, the U-Net model is employed to predict the probability of pixel to belong to vein pattern. Thirdly, we propose a robust scheme to encode the vein patterns. Finally, experimental results on the public dataset demonstrate the proposed approach’s effectiveness.

References
[1] Su J.F, Yang Y, Yang. (2018) Measuring knowledge diffusion efficiency in R&D networks. Knowledge management research and practice. 16(2):208-219.
[2] Kasban H. (2016) Fingerprints verification based on their spectrum J. Neuro computing, 171: 910–920.
[3] Wildes, R. P. (1997). Iris recognition: an emerging biometric technology. Proceedings of the IEEE, 85(9): 1348-1363.
[4] Isaac E R H P, Elias S, Rajagopalan S, et al. (2019) Gait Verification System Through Multiperson Signature Matching for Unobtrusive Biometric Authentication J. Journal of Signal Processing Systems, 91(2): 147-161.
[5] Wang F, Cheng J, Liu W, et al. (2018) Additive margin softmax for face verification J. IEEE Signal Processing Letters, 25(7): 926-930
[6] Qin H, El-Yacoubi M A. (2017) Deep representation-based feature extraction and recovering for finger-vein verification J. IEEE Transactions on Information Forensics and Security, 12(8): 1816-1829.
[7] Han W Y, Lee J C. (2012) Palm vein recognition using adaptive Gabor filter J. Expert Systems with Applications, 39(18): 13225-13234.
[8] Kang W, Liu Y, Wu Q, et al. (2014) Contact-free palm-vein recognition based on local invariant features J. PloS one , 9(5): e97548.
[9] Zhang Y B, Li Q, You J, et al. (2007) Palm vein extraction and matching for personal authentication. In: International Conference on Advances in Visual Information Systems, pp.154-164.
[10] Krizhevsky A, Sutskever I, Hinton G E. (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097-1105.
[11] Oquab M, Bottou L, Laptev I, et al. (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 1717-1724.
[12] Gatys L A, Ecker A S, Bethge M. (2016) Image style transfer using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 2414-2423.
[13] Qin H, El Yacoubi M A, Lin J, et al. (2019). An iterative deep neural network for hand-vein verification J. IEEE Access, 7: 34823-34837.
[14] Ronneberger O, Fischer P, Brox T. (2015) U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp.234-241.
[15] Miura N, Nagasaka A, Miyatake T. (2004) Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification J. Machine vision and applications, 15(4): 194-203
[16] Song W, Kim T, Kim H C, et al. (2011) A finger-vein verification system using mean curvature J. Pattern Recognition Letters, 32(11): 1541-1547
[17] Miura N, Nagasaka A, Miyatake T. (2007) Extraction of finger-vein patterns using maximum curvature points in image profiles J. IEICE TRANSACTIONS on Information and Systems, 90(8): 1185-1194
[18] Han, W. Y., & Lee, J. C. (2012). Palm vein recognition using adaptive Gabor filter. Expert Systems with Applications, 39(18): 13225-13234.
[19] Hamming, R.W. (1950) Error detecting and error correcting codes. Bell System technical journal, 29: 147–160.