

# Unsupervised Palm Vein Image Segmentation

Ekaterina Safronova<sup>[0000-0001-7473-0178]</sup> and Elena Pavelyeva<sup>[0000-0002-3249-2156]</sup>

Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University,  
Moscow, Russia  
katyasafit@gmail.com, paveljeva@yandex.ru

**Abstract.** In this article the new hybrid algorithm for palm vein image segmentation using convolutional neural network and principal curvatures is proposed. After palm vein image preprocessing vein structure is detected using unsupervised learning approach based on W-Net architecture, that ties together into a single autoencoder two fully convolutional neural network architectures, each similar to the U-Net. Then segmentation results are improved using principal curvatures technique. Some vein points with highest maximum principal curvature values are selected, and the other vein points are found by moving from starting points along the direction of minimum principal curvature. To obtain the final vein image segmentation the result of intersection of the principal curvatures-based and neural network-based segmentations is taken. The evaluation of the proposed unsupervised image segmentation method based on palm vein recognition results using multilobe differential filters is given. Test results using CASIA multi-spectral palmprint image database show the effectiveness of the proposed segmentation approach.

**Keywords:** Biometrics · Image Segmentation · Palm Vein Recognition · Unsupervised Learning · Principal Curvatures.

## 1 Introduction

Nowadays information security plays crucial role in human life and, as it turned out, accustomed keys and passwords are not reliable enough. Instead, the biometric characteristics that uniquely identify a person from an entire population based on intrinsic physical or behavioral traits [1], provide stable and safe data protection. Biometrics recognizes individuals based on these characteristics.

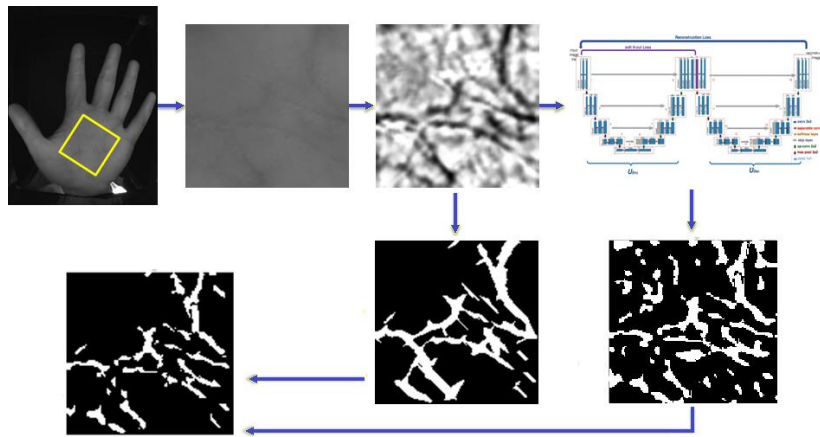
One of the most advanced and progressive personal identification technologies is palm vein recognition. Veins are usually not visible to others that provides low risk of fake or theft. As for other important advantages, vein patterns are quite unique to the owners, image acquisition does not require physical contact and the system can be made compact. Deoxygenated hemoglobin in the vein blood absorbs near infrared light so infrared camera captures images containing veins.

---

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Palm vein recognition algorithm consists of several steps. Firstly, the region of interest (ROI) is extracted and segmentation of ROI image into two classes is performed. Vein points are marked in white while the other points are marked in black. The second step, feature vector extraction, represents the main difference between existing approaches. Since vein recognition is relatively young study, some feature extraction methods could be derived from other biometric recognition algorithms based on statistical information [2], image key points [3, 4, 5], subspace-based methods [6], phase based methods [7, 8], etc. Some approaches were developed specifically for vein recognition [9]. The last algorithm step, image matching, is based on feature vector type. At this step the distance between palm vein images is calculated. Much recent work has been focused on employing deep convolutional neural networks (CNN) in biometrics. Deep learning methods can be applied to any step of palm vein recognition algorithm [10, 11, 12, 13].

In this paper we propose a hybrid approach based on unsupervised machine learning and mathematical methods to obtain good vein segmentation. The problem of unsupervised image segmentation is one of the major challenge in computer vision which has been deeply researched. The range of well-known techniques for solving this issue contains normalized-cuts [14], Markov random field-based methods [15], CNN-based approaches [16], etc. However, the results of applying of these methods may include inaccuracy due to specific features of a technique and lack of correct ground truth. So the mathematical methods shall control the results of CNN. In this paper we propose a hybrid segmentation method consisting of two approaches: based on CNN and principal curvatures (Fig.1).



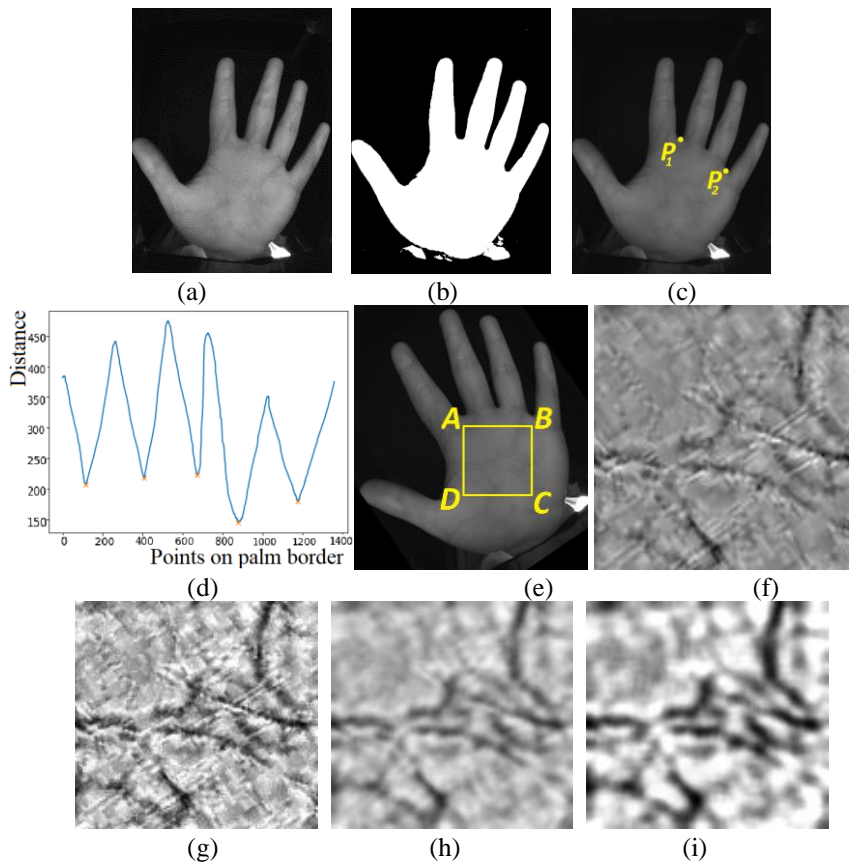
**Fig. 1.** The scheme of the proposed palm vein segmentation algorithm.

The rest of this paper is organized as follows. In Section 2 the palm vein image preprocessing and ROI extraction algorithms are described. The vein structure extraction is described in Section 3 where Subsections 3.1 and 3.2 present principal curvatures and CNN-based segmentation algorithms, the hybrid approach is described in Subsection 3.3. The evaluation of the proposed unsupervised image segmentation method based on

multilobe differential filters for palm vein feature extraction is described in Section 4. The experimental results for images from CASIA multi-spectral palmprint image database [17] are given in Section 5. Finally, Section 6 concludes this paper.

## 2 Palm vein image preprocessing

The proposed palm vein region of interest (ROI) detection and enhancement scheme is illustrated in Fig. 2.



**Fig. 2.** Illustration of palm vein ROI extraction and preprocessing: (a) original palm image, (b) binary hand image, (c) points  $P_1$  and  $P_2$ , (d) function that represents the distance between the center of palm and all points on the hand contour, (e) ROI - the square region of interest on the rotated image. (f) ROI image with uniform illumination. (g) ROI after CLAHE, (h) ROI after NLM, (i) final ROI image.

First, hand boundary is detected by OTSU binarization algorithm [18] and points between the fingers are found as points where local minimum of Euclidean distances between the center of palm and all points on the hand contour is reached. The points

between index and middle fingers,  $P_1$ , and forth and little fingers,  $P_2$ , can be taken as landmarks for extraction of square ROI (Fig. 2 d) [19]. To eliminate the influence of palm rotation, the image is rotated to the angle  $\theta$  which is the angle between the line  $P_1P_2$  and the horizontal line. To reduce the non-uniform illumination appearing in palm vein images the background is subtracted and the histogram is stretched (Fig. 2 f). To emphasize vein structure contrast-limited adaptive histogram equalization (CLAHE) technique [20] is used (Fig. 2 g). After contrast enhancement all image details including noise and glares are sharper. In order to smooth the undesirable details, non-local means (NLM) algorithm [21] is used to reduce noise (Fig. 2 h). NLM smoothies also veins a little so CLAHE is applied again to obtain distinguishable veins (Fig. 2 i). Fig. 2 shows the ROI of palm vein image and the results of preprocessing algorithm. After preprocessing veins become sharper and more distinguishable [22].

### 3 Vein structure extraction

#### 3.1 Principal curvatures

The next step is the vein structure extraction. Consider an image as a surface in a three-dimensional space, where the brightness value of the pixels is the z-coordinate. We are going to extract vein structure using principal curvatures method [22].

Let  $L(x, y)$  denote the image intensity at the pixel position,  $G(x, y)$  be the image gradient vector. Then the normalized gradient after a hard thresholding is defined as:

$$G_\gamma(x, y) = \begin{cases} \frac{G(x, y)}{\|G(x, y)\|}, & \|G(x, y)\| \geq \gamma \\ 0, & \|G(x, y)\| < \gamma \end{cases}, \quad (1)$$

where  $\gamma$  is a threshold level. In the experiments we use  $\gamma = 4$ . The normalized gradient field contains noisy components so we smooth it with Gaussian function  $H(x, y)$ :

$$H_\gamma(x, y) = G_\gamma(x, y) * H(x, y). \quad (2)$$

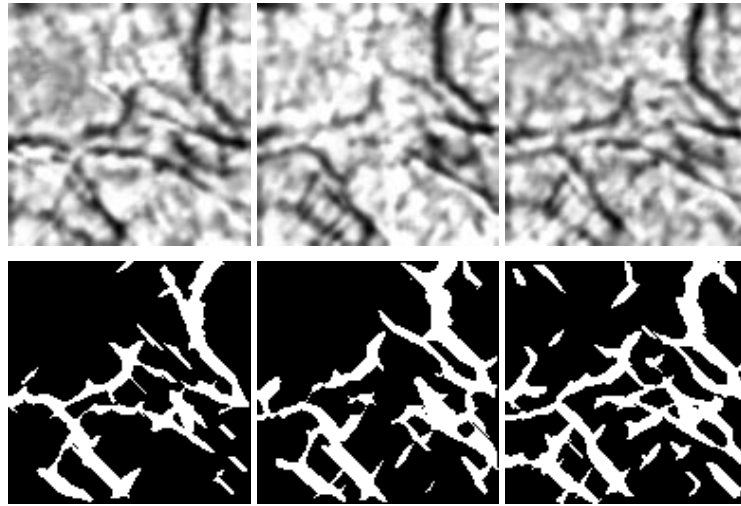
Let  $H_\gamma(x, y) = (h_x(x, y), h_y(x, y))$ . The local shape characteristics of an image at a point  $(x, y)$  can be described by the Hessian matrix  $H_S(x, y)$ .

$$H_S(x, y) = \begin{pmatrix} \frac{\partial h_x(x, y)}{\partial x} & \frac{\partial h_x(x, y)}{\partial y} \\ \frac{\partial h_y(x, y)}{\partial x} & \frac{\partial h_y(x, y)}{\partial y} \end{pmatrix}. \quad (3)$$

Let  $\lambda_1, \lambda_2$  be the eigenvalues and  $v_1, v_2$  be the corresponding eigenvectors of  $H_S(x, y)$ ,  $|\lambda_1| > |\lambda_2|$ . Then two principal directions, the directions of the maximum and minimum curvatures, are determined by two eigenvectors  $v_1$  and  $v_2$ . Consequently, two eigenvalues  $\lambda_1, \lambda_2$  represent the principal curvatures (the curvatures along the principal directions) [9]. The tubular-shaped regions have maximum principal curvature  $\lambda_1$  higher than other regions and vector  $v_1$  is directed across tubular direction, vector  $v_2$  – along tubular direction [23].

In order to catch veins of different widths, consider the set of parameters  $\sigma$  for the Gauss function:  $\sigma_0, \dots, \sigma_{n-1}$ , where  $n = 10$ ,  $\sigma_i = \sigma_0 \cdot \sqrt[4]{2^i}$ ,  $\sigma_0 = 2$ ,  $i = 0, 1, \dots, 9$ . For each value of  $\sigma$  the Hessian matrix is constructed, at each point the maximum positive eigenvalue  $\lambda_1$  and an eigenvector  $v_2$  corresponding to  $\lambda_2$  are calculated. Then, at each point of the image, the largest value of  $\lambda_1$  over all  $\sigma$  and the corresponding vector  $v_2$  are taken.

We select points with highest maximum principal curvature values as points that certainly belong to veins. The other vein points can be found [22] from starting points by moving along direction of vector  $v_2$  by  $|\lambda_1|$ . The results of this approach is shown in Fig.3.



**Fig. 3.** Vein structures extraction. First row shows ROI for different palm vein images of one person, second row shows found vein structure.

### 3.2 Unsupervised convolutional neural network

As we do not have the ground truth for the task of ROIs segmentation, the unsupervised method is required, so the approach based on W-Net architecture [16] is proposed. The authors of W-Net method present a new architecture which ties two fully convolutional network (FCN) architectures, each similar to the U-Net [24], together into a single autoencoder (Fig. 4). The first FCN encodes an input image into a k-way soft segmentation:  $U_{Enc}: \mathbb{R}^{H \times W \times 3} \rightarrow \mathbb{R}^{H \times W \times K}$ , where  $H \times W$  denotes a size of input image,  $U_{Enc}(x)_{ijk} = p(x_{ij} = A_k) \in [0, 1]$  measures the probability of pixel  $x_{ij}$  belonging to class  $k$  ( $A_k$  is set of pixels in segment  $k$ ). The second FCN, decoder, reverses this process, going from the segmentation layer back to a reconstructed image:  $U_{Dec}: \mathbb{R}^{H \times W \times K} \rightarrow \mathbb{R}^{H \times W \times 3}$  (Fig.4).

The both reconstruction errors of the autoencoder and soft normalized cut loss function on the encoding layer are used during training. The reconstruction loss is standard

for training encoder-decoder architecture and can be defined as:  $J_{reconstr} = \|x, U_{Dec}(U_{Enc}(x))\|_2^2$ .

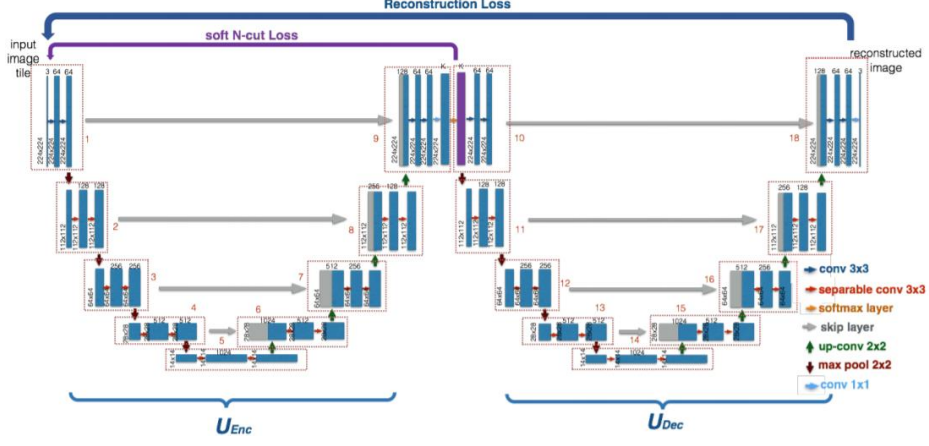


Fig. 4. W-Net architecture.

The output of the  $U_{Enc}$  is a normalized dense prediction. By taking the argmax, we can obtain a  $K$ -class prediction for each pixel and compute the normalized cut loss as following [14]:

$$Ncut_K(V) = \sum_{k=1}^K \frac{cut(A_k, V - A_k)}{assoc(A_k, V)} = \sum_{k=1}^K \frac{\sum_{u \in A_k, v \in V - A_k} w(u, v)}{\sum_{u \in A_k, t \in V} w(u, t)}, \quad (4)$$

where  $A_k$  is set of pixels in segment  $k$ ,  $V$  is the set of all pixels, and  $w$  measures the weight between two pixels.

However, since the argmax function is non-differentiable, it is impossible to calculate the corresponding gradient during backpropagation. Instead, it is proposed to use a soft version of the  $Ncut$  loss which is differentiable [16]:

$$J_{soft-Ncut}(V, K) = \sum_{k=1}^K \frac{cut(A_k, V - A_k)}{assoc(A_k, V)} = K - \sum_{k=1}^K \frac{assoc(A_k, A_k)}{assoc(A_k, V)} = K - \sum_{k=1}^K \frac{\sum_{u \in V, v \in V} w(u, v) p(u=A_k) p(v=A_k)}{\sum_{u \in A_k, t \in V} w(u, t) p(u=A_k)} = K - \sum_{k=1}^K \frac{\sum_{u \in V} p(u=A_k) \sum_{v \in V} w(u, v) p(v=A_k)}{\sum_{u \in V} p(u=A_k) \sum_{t \in V} w(u, t)}, \quad (5)$$

where  $p(u = A_k)$  measures the probability of node  $u$  belonging to class  $k$  that is directly computed by the encoder. The weight matrix  $W$  for  $J_{soft-Ncut}$  is defined as:

$$w_{i,j} = e^{-\frac{\|F(i) - F(j)\|_2^2}{\sigma_f^2}} * \begin{cases} e^{-\frac{\|X(i) - X(j)\|_2^2}{\sigma_x^2}} & \text{if } \|X(i) - X(j)\|_2 < r, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where  $X(i)$  and  $F(i)$  are the spatial location and pixel value of node  $i$ , respectively. Since the size of our ROI images is  $128 \times 128$  which is smaller than in the original work [16], the depth of W-Net was decreased in our experiments, as it is shown in Fig. 5. We use  $U_{Enc}: \mathbb{R}^{128 \times 128 \times 1} \rightarrow \mathbb{R}^{128 \times 128 \times K}$  and  $U_{Dec}: \mathbb{R}^{128 \times 128 \times K} \rightarrow \mathbb{R}^{128 \times 128 \times 1}$ .

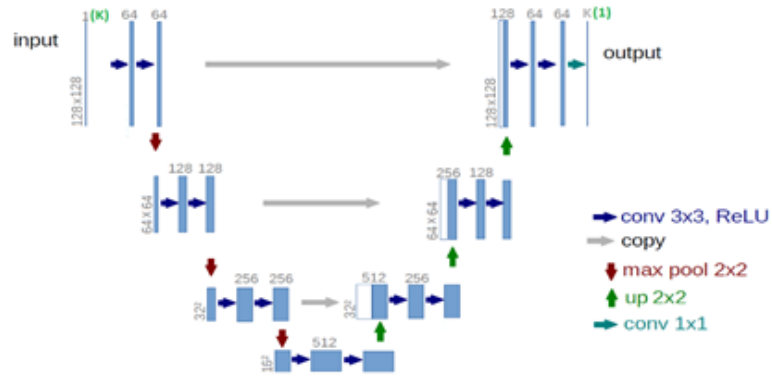


Fig. 5. The modified FCN for W-Net architecture.

As vein images have several semantic classes, such as veins with different intensity, background, skin wrinkles, the neural network was applied for the overclustering,  $K = 16$ . The training dataset contains 120 images. In Fig.6 one can see the results of this approach: the first column in the figure shows input images (Fig. 6 a) while the second one illustrates their reconstruction (Fig. 6 b); the third column presents the result of overclustering (Fig. 6 c). After unification of classes corresponding to veins the obtained vein image binarization is shown in the fourth column (Fig. 6 d).

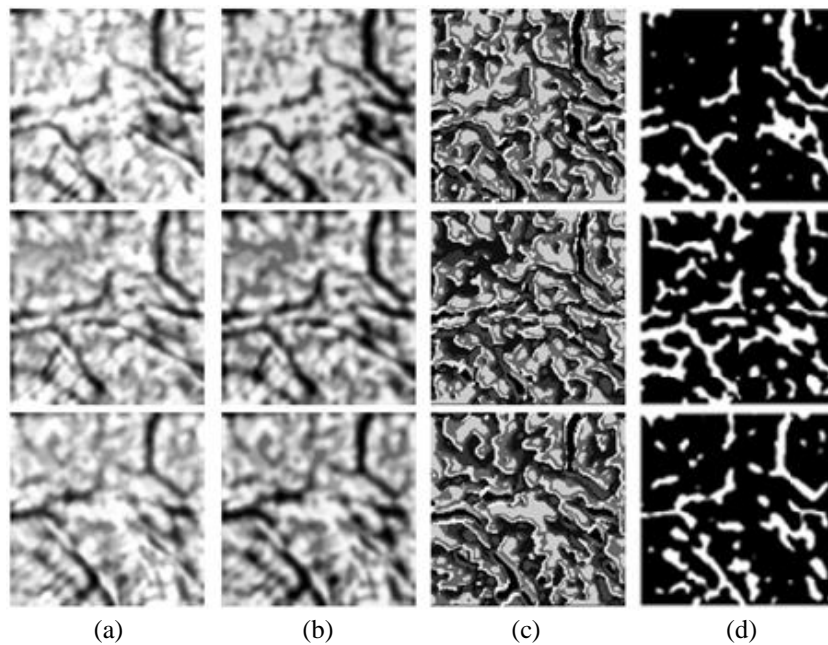
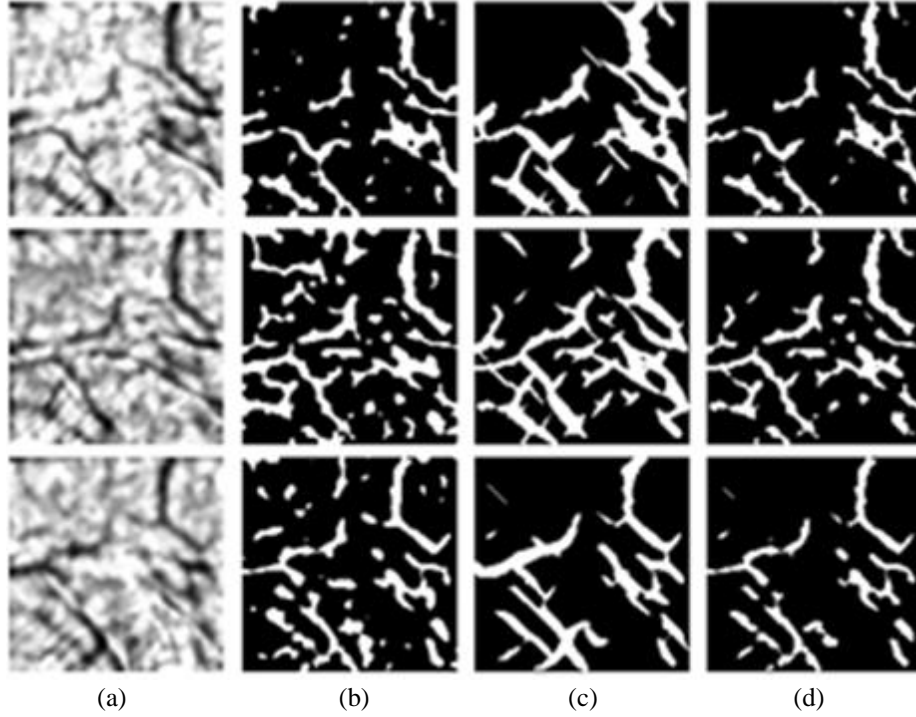


Fig. 6. (a) Input ROI images; (b) The result of CNN reconstruction; (c) The result of overclustering: each color corresponds to its own class; (d) The result of vein image segmentation using W-Net.

### 3.3 Hybrid segmentation method

Both principal curvatures-based and CNN-based approaches provide oversegmentation. To find the final vein mask the result of intersection of the principal curvatures-based and CNN-based vein segmentations is taken (Fig. 7).



**Fig. 7.** (a) Input ROI images; (b) Segmentation results using CNN; (c) Segmentation results using principal curvatures approach; (d) The hybrid palm vein image segmentation.

## 4 Evaluation of the image segmentation method

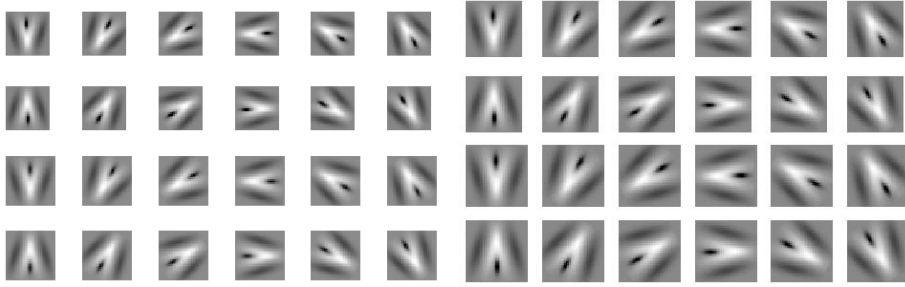
There are no ground truth masks for CASIA palmprint image database [17]. To evaluate the proposed unsupervised segmentation method, we use multilobe differential filters (MLDF) [25] that highlight vein branch points (Fig. 8) for palm vein feature extraction and normalized root-mean-square error for feature maps matching [22]. Mathematically the MLDFs are given as follows:

$$MLDF = C_p \sum_{i=1}^{N_p} \frac{1}{\sqrt{2\pi\sigma_{pi}}} e^{-\frac{(x-\mu_{pi})^2}{2\sigma_{pi}^2}} - C_n \sum_{i=1}^{N_n} \frac{1}{\sqrt{2\pi\sigma_{ni}}} e^{-\frac{(x-\mu_{ni})^2}{2\sigma_{ni}^2}}, \quad (7)$$

where the variables  $\mu$  and  $\sigma$  denote the central positions and the scales of a 2D Gaussian filters respectively,  $N_p$  denote the number of positive lobes, and  $N_n$  denote the number



of negative lobes. Constant coefficients  $C_p$  and  $C_n$  are used to ensure zero sum of the MLDF. We take the convolution results of the ROI images in vein points obtained after vein image segmentation with the proposed MLDF kernels to obtain the feature maps of vein images. In order to provide slight translation and rotation invariance matching, the normalized root-mean-square error (NRMSE) [26] is proposed for feature maps matching [22].



**Fig. 8.** Multilobe differential filters for vein image analysis.

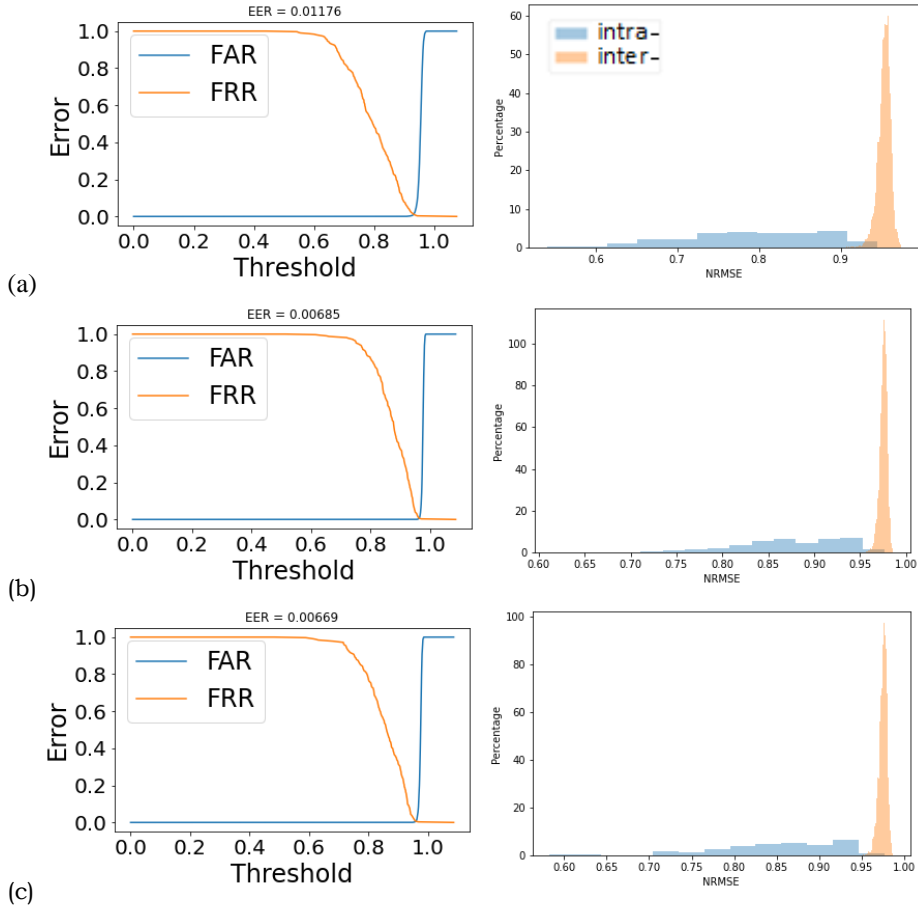
Given the intra- and interclass vein matching results, the recognition performance is measured by the following indicators: the distribution of genuine and impostor scores, False Acceptance Rate (FAR), False Reject Rate (FRR) and Equal Error Rate (EER) – the cross-over error rate when FAR is equal to FRR. Lower EER means higher accuracy of a biometric matcher.

## 5 Experimental results

Experimental results using CASIA Multi-Spectral Palmprint Image Database [17] are presented. The database contains 7200 palm images captured from 100 different people using a self-designed multiple spectral imaging device. Each sample contains six palm images which are captured at the same time with six different electromagnetic spectrums. Each hand of any person in the database is represented by six images at one wavelength. In our study the images from CASIA database obtained at 850 nm are taken.

The CNN model was implemented in PyTorch [27] framework and trained for 100 epochs with Google Colaboratory with a batch size of 16 using Adam optimizer [28] with the learning rate of 0.001. In order to train W-Net we randomly selected 20 hands from the dataset and took all 6 corresponding images, so we got 120 images in training set.

To test the proposed hybrid segmentation method, the recognition results using a part of CASIA database are presented in Fig. 9. Recognition results after image segmentation with principal curvatures and without W-Net based CNN (Fig. 3) is shown in Fig.9 a. Recognition results after image segmentation with W-Net based CNN and without principal curvatures (Fig. 6) is shown in Fig.9 b. Recognition results after proposed hybrid segmentation (Fig. 7) is shown in Fig.9 c.



**Fig. 9.** Illustrations of FAR and FRR curves with EER (the left column) and the distribution of genuine and impostor scores (the right column) on validation set using different segmentation methods: (a) the principal curvature approach; (b) the CNN-based approach; (c) the hybrid method.

## 6 Conclusion

In this article the new palm vein image segmentation method based on principal curvatures and unsupervised convolutional neural network is proposed. It is shown that the method based on principal curvatures improve segmentation results obtained by CNN. Experimental results using CASIA multi-spectral palmprint image database are presented.

## References

1. Jain, A. K., Bolle, R., Pankanti, S.: Biometrics: personal identification in networked society, Vol. 479. Springer Science & Business Media (2006).
2. Rosdi, B. A., Shing, C. W., Suandi, S. A.: Finger vein recognition using local line binary pattern. *Sensors* 11(12), 11357-11371 (2011).
3. Matsuda, Y., Miura, N., Nagasaka, A., Kiyomizu, H., Miyatake, T.: Finger-vein authentication based on deformation-tolerant feature-point matching. *Machine Vision and Applications* 27(2), 237-250 (2016).
4. Protsenko, M. A., Pavelyeva, E. A.: Iris image key points descriptors based on phase congruency. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42(2/W12), 167-171 (2019).
5. Wang, L., Leedham, G., Cho, D. S. Y.: Minutiae feature analysis for infrared hand vein pattern biometrics. *Pattern recognition* 41(3), 920-929 (2008).
6. Wu, J. D., Liu, C. T.: Finger-vein pattern identification using principal component analysis and the neural network technique. *Expert Systems with Applications* 38(5), 5423-5427 (2011).
7. Han, W. Y., Lee, J. C.: Palm vein recognition using adaptive Gabor filter. *Expert Systems with Applications* 39(18), 13225-13234 (2012).
8. Pavelyeva, E. A.: Image processing and analysis based on the use of phase information. *Computer Optics* 42(6), 1022-1034, 2018.
9. Choi, J. H., Song, W., Kim, T., Lee, S. R., Kim, H. T.: Finger vein extraction using gradient normalization and principal curvature. In: *Image Processing: Machine Vision Applications II*, pp. 725111. International Society for Optics and Photonics (2019).
10. Jha, R. R., Thapar, D., Patil, S. M., Nigam, A.: Ubsegnet: Unified biometric region of interest segmentation network. In: *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 923-928. IEEE (2017).
11. Lefkovits, S., Lefkovits, L., Szilágyi, L.: Applications of different CNN architectures for palm vein identification. In: *International Conference on Modeling Decisions for Artificial Intelligence*, pp. 295-306. Springer, Cham (2019).
12. Thapar, D., Jaswal, G., Nigam, A., Kanhangad, V.: PVSNet: Palm Vein Authentication Siamese Network Trained using Triplet Loss and Adaptive Hard Mining by Learning Enforced Domain Specific Features. In: *2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, pp. 1-8. IEEE (2019).
13. Wang, J., Yang, K., Pan, Z., Wang, G., Li, M., Li, Y.: Minutiae-Based Weighting Aggregation of Deep Convolutional Features for Vein Recognition. *IEEE Access* 6, 61640-61650 (2018).
14. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence* 22(8), 888-905 (2000).
15. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging* 20(1), 45-57 (2001).
16. Xia, X., Kulis, B.: W-Net: A deep model for fully unsupervised image segmentation. *arXiv preprint arXiv:1711.08506* (2017).
17. CASIA Multi-Spectral Palmprint Image Database, <http://biometrics.idealtest.org/>.
18. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* 9(1), 62-66 (1979).
19. Lin, C. L., Chuang, T. C., Fan, K. C.: Palmprint verification using hierarchical decomposition. *Pattern Recognition* 38(12), 2639-2652 (2005).

12 E. Safronova, E. Pavelyeva

20. Zuiderveld, K.: Contrast limited adaptive histogram equalization. *Graphics gems*, 474-485 (1994).
21. Buades, A., Coll, B., Morel, J. M.: A non-local algorithm for image denoising. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol.2, pp. 60-65. IEEE (2005).
22. Safronova, E. I., Pavelyeva, E. A.: Palm Vein Recognition Algorithm using Multilobe Differential Filters. In: Proceedings of 29-th International Conference on Computer Graphics and Vision GraphiCon, vol.1, pp. 117-121 (2019).
23. Renault, C., Desvignes, M., Revenu, M.: 3D curves tracking and its application to cortical sulci detection. In: Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101), vol.2, pp. 491-494. IEEE (2000).
24. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. Springer, Cham (2015).
25. Sun, Z., Tan, T.: Ordinal measures for iris recognition. *IEEE Transactions on pattern analysis and machine intelligence* 31(12), 2211-2226 (2008).
26. Fienup, J. R.: Invariant error metrics for image reconstruction. *Applied optics* 36(32), 8352-8357 (1997).
27. Paszke, A. et al. Automatic differentiation in pytorch (2017).
28. Kingma, D. P., Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).