International Journal of Innovative Computing, Information and Control
Volume **21**, Number **5**, October **2025**

ICIC International ©2025 ISSN 1349-4198
pp. **1333–1351**

IMPROVED PALM VEIN FEATURE SEGMENTATION ALGORITHM BASED ON DEEP ADVERSARIAL LEARNING

Shuqiang Yang[1,2], Zhaodi Wang[2,*], Huafeng Qin[3], Haofei Xi[3] and Junqiang Wang[2]

[1]School of Information and Control Engineering
China University of Mining and Technology
No. 1, Daxue Road, Nanhu District, Xuzhou 221116, P. R. China

[2]College of Physics and Electronic Information
Luoyang Normal University
No. 6, Jiqing Road, Yibin District, Luoyang 471934, P. R. China
{ joanhn; wangjunqiang }@lynu.edu.cn
[*]Corresponding author: wangzhaodi@lynu.edu.cn

[3]Chongqing Key Laboratory of Intelligent Perception and BlockChain Technology
Chongqing Technology and Business University
No. 19, Xuefu Avenue, Nan'an District, Chongqing 400067, P. R. China
qinhuafengfeng@163.com; xihaofei@ctbu.edu.cn

Received March 2025; revised July 2025

Abstract. *Deep learning model training on huge data is capable of achieving robust feature representation capacity. However, there are small registered samples for vein recognition task, which reduce the performance of deep learning model due to overfitting. To address this problem, this paper proposes a palm vein feature segmentation algorithm, named by AdveinSeg for vein recognition. Firstly, a convolutional neural network-based model is established as generator to produce vein images and gold standard (Ground Truth). Then, we build a U-Net network-based model for vein texture segmentation. Finally, the generation model is combined with the segmentation model to obtain a vein pattern segmentation model, which is trained in adversarial way. To prevent semantic collapse of vein images, cosine similarity is added to regularize the transformed sample pairs. Experimental results on three public palm vein datasets, CASIA, VERA and PolyU, show that the deep adversarial learning-based palm vein feature segmentation algorithm outperforms existing vein segmentation methods with equal error rate (EER) of 0.33%, 0.59% and 0.60%, respectively.*

Keywords: Palm vein segmentation, Deep adversarial learning, AdveinSeg algorithm, Image segmentation

1. **Introduction.** With the progress of science and technology, information technology and computer technology, information security has become a hot topic in access control and user privacy management. The traditional method of identity authentication is vulnerable to exploitation by unauthorized individuals who may obtain passwords or ID cards. Furthermore, other authentication methods are not secure. Therefore, biometric-based identity authentication technology has been widely investigated and applied over recent years. This is due to the fact that biometric identification technology offers a secure, accurate and convenient identification mechanism. Currently, biometric identification technology is replacing some traditional identification technologies in an increasing number of fields. It is also becoming a key research area in the field of biometric identification.

DOI: 10.24507/ijicic.21.05.1333

Biometric modality can currently be categorized into two main types: extrinsic and intrinsic. Extrinsic biometric features include facial characteristics, fingerprints, and voiceprints, whereas intrinsic features refer to attributes such as vein patterns. Extrinsic characteristics are generally more susceptible to environmental interference and potential forgery compared to internal characteristics. For example, facial recognition accuracy is compromised by variations in lighting, facial expressions, and head posture. Similarly, commonly used fingerprints are vulnerable to forgery and its recognition accuracy will reduce when it is contaminated by water or foreign substances. Moreover, fingerprint recognition requires physical contact, which may be perceived as intrusive. Voiceprint recognition, on the other hand, can exhibit variability due to physiological, psychological, and environmental factors. As voiceprint recognition systems face challenges from sophisticated attempts to mimic voiceprints using high-definition audio, it is increasingly necessary to integrate this technology with other biometric modalities to enhance accuracy and security. In contrast, vein recognition offers distinct advantages due to its reliance on the differential absorption of near-infrared light by hemoglobin in blood vessels compared to surrounding tissues. This process generates high-contrast images of vein patterns based on data preprocessing, revealing unique vein distribution characteristics. Vein recognition is particularly competitive among biometric modality for its non-contact nature, capability for living body detection, and uniqueness. The segmentation of veins as an essential step in vein recognition involves the precise isolation of blood vessels and extraction of regions of interest from vein images. This step is crucial for improving the processing efficiency of biometric authentication tasks, which has led to growing research interest in recent years. Additionally, palm vein patterns, which contain more information than finger veins, offer greater reliability for identity verification, information protection, and security applications.

1.1. **Related work.** Venous blood vessels are hidden under the skin of the human body and their distribution is well connected, which is difficult to observe under visible light and requires the help of infrared light of about 850 nm wavelength to capture the venous images, which leads to the advantages of stability, viability and inability to forge the venous features. In some medical research works, vein features have been shown to have a unique structure for each individual and are therefore unique. However, vein recognition is still challenging because palm vein images are susceptible to ambient lighting, rotation angle, and individual differences in acquisition equipment during the acquisition process, resulting in poor contrast, uneven brightness, and unclear vein texture in the acquired palm vein images, which increases the difficulty of image segmentation. In order to solve this problem, many methods have been proposed, which are roughly divided into three main categories according to the segmentation methods.

1) Local Feature-Based Segmentation: This type of method is based on the local features of the image, and typical methods are threshold segmentation [1], edge detection [2], region growth [3], and feature space clustering [4]. Threshold segmentation relies heavily on selecting the optimal threshold. For example, Wang et al. [5] introduced an improved algorithm to find the best multi-threshold combination, enabling effective segmentation into meaningful regions. Edge detection is an image segmentation method using classical detection operators. In [6], Tchinda et al. combined edge detection filters with neural networks to enhance segmentation accuracy in blood vessel images, first detecting edges and then refining them through neural networks. Region growth is based on local image features and often requires prior knowledge. Raja et al. [7] proposed an enhanced method that optimizes seed selection and uses Tsallis entropy, achieving accurate segmentation in contrast-enhanced medical images. Feature space clustering groups similar pixels in

feature space. In [8], Wisaeng presented a method that combines weighted kernel fuzzy C-means (WKFCM) clustering with function-based expansion. This approach captures nonlinear relationships and emphasizes important features, resulting in precise segmentation of vessels.

2) Theory-Based Segmentation: This kind of segmentation method is based on specific mathematical theories, such as morphological theory [9], genetic algorithm theory [10], and fuzzy graph theory [11]. The method based on morphology theory completes the segmentation task by analyzing the topological structure of the image. For example, Tian et al. [12] presented a segmentation technique that combines an improved Frangi filter with mathematical morphology. This approach enhances the edge features of retinal vessels through morphological operations, thereby achieving precise segmentation of vascular structures within images. Additionally, the method based on genetic algorithm theory, which is a global optimization technique simulating the process of biological evolution, is employed to determine the optimal threshold or parameters for image segmentation. In [13], Bahadure et al. illustrated the application of genetic algorithms to optimize parameter selection during the segmentation process. By harnessing the robust search capabilities of genetic algorithms, this method effectively determines optimal thresholds and parameters, leading to accurate segmentation and classification of brain tumors in MRI (Magnetic Resonance Imaging) images. The method based on graph theory converts the segmentation problem into graph partition, and completes the segmentation process by optimizing the objective function. [14] introduced a region and boundary aggregation method based on graph theory, which constructs a graph model of the image and aggregates region and boundary information. This method successfully integrates structural features of images through graph-theoretic techniques, resulting in accurate segmentation of biomedical images.

3) Deep Learning-Based Segmentation: With the development and introduction of deep learning theory, the field of computer vision has made breakthrough progress. Convolutional neural network (CNN) has become an important tool in image processing. Deep learning techniques exploit semantic information from images to achieve precise semantic segmentation. Several deep learning-based semantic segmentation methods have been developed, including fully convolutional network (FCN), pyramid scene parsing network (PSPNet), DeepLab, Mask R-CNN, SegFormer, and U-Net. For instance, the fully convolutional network (FCN) introduced by Long et al. [15] can be trained end-to-end, utilizing shared image information between the up-sampling and down-sampling paths to enhance segmentation. [16] employs a fully convolutional network that avoids fully connected layers and uses skip connections to merge image features at different scales, allowing for pixel-wise predictions. However, the results produced by FCN lack refinement and do not adequately consider the relationships between pixels. SegNet [17] is built on an Encoder-Decoder structure. It encodes the input image into a low-dimensional representation, and then reconstructs the image using directional invariance in the decoder to produce the segmented output. DeepLab [18] integrates deep convolutional neural networks with probabilistic map models for semantic segmentation, allowing for pixel-wise classification. Its key innovation is the combination of conditional random fields with deep convolutional networks, utilizing dilated convolutions to control the receptive field size. However, dilated convolutions struggle with capturing relationships between objects of different sizes, leading to challenges in accurately segmenting small objects. Mask R-CNN [19] builds on Faster R-CNN by adding a mask prediction branch, excelling in instance segmentation, object detection, and pixel-level segmentation. Mask Scoring R-CNN [20] improves the quality of predicted masks by incorporating Mask-IoU, though this adds complexity to the network structure. Gated-SCNN [21] introduces a two-stream convolutional neural

network for semantic segmentation, featuring a separate branch called shape flow to learn edge information. This enhancement allows the model to accurately predict object edges and significantly improves segmentation for small and thin objects. SegFormer [22] integrates a hierarchical structure into the Transformer model, extracting information at different scales and combining the Transformer with lightweight MLPs. This approach achieves superior results on datasets like ADE 20k and Cityscapes. U-Net [23], a variant of FCN, utilizes an Encoder-Decoder structure. The encoder reduces the spatial dimensions of the image, while the decoder gradually restores details. The interconnected layers allow the decoder to effectively repair and reconstruct object details. This method is particularly suitable for small datasets, making it effective for image segmentation with limited palm vein data samples.

1.2. **Motivation.** In the aforementioned work, local feature-based segmentation methods are simple and easy to implement, and do not require a large number of training samples, but the segmentation effect for complex backgrounds is poor, and due to lack of global information, it is easy to over-segment or under-segment. The threshold segmentation method [24] struggles to find an appropriate global threshold because of the small gray-level difference between the vein and background, leading to inaccurate results. Edge detection methods [25] can detect vein edges, but for palm vein images, small vessels and branches are easily affected by noise, resulting in broken or incomplete edges. The region growth method [26] relies on seed points, and if they are placed in non-uniform areas or are affected by noise, it can lead to incorrect growth direction and affect segmentation accuracy. Although feature space clustering methods can differentiate vein and non-vein regions to some extent, the complexity of vein structures and diverse background changes in palm vein images may cause misclassification of non-vein structures, resulting in segmentation errors [27].

Theory-based segmentation methods require a strong mathematical foundation, and they can handle complex backgrounds and irregular shapes effectively under certain conditions [28]. However, when applied to large datasets, they tend to have high computational costs, slow convergence, and low efficiency. For example, morphological theory-based methods [29] enhance vein features and reduce noise through operations like dilation and erosion, but the results depend heavily on the shape and size of the chosen structural elements. This is particularly challenging for palm vein images, where vein shapes and distributions vary significantly, requiring careful parameter adjustment. Genetic algorithm-based methods [30], which use global optimization to find the best segmentation parameters, are suitable for complex images but are not ideal for real-time applications in palm vein segmentation due to high computational complexity and slow convergence. Graph theory-based methods [31] often involve complex optimization problems with computationally intensive objective functions, leading to long processing times and high computational costs for palm vein images. The fine branching and complex topology of veins make graph theory methods prone to mis-segmentation, especially when vessels overlap or intersect. Additionally, the effectiveness of these methods relies heavily on appropriate parameter settings, which is complicated by individual differences and varying imaging conditions in vein images, further increasing the difficulty of the segmentation process.

Compared to the first two methods, deep learning-based segmentation methods can extract rich information when trained on large datasets in an end-to-end manner, resulting in better performance. FCN [32] is used for initial segmentation; PSPNet [33] or DeepLab [34] is preferred for finer tasks due to their superior detail capture. Mask R-CNN [35] is ideal for distinguishing multiple target regions because of its instance-level segmentation

ability. U-Net [36] helps recover image details and boundaries through skip connections and the symmetric structure between the encoder and decoder. Additionally, U-Net can train effective models on relatively small datasets [37], which is particularly beneficial for palm vein segmentation, as these datasets are usually small. U-Net also has good generalization ability [38], maintaining stable performance across different vein images and delivering reliable results even with individual variations and changing imaging conditions. These features make U-Net a highly effective tool for palm vein image segmentation. However, palm vein images are affected by light, rotation angle, and collection equipment [39], leading to issues like small datasets, low contrast, and background noise. Additionally, deep learning models require large amounts of training data, but in practice, privacy concerns make users hesitant to provide extensive data. This results in fewer registration samples per person, limiting the model's performance in vein segmentation and leading to overfitting. How to segment the palm vein features effectively is an urgent problem to be solved in the case of limited vein data samples.

Despite their respective advantages, a common shortcoming across all three categories is the limited ability to accurately segment palm veins under small-sample, noisy, or low-contrast conditions while preserving structural integrity. Traditional methods fail to generalize across image variations; theory-driven models are computationally demanding; and deep learning models require abundant training data and often neglect vascular boundary precision. Therefore, a critical research gap remains: how to design a segmentation framework that simultaneously ensures generalization in small-data regimes and accurately preserves the detailed morphology of palm veins under challenging imaging conditions. Addressing this dual challenge is crucial for advancing palm vein recognition systems towards practical and robust deployment.

1.3. **Our work.** Inspired by the above research, we propose AdveinSeg, a deep adversarial learning-based algorithm for palm vein feature segmentation, as shown in Figure 1. First, we label the images with probability maps accumulated by four traditional segmentation methods. To realize vein data augmentation, we construct Pix2Pix convolution model to transform the palm vein images and gold standard images to generate diverse vein images and gold standard images; Then we construct a U-Net segmentation model and train the segmentation network to minimize the cross-entropy loss by adversarial learning with the transformed palm vein dataset. Finally, we establish an adversarial learning-based segmentation framework by combining the transformation and segmentation models. In order to prevent semantic collapse of vein images, cosine similarity is added to regularize the transformed sample pairs. We conduct rigorous experiments on three public palm-vein datasets, and the experimental results show that the proposed AdveinSeg algorithm can reduce the equal error rate of existing vein segmentation and outperforms existing vein segmentation algorithms.

The main contributions of this paper are as follows.

- We design a deep adversarial learning-based palm vein segmentation algorithm, called AdveinSeg. This algorithm transforms palm vein images using a Pix2Pix image transformation network and performs segmentation with a U-Net network. We introduce an adversarial framework that jointly trains both the image segmentation and image transformation networks. Through adversarial learning, the segmentation network's training loss is refined, allowing the model to achieve optimal feature representation.
- In order to enhance model generalization, we use adversarial learning to generate diverse vein images and gold standards. We apply a conditional GAN (cGAN) for image transformation, where the generator is based on a U-Net network and the

discriminator uses a convolution-based PatchGAN. This adversarial network setup transforms palm vein images and gold standards, producing diverse vein images and augmenting the dataset.

- To prevent overfitting, we design the AdveinSeg loss function and optimization of AdveinSeg model by using the method of gradient descent and gradient ascent updating alternately. We adopt adversarial learning method, which takes the training sample set obtained by image transformation network as the opponent of target network, and prevents the learning representation from overfitting the target network through supervised adversarial learning.
- To evaluate the performance of AdveinSeg, we conducted extensive experiments on three public palm vein databases and compared our method with traditional image segmentation approaches. The experimental results demonstrate that the deep adversarial learning-based algorithm outperforms existing vein segmentation methods. Specifically, AdveinSeg achieves superior smoothness and continuity in vein features, a lower equal error rate, and overall better segmentation performance.

2. **The Proposed Approach.** In this section, we detail the proposed AdveinSeg, a palm vein feature segmentation algorithm based on deep adversarial learning, which consists of two main parts: an image transformation module and an image segmentation module, as shown by the architecture of AdveinSeg in Figure 1. First, we establish a convolutional neural network to transform palm vein images and their gold standards, generating diverse vein images and corresponding gold standards. Next, we build a vein segmentation model based on the U-Net network. Finally, we combine the transformation model with the segmentation model to create a vein segmentation framework based on adversarial learning. In this framework, the transformation module generates challenging samples using a convolutional neural network, increasing the difficulty for the segmentation model. The segmentation module then learns more robust feature representations from these challenging samples, improving the generalization ability of the entire model. To prevent semantic collapse of the vein images, cosine similarity is applied to regularizing the transformed sample pairs.

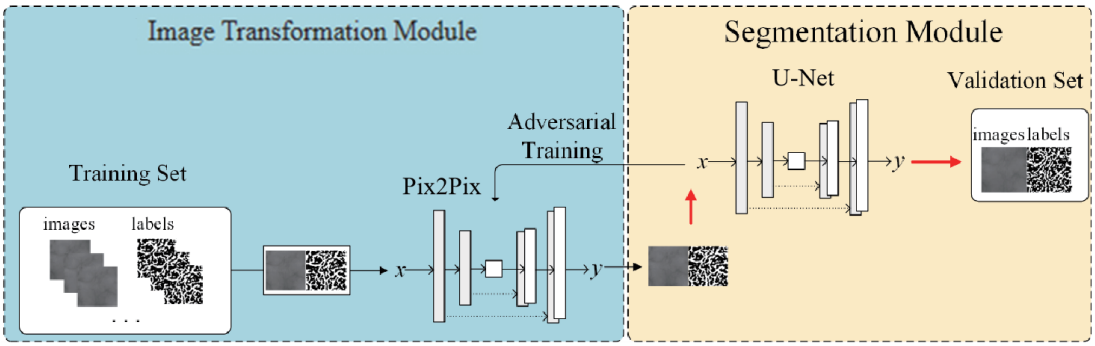


FIGURE 1. Palm vein feature segmentation algorithm based on deep adversarial learning

2.1. **Image labelling.** Pix2Pix is a cGAN-based adversarial learning network that, given a real vein image and corresponding semantic labels, is able to generate a real image under supervised learning of the labels. In order to generate a variety of vein images and gold standard, and improve the generalization ability of segmentation model, we use automatic labeling method to make vein labels, as shown in Figure 2. In this paper, we use four traditional segmentation algorithms, namely mean curvature [40], region growth

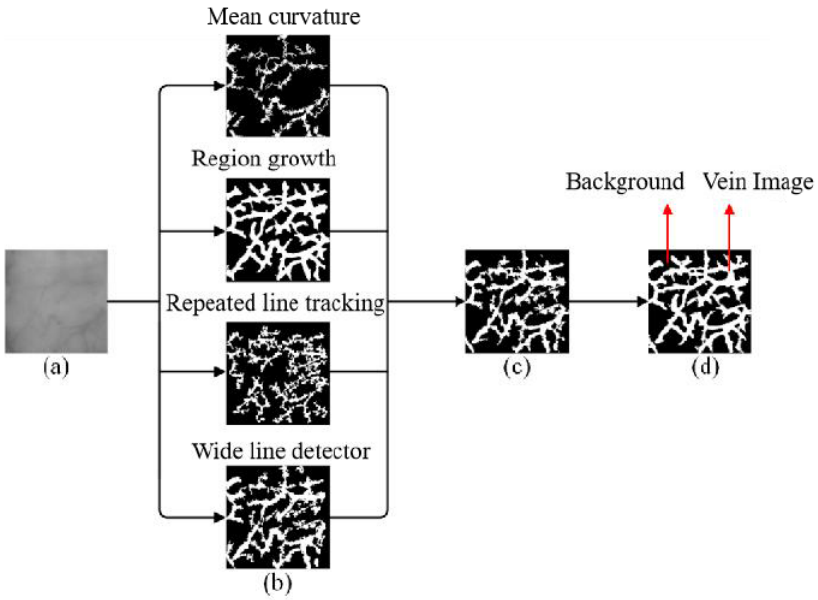


FIGURE 2. Image labelling: (a) Original image; (b) four vein extraction methods; (c) probability map; (d) labelled image

[41], repetitive linear tracking [42] and wide line detector [43], to segment the vein and background in the palm vein image, respectively. The resulting binary images from each method are combined into a probability map, which is then binarized by Equation (1) to create the final labels, with black background pixels set to 0 and white vein pixels set to 1.

$$L(i,j) = \begin{cases} 255 & \sum_{k=1}^{K} l_k(i,j) \geq \dfrac{K}{2} \\ 0 & \sum_{k=1}^{K} l_k(i,j) < \dfrac{K}{2} \end{cases} \tag{1}$$

where K is the number of existing vein segmentation methods, and the value in this paper is 4, $l_k(i,j)$ is the jth segmented image in the ith class obtained by k existing segmentation methods, and $L(i,j)$ represents the jth label in the ith class.

2.2. **Pix2Pix-based augmentation.** Direct training on limited palm vein data is prone to overfitting. To alleviate this, we adopt classical image augmentation techniques, such as random horizontal/vertical flipping and brightness adjustment, which help the neural network to learn more diverse feature distributions. To further improve generalization and generate challenging samples, we employ a conditional Generative Adversarial Network (cGAN) following the Pix2Pix framework [44] in the image transformation module. In this framework, the generator adopts a U-Net structure, and the discriminator is a convolutional PatchGAN. The generator is responsible for learning a mapping from grayscale palm vein images to their corresponding binary segmentation maps. This mapping preserves spatial structure and vein continuity by leveraging skip connections inherent in the U-Net design. The image transformation network ensures consistency between input and output image structures, facilitating realistic image generation that mimics the distribution of real palm vein data.

Note that structural details of the U-Net are discussed in Section 2.3, as the same architecture is employed in the segmentation network. Here, its role is to support data-level augmentation via image synthesis.

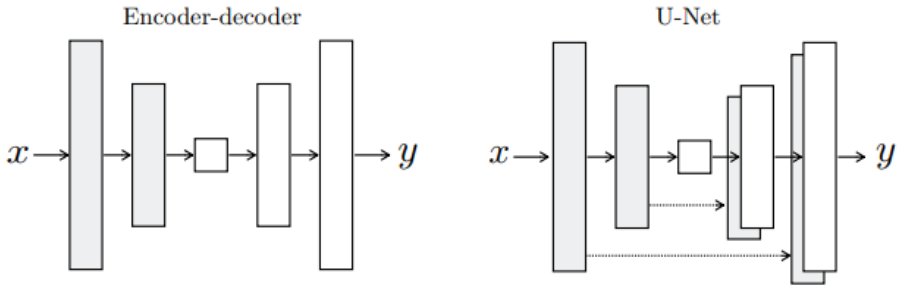


FIGURE 3. Encoder-decoder structure and U-Net

2.3. **U-Net based segmentation.** As discussed in Section 2.2, the image transformation network uses a U-Net generator to synthesize new grayscale/binary vein image pairs for data augmentation. However, such augmentation alone is insufficient to fully cover the diversity of palm vein distributions. Therefore, we further propose a deep adversarial learning-based segmentation framework – AdveinSeg – which integrates both image transformation and segmentation in a unified U-Net-based structure. The detailed structure and function of the U-Net employed in segmentation are presented below. The image transformation network generates palm vein images based on varying inputs, but it cannot fully cover all positions within the data distribution space. Despite the use of data augmentation during image transformation, training the segmentation network remains challenging due to the inherent limitations in the diversity and quality of the training data. To address this, we propose a novel deep adversarial learning-based palm vein feature segmentation algorithm, termed AdveinSeg, which jointly optimizes both the segmentation network and the image transformation network.

Deep adversarial learning is a machine learning framework in which two neural networks – the generator and the discriminator – are trained in opposition. The generator's role is to produce feature maps or images that closely resemble real data, while the discriminator's task is to distinguish between the generated images and real images. This adversarial process fosters mutual improvement of both networks: the generator learns to produce more realistic outputs, and the discriminator becomes better at distinguishing between real and generated data. In the context of our method, deep adversarial learning refines the segmentation output by ensuring that the generated feature maps more closely align with the ground truth palm vein features. This adversarial training process enables the segmentation network to learn more robust, discriminative features, thereby improving overall segmentation accuracy.

The segmentation and image transformation networks in our approach are based on the U-Net architecture, which is specifically designed for pixel-level segmentation tasks. The U-Net structure consists of three primary stages: feature extraction, concatenation, and up-sampling.

During the feature extraction, spatial information is progressively reduced while enhancing the relevant palm vein features. This is accomplished through two consecutive 3×3 convolutional operations (without padding), followed by a ReLU activation function and 2×2 max-pooling for down-sampling with a stride of 2. At each down-sampling stage, the number of feature channels is doubled to capture increasingly complex and abstract features from the input data. This design allows the network to learn multi-scale features that are crucial for accurate segmentation. The image transformation network plays a pivotal role in improving feature representation by generating new images based on varying inputs. This ensures a more comprehensive coverage of the data distribution

space, thereby augmenting the training set. During training, the adversarial process prevents overfitting to specific data patterns, making the model more adaptable and capable of generalizing to unseen data. In the up-sampling stage, the feature map is up-sampled using transposed convolutions (2×2 "up-convolutions") to reduce the number of feature channels. The up-sampled feature map is then concatenated with the corresponding feature map from the down-sampling path, followed by two additional 3×3 convolutions, each followed by a ReLU activation. This concatenation ensures that both low-level and high-level features are preserved and combined, which enhances the model's ability to make accurate pixel-wise decisions.

One common challenge during convolution operations is the loss of border pixels, which occurs due to the nature of convolutional kernels. This issue is addressed by implementing a clipping mechanism, which ensures that the border pixels are properly handled. Finally, the output layer consists of a 1×1 convolution that classifies each 64-dimensional feature vector, effectively distinguishing between palm veins and the background.

By combining the powerful segmentation capabilities of U-Net with the robustness of deep adversarial learning, AdveinSeg is able to produce highly accurate and consistent palm vein feature segmentation results. The adversarial training approach ensures that the method is resilient to variations in image quality, and can handle diverse palm vein patterns, making it an effective solution for challenging biometric applications.

2.4. **Adversarial segmentation.** In segmentation module, shown as in Figure 1, $\mathbb{X} = \{x_{ck} | c = 1, \ldots, C; k = 1, \ldots, K\}$ is the palm vein grey scale image dataset, where x_{ck} denotes the Kth palm vein grey scale image of class ith. $\mathbb{Y} = \{y_{ck} | c = 1, \ldots, C; k = 1, \ldots, K\}$ is the palm vein labelled image dataset, where y_{ck} denotes the Kth palm vein binary image of class ith. The converted palm vein image dataset $\mathcal{D}$ can be obtained based on the image transformation network. Based on the transformed palm vein dataset $\mathcal{D}$, the segmentation network is trained to minimize cross-entropy loss. The image transformation network generates the palm vein dataset $\mathcal{D}$ using a gray image and its corresponding binary image as inputs. This process aims to reach an equilibrium where the feature representation ability of the model is optimized by addressing the training loss of the segmentation network through adversarial learning.

1) Adversarial Loss: The loss function is defined as $L_{ce}(\cdot)$, the image transformation network as $F(\cdot)$ and the segmentation network as $S(\cdot)$. The loss function for jointly training the segmentation network $S(\cdot)$ and the image transformation network $F(\cdot)$ for adversarial learning by converting the palm vein grey scale image and the binary image into the sample data space is defined as Equation (2).

$$S^*, F^* = \arg \min_S \max_F \left[\mathbb{E}_{x_{ck}, y_{ck}} (L_{ce}(\phi_F(x_{ck}, y_{ck}))) + \mathbb{E}_{x'_{ck}, y'_{ck}} \left(L_{ce} \left(\phi_S \left(x'_{ck}, y'_{ck} \right) \right) \right) \right] \tag{2}$$

where x_{ck} represents the Kth palm vein image of class ith, and y_{ck} is the corresponding binary image. x'_{ck} is the transformed version of the Kth palm vein image of class ith, obtained through the image transformation network $F(\cdot)$, while y'_{ck} is the transformed binary image. The predicted labels for each query in the current mini-batch are determined by the segmentation network $S(\cdot)$, and $S(\cdot)$ is optimized based on these predictions. The loss function $L_{ce} \left(\phi_S \left(x'_{ck}, y'_{ck} \right) \right)$ is calculated using the transformed palm vein image x'_{ck} and its corresponding binary image y'_{ck}. To optimize the parameter S, $\phi_F(\cdot)$ introduces noise in the output images, challenging the network. However, this can result in images with missing semantic details. To address this, cosine similarity is added as a regularization term to control the output of the image transformation network, and the loss function is modified as Equation (3) accordingly.

$$S^*, F^* = \arg\min_S \max_F \left[\mathbb{E}_{x_{ck},y_{ck}}(L_{ce}(\phi_F(x_{ck}, y_{ck}))) - \lambda_1 L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) - \lambda_2 L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right) + \mathbb{E}_{x'_{ck},y'_{ck}}\left(L_{ce}\left(\phi_S\left(x'_{ck}, y'_{ck}\right)\right)\right)\right] \tag{3}$$

where *cosine* is the cosine similarity function, and λ_1 and λ_2 are scaling factors, in the experiment $\lambda_1 = \lambda_2 = 0.5$, $x'_{ck} = \phi_F(x_{ck})$, $y'_{ck} = \phi_F(y_{ck})$.

2) Model Optimization: As shown in current adversarial learning methods, finding saddle point solutions (S^*, F^*) in the above equation is challenging. Typically, gradient descent and gradient ascent are used to alternately update S and F. For instance, the optimization of S can be reformulated as a minimization problem in Equation (4).

$$S^* = \arg\min_S \left[\mathbb{E}_{x_{ck},y_{ck}}(L_{ce}(\phi_F(x_{ck}, y_{ck}))) - \lambda_1 L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) - \lambda_2 L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right) + \mathbb{E}_{x'_{ck},y'_{ck}}\left(L_{ce}\left(\phi_F\left(x'_{ck}, y'_{ck}\right)\right)\right)\right] \tag{4}$$

where $\lambda_1 = \lambda_2 = 0.5$. The problem in the above equation is usually solved by vanilla SGD with learning rate α and batch size N, the training process for each batch can be expressed as Equation (5).

$$S(t+1) = S(t) - \alpha\frac{1}{N}\left[\sum_{n=1}^{N} \nabla_S \left[L_{ce}(\phi_F(x_{ck}, y_{ck})) - \lambda_1 L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) - \lambda_2 L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right) + \mathbb{E}_{x'_{ck},y'_{ck}} L_{ce}\left(\phi_S\left(x'_{ck}, y'_{ck}\right)\right)\right]\right] \tag{5}$$

Since the cosine similarity is independent of S, Equation (5) is replaced by

$$S(t+1) = S(t) - \alpha\frac{1}{N}\left[\sum_{n=1}^{N} \nabla_S \left[L_{ce}(\phi_F(x_{ck}, y_{ck})) + \mathbb{E}_{x'_{ck},y'_{ck}}\left(L_{ce}\left(\phi_S\left(x'_{ck}, y'_{ck}\right)\right)\right)\right]\right] \tag{6}$$

The training process uses gradient calculations averaged over more than N instances to reduce gradient variance, leading to faster convergence of the target network. However, this approach is prone to overfitting due to limited training data. To address this, the training set generated by the adversarial learning-based image transformation network is used as an adversary to the target network, resulting in a minimax problem for the self-trained network. This self-supervised objective is designed to be challenging enough to prevent the model from overfitting, and can be mathematically formulated as the maximization problem in Equation (7).

$$F^* = \arg\max_F \left[\mathbb{E}_{x_{ck},y_{ck}}(L_{ce}(\phi_F(x_{ck}, y_{ck}))) - \lambda_1 L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) - \lambda_2 L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right) + \mathbb{E}_{x'_{ck},y'_{ck}}\left(L_{ce}\left(\phi_S\left(x'_{ck}, y'_{ck}\right)\right)\right)\right] \tag{7}$$

To solve the above problem, a gradient ascent method is used for updating with a learning rate of β. The parameter updating rule is defined as in Equation (8).

$$F(t+1) = F(t) + \beta\frac{1}{N}\left[\sum_{n=1}^{N_1} \nabla_F [L_{ce}(\phi_F(x_{ck}, y_{ck}))] - \lambda_1 L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) - \lambda_2 L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right) + \mathbb{E}_{x'_{ck},y'_{ck}}\left(L_{ce}\left(\phi_S\left(x'_{ck}, y'_{ck}\right)\right)\right)\right] \tag{8}$$

Intuitively, the optimization of Equation (7) can be understood as follows: for the image transformation F, when $L_{ce}(\phi_F(x_{ck}, y_{ck}))$ is maximized and $L_{ce}\left(cosine\left(x_{ck}, x'_{ck}\right)\right) +$

$L_{ce}\left(cosine\left(y_{ck}, y'_{ck}\right)\right)$ is minimized, the loss of the segmentation network is maximized. This tends to push the generated samples further away. Samples from the same class are generated as challenging examples, following the trajectory of segmentation network updates under cosine similarity constraints, while remaining within recognizable limits and at a controlled distance from the original image. The palm vein feature segmentation algorithm based on deep adversarial learning is summarized in Table 1.

Table 1. Description of the AdveinSeg algorithm

Algorithm. Joint training of segmentation network and image transformation network

Input:

The set of palm vein grayscale images $\mathbb{X} = [x_{11}, x_{12}, \ldots, x_{CL}]$; The set of palm vein binary images $\mathbb{Y} = [y_{11}, y_{12}, \ldots, y_{CL}]$; The image transformation network ϕ_F; The segmentation network ϕ_S;

Output: S^*, F^*

For $1 \leq e \leq epoch$;
  Input palm vein grayscale images set $\mathbb{X}$ and paired binary images set $\mathbb{Y}$ to the trained ϕ_F to obtain the training samples X' and Y'. Then we obtain the augmentation dataset $\mathcal{D} = (X', Y')$; We randomly select N samples from $\mathcal{D}$ to construct mini-batches.
  For $1 \leq t_1 \leq T_1$;
    Update $S(t+1)$ according to Equaiton (6);
  end
  For $1 \leq t_2 \leq T_2$;
    Update $F(t+1)$ according to Equation (8);
  end
end

3. **Results and Analysis.** In this section, we evaluate the performance of the proposed method through experiments on three public palm vein datasets. To test the capability of the AdveinSeg algorithm, we compare it quantitatively with traditional image segmentation algorithms, as well as deep learning-based U-Net and UNet 3+ networks. We calculate the EER and plot ROC curves for comparison. All experiments were conducted using PyTorch on NVIDIA GeForce GTX 3090 GPUs.

3.1. **Datasets.** The datasets used in this study are described as follows.

Dataset A: The CASIA dataset [45] includes palm vein images from 100 volunteers, collected in two stages. Three images were captured from each hand during both stages, with the left and right palms as one class, resulting in 200 classes (100 subjects × 2 palms × 3 images × 2 sessions), totaling 1,200 palmar vein images.

Dataset B: The VERA Palm Vein dataset [46], created by the Idiap Group, contains 2,200 images from 110 volunteers across two sessions. Each volunteer's left and right palms were photographed, capturing five images per palm per session, leading to 20 images per volunteer.

Dataset C: The PolyU Multispectral Palmprint Database from Hong Kong Polytechnic University [47] consists of 6,000 palmprint images (250 subjects × 2 palms × 6 images × 2 sessions) collected using near-infrared light. Images were taken from 250 volunteers in two phases, approximately nine days apart, with each volunteer providing 24 images (2 palms × 6 images × 2 sessions), with 6 images per hand per session.

The CASIA, VERA, and PolyU datasets were selected to evaluate the proposed deep adversarial learning-based palm vein feature segmentation method under diverse conditions commonly encountered in palm vein recognition. Each dataset offers unique advantages for a comprehensive assessment of the method's performance. The CASIA dataset provides a large and diverse set of palm vein images, essential for training deep learning models and testing their generalization across subjects and imaging conditions. The VERA dataset, with its variety of images captured under different hand positions and lighting, is ideal for evaluating the robustness of the method in real-world, variable settings. The PolyU dataset includes multispectral palm vein images, which introduce challenges related to image quality and near-infrared imaging, testing the algorithm's adaptability to different imaging modalities. By using these well-established, publicly available datasets, we ensure a thorough evaluation of our method across a wide range of real-world conditions.

3.2. **Experimental setting.** To verify the effectiveness of the proposed method, each palm vein dataset is divided into two subsets. For Dataset A, which includes 200 classes, the palm vein images from the first 100 classes in the first stage are used for training, while the images from the last 100 classes in the second stage are used for testing. This results in a training set A1 with 600 images (100 classes $\times$ 6 images) and a test set A2 with 600 images (100 classes $\times$ 6 images). Similarly, Dataset B is divided into training set B1 with 1,200 images (120 classes $\times$ 10 images) and test set B2 with 1,000 images (100 classes $\times$ 10 images). For Dataset C, the training set C1 contains 4,800 images (400 classes $\times$ 12 images), while the test set C2 has 1,200 images (100 classes $\times$ 12 images).

The method alternately updates the weights of the segmentation network and the image transformation network until convergence. The final model is then used for palmar vein segmentation. In the experiments, we record the recognition accuracy of various segmentation algorithms on test set A2 from Dataset A, test set B2 from Dataset B, and test set C2 from Dataset C to evaluate the performance of the proposed method. The equal error rate (EER) is used as the performance metric.

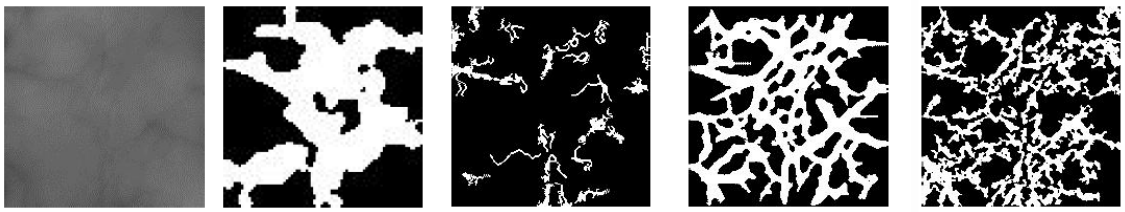
$$FAR = \frac{N_{FT}}{N_F} \tag{9}$$

$$FRR = \frac{N_{TF}}{N_T} \tag{10}$$

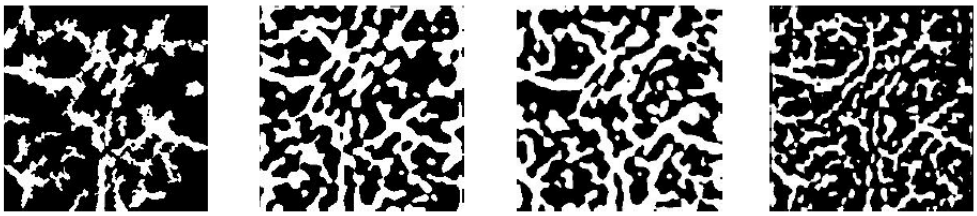
Let N_F denote the number of samples that are false, N_T denote the number of samples that are true, N_{FT} denote the number of samples that are false but mistaken for true, and N_{TF} denote the number of samples that are true but mistaken for false. The equal error rate (EER) is the point at which the false rejection rate (FRR) and false acceptance rate (FAR) are equal. A lower EER indicates better authentication accuracy for the model. The genuine acceptance rate (GAR) can be derived from the rejection rate, and different FAR and GAR values can be obtained by adjusting the threshold. ROC curves, or receiver operating characteristic curves, are used to evaluate the accuracy of predictions between two classes (X and Y). The ROC curve divides the graph into two parts based on its position. The area under the curve (AUC) represents the model's predictive accuracy: a higher AUC value indicates a larger area under the curve and, consequently, better prediction accuracy. The closer the curve is to the top left corner, the higher the prediction accuracy.

3.3. **Visual assessment.** To evaluate the effectiveness of the proposed method, we compare it with several existing segmentation techniques: Difference curvature [48], mean curvature [40], region growth [41], repeated line tracking [42], wide line detector [43], U-Net [23], and UNet 3+ [49]. Figures 4, 5, and 6 display the segmentation results for

each method applied to the CASIA, VERA, and PolyU palm vein datasets, respectively. It can be seen that the adversarial segmentation method proposed in this paper (AdveinSeg) can effectively extract palm vein features and remove noise in the background, and the extracted palm vein features have smooth and continuous characteristics. In contrast, the palm vein features extracted by differential curvature method have partial matching in relative position, but the vein details are seriously missing, some background regions are identified as palm vein features, and the segmentation results are not fine and some positions are inaccurate. The continuity of palm vein extracted based on mean curvature method is very poor, the extracted palm vein features are far from each other, lack of integrity and continuity, only part of the palm vein features are extracted at the edge, and the main features of palm vein are not extracted. Although the palm vein features

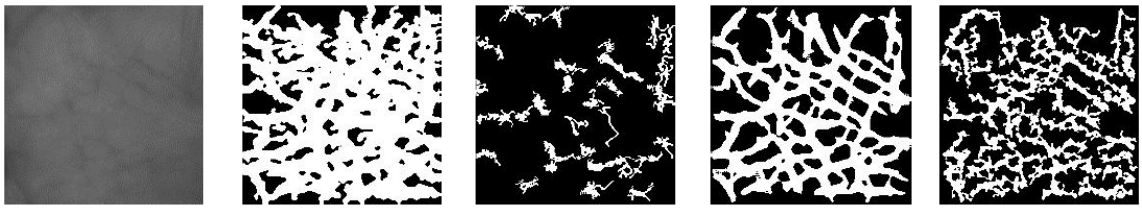


(a) Source image (b) Difference curvature (c) Mean curvature (d) Region growth (e) Repeated line tracking

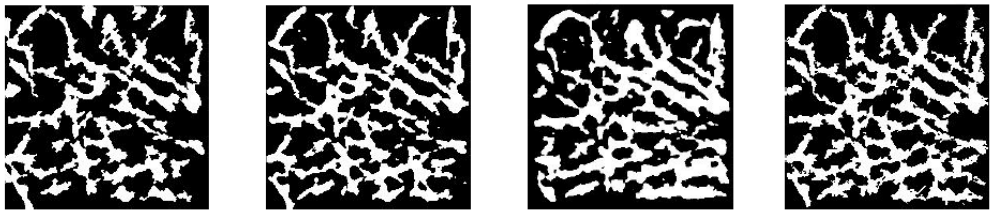


(f) Wide line detector (g) U-Net (h) UNet 3+ (i) AdveinSeg

FIGURE 4. Segmentation results of different methods on Dataset A

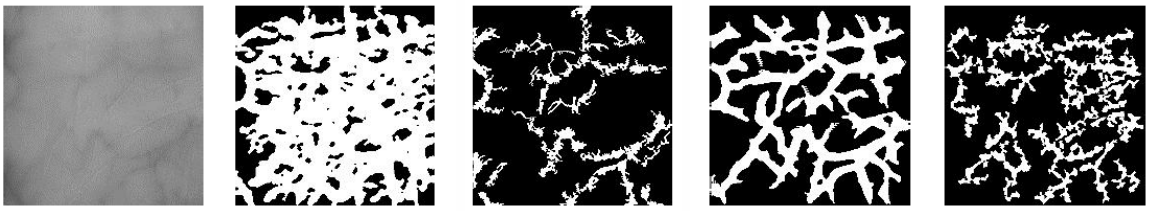


(a) Source image (b) Difference curvature (c) Mean curvature (d) Region growth (e) Repeated line tracking

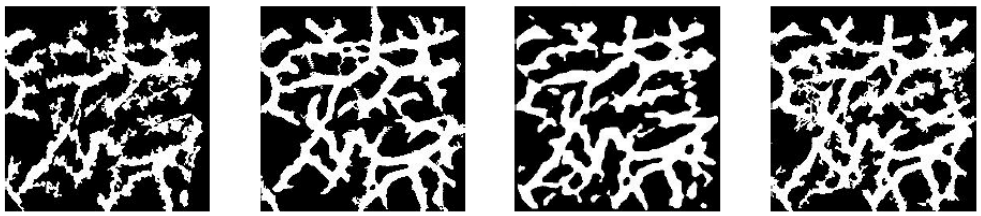


(f) Wide line detector (g) U-Net (h) UNet 3+ (i) AdveinSeg

FIGURE 5. Segmentation results of different methods on Dataset B



(a) Source image (b) Difference curvature (c) Mean curvature (d) Region growth (e) Repeated line tracking



(f) Wide line detector (g) U-Net (h) UNet 3+ (i) AdveinSeg

FIGURE 6. Segmentation results of different methods on Dataset C

extracted by region growing method are smooth, the extracted palm vein features are not fine enough, and a few non-vein parts are identified as palm veins, and the noise interference is strong. The palm vein features extracted by the repetitive linear tracking method have the initial shape of palm vein, but there are still many burrs and lack of smoothness, which will affect the recognition performance of palm vein. The palm vein feature extracted by the method based on wide line detection has the initial shape of palm vein, but there are some burrs, but compared with repeated linear tracking, there are less burrs, and the palm vein feature is not smooth and continuous, and the palm vein extracted in some parts does not form intersection points. The palm vein features extracted based on U-Net and UNet 3 + have some continuity, but lack some smoothness. The overall segmentation results are better than those of previous segmentation methods, and the main parts of the vein are correctly identified. The overall results are close to those obtained by this method. Compared with other methods, the AdveinSeg method proposed in this paper can extract better palm vein features in smoothness and continuity.

3.4. **Quantitative assessment.** In the previous section, we visually evaluated the proposed method alongside existing segmentation algorithms. In this section, we quantitatively compare and analyze the recognition performance of different segmentation algorithms. Table 2 summarizes the eight segmentation methods discussed in this paper.

TABLE 2. EER for different data sets (%)

Segmentation method	Dataset A	Dataset B	Dataset C
Difference curvature	2.5	6.07	6.55
Mean curvature	22.37	15.62	12.17
Region growth	1.68	5.81	0.91
Repeated line tracking	10.87	10.24	7.88
Wide line detector	0.67	5.84	0.77
U-Net	0.43	5.36	0.70
UNet 3+	0.41	0.77	0.68
AdveinSeg	0.33	0.59	0.60

For Dataset A, the test set consists of 600 images from 100 palm categories (100 palms × 3 images × 2 stages). When matching any two images within the same class, the matching score is counted as a true score, while matching images from different classes is counted as a false score. This results in 600 true scores (100 × 6) and 59,400 false scores (100 × 99).

For Dataset B, the test set includes 1,000 images from 100 palms. Matching scores from the same class yield 1,000 true scores (100 × 10), while matching images from different classes results in 99,000 false scores (100 × 990).

For Dataset C, the test set contains 1,200 images from 100 palms, with 1,200 true scores (100 × 12) and 118,800 false scores (100 × 990). The EER values and ROC curves can be derived from these true and false scores.

Figure 7 shows the ROC curves for the three datasets. The FRR values of the proposed method were higher than those of the other methods in the three public datasets. As shown in Table 2, the equal error rate (EER) of the proposed method is lower than that

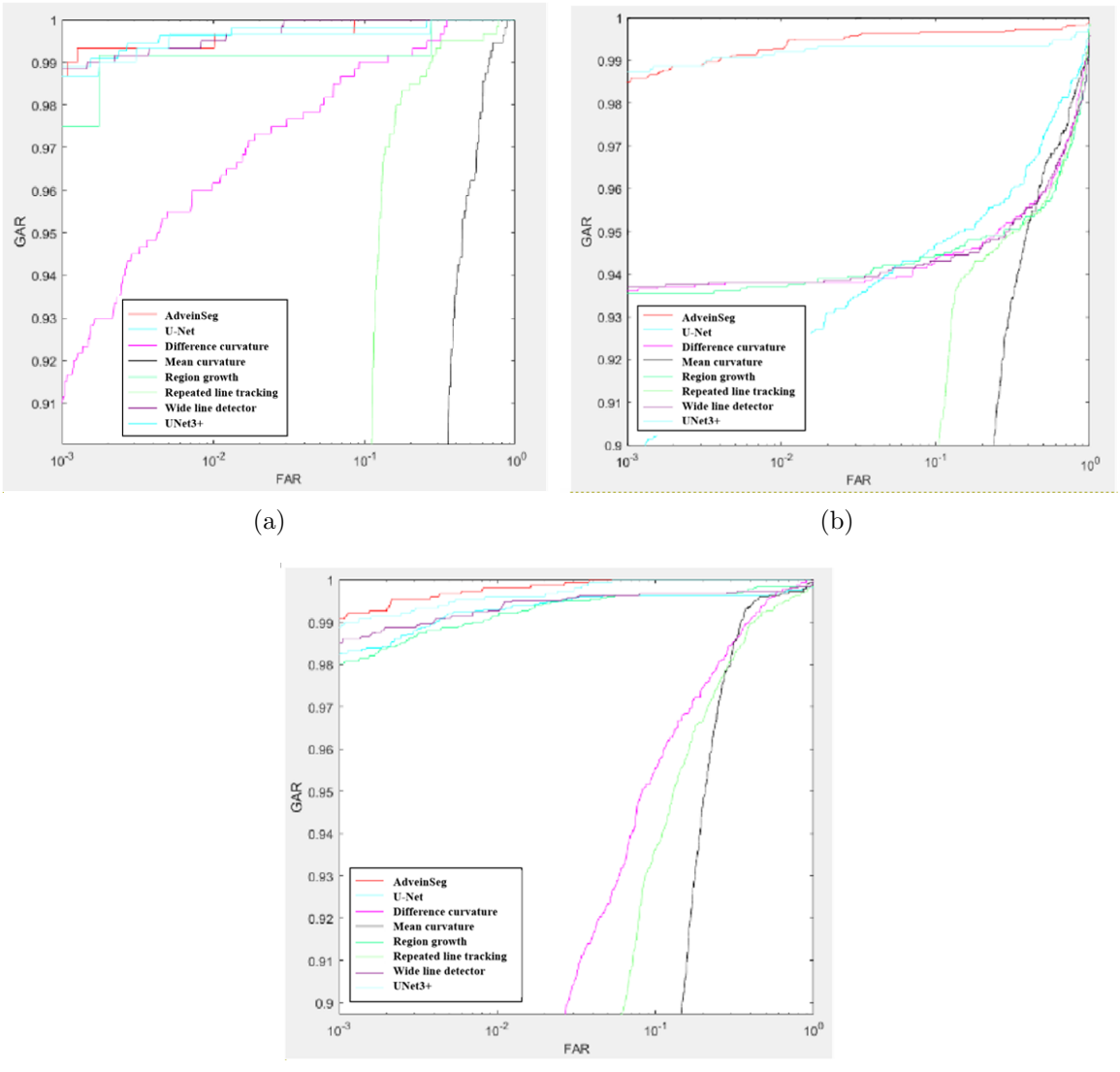


FIGURE 7. (color online) ROC curve: (a) Dataset A; (b) Dataset B; (c) Dataset C

of the comparative methods. This comparison of EER values and ROC curves across the three datasets demonstrates the effectiveness of the proposed method. The reasons for this improvement are as follows.

1) Traditional segmentation methods rely on mathematical models to assume the distribution of palm veins, leading to ineffective feature extraction, increased noise, and poor continuity in the segmentation results. In contrast, deep learning methods, while also assuming a distribution, enable neural networks to learn deep features of palm veins. Even with a small number of mislabeled samples in the training set, the model can still effectively learn palm vein features after sufficient training.

2) U-Net and UNet 3+ networks predict the probability of pixels belonging to palm vein features but often overlook the relationships between pixels, which can result in issues with continuity and smoothness of the extracted features. Adversarial segmentation does not rely on assumptions and directly utilizes distribution sampling, allowing for the extraction of deeper features. The goal of palm vein feature segmentation based on deep adversarial learning is to ensure that the segmentation model is insensitive to whether the image is from the source or target domain, allowing the abstract features extracted from the input image to remain unaffected by inter-domain differences.

4. **Conclusions.** In this study, we propose a novel palm vein feature segmentation framework, AdveinSeg, based on deep adversarial learning. The method integrates an image transformation module and a segmentation network, both built upon the U-Net architecture, and jointly trained in an adversarial manner. Specifically, the transformation network synthesizes diverse and challenging palm vein image pairs, thereby augmenting the training data and enhancing the robustness of the segmentation network. To further preserve semantic consistency between transformed images and their corresponding labels, a cosine similarity constraint is introduced. This design encourages the segmentation model to extract more discriminative and generalizable features, especially under conditions of limited annotated data. Extensive experiments on three publicly available datasets demonstrate that AdveinSeg achieves superior segmentation performance compared with a range of state-of-the-art traditional and deep learning-based approaches. The improvements are consistent across multiple evaluation metrics, confirming the effectiveness of the proposed adversarial training strategy in addressing data scarcity and improving feature representation quality.

In subsequent research, we plan to explore a unified architecture where the transformation and segmentation networks share a common encoder, reducing computational overhead while maintaining mutual benefit. Additionally, we aim to investigate the application of the proposed framework to broader biometric segmentation tasks and other real-world domains such as medical imaging and identity verification systems [50], thereby demonstrating its scalability and adaptability in diverse practical scenarios.

REFERENCES

[1] Y. Li, D. Zhao, C. Ma et al., CDRIME-MTIS: An enhanced RIME optimization-driven multi-threshold segmentation for COVID-19 X-ray images, *Computers in Biology and Medicine*, vol.169, 107838, 2024.

[2] R. Zulunov, B. Soliyev, A. Kayumov et al., Detecting mobile objects with AI using edge detection and background subtraction techniques, *E3S Web of Conferences*, vol.508, 03004, 2024.

[3] Y. Li, Z. Li, Z. Guo et al., Infrared small target detection based on adaptive region growing algorithm with iterative threshold analysis, *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[4] A. M. Ikotun, A. E. Ezugwu, L. Abualigah et al., K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data, *Information Sciences*, vol.622, pp.178-210, 2023.

[5] Z. Wang et al., Multi-threshold segmentation of breast cancer images based on improved dandelion optimization algorithm, *The Journal of Supercomputing*, vol.80, pp.3849-3874, 2024.
[6] B. S. Tchinda et al., Retinal blood vessels segmentation using classical edge detection filters and the neural network, *Informatics in Medicine Unlocked*, vol.23, 100521, 2021.
[7] N. S. M. Raja et al., Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation, *Journal of Ambient Intelligence and Humanized Computing*, pp.961-972, 2024.
[8] K. Wisaeng, Retinal blood-vessel extraction using weighted kernel fuzzy C-means clustering and dilation-based functions, *Diagnostics*, vol.13, no.3, 342, 2023.
[9] A. Kulshreshtha and A. Nagpal, Brain image segmentation using variation in structural elements of morphological operators, *International Journal of Information Technology*, vol.15, no.4, pp.2283-2291, 2023.
[10] W. Kong, J. Chen, Y. Song et al., Sobel edge detection algorithm with adaptive threshold based on improved genetic algorithm for image processing, *International Journal of Advanced Computer Science and Applications*, vol.14, no.2, 2023.
[11] A. Bhattacharya and M. Pal, A fuzzy graph theory approach to the facility location problem: A case study in the Indian banking system, *Mathematics*, vol.11, no.13, 2992, 2023.
[12] F. Tian et al., Blood vessel segmentation of fundus retinal images based on improved frangi and mathematical morphology, *Computational and Mathematical Methods in Medicine*, vol.1, 4761517, 2021.
[13] Bahadure, N. Bhaskarrao, A. K. Ray and H. P. Thethi, Comparative approach of MRI-based brain tumor segmentation and classification using genetic algorithm, *Journal of Digital Imaging*, vol.31, pp.477-489, 2018.
[14] Y. Meng et al., Graph-based region and boundary aggregation for biomedical image segmentation, *IEEE Transactions on Medical Imaging*, vol.41, no.3, pp.690-701, 2021.
[15] J. Long, E. Shelhamer and T. Darrell, Fully convolutional networks for semantic segmentation, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.3431-3440, 2015.
[16] W. Kang, H. Liu, W. Luo and F. Deng, Study of a full-view 3D finger vein verification technique, *IEEE Transactions on Information Forensics and Security*, vol.15, pp.1175-1189, 2020.
[17] T. Agrawal and P. Choudhary, COVID-SegNet: Encoder-decoder-based architecture for COVID-19 lesion segmentation in chest X-ray, *Multimedia Systems*, vol.29, no.4, pp.2111-2124, 2023.
[18] J. Yang, J. Tu, X. Zhang et al., TSE DeepLab: An efficient visual transformer for medical image segmentation, *Biomedical Signal Processing and Control*, vol.80, 104376, 2023.
[19] K. He, G. Gkioxari, P. Dollár and R. B. Girshick, Mask R-CNN, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.42, pp.386-397, 2017.
[20] Z. Huang, L. Huang, Y. Gong et al., Mask scoring R-CNN, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.6409-6418, 2019.
[21] T. Takikawa, D. Acuna, V. Jampani et al., Gated-SCNN: Gated shape CNNs for semantic segmentation, *Proc. of the IEEE/CVF International Conference on Computer Vision*, pp.5229-5238, 2019.
[22] E. Xie, W. Wang, Z. Yu et al., SegFormer: Simple and efficient design for semantic segmentation with transformers, *Advances in Neural Information Processing Systems*, vol.34, pp.12077-12090, 2021.
[23] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *arXiv Preprint*, arXiv: abs/1505.04597, 2015.
[24] S. S. Al-Amri and N. V. Kalyankar, Image segmentation by using threshold techniques, *arXiv Preprint*, arXiv: 1005.4020, 2010.
[25] R. Muthukrishnan and M. Radha, Edge detection techniques for image segmentation, *International Journal of Computer Science & Information Technology*, vol.3, no.6, 259, 2011.
[26] A. Kerkeni et al., A coronary artery segmentation method based on multiscale analysis and region growing, *Computerized Medical Imaging and Graphics*, vol.48, pp.49-61, 2016.
[27] W. Kim, A. Kanezaki and M. Tanaka, Unsupervised learning of image segmentation based on differentiable feature clustering, *IEEE Transactions on Image Processing*, vol.29, pp.8055-8068, 2020.
[28] S. K. Mamatha, H. K. Krishnappa and N. Shalini, Graph theory based segmentation of magnetic resonance images for brain tumor detection, *Pattern Recognition and Image Analysis*, vol.32, no.1, pp.153-161, 2022.
[29] M. Xiao et al., SAR image segmentation and target detection based on mathematical morphology, *2021 2nd China International SAR Symposium (CISS)*, 2021.
[30] K. Jagadeesh and A. Rajendran, Improved model for genetic algorithm-based accurate lung cancer segmentation and classification, *Computer Systems Science & Engineering*, vol.45, no.2, 2023.

[31] X. Guo et al., A novel method to model hepatic vascular network using vessel segmentation, thinning, and completion, *Medical & Biological Engineering & Computing*, vol.58, pp.709-724, 2020.
[32] Y. Wen, L. Zhang, X. Meng et al., Rethinking the transfer learning for FCN based polyp segmentation in colonoscopy, *IEEE Access*, vol.11, pp.16183-16193, 2023.
[33] S. Samudrala and C. K. Mohan, Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet, *Multimedia Tools and Applications*, vol.83, no.15, pp.46037-46063, 2024.
[34] Q. Zeng, L. Zhang, Y. Wang et al., DeepLab-Rail: Semantic segmentation network for railway scenes based on encoder-decoder structure, *Journal of Electronic Imaging*, vol.33, no.4, 043038, 2024.
[35] C. Tang, D. Chen, X. Wang et al., A fine recognition method of strawberry ripeness combining Mask R-CNN and region segmentation, *Frontiers in Plant Science*, vol.14, 1211830, 2023.
[36] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015): The 18th International Conference*, Munich, Germany, 2015.
[37] F. Marattukalam and W. H. Abdulla, Segmentation of palm vein images using U-Net, *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2020.
[38] Y. Wu et al., Transformer-based 3D U-Net for pulmonary vessel segmentation and artery-vein separation from CT images, *Medical & Biological Engineering & Computing*, vol.61, no.10, pp.2649-2663, 2023.
[39] E. H. Salazar-Jurado, R. Hernández-García, K. Vilches-Ponce et al., Towards the generation of synthetic images of palm vein patterns: A review, *Information Fusion*, vol.89, pp.66-90, 2023.
[40] W. Song, T. Kim, H. C. Kim et al., A finger-vein verification system using mean curvature, *Pattern Recognition Letters*, vol.32, no.11, pp.1541-1547, 2011.
[41] H. Qin, L. Qin and C. Yu, Region growth-based feature extraction method for finger-vein recognition, *Optical Engineering*, vol.50, no.5, 057208-057208-8, 2011.
[42] A. Kumar and Y. Zhou, Human identification using finger images, *IEEE Transactions on Image Processing*, vol.21, no.4, pp.2228-2244, 2011.
[43] B. Huang, Y. Dai, R. Li et al., Finger-vein authentication based on wide line detector and pattern normalization, *2010 20th International Conference on Pattern Recognition*, pp.1269-1272, 2010.
[44] P. Isola, J.-Y. Zhu, T. Zhou and A. A. Efros, Image-to-image transformation with conditional adversarial networks, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.5967-5976, 2017.
[45] Z. Sun, T. Tan, Y. Wang et al., Ordinal palmprint representation for personal identification, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
[46] P. Tome and S. Marcel, On the vulnerability of palm vein recognition to spoofing attacks, *2015 International Conference on Biometrics (ICB)*, pp.319-325, 2015.
[47] D. Zhang, Z. Guo, G. Lu et al., An online system of multispectral palmprint verification, *IEEE Transactions on Instrumentation and Measurement*, vol.59, no.2, pp.480-490, 2009.
[48] H. Qin et al., Finger-vein verification based on multi-features fusion, *Sensors*, vol.13, no.11, pp.15048-15067, 2013.
[49] H. Huang et al., UNet 3+: A full-scale connected UNet for medical image segmentation, *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.
[50] Z. Wang, S. Yang, H. Qin, Y. Liu and J. J. E. Wang, MixCFormer: A CNN-transformer hybrid with mixup augmentation for enhanced finger vein attack detection, *Electronics*, vol.14, 362, 2025.

Author Biography

Shuqiang Yang received B.Sc. degree from Luoyang Normal University, China, in 2000, and received M.Sc. degree from Chongqing University of Technology, China, in 2009. He is pursuing the Doctor's degree at China University of Mining and Technology, China. Now he is a professor in College of Physics and Electronic Information, Luoyang Normal University, China. His main research fields and specialties are signal acquisition processing, IoT wireless network, automotive electronics, automatic control and biometric recognition.

Zhaodi Wang received the B.S. degree in Electronic Information Science and Technology from Henan University of Technology, China, in 2011, and the M.S. degree in Circuits and Systems from Zhengzhou University, China, in 2014, and is currently a lecturer in the College of Physics and Electronic Information, Luoyang Normal University, China. Her main research areas are digital image and video processing, deep learning, and biometric recognition, and she has published more than 20 papers in journals and conferences.

Huafeng Qin received B.Sc. degree in School of Mathematics and Physics and M.Eng. degree in College of Electronic and Automation from Chongqing University of Technology, China, in 2009, and a Ph.D. degree in College of Opto-Electronic Engineering from Chongqing University, China, in 2012. Currently, he is a professor with the Chongqing Key Laboratory of Intelligent Perception and BlockChain Technology at Chongqing Technology and Business University, China. His main research areas are digital image and video processing, deep learning and biometric recognition.

Haofei Xi received the Bachelor's degree in Computer Science and Technology from Yan'an University, China, in 2020. He is pursuing the Master's degree in the Chongqing Key Laboratory of Intelligent Perception and BlockChain Technology at Chongqing Technology and Business University, China. His research interests include palm-vein identification and machine learning.

Junqiang Wang received B.Sc. degree from Luoyang Normal University, China in 2002, and received M.Sc. degree from Chongqing University, China in 2009. Currently, he is an associate professor at the College of Physics and Electronic Information at Luoyang Normal University, China. His main areas of research and expertise include Industrial Internet of Things, new energy thermal management technology, automatic control, and biometric recognition.