ITC 4/54
Information Technology and Control
Vol. 54 / No. 4/ 2025
pp. 1408-1427
DOI 10.5755/j01.itc.54.4.39124

Omni-direction Palm Vein Recognition Technology with Multi-Dimensional Image Quality Feedback Network

Received 2024/10/15 | Accepted after revision 2025/04/26

HOW TO CITE: Xu, C., Wei, H., Lin, L., Yang, G., Chu, J. (2025). Omni-direction Palm Vein Recognition Technology with Multi-Dimensional Image Quality Feedback Network. *Information Technology and Control*, 54(4), 1408-1427. https://doi.org/10.5755/j01.itc.54.4.39124

Omni-direction Palm Vein Recognition Technology with Multi-Dimensional Image Quality Feedback Network

Chen Xu, Huihui Wei, Longqing Lin, Genke Yang*, Jian Chu

Ningbo Artificial Intelligence Insititute, Shanghai Jiao Tong University and the Department of Automation, Shanghai Jiao Tong University

Corresponding author: gkyang@sjtu.edu.cn

The advantages of non-contact and embedded palm vein recognition technology in terms of security, privacy, and convenience are highlighted. Performance factors include error rate, rejection rate, and real-time processing, while engineering factors include small size, scalability, and cost-effectiveness. However, the non-contact sampling mode is susceptible to uncontrollable factors such as palm placement and posture, as well as external environmental influences like complex backgrounds and strong lighting. This paper proposes an omni-direction palm vein recognition (oPVR) method that integrates geometric and brightness multi-dimensional image quality feedback (MIQF), including integrity judgment, scale assessment and rotation angle detection. The oPVR system employs a CNN-based hand detection network for initial palm screening, reducing the false detection rate under complex backgrounds. The palm single-shot multibox detection (P-SSD) method enhances accuracy by effectively extracting regions of interest (ROI) even when fingers are close together or under varying palm angles and background interference. Additionally, the vein feature extraction network based on MobileNet improves recognition rate under weak vein conditions and across diverse demographics. The MIQF model delivers real-time quality assessment result and feedback for registration process, enhancing recognition adaptability and registration validity under various postures and lighting conditions. The accuracy, EER, FAR (@FRR=0.01) of the oPVR system under SJTU Real-Senario Palm Vein Database are 98.9%, 0.014, 0.123×10^{-3}, while the accuracy improves by 0.475%, EER decreases by 54.2%, FAR decreases by 82.11% with the implementation of the MIQF process. These enhancements effectively improve recognition performance against environmental disturbances and pose variations. Based on a real-world database and a

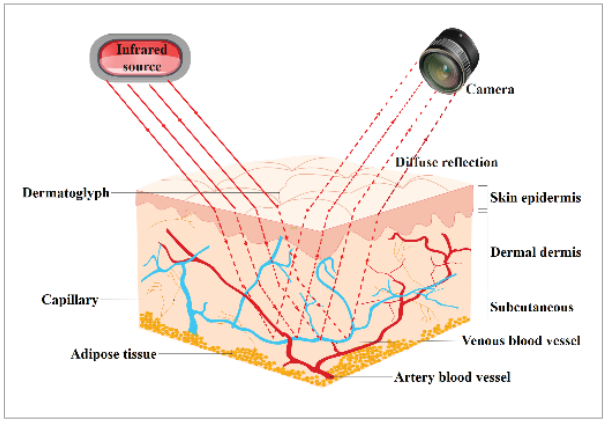
real-time demonstration system, a series of comparative tests and ablation experiments were conducted to validate the algorithm's speed and usability for large-scale applications.

KEYWORDS: Multi-dimensional image quality feedback, Omni-direction palm vein recognition, Region of interest extraction, Vein feature extraction, Reflective imaging device.

1. Introduction

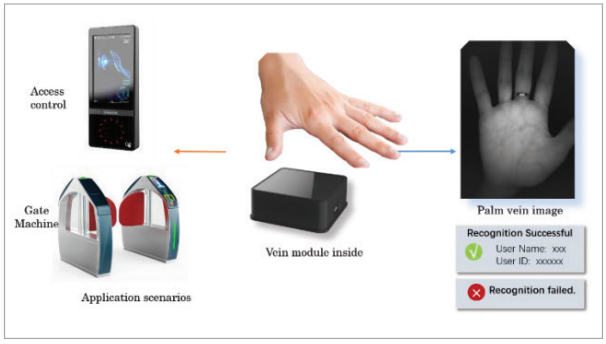
The emergence of identity authentication is based on a security awareness of individuals and social organizations. There are three basic methods of identity authentication: (1) based on information secrecy, such as passwords, (2) based on trusted objects, such as cards, (3) based on biological characteristics, such as faces. Traditional identity authentication faces issues like loss, duplication, cracking, and forgery. Surface-based biometric features like fingerprints and faces, also pose security risks of being acquired and forged, posing technical challenges in terms of authenticity and usability. On the other hand, from the perspective of mass demand, data security and privacy boundaries are becoming increasingly important [15]. Therefore, vein recognition has become an important direction for leading biometric technology to address the above pain points [38]. Among them, palm vein recognition, as a branch of vein recognition technology, has more research significance due to its advantages such as non-contact usage, larger recognition area and higher recognition accuracy [19].

Figure 1
Principle of Palm Vein Image Acquisition.



Palm vein recognition is a method of identifying individuals based on the unique characteristics of the veins in the palm of the hand. It utilizes the difference in absorption rates of deoxyhemoglobin in veins and other physiological tissues to form an image using specific wavelengths of near-infrared light to illuminate the palm [34]. As depicted in Figure 1, when the light passes through the epidermis and scatters in the subcutaneous tissue, the near-infrared light is extensively absorbed by deoxyhemoglobin in the venous blood, causing the vein patterns to appear as dark shadows on the image sensor. Palm vein recognition technology features the following characteristics: (1) being intrinsic features, difficult to forge, more secure than fingerprint and palm print recognition, (2) being short-range recognition, resistant to interference, less affected by environmental factors like backlighting and nighttime than facial recognition, (3) convenient and good personal privacy protection. It meets the needs of more real-life scenarios (see Figure 2).

Figure 2
The palm vein module and algorithm in this paper has been used in real scenarios as access control and gate machine.



Palm vein recognition systems have two modes of acquisition: contact and contactless [34]. The former requires the user to place the palm on a fixed position, while the latter allows for a convenient, hygienic and traceless usage without touching the device. Achieving omni-direction recognition in a non-contact palm vein system has practical application value [28], as it requires accurate acquisition

and identification of palm vein structure in different directions, angles and distances, thereby improving system reliability and user experience. However, due to the flexible and unrestricted usage, challenges such as incomplete palm images, fingers close together, variations in hand shapes and scales, as well as external factors like temperature, lighting, obstruction and complex backgrounds can affect vein visibility and lower recognition accuracy.

This paper proposes an omni-direction palm vein recognition (oPVR) method based on the fusion of geometric scale and brightness multi-dimensional image quality feedback (MIQF), aiming to support recognition with planes rotation of 360, pitch from -45 to 45, yaw from -45 to 45, and a distance of 5-20cm. Utilizing a large dataset of palm vein images collected in real application scenarios as the test set, the paper presents system software and hardware design solutions, developing a demonstration system for palm vein recognition with high precision, omnidirectional recognition, and high speed capabilities.

The main contributions of the proposed method include: 1) High-accuracy palm vein recognition algorithm: The oPVR mainly consists of palm single-shot multibox detection (P-SSD) algorithm, lightweight feature extraction network based on MobileNet, and MIQF algorithm with detailed experimental analysis for each part, achieving a notable FAR of 0.00022 @ FRR=0.01 on the SJTU Real-Scenario Palm Vein Database. 2) The MIQF algorithm comprehensively assesses image quality by considering integrity, scale, rotation angle, and brightness from multi-dimensional information, effectively adapting to palm postures and interference backgrounds. 3) The P-SSD model realizes omnidirectional recognition and addresses the challenges of ROI extraction accuracy under conditions of fingers being close together, palm angle deviations, and background interference. 4) Full-process palm vein collection and recognition system: Design the system to capture palm vein image in real-time, facilitating a streamlined process that includes image processing, data upload, personnel registration, feature comparison and recognition result output. 5) Low-cost and low-power reflective near-infrared collection device: The prototype of this system, based on real-world requirements, achieving a product-level standard.

The organization of this article is as follows: Section 2 outlines the research status, Section 3 introduces the hardware and software design scheme of the system, Section 4 describes the design method of the oPVR and Section 5 presents the experimental analysis results.

2. Related Works

In 1991, MacGregor and Welford [38] scientifically elaborated on palm veins as a biological feature for identification. Early vein authentication methods primarily relied on feature engineering. This approach involved extracting distinguishable features from pre-processed vein images, measuring the similarity of features in images to achieve authentication. Palm vein algorithms mainly involve ROI extraction and feature extraction. For palm ROI localization, specific areas such as the base of the fingers, fingertips, and the midpoint of the wrist are identified as ROIs [39], [6]. In 2021, Ananthi et al. [1] utilized an improved boundary rectangle strategy to extract the entire palm region, enhancing veins using Gaussian difference and histogram equalization methods to emphasize vein patterns further. Regarding feature extraction, the following methods are commonly used: 1) Template matching algorithm based on vein pattern, extracting vein line information from the original image, binarizing it to obtain template features, and calculating similarity using template convolution matching [13], [24]. 2) Feature matching method based on key points, extracting specific points such as vein endpoints, bifurcations, and corners from the original image using algorithms like SIFT, SURF, and Minutiae features [30], [37]. 3) Subspace-based image classification algorithm involves selecting a certain number of vein samples, enhancing or reducing their dimensions, training a classifier model, and using the trained model for image classification as well as recognition [26], [25].

Deep learning methods are gradually being mainstream in the research of palm vein recognition. Abdallah et al. [11] proposed a palm vein feature extraction method based on PCANet. Zhang et al. [36] introduced a palm print and palm vein recognition network called PalmRCNN based on an improved Inception_ResNet_v1. Bhilare et al. [3] proposed a

single-sensor recognition system based on a multi-scale deep pyramidal approach, significantly improving recognition accuracy and efficiency. D. Thapar et al. [29] designed a palm vein authentication neural network named PVSNet using an encoder-decoder architecture in 2019, achieving end-to-end palm vein authentication. Lefkovits et al. [16] applied four CNN models for palm vein recognition, including AlexNet, VGG-16, ResNet-50, and SqueezeNet, showing high recognition accuracy for almost every CNN-based method under PUT database [12]. Chantaf [4] utilized Inception_v3 and SmallerVggNet for palm vein recognition. In 2020, Felix [7] and Wang [32] presented palm vein feature extraction algorithm based on a U-Net network, integrating multiple segmentation methods to annotate palm vein images. Kuzu et al. [14] employed a cascaded supervised CNN trained with an unsupervised autoencoder, introducing a new dense connected convolutional autoencoder on the backbone CNN. This architecture enhanced the discriminative ability of features generated from dorsal and palm vein patterns. In 2021, Ahmad et al. [5] implemented lightweight CNN-based palm vein recognition network and Felix Olanrewaju [2] combined the Binary Statistical Image Features (BSIF) method with CNN model, achieving good results in

Table 1
Analysis of Research status on Palm Vein Recognition Algorithms.

Paper	Backbone	Database	Input size	Performance	Year
[36]	DCNN Inception-ResNet	Tongji	160×160	ACC=100% EER=2.30%	2018
[3]	DCNN LBP	CASIA	150×150	ACC=98% EER=2.61%	2018
[29]	Siamese network	PolyU	128×128	EER=0.66% CRR(Correct Recognition Rate)=98.78%	2019
[16]	CNN	PUT [37]	224×224	ACC=95%	2019
[4]	SmallerVggNet	private	96×96	ACC= 93.2%	2020
[7]	U-Net	PolyU	-	ACC=70%	2020
[32]	U-Net	CASIA	512×512	EER=0.47%	2019
[14]	Densenet	PolyU	-	EER=0.194%	2020
[5]	CNN	CASIA	160×160	EER=0.0556%	2021
[2]	CNN AlexNet	CASIA	128×128	ACC=99.83%	2021
[21]	CNN	CASIA	40×40	ACC=99.4% EER=0.0683%	2021
[22]	VGG-16	PolyU	224×224	EER=0.026%	2021
[9]	U-Net ResNet	CASIA	112×112	ACC=100% EER=0.018%	2023
[8]	SVM	PUT	-	ACC=95.50% EER=0.38%	2023
[27]	CNN	self-built	-	EER =0.0467%	2023
[13]	EfficientNet	CASIA		EER=0.0541%	2023
[24]	CNN Self-Attention Module	CASIA	-	ACC=97% EER=3.50%	2024
[18]	Lightweight CNN	CASIA	224×224	EER=1.80%	2024

a contactless palm vein database. Obayya et al. [21] proposed a palm vein recognition network architecture using Bayesian optimization. However, generating vein enhancements for each ROI image at different scales is time-consuming and impractical in real scenarios. Pan et al. [22] utilized a multi-scale deep representation aggregation to remove noisy features in pre-trained CNNs and refined feature maps using a local average thresholding method, in order to improve feature quality and alleviate limitations posed by small training database samples. In 2023, Htet [9] proposed an attention gate-enhanced segmentation model based on U-Net for extracting palm vein patterns, filtering out irrelevant and noise pixels to achieve recognition. The channel attention module enhances useful information in feature maps, and to some extent addresses issues related to rotation and scale transformation. However, the use of poorly performing initial layers in skip connections may lead to redundant low-level feature extraction. Gurunathan et al. [8] proposed an advanced algorithm based on Support Vector Machine (SVM) classifiersfor analyzing and classifying palm vein images. Meanwhile, advancements in non-contact palm vein recognition have focused on integrating self-learning features, as demonstrated by Sandhya et al.'s [27] enhanced Elman Neural Network, and Musunuru et al.'s [13] improved CNN approach. In 2024, Liao [24] proposed an innovative method for enhancing the deep learning model's ability for vein features extraction based on Gabor Attention Aggregation network.

With the rapid growth in palm vein applications requiring, the research on image quality feedback methods has become crucial. While some progress has been made in palm vein recognition based on deep learning, the accuracy of algorithms remains limited due to the small size of the general non-contact palm vein database and its low application rate. Palm vein recognition commonly faces quality issues during data collection and authentication phases. Currently, there are few studies on image quality feedback for palm vein systems. In 2015, Peng [23] proposed an evaluation method for palm vein image quality using 2D image entropy and local 2D entropy, accelerated by FPGA platform. The capturing device controls NIR light source intensity based on image quality to adjust device parameters. If the quality of the obtained image does not meet requirements, the device recaptures the image until a high-quality image is obtained. In 2017, Wang [31] introduced an algorithm that evaluates palm vein image quality based on the tilt angle of the hand and the fusion of two standards, namely clarity and brightness uniformity. In 2024, Luo et al. [18] proposed a new method focusing on recognition accuracy under unconstrained and weak-cooperative conditions, advancing further developments in the field. To facilitate summarization and referencing, Table 1 summarizes these main research findings and presents key data such as algorithm framework and recognition results. This paper aims to solve the interference issues such as palm posture and environmental impacts on the palm vein recognition system through image quality feedback method, thereby improving the accuracy and user experience of the system.

3. Hardware and Software Solution

3.1. Equipment and System Architecture

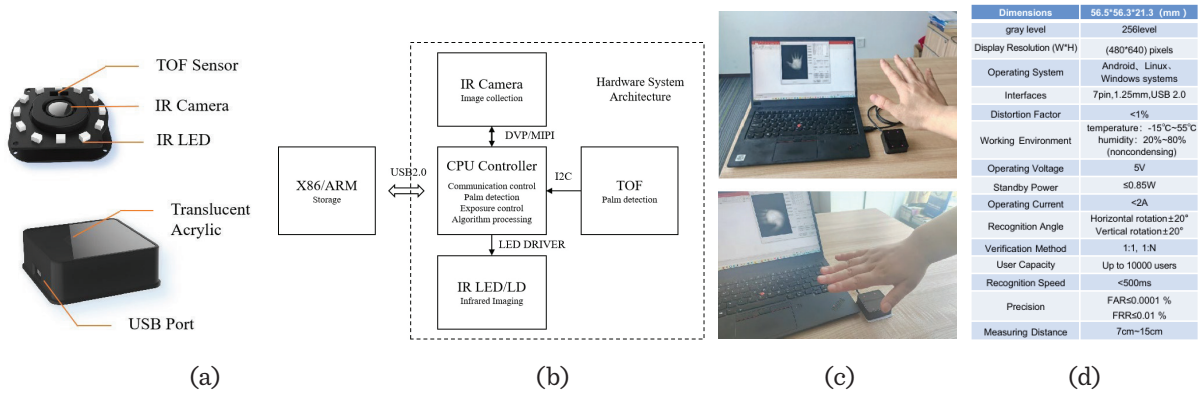
Develop a high-precision palm vein recognition device, including sensing and interaction modules, optical imaging modules, as well as computing and communication modules [8]. The device adopts a reflective imaging structure, including core components such as an infrared filter, light source driver, and main control circuit, as shown in Figure 3. The infrared filter only allows specific wavelengths of near-infrared light to pass through, filtering out ambient light to enhance imaging contrast. The light source driver controls the brightness of the near-infrared light source. The main control circuit is responsible for controlling light intensity, processing image acquisition and transmission, with an image resolution of no less than 1 million pixels. The low power consumption feature is mainly achieved by adding a Time-of-Flight (TOF) sensor to detect object distance in real time. When the object distance is between 5-20cm, the light source and camera are activated to capture images. The supplementary lighting module mainly consists of 12 infrared LEDs arranged in a circular array, which can be uniformly illuminated in intervals of 2, 3, 4, 6, and 12 LEDs based on ambient light intensity. This design reduces power consumption and significantly saves computational time. Also, the equipment connects to a

host computer equipped with an i5 processor and 4GB of RAM. Additionally, storage requirements should be based on the data volume, which is approximately 2KB per feature. For the computer, the storage space is sufficient.

To evaluate the performance of the palm vein recognition device under various environmental conditions, the experiment was conducted using the same device and tester, with palm placed naturally above the device according to standard requirements. The

Figure 3

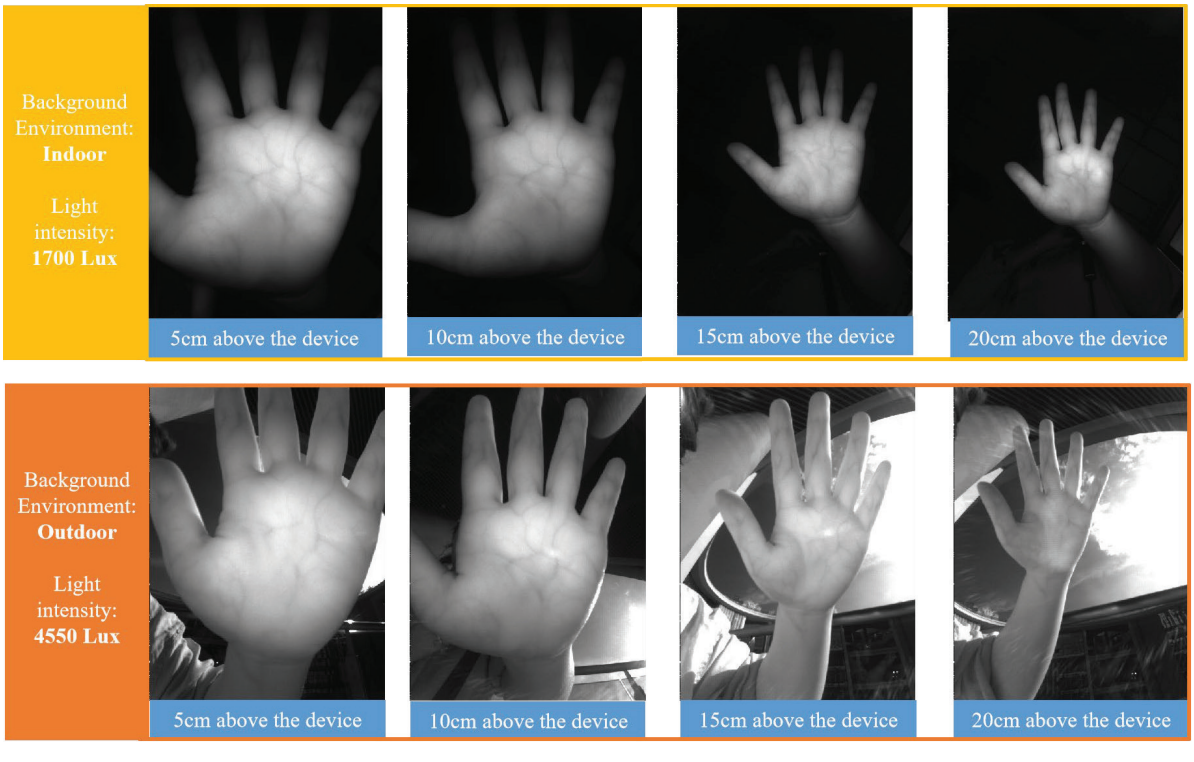
Hardware design: (a) Hardware device, (b) Device usage, (c) Hardware system architecture, and (d) Specifications.



Dimensions	56.5*56.3*21.3 (mm)
gray level	256level
Display Resolution (W*H)	(480*640) pixels
Operating System	Android、Linux、Windows systems
Interfaces	7pin,1.25mm,USB 2.0
Distortion Factor	<1%
Working Environment	temperature: -15℃~55℃ humidity: 20%~80% (noncondensing)
Operating Voltage	5V
Standby Power	≤0.85W
Operating Current	<2A
Recognition Angle	Horizontal rotation±20° Vertical rotation±20°
Verification Method	1:1, 1:N
User Capacity	Up to 10000 users
Recognition Speed	<500ms
Precision	FAR≤0.0001 % FRR≤0.01 %
Measuring Distance	7cm~15cm

Figure 4

Comparison of palm vein samples collected by the same tester at different light intensities and heights.



output image quality was evaluated to verify the parameters of imaging exposure algorithm. The experiment was conducted indoors and outdoors, with luminous flux measurements taken using a detector. The device's performance was also tested at different distances (5cm, 10cm, 15cm, and 20cm) above the device. The results, shown in Figure 4, demonstrate the device's ability to capture high-quality palm vein images in different environmental lighting conditions and at various distances. This will provide valuable insights into its potential applications in real-world scenarios.

Figure 5
Test software interface diagram.

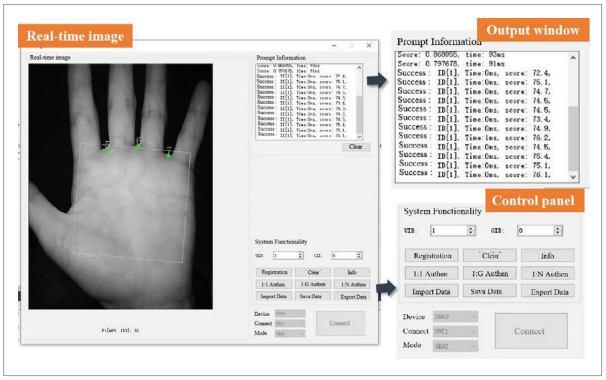
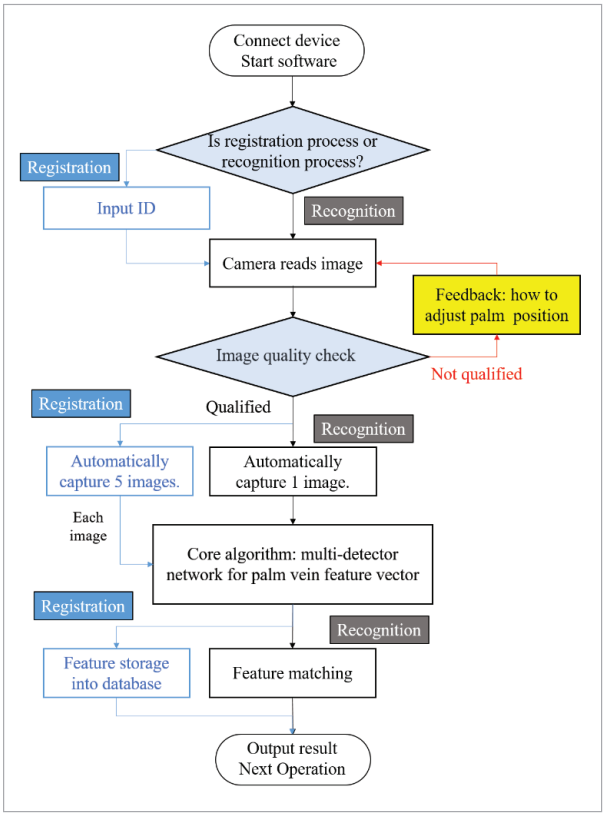


Figure 6
System flow.



3.2. Software Framework Process

Design an interactive software based on the QT framework with image acquisition, registration, and recognition functions. The software can real-time display the collected palm vein images, recognition time, score, and other information, that interface is shown in Figure 5. When in use, the user should naturally place their palm parallel to the outer surface of the module, and an optimal recognition distance of 5 to 20cm. The registration and recognition process, as shown in Figure 6, includes important steps such as image quality check, and feature detection network.

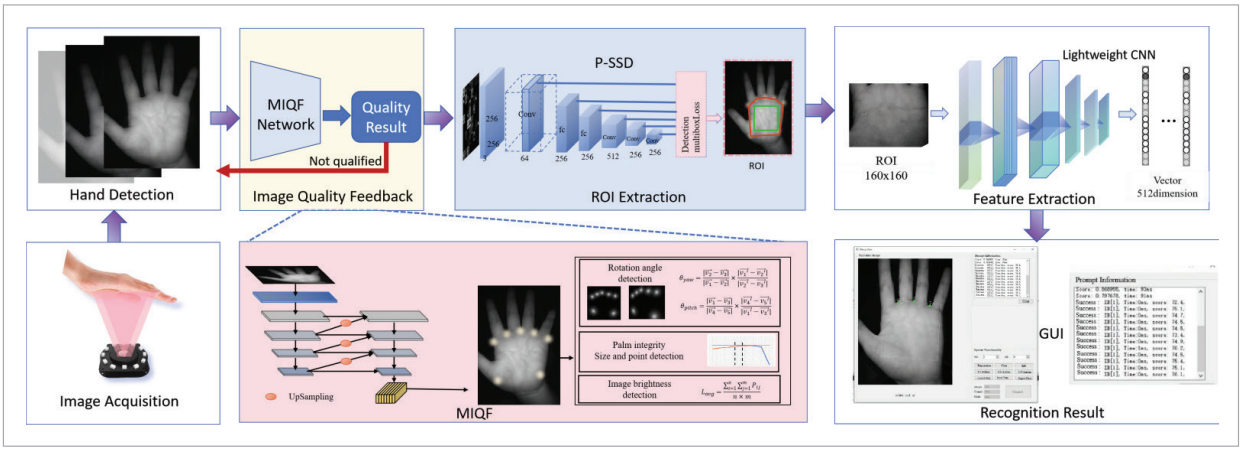
4. Omni-direction Palm Vein Recognition Algorithm (oPVR)

4.1. Palm Vein Detection System Structure

Supported by the hardware acceleration with NPU chips, this oPVR system can be used in embedded scenarios, which utilizes a combination of multiple deep networks separately for hand detection, image quality feedback, ROI extraction and feature extraction algorithms, see in Figure 7.

The hand detection network confirms the presence of palms in the scene, delineates a more precise set of objects to be recognized before the identification phase to enhance detection speed. In the image quality feedback loop, MIQF is utilized for integrity, scale, rotation and brightness assessment based on Heatmap network, effectively enhancing recognition adaptability and registration validity. The palm single-shot multibox detection (P-SSD) is used for ROI extraction to locate palm areas and extract key points, followed by cropping and normalization. P-SSD model addresses the challenges of ROI extraction accuracy under conditions of fingers being close together, palm angle deviations, and background interference. In feature extraction, the lightweight CNN is selected based on

Figure 7
Omni-direction Palm Vein Recognition Algorithm Structure.



the characteristics of sample data for large-scale data training. The parameters of multi-layer networks are adjusted, with ROI images as input and 512-dimensional latent feature vectors as output. Various experiments are conducted including loss function evaluation and resolution testing. Experiment shows that the vein feature extraction network based on MobileNet enhances the recognition rate under weak vein conditions and across different demographics. Feature matching is performed by evaluating the similarity of two feature vectors using Euclidean distance.

4.2. Multi-dimentional Image Quality Feedback (MIQF) Algorithm

The core of MIQF algorithm is image quality assessment. In practical system applications, issues such as incomplete palms, large-scale deformations, rotational offsets and brightness differences in images have a significant impact on recognition accuracy. Therefore, a general quality assessment model for palm infrared images is proposed—the 'quadruple' model. The quadruple consists of **integrity**, **scale**, **rotation** and **brightness.** Thus, the problem of image quality assessment can be simplified into a quadruple, represented as:

$$\text{Assess}=\{\text{I,S,R,B}\} \quad (1)$$

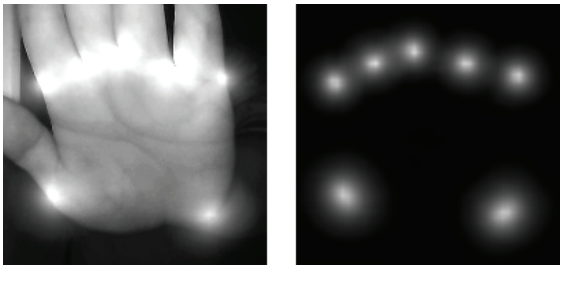
In MIQF, the Heatmap network is crucial as it utilizes image's heatmap to acquire the key points of the palm, an example is shown in Figure 8. The response values of each point in the heatmap, along with the area formed by the points and the angular scale variation of the points, are computed to assess changes in geometric features.

We introduce a residual connection to better learn features in Heatmap, and the obtained features F_n can be formulated as:

$$F_n = Conv(Z_n) + Conv(UpSampling(Z_{n-1})) \quad (2)$$

Z_n represents the feature maps output at different stages of the backbone network where n denotes different stages. F_n refers to the multi-scale feature maps obtained after the top-down path and lateral connections fusion in the Feature Pyramid Network (FPN) while n represents different scales.

Figure 8
Palm heatmap and key-point detection.



For the quadruple {I, S, R, B}, it is necessary to first obtain the quality thresholds θ_I, scale threshold $[\theta_s,\theta_{roi}]$, rotation threshold $[\bar{\theta}_{yaw},\bar{\theta}_{pitch}]$, brightness threshold θ_L}. Based on this, quality assessment can be performed on the input image, and compared with the thresholds to make judgments. The following sections will introduce the methods for solving the quadruple thresholds and assessing input quality, along with the experimental results in 5.2.2.

Integrity detection. The original image is input into the network, and the output is key landmark points. If these points have response values in the heatmap that exceed a certain threshold, they are considered as complete palm. To determine the threshold θ_I, we can use the following formula:

$$\bar{\theta}_I = argmax\left(\left(\frac{KP}{K} - \frac{KN}{N}\right)|\theta_I\right), \tag{3}$$

where $\bar{\theta}_I$ is the optimal choice for θ_I. KP and KN represent the pass numbers for complete and incomplete palm detection across different threshold θ_I values, while K and N denote the total numbers of complete and incomplete palms, respectively.

$$y|\bar{\theta}_I = \sigma(ReLU(w_1 F_n + b_1)w_2 + b_2). \tag{4}$$

Here, F_n is the input, while F_n is the output of the heatmap network. y represents the integrity of the image, with a value range of [0, 1]. w_1, w_2 are the weight matrices, and b_1, b_2 are the bias terms.

Scale detection. Based on the number and position of feature points from the heatmap, the scale and effective area of the palm can be judged, serving as one of the quality evaluation indicators.

$$P = \frac{TP}{T}, R_s = \frac{A}{S}, \mathrm{P}_{roi} = \frac{T\mathrm{P}_{roi}}{T} \tag{5}$$

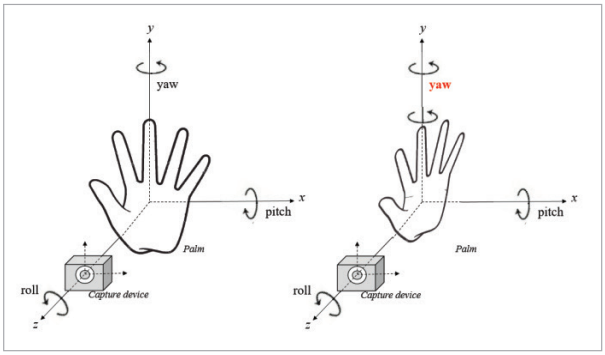
$$\mathrm{Pass} = \begin{cases} Yes & if\ R_s > \theta_s\ and\ P = 1 \\ Yes & if\ R > \theta_s\ and\ P < 1\ and\ \mathrm{P}_{roi} > \theta_{roi}, \\ No & otherwise \end{cases} \tag{6}$$

where P represents the extraction rate of feature points, R_s represents the ratio of the area formed by feature points to the total area, P_{roi} represents the probability of obtaining palm ROI, θ, θ_{roi} represents the threshold for R_s and P_{roi}, TP represents the number of extracted feature points, T represents the total number of feature points, A represents the area formed by feature points, S represents the total area, TP_{roi} represents the number of obtained points of the palm ROI, and T_{roi} represents the total number.

Rotation detection. Using the heatmap model to extract seven key points on the palm, calculation of the rotation angle is based on five of these points. The algorithm is capable of adapting to roll direction rotation, achieving 360-degree omnidirectional recognition. Detecting pitch and yaw rotation angles is beneficial for the algorithm to perform rotation correction and quality assessment. Rotation angles are defined as shown in Figure 9.

Figure 9

Rotation angle explanation.

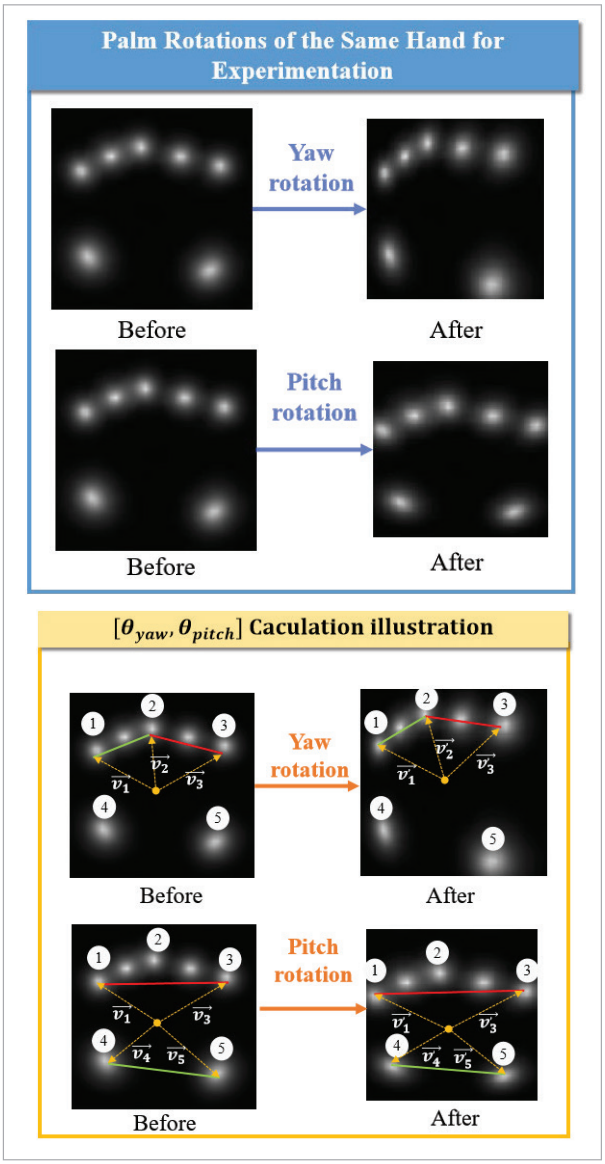


When the palm rotates in the yaw direction as shown in Figure 10, due to perspective transformation, it can be observed that the distance between point 2 and point 3 on the image will change compared to the distance between point 1 and point 2. Similarly, when the palm undergoes pitch rotation, the distance from point 1 to point 3 divided by the distance from point 4 to point 5 on the image will change. $\theta_{yaw},\theta_{pitch}$ can be used to determine the magnitude of yaw and pitch rotation separately. $\bar{\theta}_{yaw},\bar{\theta}_{pitch}$ are the rotation thresholds determined using the mean values of the experimental dataset, while N_i and N_j represent the number of samples in the dataset.

Here, $\overrightarrow{v_1}$, $\overrightarrow{v_2}$, $\overrightarrow{v_3}$, $\overrightarrow{v_4}$ and $\overrightarrow{v_5}$ represent the vectors from the center to points 1, 2, 3, 4 and 5, respectively. Using the following formulas, the values of θ_{yaw} and θ_{pitch} can be calculated.

$$\theta_{yaw} = \frac{|\overrightarrow{v_2} - \overrightarrow{v_3}|}{|\overrightarrow{v_1} - \overrightarrow{v_2}|} \times \frac{|\overrightarrow{v_1'} - \overrightarrow{v_2'}|}{|\overrightarrow{v_2'} - \overrightarrow{v_3'}|} \tag{7}$$

Figure 10
Palm rotation calculation illustration.



$$\theta_{pitch} = \frac{|\overrightarrow{v_1} - \overrightarrow{v_3}|}{|\overrightarrow{v_4} - \overrightarrow{v_5}|} \times \frac{|\overrightarrow{v_4'} - \overrightarrow{v_5'}|}{|\overrightarrow{v_1'} - \overrightarrow{v_3'}|} \tag{8}$$

$$\bar{\theta}_{yaw} = \frac{1}{N_i} \sum_{i=1}^{N_i} \theta^i{}_{yaw} \, (i = 1,2,\ldots,N_i) \tag{9}$$

$$\bar{\theta}_{pitch} = \frac{1}{N_j} \sum_{j=1}^{N_J} \theta^j{}_{pitch} (j = 1,2,\ldots,N_j) \tag{10}$$

Brightness detection. The formula for calculating average brightness is as follows, where n and m represent the height and width of the image, P_{ij} denotes the grayscale value of the pixel (x_i,y_i) at the coordinates in the image. L_{avg} represents the average brightness and is compared with the brightness threshold θ_L. If the average brightness is lower than threshold, the system will provide feedback to the light source module. In this case, the lighting can be supplemented by increasing the light intensity or the number of lights.

$$L_{avg} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij}}{n \times m} \tag{11}$$

The four-tuple image quality evaluation metrics can be employed individually or collectively to assess the quality of the input image. Based on the evaluation results, the input image can either be processed further through ROI extraction and feature extraction or be returned to the input acquisition device for re-acquisition.

4.3. Palm ROI SSD Model(P-SSD)

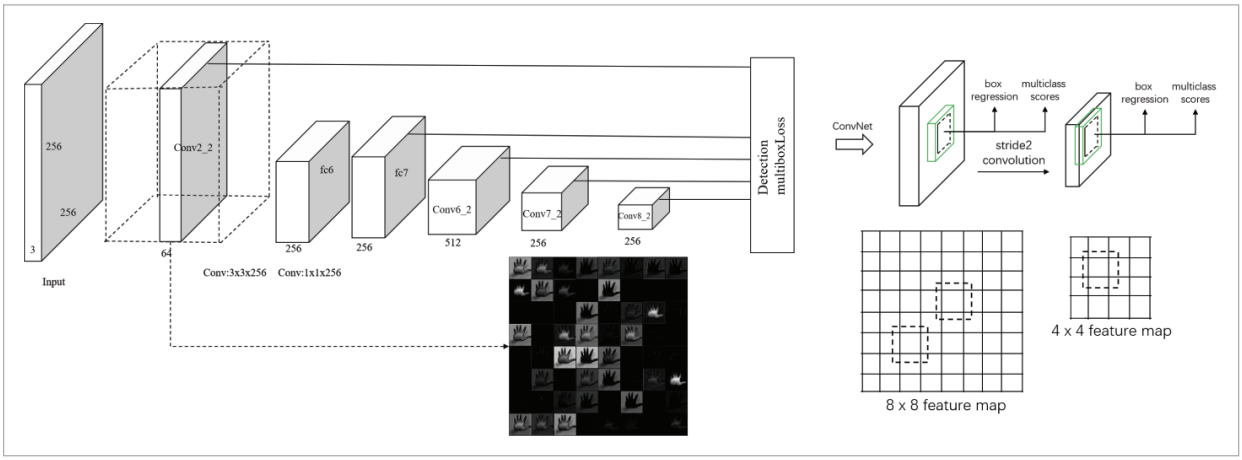
The palm image undergoes non-rigid transformations such as deformation and rotation. A method is proposed based on a single-shot multibox detector (SSD) to locate arbitrary quadrilateral boxes, allowing for the customization of additional key points in the network, as shown in Figure 11.

The SSD model is an advanced end-to-end deep learning framework for object detection, utilizing anchor box regression and classification to detect target objects across different scales. This system utilizes improved SSD model that extends from rectangles to locating arbitrary quadrilaterals while allowing for the addition of custom key points, such as connection points between two fingers. This approach enhances the localization performance in complex scenarios and gestures, providing more accurate detection results of hand regions. A 256x256x3 input image undergoes convolution, pooling, and activation to obtain feature maps. By selecting feature maps from different layers of the network with varying sizes, objects of different scales can be detected.

Assuming there are K anchor boxes at a certain position on the feature map. For each anchor box, C class scores are predicted, along with the position parameters (x,y,w,h) of the predicted box p. There-

Figure 11

Improved SSD model structure. The input image size is 256x256x3. View the results of a specific convolutional layer conv2_2.



fore, there will be $(C+4)\times K$ prediction values at each position. The final predicted position (x_i,y_i,w_i,h_i) is determined by selecting the predicted box p with the highest confidence score $\hat{c}^p$ at each position, where c_j represents the classification score, and the prediction values are (c_j,w_i,h_i,x_i,y_i), with i from 1 to K and j from 1 to C.

The SSD loss function is defined as shown in Formula 12, where the confidence loss $L_{conf}(x,c)$ is the softmax loss with c as the class, and the localization loss $L_{loc}(x,l,g)$ is the Smooth L1 loss. The difference l between the predicted box p and the anchor box d approaches the difference g between the ground truth box g and the anchor box d. The anchor box is composed of four parameters (x,y,w,h) representing the coordinates and dimensions of the target box. N is the number of positive samples based on anchor boxes, i is the count of positive sample anchor boxes (predicted boxes), j is the count of ground truth boxes, and k is the class.

$$L(x,c,l,g)=\frac{1}{N}\left(L_{conf}(x,c)+\alpha L_{loc}(x,l,g)\right) \quad (12)$$

$$L_{conf}(x,c)=-\sum_{i\in Pos}^{N} x_{ij}^{p}\, log\left(\hat{c}_i^p\right)-\sum_{i\in Neg}\log(\hat{c}_i^0) \quad (13)$$

$$L_{loc}(x,l,g)=\sum_{i\in Pos}^{N}\sum_{m\in\{cx,cy,w,h\}} x_{ij}^{k}\,\mathrm{smooth}_{L1}\left(l_i^m-\hat{g}_j^m\right) \quad (14)$$

$$\hat{c}_i^p=\frac{\exp\left(c_i^p\right)}{\sum_p \exp\left(c_i^p\right)} \quad \hat{g}_j^{cx}=\frac{g_j^{cx}-d_i^{cx}}{d_i^w},\hat{g}_j^{cy}=\frac{g_j^{cy}-d_i^{cy}}{d_i^h} \quad (15)$$

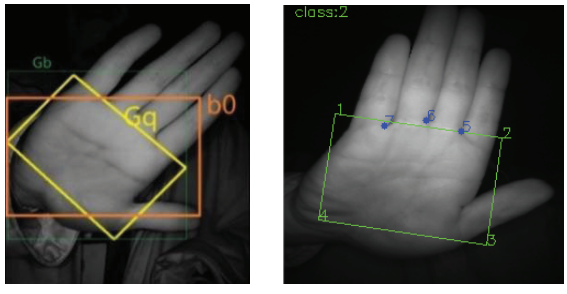
$$\hat{g}_j^w=\log\left(\frac{g_j^w}{d_j^w}\right),\hat{g}_j^h=\log\left(\frac{g_j^h}{d_j^h}\right) \quad (16)$$

$$\mathrm{smooth}_{L1}(x)=\begin{cases}0.5x^2 & if|x|<1\\ |x|-0.5 & otherwise\end{cases} \quad (17)$$

The SSD predicts rectangular boxes, which may include unnecessary background areas in cases of rotated objects. The coordinates (x,y,w,h) lack practical meaning as shown in Figure 12(a). To address this, an enhanced SSD model is proposed to locate

Figure 12

(a) Anchor illustration diagram (b) P-SSD add key-point: G_{q5},G_{q6},G_{q7}.



(a) (b)

arbitrary quadrilateral boxes using point coordinates $(x_{01},y_{01},x_{02},y_{02},x_{03},y_{03},x_{04},y_{04})$ and dimensions (w_0,h_0) to represent anchor box d_0. Here, $d_{0m}=(x_{0m},y_{0m})$ denotes the coordinates of a vertex of the quadrilateral box, where m = {1, 2, 3, 4}. The actual values are represented by the bounding rectangle Gband and four point coordinates G_q, resulting in (x_0,y_0,w_0,h_0).

Assuming single-object single-category recognition, the modified localization loss is described as follows:

$$L_{loc}(l,g)=\sum_{i\in Pos}^{N}\sum_{m\in\{1,2,3,4\}}\text{smooth}_{L1}\left(l_i^m-g^m\right) \quad (18)$$

$$g^m=\frac{Gq^m-d_i^{0m}}{w_i^0},\ l^m=\frac{p^m-d_i^{0m}}{h_i^0} \quad (19)$$

The predicted box $p=(x_1,y_1,x_2,y_2,x_3,y_3,x_4,y_4)$ can be represented as follows:

$$x_m=x_{0m}+w_0\times\Delta x_m \quad (20)$$

$$y_m=y_{0m}+h_0\times\Delta y_m(m=\{1,2,3,4\}) \quad (21)$$

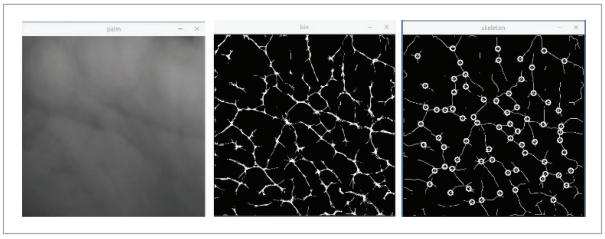
Here, (x_m, y_m) represents the coordinates of each vertex of the predicted box p, (x_{0m}, y_{0m}) represents the coordinates of the vertex of anchor box d_0, w_0 and h_0 represent the width and height of anchor box d_0, and $(\Delta x_m, \Delta y_m)$ represent the offsets of each corner of the box.

Based on a multi-task fusion object detection network, the number of key points can be increased. Therefore, to add key points detection required for subsequent tasks in object detection and classification, key points G_{q5},G_{q6},G_{q7} are added as shown in Figure 12(b), with a reference point selected as the center point of anchor box (x_0, y_0). Hence, the range of index m is from 1 to 7, where $d_{05}=d_{06}=d_{07}=(x_0,y_0)$.

4.4. Feature Extraction Network

Currently, most palm vein recognition algorithms adopt feature point matching methods such as SIFT and SURF features, with feature sizes typically around a few kilobytes. A single palm contains hundreds of feature points (see Figure 13), leading to a high computational load in the feature matching process. Additionally, traditional algorithms tend to focus on expressing detailed information, lacking in representing macro information. This deficiency results in increased rejection rates when there are significant changes in the collected image data due to palm movement, rotation, or deformation. By utilizing deep learning for feature extraction, it is possible to abstract features more effectively, encompassing a broader range of macro information, thus providing better responses to complex scenarios and different hand gestures.

Figure 13
Feature map and key points.



The recognition experiment mainly involves designing a feature extraction network to obtain the rules for extracting image features from the training set, thus achieving better feature extraction results in the test set. Based on the pre-existing ROI extraction network, the palm ROI image is input, and the feature extraction network outputs the feature vector of the image.

Table 2
Net body architecture.

Type	Input size	operator	t	c	n
conv1	160x160x3	conv3x3	-	64	1
conv2_dw	80x80x64	depth wise conv3x3	-	64	1
conv_23	80x80x64	bottleneck	2	64	1
conv_3	40x40x64	bottleneck	2	128	4
conv_34	40x40x64	bottleneck	4	128	1
conv_4	20x20x128	bottleneck	2	128	6
conv_45	20x20x128	bottleneck	4	128	1
conv_5	10x10x128	bottleneck	2	128	2
conv_56	10x10x128	bottleneck	4	128	1
conv_6	5x5x128	bottleneck	2	128	2
conv_7	5x5x128	conv1x1	-	512	1
conv_linear	5x5x128	Linear GD-Conv5x5	-	512	1

The model structure, parameter adjustments, and training strategies are designed based on the characteristics of palm images under near-infrared light. Considering the requirements of real application scenarios, a model based on the lightweight MobileNet backbone network is designed, which significantly reduces model size and computational complexity while maintaining high accuracy. Net body architecture and loss functions are listed in Tables 2-3.

Table 3
Loss function for test and its meaning.

Loss function	Formula
Center loss	$L_C = \frac{1}{2}\sum_{i=1}^{m}\left\|x_i - c_{y_i}\right\|_2^2$
Effectively describe the intra-class variations	
LMCL	$L_{lmc} = \frac{1}{N}\sum_i -\log\frac{e^{s(\cos(\theta_{y_i,i})-m)}}{e^{s(\cos(\theta_{y_i,i})-m)} + \sum_{j\neq y_i} e^{s\cos(\theta_{j,i})}}$
Maximize between-class variance and minimize within-class variance	
AAML	$\overline{L_{aam}} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))} + \sum_{j=1,j\neq y_i}^{n} e^{s\cos\theta_j}}$
Improve inter-class separability while enhancing intra-class compactness and inter-class differences	

5. Experimental Analysis

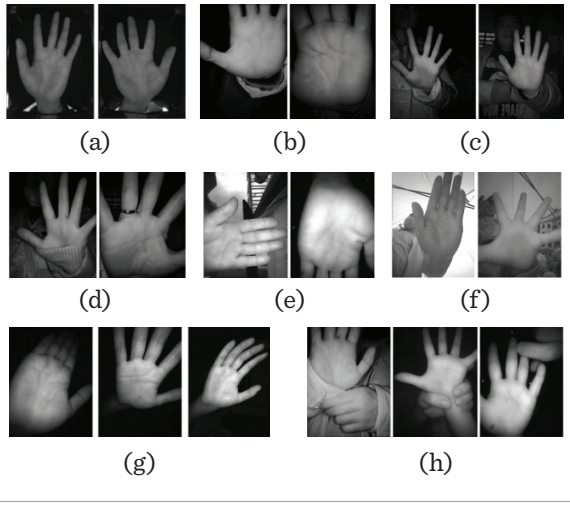
5.1. Database

Due to the insufficiency of palm vein sample data in publicly available datasets like The CASIA Multi-Spectral Palmprint Image Database, which contains 2400 palm vein sample images (acquired with samples under 850nm and 940nm light sources), it fails to meet the industry's needs for real-world applications. Therefore, we've established a new dataset, named SJTU Real-Senario Palm Vein Image Database, covering samples from multiple ethnicities, various scenarios, and diverse terminals.

Whether it is the ROI extraction network or the feature extraction network, it is crucial to include a wide range of sample data from different environments and poses during the training process to enhance the system's adaptability. The experiment data consists of two main parts. The first part is from data collection, where the hardware device is used in different scenarios (such as indoor, outdoor, with or without background). Users are required to place their palms naturally on the device in multiple captures, moving within a range of 5-20cm, rotating 360°, and tilting the palm ±45°. The second part is from practical applications, such as data from the access control device formed by the hardware module. SJTU Real-Senario Palm Vein Image Database has a total of 300,000 original training data were collected, augmented to 1 million during the training process. The first test set comprises 200 IDs, with each ID having 100 captures, totaling 20,000 images for recognition experiments. The second test set consists of 12,000 palm IDs, with 5 captures per ID, totaling 60,000 images for quality feedback experiments.

Based on the self-developed hardware device, scenario verification was conducted, resulting in data categorized into seven types, with examples provided for each type.

Figure 14
Types of collected samples: (a) Single-colored background, appropriate size, clear palm, (b) Partial hand, complete palm, clear image, (c) Large movable range, clear image, (d) Obstruction or accessories, (e) Different angles, (f) Complex background: including but not limited to environmental brightness, external objects, self-interference, etc., (g) Different postures: fingers clenched, bent; palm tilted, deformed, etc., and (h) Multiple hands coexist.



Under the test data set, recognition accuracy and acceptance rate indicators are obtained through experiments. Accuracy indicates the probability of correct identification in all comparisons, where matches are correctly identified as matches and non-matches as non-matches. FRR (False Rejection Rate) is the probability that the system mistakenly rejects a legitimate user, while FAR (False Acceptance Rate) is the probability that the system mistakenly accepts an unauthorized user. EER (E qual Error Rate) is the error rate at the point where FAR and FRR are equal. Validation Rate is the proportion of legitimate users that the system correctly verifies and accepts.

5.2. Experiment Analysis

5.2.1. Performance of P-SSD

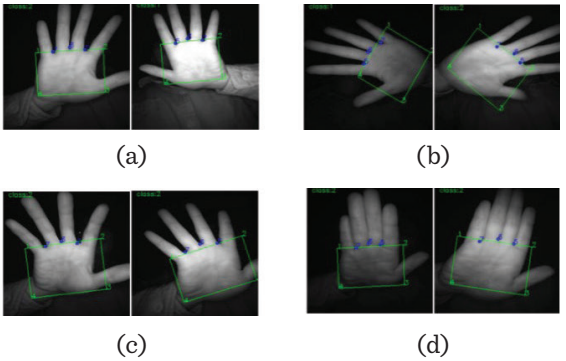
To verify the performance of P-SSD algorithm, experiments were conducted to test the effectiveness of ROI localization and key-point positioning. The evaluation metric Intersection over Union (IoU) was used to compare predicted boxes with ground truth boxes, measuring the degree of overlap between them. During the training of the ROI model, a confidence threshold of 0.98 was set, ensuring that only output bounding boxes with prediction scores higher than this threshold of 0.79 would be further processed. This helps to ensure high-confidence predictions and reduce the chances of false positives. The measured localization error of this system is 0.21, indicating the average distance between predicted bounding boxes and ground truth boxes. A lower localization error implies higher accuracy in locating regions of interest. The difference between the Test1 and Test2 datasets lies in the orientation of the palm, with Test1 samples having orientations ranging approximately from -90° to 90°, while Test2 samples have orientations ranging from -45° to 45°. For Test1-1 and Test2-1, the confidence threshold is set at 0.98, while for Test1-2 and Test2-2, the confidence threshold is set at 0.79.

Table 4
Experiment results of recognition rate.

Dataset	Left hand	Right hand	Average rate
Test1-1	0.75	0.88	0.82
Test1-2	0.79	0.95	0.87
Test2-1	0.83	0.94	0.89
Test2-2	0.87	0.98	0.92

The experimental results demonstrate the algorithm's ability to accurately locate ROI and extract key points. Even in cases of large angular deviation or deformation of the palm, precise positioning can still be achieved. Additionally, it can detect key points of fingertips even when fingers are closely aligned.

Figure 15
Recognition output. Class 1 means left hand, and class 2 means right hand. ROI area displayed with green anchors, three key points G_{q5},G_{q6},G_{q7} were also identified and marked. (a) Output graphs of appropriate palm, (b) Output graphs of palm in different angles, (c) Output graphs of palm with large angle deviation or deformation, and (d) Output graphs of key point detection in the case of fingers being close together.



5.2.2. Determining the Core Parameters of MIQF

a Determining integrity threshold θ_I

Using 105 complete palm samples and 100 incomplete palm samples, a palm integrity discrimination experiment was conducted. The key point positions and quantities were used as the threshold criteria for filtering incomplete palm samples. Thresholds were set at 0.1-0.8, with intervals of 0.1. Through experimentation, it was found that setting the threshold θ_I between 0.4-0.6 yielded a good effect in filtering incomplete palm samples without affecting the normal palm input.

Table 5
Experiment for the threshold of palm integrity detection algorithm.

θ_I	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
KP (K=105)	105	105	105	105	105	103	96	2
KN (N=100)	19	13	10	4	2	1	0	0

b Evaluation of rotation threshold $[\bar{\theta}_{yaw}, \bar{\theta}_{pitch}]$

Filter palm samples with large rotation angles without quantifying the specific rotation angles of the palms. Given that in practical use, the rotation of the palm is within the controllable range of individuals, the degree of rotation in brushing palm movements generally does not exceed a certain limit. To determine $\bar{\theta}_{yaw}$, divide the dataset of 300 groups into 3 groups: clearly left-rotated samples, clearly right-rotated samples, and samples with controllable deviation, each group containing 100 samples. Similarly, to calculate $\bar{\theta}_{pitch}$, the dataset is classified into significant upward rotation, significant downward rotation, standard rotation.

Table 6

Result of average deflection rate under different angle rotation data sets.

Data set	Over-left rotation	Standard	Over-right rotation
$\bar{\theta}_{yaw}$	0.737	0.980	1.241
Data set	**Over-upward rotation**	**Standard**	**Over-downward rotation**
$\bar{\theta}_{pitch}$	0.88	1.03	1.30

Experimental data has shown a correlation between and the direction and angle of palm rotation, serving as a screening reference indicator. Under normal conditions, the palm is typically positioned with $\bar{\theta}_{yaw}$ ranging from 0.85 to 1.1 and $\bar{\theta}_{pitch}$ ranging from 0.9 to 1.2 in different scenarios.

Due to variations in the actual width of the palm and finger spacing among individuals, as well as the palm's non-rigid nature and inherent flexibility, these factors can all influence the actual results. Therefore, this method may exclude palms with larger rotation angles, which makes it challenging to accurately capture the angles change.

Additionally, since this system adopts an omnidirectional registration mode, screening for rotation in the roll direction is not performed.

5.2.3. Performance of Different Loss Functions on oPVR

Experimental tests were conducted to evaluate the effectiveness of feature extraction network loss functions, including Center loss, Large Margin Cosine Loss (LMCL), and Additive Angular Margin Loss (AAML). Similarity was evaluated based on Euclidean distance, with input being images of 160x160 resolution and output as a 512-dimensional feature vector to determine the misidentification rate.

Table 7

Recognition experiment under different loss functions.

	Center loss	LMCL	AAML
Accuracy	0.981	0.991	0.993
ValidationRate@ FAR=0.001	0.932	0.985	0.988

Experimental results show that the loss function AAML performs the best with a precision of 99.3%, a false acceptance rate below one thousandth, and a recognition rate of 98.8%. The AAML loss function normalizes feature vectors and enhances the additive angular margin, improving inter-class separability while reinforcing intra-class compactness and inter-class differences. Although based on a face recognition model, it has demonstrated good discriminative capability in palm vein recognition models, further enhancing discriminative ability and training stability.

5.2.4. Performance of Different Image Resolutions on oPVR

Considering that the image resolution for face recognition algorithms using MobileNet is 112x112, therefore experiments are conducted on the influence of input image resolution on palm vein recognition algorithms.

Table 8

Recognition experiment under different image resolusion.

	112x112	160x160
Accuracy	0.984	0.993
ValidationRate@ FAR=0.001	0.956	0.988

The experiment demonstrates that palm vein recognition performs better at a resolution of 160x160 compared to 112x112. Further increasing the resolution has little impact on recognition rates and adds to computational costs. Therefore, the 160x160 resolution is chosen.

5.2.5. Testing System Real-time Performance

To assess the real-time performance of the system, 1:1 experiments and 1:N experiments were conducted.

1:1 Experiment: (1) Register a palm vein image, (2) Define the time taken by the device from capturing the palm vein image to returning the result as one instance of 1:1 recognition, (3) Conduct 40 palm vein 1:1 experiments for the registered palm, (4) 1:1 Recognition Average Response Time=Total time for 40 palm vein 1:1 experiments/40.

1:N Experiment: (1) Register 10380 palm vein images in the test database, (2) Define the time taken by the device from capturing the palm vein image to returning the result as one instance of 1:N recognition, (3) Conduct 12 palm vein 1:N experiments on a single registered palm, (4) 1:N Recognition Average Response Time=Total time for 12 palm vein 1:N experiments/12.

Table 9
Recognition experiment for response time for 1:1 experiment and1:N experiment.

	1:1	1:N
Average Response Time(ms)	37	245

Image collection and upload stages are the most time-consuming step for 1:1 recognition, while feature comparison stage is the most time-consuming step for 1:N recognition. And the feature comparison time being directly related to the size of the information database. Experimental results have shown that the system is able to meet the demands of real-world applications effectively with good real-time performance.

Given the extensive use of P-SSD, MIQF and Feature Extraction Network (based on MobileNet), the complexity is listed as below:

Table 10
Complexity analysis for oPVR.

Component	Parameters	Million Mult-adds	FLOPs	AVG Latency
P-SSD	2.8MB	460M	1120M	45ms
MIQF	-	-	-	5-10ms
MobileNet	2.4MB	420M	960M	32ms
Entire System	**5.2MB+**	**880M**	**2080M**	**<77ms**

It is shown through analysis that this set of algorithms supports comparisons not on a host computer but also in embedded palm vein modules. For example, the main processor of the chip adopted by embedded module is the RISCV C906 @1.0Ghz, and the coprocessor is the RISCV C906 @700Mhz, equipped with a TPU computing power of 0.3TOPS. The flash memory size of the hardware is 16 megabytes (MB), which is more than sufficient to meet the entire system's requirement of 5.2MB+.

5.2.6. Comparison Experiment

The proposed method was tested on the publicly available CASIA dataset, which consists of 2400 palm vein samples. 1920 samples were used for training, and 480 samples for testing. Due to the limited data, data augmentation techniques such as rotation, brightness variation, simulated lens distortion, and shadows were applied to the original dataset, increasing it to 700,000 samples for training. The system showed promising results on the public dataset and proved its universality.

Table 11
Evaluate the performance of this method on the CASIA multispectral dataset.

Performance on CASIA multispectral dataset	Value
Accuracy (%)	99.89
Equal Error Rate (EER)	0.020
ValidationRate@FAR=0.001	0.920

Comparing with different methods based on CASIA, as shown in Table 12, our method demonstrates superior recognition performance. This is attributed to the enhanced accuracy in ROI extraction achieved by incorporating a palm detection network, thereby resulting in an overall improved recognition network efficiency. While other methods in comparative experiments exhibit good performance on public datasets, their adaptability and anti-interference capabilities in real-world scenarios have not been adequately validated. The approach and system we propose are specifically designed for real-world scenarios, making the data collected in such scenarios more persuasive.

Table 12
Comparison of recognition performance between different paper methods based on the CASIA dataset and the method proposed in this paper.

Method	Year	Accuracy (%)	EER
PCANet with deep learning [11]	2017	96.50	0.949
PVSNet [29]	2018	85.16	3.170
LBP with DCNN [3]	2018	98	0.0261
UNet [4]	2019	96.33	0.470
CNN with Bayesian optimization [5]	2020	99.40	0.068
CNN with SIFT and Gabor filter [37]	2022	99.73	0.026
Lightweight CNN [18]	2024	/	0.018
CNN Self-Attention Module [24]	2024	97	0.035
The proposed method	2024	99.89	0.020

5.2.7. Ablation Experiment: Identification Experiment on Different Feedback Methods for Input Quality

Integrity palm inspection method, such as A. Brightness inspection, calculating the average brightness value of the image, selecting images with an average brightness value between 90-200, rotation angle inspection, calculating θ_{yaw}, θ_{pitch}, selecting images within 0.9-1.1; scale inspection, combined with integrity judgment, calculating the palm area to prevent palms that are too small.

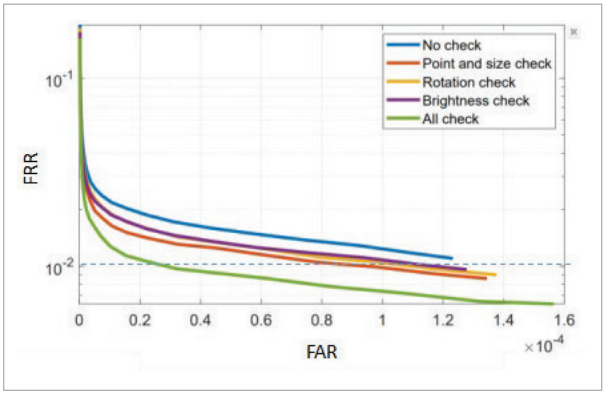
Based on the above quality assessment methods, five sets of control experiments were conducted. In the input stage, brightness inspection, rotation angle inspection, scale inspection was added respectively, and after all quality inspection feedback steps, data filtering was performed (each ID has 5 images, selecting the best image for input, rejecting input if none are suitable, removing from the test dataset), obtaining the false recognition rate-rejection rate results of the five experiments, as shown in the figure.

Through experiments, it has been found that adding the quality assessment and feedback step significantly improves recognition efficiency. Increasing check methods for brightness, rotation and completeness all lead to improvements in recognition results. The optimal outcome is achieved by combining these methods. The oPVR system achieved an accuracy of 98.9%, EER of 0.014 and FAR of 0.00123 using the SJTU Real-Scenario Palm Vein Database. With FRR equals 0.01, the implementation of the MIQF check process leads to an 82.11% improve-

Table 13
The effect of different image quality check algorithm on system recognition rate.

Image quality check	No check	Brightness check	Rotation check	Integrity & scale check	All check
FAR ($\times 10^{-3}$) @FRR=0.01	0.123	0.112	0.098	0.078	0.022
Accuracy (%)	98.90	99.04	99.10	99.14	99.37
EER	0.014	0.0098	0.012	0.0088	0.0064

Figure 16
FRR-FAR curve for recognition systems under seperate or total image quality feedback methods.



ment in FAR compared to situations without checks. While, accuracy increases by 0.475%, while the EER decreases by 54.2%. These enhancements significantly bolster recognition performance against environmental disturbances and pose variations.

As a conclusion, conducting image quality checks enables the early screening of poorly registered samples to ensure that registered samples have complete palm prints, effective brightness, feasible scale, and controllable rotation. This ensures that registered samples contain sufficient information to achieve a high-precision, wide-ranging identification experience in identity authentication.

6. Conclusion

Based on the reflective infrared imaging embedded device, this article introduces a multi-dimensional image quality feedback method, enabling the realization of an omni-direction palm vein recognition system for real-world applications.

The technological innovations mainly focus on the following aspects: Innovation 1: Introducing an image quality feedback algorithm that includes various checks such as brightness, scale, rotation angle, and integrity of the palm, effectively improving recognition performance. Innovation 2: Proposing the P-SSD model which realizes omnidirectional recognition and addresses the challenges of ROI extraction accuracy under conditions of fingers being close together, palm angle deviations and background interference. Innovation 3: High-accuracy oPVR algorithm mainly consists of P-SSD algorithm, lightweight feature extraction network based on MobileNet and MIQF algorithm to enhance accuracy and reduce computation time. Innovation 4: Design full-process palm vein collection and recognition system that includes image processing, data upload, personnel registration, feature comparison and recognition result output. Innovation 5: Implementing a low-cost, low-power embedded hardware design for palm vein recognition devices that integrate data collection and identification.

Through experimental analysis, the new algorithm has shown unique advantages in adapting to complex scenarios and gestures, enhancing the accuracy, speed, and user experience of palm vein recognition systems.

In the era of informatization, the verification and protection of individual information are gaining more prominence, posing higher demands on identity information authentication technology and products [35], [33]. In the future, efforts will be directed towards comprehensive application scenarios of 'palm vein recognition+', constructing a multi-dimensional, stereoscopic, and all-encompassing information data application platform to enhance the precision, confidentiality, and adaptability of identity authentication technology products, serving economic and social development [10].

Acknowledgement

The research was partially funded by National Major Science and Technology Projects of China (2020YFB1711204).

References

1. Ananthi, G., Sekar, J. R., Arivazhagan, S. Ensembling Scale Invariant and Multiresolution Gabor Scores for Palm Vein Identification. Information Technology and Control, 51(4), 704-722, 2022, doi: 10.5755/j01.itc.51.4.30858. https://doi.org/10.5755/j01.itc.51.4.30858
2. Babalola, F. O., Bitirim, Y., Toygar, Ö. Palm Vein Recognition Through Fusion of Texture-Based and CNN-Based Methods. Signal, Image and Video Processing, 15(3), 459-466, 2021, doi: 10.1007/s11760-020-01765-6. https://doi.org/10.1007/s11760-020-01765-6
3. Bhilare, S., Jaswal, G., Kanhangad, V., Patil, P. Single-Sensor Hand-Vein Multimodal Biometric Recognition Using Multiscale Deep Pyramidal Approach. Machine Vision and Applications, 29, 1269-1286, 2018, doi: 10.1007/s00138-018-0959-2. https://doi.org/10.1007/s00138-018-0959-2
4. Chantaf, S., Hilal, A., Elsaleh, R. Palm Vein Biometric Authentication Using Convolutional Neural Networks. Proceedings of the 8th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, Springer, 146,

352-363, 2020, doi: 10.1007/978-3-030-21005-2_34. https://doi.org/10.1007/978-3-030-21005-2_34

5. Chen, Y. Y., Hsia, C., Chen, P. Contactless Multispectral Palm-Vein Recognition with Lightweight Convolutional Neural Network. IEEE Access, 9, 149796-149806, 2021, doi: 10.1109/ACCESS.2021.3124631. https://doi.org/10.1109/ACCESS.2021.3124631
6. Elnasir, S., Shamsuddin, S. M. Proposed Scheme for Palm Vein Recognition Based on Linear Discrimination Analysis and Nearest Neighbour Classifier. 2014 International Symposium on Biometrics and Security Technologies (ISBAST), Kuala Lumpur, Malaysia, 67-72, 2014, doi: 10.1109/ISBAST.2014.7013096. https://doi.org/10.1109/ISBAST.2014.7013096
7. Felix, M., Abdulla, W. H. Segmentation of Palm Vein Images Using U-Net. Proceedings of the 2020 APSIPA ASC, Auckland, New Zealand, 7(10), 64-70, 2020.
8. Gurunathan, V., Sudhakar, R., Sathiyapriya, T., Sureka, N., Suhita, S., Sagar, P. A. Palm Vein Biometric System Using Support Vector Machine Classifier. Soft Computing for Security Applications, Springer, 1449, 2023, doi: 10.1007/978-981-99-3608-3_17. https://doi.org/10.1007/978-981-99-3608-3_17
9. Htet, A. S., Lee, H. J. Contactless Palm Vein Recognition Based on Attention-Gated Residual U-Net and ECA-ResNet. Applied Sciences, 13, 6363, 2023, doi: 10.3390/app13116363. https://doi.org/10.3390/app13116363
10. Irshad, A., Ch, S., Shafiq, M., Khan, R. An Improved Biometric Multi-Server Authentication Scheme for Chang et al.'s Protocol. Information Technology and Control, 48(2), 211-224, 2019, doi: 10.5755/j01.itc.48.2.17417. https://doi.org/10.5755/j01.itc.48.2.17417
11. Jiang, R., Zhang, Z., Xie, Z., Huang, Y., Zhao, Y. Improving Biometric Identification Performance Using PCANet Deep Learning and Multispectral Palmprint. Springer, 51-69, 2017, doi: 10.1007/978-3-319-47301-7_3. https://doi.org/10.1007/978-3-319-47301-7_3
12. Kabaciński, R., Kowalski, M. Human Vein Pattern Segmentation from Low-Quality Images: A Comparison of Methods. Image Processing and Communications Challenges, Springer, 84, 105-112, 2018, doi: 10.1007/978-3-642-16295-4_12. https://doi.org/10.1007/978-3-642-16295-4_12
13. Kang, W., Wu, Q. Contactless Palm Vein Recognition Using a Mutual Foreground-Based Local Binary Pattern. IEEE Transactions on Information Forensics and Security, 9(11), 1974-1985, 2014, doi: 10.1109/TIFS.2014.2361020. https://doi.org/10.1109/TIFS.2014.2361020
14. Kuzu, R. S., Maiorana, E., Campisi, P. Vein-Based Biometric Verification Using Densely-Connected Convolutional Autoencoder. IEEE Signal Processing Letters, 27, 1869-1873, 2020, doi: 10.1109/LSP.2020.3030533. https://doi.org/10.1109/LSP.2020.3030533
15. Lee, C., Li, C., Chen, Z., Li, X. A Biometric-Based Authentication and Anonymity Scheme for Digital Rights Management System. Information Technology and Control, 47(2), 262-274, 2018, doi: 10.5755/j01.itc.47.2.18506. https://doi.org/10.5755/j01.itc.47.2.18506
16. Lefkovits, S., Lefkovits, L., Szilágyi, L. Applications of Different CNN Architectures for Palm Vein Identification. Modeling Decisions for Artificial Intelligence, Springer, 11(676), 295-306, 2019, doi: 10.1007/978-3-030-26773-5_26. https://doi.org/10.1007/978-3-030-26773-5_26
17. Liao, H., Jin, X., Zhu, H., Fu, Y., El Yacoubi, M. A., Qin, H. GAN ET: Gabor Attention Aggregation Network for Palmvein Identification. 2024 HSI, Paris, France, 1-6, 2024, doi: 10.1109/HSI61632.2024.10613571. https://doi.org/10.1109/HSI61632.2024.10613571
18. Luo, D., Qiao, Y., Xie, D., Zhang, S., Kang, W. Palm Vein Recognition Under Unconstrained and Weak-Cooperative Conditions. IEEE Transactions on Information Forensics and Security, 19, 4601-4614, 2024, doi: 10.1109/TIFS.2024.3378427. https://doi.org/10.1109/TIFS.2024.3378427
19. Macgregor, P., Welford, R. Veincheck: Imaging for Security and Personnel Identification. Advanced Imaging, 6(7), 52-56, 1991.
20. Musunuru, R. G., Sivaprakasam, D. T., Kishore, D. G. Contactless Palm Vein Recognition System with Integrated Learning Approach System. International Journal of Advanced Computer Science and Applications, 2023. https://doi.org/10.14569/IJACSA.2023.0141146
21. Obayya, M. I., El-Ghandour, M., Alrowais, F. Contactless Palm Vein Authentication Using Deep Learning with Bayesian Optimization. IEEE Access, 9, 1940-1957, 2021. https://doi.org/10.1109/ACCESS.2020.3045424
22. Pan, Z., Wang, J., Wang, G., Zhu, J. Multi-Scale Deep Representation Aggregation for Vein Recognition. IEEE Transactions on Information Forensics and Security, 16, 1-15, 2021. https://doi.org/10.1109/TIFS.2020.2994738
23. Peng, C., Xiaoke, L., Feng, J., Jinzhuan, S. Research of High Quality Palm Vein Image Acquisition and ROI Extraction. Chinese Journal of Sensors and Actuators, 7(13), 2015.

24. Qin, H., Xi, H., Li, Y., Wang, J. Adversarial Learning-Based Data Augmentation for Palm-Vein Identification. IEEE Transactions on Circuits and Systems for Video Technology, 34(6), 4325-4341, 2023. https://doi.org/10.1109/TCSVT.2023.3334825

25. Qino, H. F., Zhang, Y., Li, X., Wu, J., Chen, Z. Multi-Scale and Multi-Direction GAN for CNN-Based Single Palm-Vein Identification. IEEE Transactions on Information Forensics and Security, 16(1), 2652-2666, 2021, doi: 10.1109/TIFS.2021.3059340. https://doi.org/10.1109/TIFS.2021.3059340

26. Ribaric, S., Fratric, I., Pavesic, N. A Biometric Identification System Based on Eigenpalm and Eigenfinger Features. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(11), 1698-1709, 2005, doi: 10.1109/TPAMI.2005.209. https://doi.org/10.1109/TPAMI.2005.209

27. Sandhya, T., Reddy, G. S. An Optimized Elman Neural Network for Contactless Palm-Vein Recognition Framework. Wireless Personal Communications, 131, 2773-2795, 2023, doi: 10.1007/s11277-023-10579-x. https://doi.org/10.1007/s11277-023-10579-x

28. Sengupta, A., Chaurasia, R., Reddy, T. Contact-Less Palmprint Biometric for Securing DSP Coprocessors Used in CE Systems. IEEE Transactions on Consumer Electronics, 67(3), 202-213, 2021, doi: 10.1109/TCE.2021.3105113. https://doi.org/10.1109/TCE.2021.3105113

29. Thapar, D., Jaswal, G., Nigam, A., Kanhangad, V. PVSNet: Palm Vein Authentication Siamese Network Trained Using Triplet Loss. ISBA, IEEE, Hyderabad, India, 2019, doi: 10.1109/ISBA.2019.8778623. https://doi.org/10.1109/ISBA.2019.8778623

30. Van, H. T., Duong, C. M., Vu, G. V., Le, T. H. Palm Vein Recognition Using Enhanced Symmetry Local Binary Pattern and SIFT Features. ISCIT, Ho Chi Minh City, Vietnam, 311-316, 2019, doi: 10.1109/ISCIT.2019.8905179. https://doi.org/10.1109/ISCIT.2019.8905179

31. Wang, C., Zeng, X., Sun, X., Dong, W., Zhu, Z. Quality Assessment on Near Infrared Palm Vein Image. 2017 YAC, IEEE, 3 July 2017. https://doi.org/10.1109/YAC.2017.7967580

32. Wang, P., Qin, H. Palm-Vein Verification Based on U-Net. IOP Conference Series: Materials Science and Engineering, 806, 012043, 2020, doi: 10.1088/1757-899X/806/1/012043. https://doi.org/10.1088/1757-899X/806/1/012043

33. Wei, F., Ma, J., Jiang, Q., Li, H., Chen, X. Cryptanalysis and Improvement of an Enhanced Two-Factor User Authentication Scheme in Wireless Sensor Networks. Information Technology and Control, 45(1), 62-70, 2016, doi: 10.5755/j01.itc.45.1.11949. https://doi.org/10.5755/j01.itc.45.1.11949

34. Wu, W., Elliott, S. J., Lin, S., Wang, X. Review of Palm Vein Recognition. IET Biometrics, 9(1), 1-10, 2020, doi: 10.1049/iet-bmt.2019.0034. https://doi.org/10.1049/iet-bmt.2019.0034

35. Yang, H., Cai, T., Zhang, C., Wang, J. Research on User Identity Authentication Method Based on Edge Computing. PCDS, Singapore, 1-4, 2024, doi: 10.1109/PCDS61776.2024.10743471. https://doi.org/10.1109/PCDS61776.2024.10743471

36. Zhang, L., Cheng, Z. X., Shen, Y., Wang, D. Q. Palmprint and Palmvein Recognition Based on DCNN and a New Large-Scale Contactless Palmvein Dataset. Symmetry, 10(4), 1-15, 2018, doi: 10.3390/sym10040078. https://doi.org/10.3390/sym10040078

37. Zhang, L., Wang, H. A Novel Finger Vein Recognition Method Based on Deep Learning. Pattern Recognition Letters, 132, 118-124, 2020, doi: 10.1016/j.patrec.2019.09.014. https://doi.org/10.1016/j.patrec.2019.09.014

38. Zhang, Y., Wu, H. Multimodal Biometric Recognition Based on Deep Learning: A Survey. IEEE Transactions on Information Forensics and Security, 16, 123-145, 2021, doi: 10.1109/TIFS.2020.2991234.

39. Zhou, Y., Kumar, A. Human Identification Using Palm-Vein Images. IEEE Transactions on Information Forensics and Security, 6(4), 1259-1274, 2011, doi: 10.1109/TIFS.2011.2158423. https://doi.org/10.1109/TIFS.2011.2158423