

Pose-Robust Finger Identification Based on Mutual Region Alignment and Minimum Convolution Point Feature

Weili Yang

Guizhou Minzu University

Dewen Kong

Guizhou Minzu University

Fujian Feng

Guizhou Minzu University

Yonghuai Yang

Guizhou Minzu University

Deng Li

Guizhou Minzu University

ZhongChi Liu

Guizhou Xingyi Electric Power Development Co., Ltd.

Dacan Luo

gmudacanluo@gmail.com

Guizhou Minzu University

Article

Keywords:

Posted Date: July 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10197073/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Pose-Robust Finger Identification Based on Mutual Region Alignment and Minimum Convolution Point Feature

Weili Yang¹, Dewen Kong¹, Fujian Feng², Yonghuai Yang¹, Li Deng², Liu ZhongChi³, and Dacan Luo^{1,*}

¹Guizhou Minzu University, School of Physics and Mechatronic Engineering, Guiyang, 550025, China

²Guizhou Minzu University, School of Data Science and Information Engineering, Guiyang, 550025, China

³Guizhou Xingyi Electric Power Development Co., Ltd., Xingyi, 562499, China

*corresponding author: gmudacanluo@gmail.com

ABSTRACT

To address the performance degradation of finger identification systems caused by longitudinal rotation during contactless finger image acquisition, this paper proposes a novel pose-robust finger identification framework integrating mutual region extraction, minimum convolution point feature learning, and weighted score fusion. Specifically, longitudinal rotation of fingers easily leads to inconsistent imaging regions and non-rigid deformations, which severely undermine the stability of feature matching. To tackle these challenges, we propose a three-stage framework: First, a mutual region extraction strategy is proposed to adaptively align inconsistent imaging areas via vertical gradient calculation and minimum distance metric, effectively eliminating content discrepancies caused by longitudinal rotation. Second, a minimum convolution point feature learning module is designed to model non-linear correlations between cross-regional features, coupled with a deformation-aware matching algorithm to enhance robustness against finger deformation, mitigating the adverse impact of pose variations. Finally, a weighted score fusion method is introduced to integrate matching results from minimum convolution point features and Uniform Local Binary Pattern(LBP) features, fully exploiting the complementary information between structural and texture features. Extensive experiments are conducted on four public datasets, SDUMLA_HMT, MMCBNU_6000, FV-USM, and LFMB-3DFB, and the results demonstrate that the proposed method achieves state-of-the-art (SOTA) performance in terms of recognition accuracy and robustness against longitudinal rotation and finger deformation.

Introduction

Driven by the increasing demand for secure and reliable identification systems, biometrics¹ has gained growing attention and become a pivotal research area in security technology. A variety of biometric traits can be utilized for identity recognition, including facial traits such as face, iris, ear, voiceprint, etc.; and hand-based traits such as fingerprint, finger vein, palmprint, palm vein, etc.. From the perspective of security and discriminability, these biometric characteristics can be categorized into extrinsic and intrinsic types: Extrinsic features(e.g., face², iris³, fingerprint⁴, palmprint⁵) are surface-visible traits with rich discriminative information but are vulnerable to forgery; intrinsic features (e.g., finger vein⁶, palm vein⁷) are internal biological structures and textures that are extremely difficult to forge without specialized equipment, though their individual discriminability is relatively limited⁸. Therefore, fusing extrinsic and intrinsic features can complement their advantages and overcome limitations, simultaneously enhancing the accuracy and security of biometric systems.

Among various biometric traits, finger-based features (e.g., fingerprint, finger vein) are more suitable for compact acquisition devices compared to other traits, making them ideal for mobile and portable applications⁹. Therefore, this paper focuses on finger biometrics, specifically combining intrinsic finger vein features and extrinsic contactless fingerprint features. However, finger-based biometric recognition faces significant challenges due to the inherent flexibility of fingers, which leads to unavoidable pose variations during contactless capture. Among these pose variations, longitudinal rotation is the most detrimental factor, causing inconsistent imaging regions between multiple captures of the same finger and inducing non-rigid deformation of finger tissues, resulting in image misalignment, loss of discriminative information, and ultimately reduced feature matching stability.

We model the finger as an ellipsoid to simplify the analysis of spatial pose changes. In contactless acquisition scenarios, unstable finger placement easily induces variations in the projected images. As observed from Figure 1(a) and (b), longitudinal rotation results in significant content inconsistency between consecutive captures of the same finger, with partial regional information missing or misaligned. Meanwhile, the flexibility of finger tissues leads to inevitable deformation under spatial

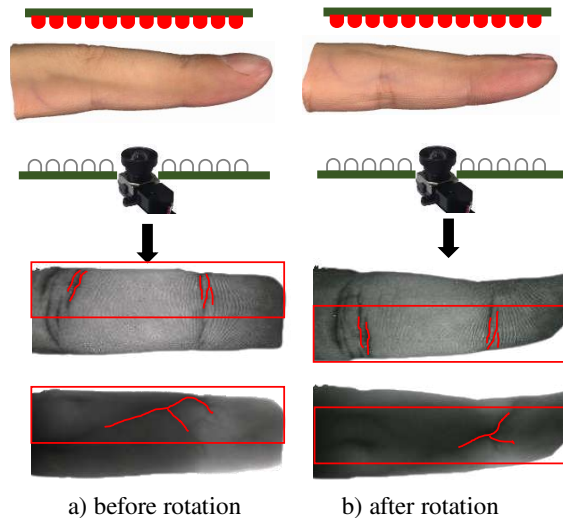


Figure 1. Finger model and the imaging plane under longitudinal rotation.

rotation, which distorts both finger vein and contactless fingerprint. Consequently, inconsistent imaging regions and finger deformation caused by longitudinal rotation are the primary bottlenecks restricting the performance of finger identification systems.

To address the above challenges, this paper proposes a novel pose-robust finger identification framework that effectively integrates mutual region extraction, minimum convolution point feature learning, and weighted score fusion. The framework systematically solves the problems of inconsistent imaging regions, deformation sensitivity, and insufficient feature complementarity. The main contributions of this work are summarized as follows:

- A mutual region extraction strategy is proposed to align pose-varying finger images by calculating vertical gradients and minimum distance metrics. This strategy adaptively segments overlapping and discriminative regions from images captured under different poses, effectively eliminating content discrepancies caused by longitudinal rotation and laying a reliable foundation for subsequent feature learning.
- A minimum convolution point feature learning module is designed to detect nonlinear structural features of finger images. Combined with a deformation-aware matching algorithm, this module enhances the system's tolerance to non-rigid finger deformation, further mitigating the adverse impact of pose variations on feature matching.
- A weighted score fusion scheme that fuse minimum convolution point features and Uniform LBP features is proposed. The minimum convolution point features focus on structural information, while Uniform LBP features excel at texture description; the weighted fusion strategy fully exploits the complementary advantages of the two feature types, significantly improving the system's identity discriminability and generalizability to complex practical scenarios.

Related Works

Finger-based biometric traits mainly consist of fingerprint and finger vein. Fingerprint features can be primarily categorized into local features and global features. Specifically, local features are capable of describing the fine-grained characteristics of fingerprints, which mainly refer to detailed feature representations such as ridge direction and ridge continuity, including endpoints, breakpoints, bridge points, bifurcations, and so on¹⁰. In contrast, global features describe the overall characteristic information of fingerprints, covering ridge pattern, pattern area, core point, delta point, and ridge count¹¹. Compared with fingerprints with abundant information, finger veins contain few texture and structural information. The feature extraction methods for finger images can be roughly divided into four categories: geometry-based methods, subspace learning-based methods, statistical-based methods, and local invariant feature-based methods. Geometry-based methods distinguish finger images of different users mainly through geometric features such as finger shape and texture topology of fingerprints and finger veins, which can be further divided into local geometry methods¹² and global topology methods^{13,14}. Subspace learning-based methods reduce the high-dimensional image data to a low-dimensional space by means of machine learning, and then perform matching using the matching scores of feature vectors between images in the low-dimensional feature subspace^{15,16}. Statistical-based methods construct the feature descriptor of an image by summarizing the grayscale variation information of the finger

image through statistical analysis^{17,18}. Local invariant feature-based methods adopt feature points extracted by scale-invariant feature algorithms such as SIFT as the feature descriptors of finger images¹⁹.

Researchers have also attempted to combine the finger feature information obtained based on feature engineering with deep learning, thereby giving full play to the advantages of both methods. Considering the excellent feature extraction performance of neural networks, Cao et al.²⁰ utilized neural networks to extract the description information of fingerprint ridges and minutiae, and then performed authentication through traditional methods based on these extracted features. In their work, Wen et al.²¹ introduced PCA filters to learn multi-level filter banks, and employed binary hashing and histogram blocks for indexing and pooling. Finally, the block-wise histograms were input into the neural network as image features for training, enabling the network to achieve high recognition accuracy even with limited training samples. Taking into account the special texture pattern of finger veins, Zhang et al.²² proposed an adaptive Gabor convolutional neural network, which replaces the convolutional kernels in the convolutional neural network with Gabor filters that excel in extracting local feature information. This method greatly reduces the network parameters and computational complexity while ensuring excellent recognition performance of the network, providing a new solution for the deployment of finger feature authentication systems on embedded devices.

Proposed Method

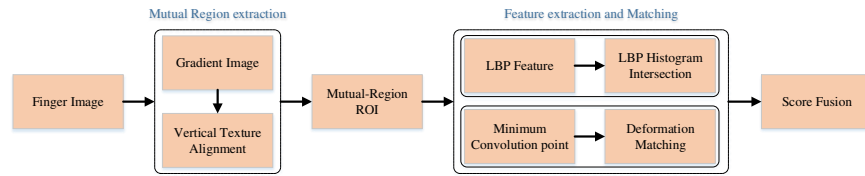


Figure 2. Overall framework of the proposed approach.

The overall framework of the proposed finger identification framework is illustrated in Figure 2, consisting of three core stages: mutual region extraction, minimum convolution point feature learning with deformation-aware matching and weighted score-level fusion. First, considering the inconsistent imaging plane caused by different finger poses, we propose a mutual region extraction method to align the imaging areas. Then, a minimum convolution point feature learning method is proposed to handle finger deformation. Subsequently, a deformation-aware point-wise matching method is developed to strengthen matching performance. Finally, to further boost recognition performance, we achieve the final recognition results by a score-level fusion strategy combining the minimum convolution point features and Uniform LBP features.

Mutual Region Extraction

For different postures in each captures, the image extracted from finger will inevitably result in some degree of deviation, which contains horizontal and vertical shift. Horizontal shift can be constrained by the acquisition device and the phalangeal joint localization²³, while vertical shift caused by longitudinal rotation is the primary factor to generate the deviation, which affects the recognition performance. To minimize the offset, we propose a method to extract the mutual region. As is shown in Figure 3, the mutual region extraction contains two principal components, i.e. vertical gradient calculation and minimum distance calculation.

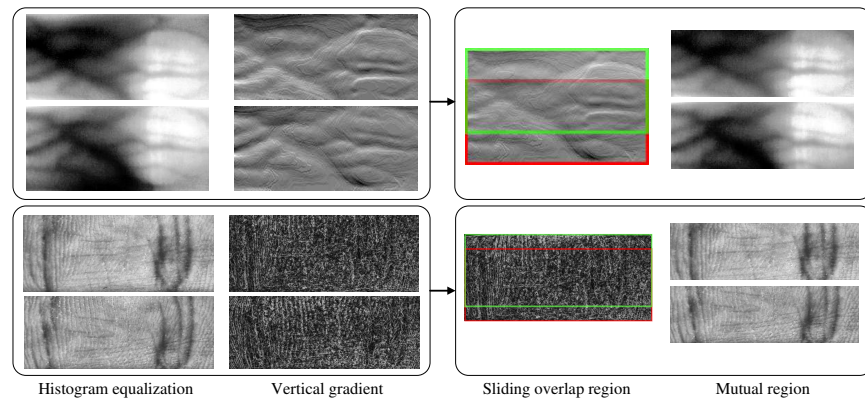


Figure 3. The block diagram of the mutual region extraction.

During contactless image acquisition, finger pose variations lead to horizontal and vertical shifts. Horizontal shifts can be constrained by acquisition devices and phalangeal joint localization, while vertical shifts caused by longitudinal rotation are the primary source of imaging inconsistency. To eliminate this inconsistency, we propose a mutual region extraction method comprising two key steps: vertical gradient calculation and minimum distance-based similarity evaluation (Figure 3).

In the capturing step, many factors can cause the low image quality which contain distortions of position, variations of illumination and the skin differences of each individual. Therefore, image pre-processing is first performed, we carry out histogram equalization algorithm to enhance image contrast, improving the distinguishability of vein patterns and fingerprint textures. After that, we put attention to the vertical gradient to deal with the vertical shift. Let $I(x, y)$ denote the gray value of pixel (x, y) after histogram equalization. The vertical gradient $G_y(x, y)$ is calculated to capture vertical intensity changes, which is defined as:

$$G_y(x, y) = I(x, y + 1) - I(x, y - 1) \quad (1)$$

After obtaining the vertical gradient images, we adopt the following scheme to extract the mutual region between two images (e.g., a template image I_1 and a probe image I_2). Keeping the template image I_1 fixed and shifting the probe image I_2 along the vertical direction. For each vertical shift, the overlapping region of the two images is extracted, and the similarity between the vertical gradient images is calculated using the minimum distance metric. To extract the mutual region effectively, we utilize the difference of vertical gradient image to evaluate the similarity, which is expressed as:

$$D(\Delta y) = \min_{x, y} |G_y^{I_1}(x, y) - G_y^{I_2}(x, y + \Delta y)| \quad (2)$$

The mutual region is then determined as the overlapping region with the minimum distance $D(\Delta y)$, which indicates the highest similarity between the vertical gradient images.

Minimum Convolution Point Feature and Deformation-Aware Matching

To address non-rigid deformation, we propose a minimum convolution point feature learning module combined with a deformation-aware matching algorithm, shown in Figure 4, consists of three core steps: key-point detection, minimum convolution point feature learning and deformation-aware matching.

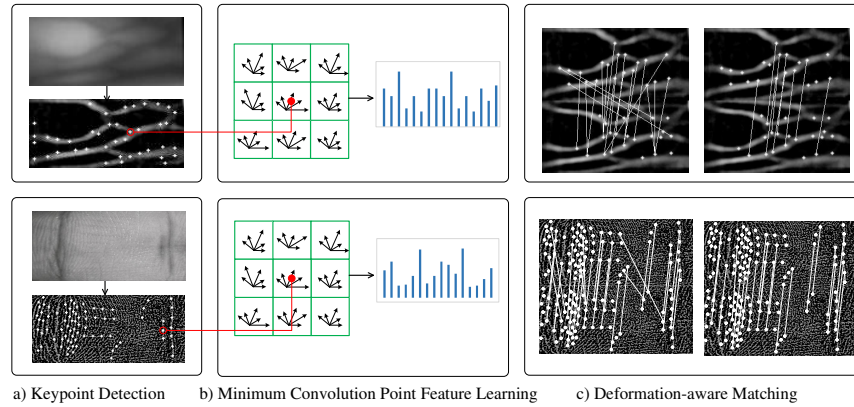


Figure 4. Framework of minimum convolution point feature and deformation-aware matching.

(1) Key-point Detection

Key-point detection is a fundamental step for reliable matching. To increase the number of discriminative key-points, we utilize non-linear edge patterns in finger images. To detect key-points robust to deformation, we utilize multi-orientation filters. As shown in Figure 5, we apply eight directional derivative filters ($0^\circ, 22.5^\circ, \dots, 157.5^\circ$) to the enhanced image. For each pixel, we compute the minimum convolution responses across all orientations. The convolution value of an edge point differs from that of a non-edge point due to the valley shape in the cross-sectional profile.

There are mainly three kinds of edge patterns. As shown in Figure 6(a), the linear edge can fully accommodate linear filters, which means the convolution value of some filters will be zero. In contrast, the non-linear edge, cannot hold the entire linear filter in its edge pattern area as shown in Figure 6(b,c), which cause the convolution values of all filter greater than zero. According to this characteristic, we can extract the non-linear points in edge area as the feature points. The minimum

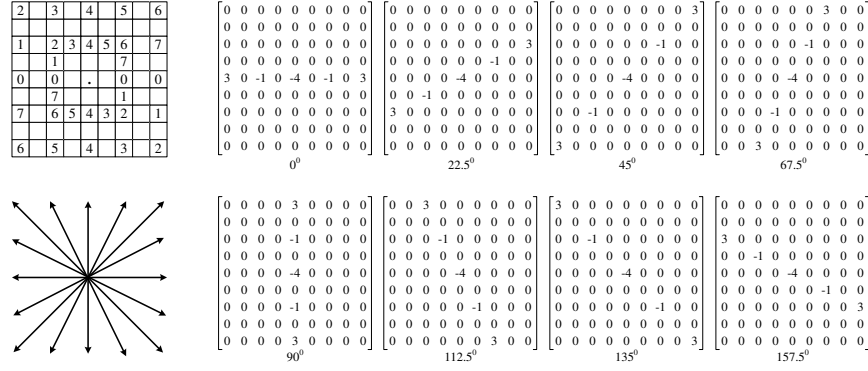


Figure 5. Different orientation of the filters.

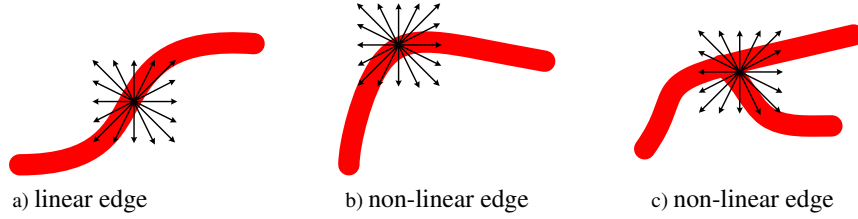


Figure 6. The filters in different vein patterns.

convolution value generated by linear filter can distinguish the non-linear point in edge pattern area and are extracted as key-points using the minimum convolution value.

Let $F_{cov}(i)$ ($i = 1, 2, \dots, 8$) denote the convolution values of a pixel with eight orientation filters, $G_{enhance}$ denote the enhance image, and G_{point} denote the feature point map, (m, n) is the position of a pixel point. The enhance image and the feature points are calculated as:

$$G_{enhance}(m, n) = \text{Max}_i(F_{cov}(i)) \quad (3)$$

$$G_{point}(m, n) = \text{Max}(\text{Min}_i(F_{cov}(i)), 0) \quad (4)$$

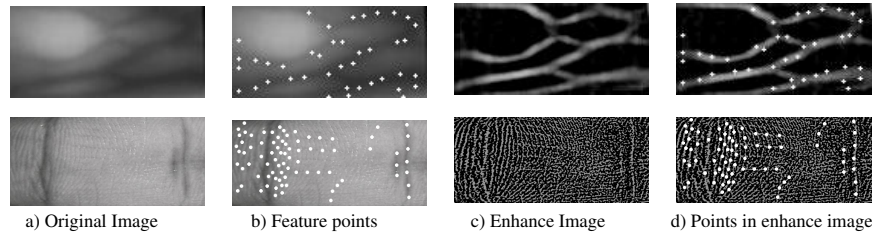


Figure 7. The key-point and edge pattern.

From Figure 6, we can see that points are produced by the detecting the non-linear edge. Furthermore, when generating the key-points, the enhance image is also obtained by the maximum corresponding value of filters. To present the effect of our proposal, we draw the points in the enhance image as shown in Figure 7.

(2) Minimum Convolution Point Feature Learning

We construct the minimum convolution point feature on each key-point, which contains the edge patterns clearly. For each key-point in the enhance image, a rectangle area around the key-point is sampled and divided into 3×3 cells. Then, the gradient of each cell is calculated, then is quantized into 5 bins (0° - 36° , 36° - 72° , ..., 324° - 360°). A 5-bin gradient histogram is generated

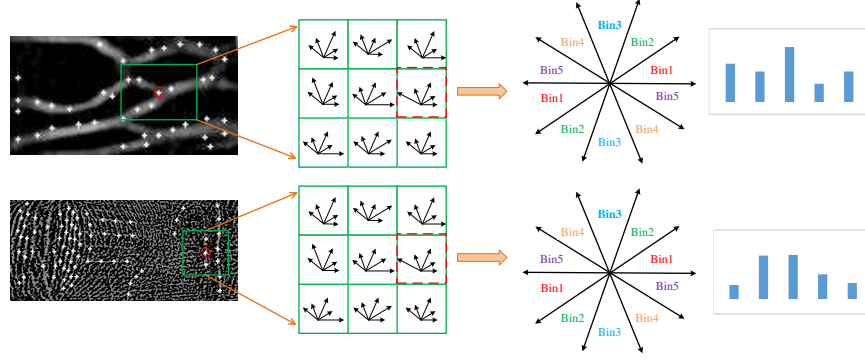


Figure 8. The generation of point descriptor.

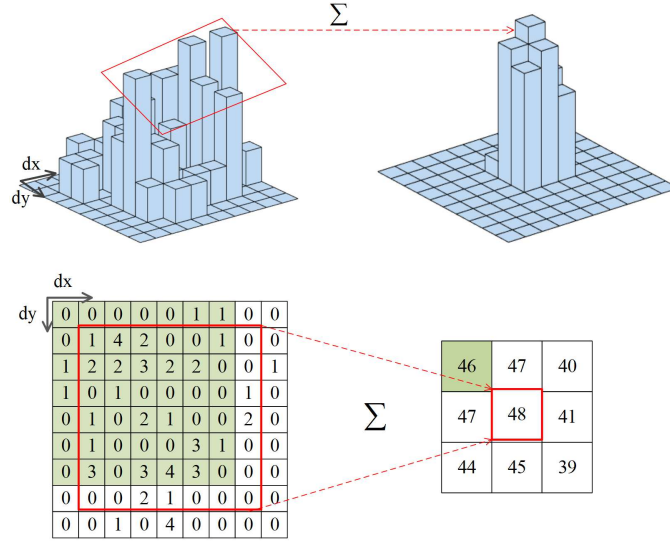


Figure 9. The statistic of histogram.

for each cell. At last, the histograms of all 9 cells are concatenated to form the final minimum convolution point feature, as illustrated in Figure 8.

(3) Deformation-aware Matching

All minimum convolution point features in probe image are matched with those in gallery image using the minimum Euclidean distance and generate initial matched paired point features. However, similar textures in different finger images may lead to false matches, which degrade recognition performance. To address this issue, some researchers have applied RANSAC to remove false matches²⁴. However, finger images exist non-rigid transformations, especially when rotating along the longitudinal axis. RANSAC is ineffective for non-rigid deformations induced by longitudinal rotation, so we propose a new deformation-aware false match removal method to improve the recognition accuracy.

Correctly matched paired point features exhibit similar orientation, which means the offsets of horizon and vertical are approximate, and these offsets can vary within a reasonable range to accommodate small deformations. Based on this observation, we construct a 2D statistic histogram of offsets along horizon and vertical orientation. Then, we accumulate the number of matched pairs within a range of offsets, which reflects the similar orientation of deviation. The maximum value in the accumulation histogram indicates the offset of most of paired point features is in this range, which denotes the correct matching. The statistic processes can be shown in Figure 9.

We suppose the window size of offsets are w in horizon and h in vertical respectively. The 2D histogram is denoted as $hist$ and H represents the final accumulation histogram. The maximum accumulation histogram can be obtained as:

$$H(i_{\max}, j_{\max}) = \text{Max} \sum_i \sum_j \text{hist}(i : i + w, j : j + h) \quad (5)$$

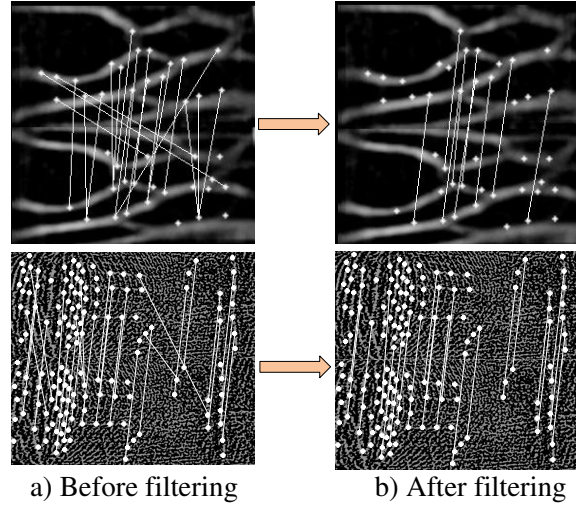


Figure 10. The removal of false point-matching.

Wherein, i_{max} and j_{max} are the index of max value in the accumulation histogram. Then, the correct matching paired point features should satisfy the following two conditions:

$$dx \in [i_{max}, i_{max} + w] \quad (6)$$

$$dy \in [j_{max}, j_{max} + h] \quad (7)$$

As shown in Figure 10, the method of mismatching removal can effectively eliminate mismatched point features. Finally, the remaining matching paired point features is the basis of evaluating similarity. Herein, we utilize the max point ratio to obtain the similarity, which is the ratio of the number of the matching paired point features to the total number of point features in the two to-be-matched images. Let $Score_{PM}$ denote the score of the deformation-aware matching, which is expressed as:

$$Score_{PM} = \frac{2m}{n_1 + n_2} \quad (8)$$

Wherein, n_1 and n_2 are the number of the two to-be matched images, m is the number of the matching paired point features.

Fusion Recognition

While the proposed minimum convolution point method effectively handles deformation, it fails to retain fine-grained texture details of the original image. In contrast, texture-based features excel at capturing texture information. To fully exploit the complementary advantages of structural and texture features, we boost up the recognition performance by fusing the minimum convolution point feature. For the texture-based feature with texture feature. As the LBP operator has strong discrimination in texture classification. Thus, we use an improved operator, Uniform LBP, in finger identification, which has the stronger ability to distinguish texture patterns. And a weighted score fusion strategy is adopted to combine the scores generated by different matchers. The block diagram of fusion recognition is shown in Figure 11.

Owing to the statistical of the LBP histogram, we can adopt histogram intersection to obtain the LBP feature matching score. The matching score is defined as follows:

$$score_{LBP} = \sum_{i=1}^m \min(H_1^{(i)}, H_2^{(i)}) \quad (9)$$

Wherein, H_1 and H_2 are the LBP feature vectors, and m is the length of the vector.

According to the weighted addition rule, the scores generated from both methods must be scale normalized to ensure a meaningful combination. We suppose that $NormScore_{LBP}$ denotes the normalized score of LBP, $NormScore_{PM}$ denotes the normalized score of minimum convolution point feature. The fusion score is calculated as follows:

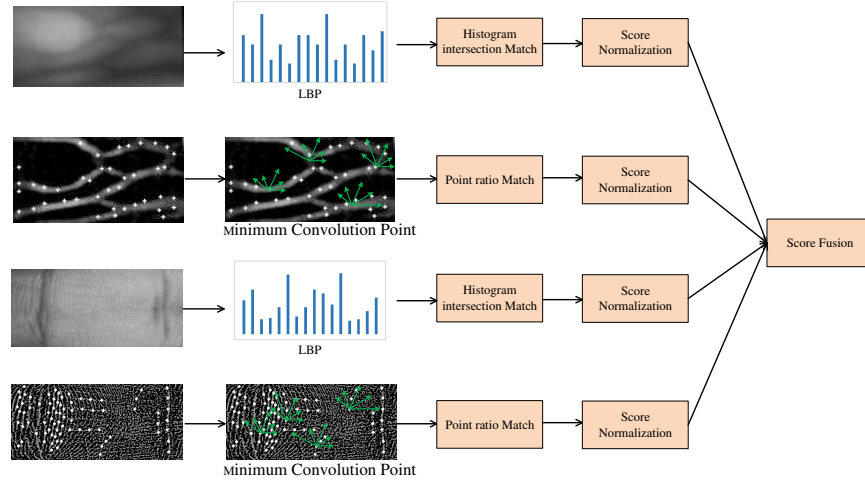


Figure 11. The fusion of recognition.

$$score = w \cdot NormScore_{PM} + (1 - w) \cdot NormScore_{LBP} \quad (10)$$

Wherein, w is the weight of the point-matching score, w ranges from 0 to 1.

Experimental Results and Discussion

In this paper, we evaluate the performance of the proposed method on three public databases: SDUMLA-HMT²⁵, MMCBNU_6000²⁶, FV-USM²⁷ and LFMB-3DFB²⁸.

The first database is the SDUMLA-HMT finger vein database, which contains 3816 images from 106 subjects, including 61 males and 45 females. Each subject was asked to provide images of the index finger, middle finger, and ring finger of both hands during the capturing process. The finger vein images were collected six times from each of the captured finger. The captured images have a resolution of 240×320. For the lack of ROI images in the database, we extract the ROI images and use the extracted ROI images in our experiments.

The second database is MMCBNU_6000 finger vein database, which consists of finger vein images from 100 volunteers, coming from 20 different countries. The finger vein images were collected ten times from each of the six fingers (index, middle, and ring fingers of both hands) for each subject. The size of image, stored in “bmp” format, is 480×640. The localized ROI images have the pixels size of 64×128. We utilize the ROI images provided in the database in our experiments.

The third database is USM finger vein database (FV-USM), which was captured from 123 subjects, containing 83 males and 40 females. Four finger veins were captured for each subject, including left index, left middle, right index and right middle fingers. The database was collected in two sessions between two weeks and each finger was captured six times in one session. As a whole, the database contains 5904 finger vein images, and we utilized the ROI image provided by the database.

The fourth database is the Large-scale Finger Multi-Biometric database and benchmark for 3D Finger Biometrics (LFMB-3DFB), which was collected from 174 volunteers using a self-designed multi-view, multi-spectral 3D finger imaging system. The database includes the index and middle fingers of both hands, totaling 695 fingers. Each finger was inserted into the imaging system 10 times for data collection, and during the capture interval, volunteers were asked to flex and relax their fingers to ensure data diversity, with the finger axial angles randomly distributed within ± 80 degrees. For each insertion, 6 contactless fingerprint images and 6 finger vein images were captured from 6 different views, resulting in a total of 41,700 contactless fingerprint images and 41,700 finger vein images. Experiments in this paper adopted the finger vein images and contactless fingerprint images from the ventral view of this dataset, with 6,950 images each. The fusion of finger vein and contactless fingerprint features was also only performed on this data subset.

To evaluate the performance of our method, we utilize some measurements, including the false accept ratio (FAR) and the false reject ratio (FRR). To consider both the FAR and FRR, we adopt equal error rate (EER) to evaluate the performance of various approaches. Herein, EER is the error rate when the false accept ratio is equal to the false reject ratio.

The Effectiveness of The Mutual Region

To verify the effectiveness of mutual region extraction in addressing imaging inconsistency induced by longitudinal rotation, we evaluate the performance on the above-mentioned four databases by utilizing the texture-based method, including LBP,

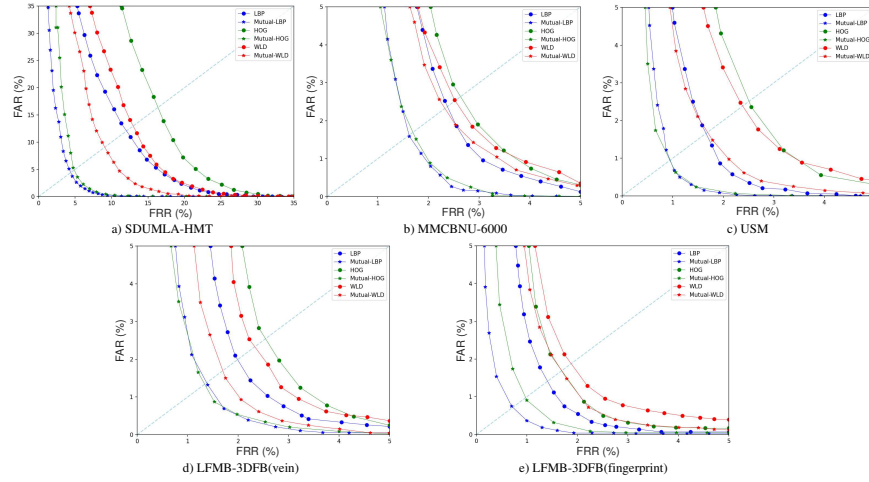


Figure 12. EER comparison of mutual region extraction in different databases.

Table 1. The performance of mutual region extraction in different databases.

Methods	SDUMLA_HMT			MMCBNU_6000			FV-USM			LFMB-3DFB(vein)			LFMB-3DFB(fingerprint)		
	LBP	WLD	HOG	LBP	WLD	HOG	LBP	WLD	HOG	LBP	WLD	HOG	LBP	WLD	HOG
Without mutual region	12.07%	12.98%	16.37%	2.37%	2.53%	2.64%	1.66%	2.36%	2.51%	1.99%	2.33%	2.54%	1.41%	1.88%	1.70%
With mutual region	4.45%	9.08%	4.94%	1.62%	2.33%	1.69%	0.94%	1.69%	0.94%	1.38%	1.69%	1.34%	0.73%	1.67%	0.99%

WLD²⁹ and HOG³⁰. As illustrated in Figure 12 and Table 1, the EERs are all reduced after integrating the mutual region extraction module, confirming that aligning overlapping discriminative regions can effectively suppress intra-class variations caused by pose shifts, with particularly significant improvements observed for LBP and HOG. Notably, the overall EERs of the SDUMLA-HMT dataset remain relatively high compared to the other datasets. Comparative analysis reveals that this is attributed to the substantial illumination variations in the SDUMLA-HMT images: even after mutual region alignment, the inherent texture variability within the aligned regions persists, limiting recognition performance. In the LFMB-3DFB dataset, texture-based features achieve superior recognition performance owing to the richer texture information in contactless fingerprint images. This observation further underscores the necessity of fusing finger vein and fingerprint modalities to leverage complementary biometric information.

The Effectiveness of Minimum Convolution Point Feature

As illustrated in Figure 13 and Table 2, the experimental results validate that the proposed minimum convolution point feature exhibits superior recognition accuracy compared to traditional statistical features (LBP, HOG, WLD) and classic point-matching methods (fuzzyEnhance-SIFT), and attains the lowest EER across the majority of the datasets. Specifically, on the SDUMLA-HMT dataset, our method yields an EER of 2.15%, which is 50.7% lower than that of fuzzyEnhance-SIFT (4.16%) and 51.6% lower than that of LBP (4.45%). This superiority originates from the proposed feature's inherent capability to model non-linear edge patterns, thereby enabling it to more robustly extract illumination-invariant structural features compared to texture-based methods, especially for datasets characterized by substantial illumination fluctuations, such as SDUMLA-HMT.

Notably, the EER of the proposed method for the FV-USM dataset (1.53%) is marginally higher than those of LBP and HOG (both 0.94%). This performance discrepancy can be ascribed to the data characteristics of FV-USM, which has minimal pose-induced intra-class deformation and high image textures, allowing texture-based features to fully capture stable texture details. Consequently, the deformation resistance advantage of our method is not fully leveraged in this context.

These in-depth comparative analyses collectively highlight the necessity of fusing structural and texture features. By harnessing the complementary strengths of structural features (robustness to pose variations and illumination fluctuations) and texture features (superior performance in stable, high-clarity scenarios), the fusion strategy can comprehensively improve the overall recognition robustness and discriminative accuracy.

In terms of computational efficiency, our method achieves a total processing time of 19ms, which is 45.7% faster than fuzzyEnhance-SIFT (35ms). This efficiency advantage makes our method suitable for real-time biometric systems, where both accuracy and speed are critical.

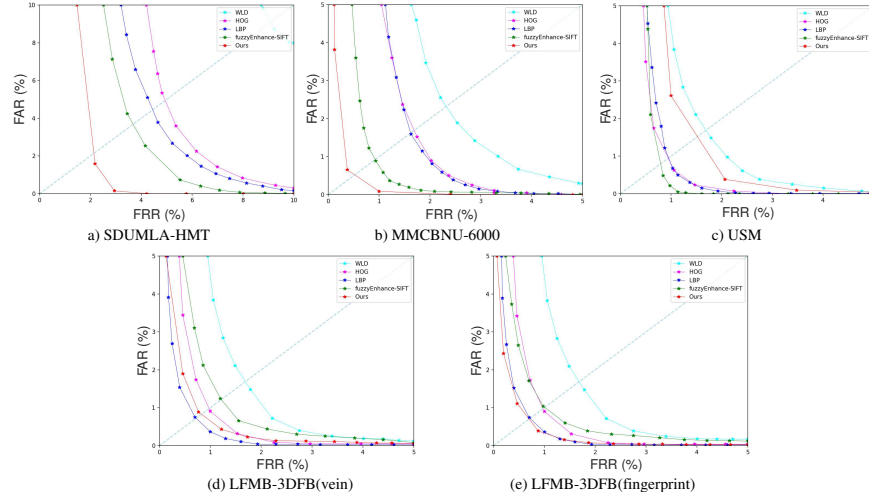


Figure 13. EER comparison of LBP, HOG, WLD and fuzzyEnhanced-SIFT.

Table 2. The performance of our proposed methods in different databases.

Methods	SDUMLA_HMT	MMCBNU_6000	FV-USM	LFMB-3DFB(vein)	LFMB-3DFB(fingerprint)
WLD	9.08%	2.33%	1.69%	1.69%	1.67%
HOG	4.94%	1.69%	0.94%	1.34%	0.99%
LBP	4.45%	1.62%	0.94%	1.38%	0.73%
fuzzyEnhance-SIFT	4.16%	1.09%	1.03%	1.22%	1.01%
Ours	2.15%	0.52%	1.53%	0.83%	0.70%

The Effectiveness of Fusion

The weighted score fusion strategy, which combines our method (structural features) and Uniform LBP (texture features), yields substantial performance improvements by harnessing the complementary advantages of the two feature types, thereby validating the rationality of multi-feature synergy.

As illustrated in Figure 14 and Table 4, the fusion strategy achieves the lowest EER across all datasets, with remarkable performance improvements compared to individual features. Specifically, in SDUMLA-HMT, the fused EER reaches 1.54%, representing a 28.4% reduction compared to the proposed structural feature alone (2.15%) and a 65.4% decrease relative to LBP (4.45%); In MMCBNU_6000, the fused EER is 0.05%, which corresponds to a 90.4% reduction versus the proposed feature (0.52%) and a 96.8% reduction versus LBP (1.62%); In FV-USM, the fused EER (0.36%) is 76.4% lower than the proposed feature (1.53%) and 61.7% lower than LBP (0.94%). Notably, even in scenarios where individual features already exhibit strong performance, the fusion strategy further mitigates residual errors: the proposed structural feature offsets LBP's susceptibility to subtle deformations, whereas LBP supplements the proposed feature's insufficiency in capturing fine-grained texture details.

In LFMB-3DFB, the fusion of vein and fingerprint modalities achieves an EER of 0.12%, which is significantly lower than the EERs of single-modal fused features (vein: 0.42%; fingerprint: 0.33%). This validates that the complementary characteristics of structural and texture features are consistent across different modalities, and the weighted fusion strategy can adaptively balance the contribution of each feature type to optimize overall recognition performance.

Comparison with The State-Of-The-Art Algorithms

As shown in Tables 5 to 8, the proposed framework can achieve satisfactory results on all public databases, and outperforms most of the existing methods, which is attributed to its systematic solution to the core challenges of finger biometric recognition: imaging region inconsistency, non-rigid deformation, and insufficient feature complementarity.

Notably, the framework exhibits consistent performance superiority across datasets with different characteristics, demonstrating its strong generalization ability. This generalization originates from the framework's emphasis on addressing universal

Table 3. Time-consuming of steps in different feature extraction and matching.

Methods	Image Enhancement	Feature extraction	Total time
fuzzyEnhance-SIFT	5ms	30ms	35ms
Ours	0ms	19ms	19ms

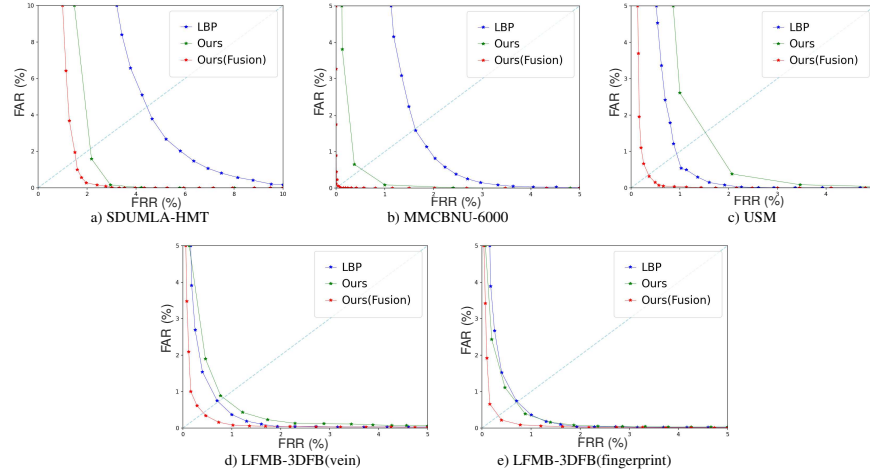


Figure 14. EER comparison of fusion recognition in different databases.

Table 4. The performance of our proposed method in different databases.

Methods	SDUMLA_HMT	MMCBNU_6000	FV-USM	LFMB-3DFB(vein)	LFMB-3DFB(fingerprint)
LBP	4.45%	1.62%	0.94%	1.38%	0.73%
Ours	2.15%	0.52%	1.53%	0.83%	0.7%
Ours(Fusion)	1.54%	0.05%	0.36%	0.42%	0.33%

Table 5. Comparison with the state-of-the-art in SDUMLA-HMT.

Reference	Approach	EER%
Vega et al. ³¹	Morphologic operation of skeletonization	27.56
Wang et al. ³²	Gabor wavelet features	2.36
Qiu et al. ³³	Dual-sliding window and pseudo-elliptical transformer	1.59
The proposed method	Ours(Fusion)	1.54

Table 6. Comparison with the state-of-the-art in MMCBNU_6000.

Reference	Approach	EER%
Xie S J et al. ³⁴	Guided filter based single scale retinex	1.52
Meng X et al. ³⁵	Local directional code	1.03
Lu Y et al. ³⁶	Histogram of salient edge orientation map	0.94
The proposed method	Ours(Fusion)	0.05

Table 7. Comparison with the state-of-the-art in FV-USM.

Reference	Approach	EER%
Wang et al. ³²	Gabor wavelet features	4.75
Asaari et al. ³⁷	Band limited phase correlation	2.34
Qiu et al. ³³	Dual-sliding window and pseudo-elliptical transformer	2.32
Qin et al. ³⁸	Deep representation-based feature extraction	1.69
The proposed method	Ours(Fusion)	0.36

Table 8. Comparison with the state-of-the-art in LFMB-3DFB.

Reference	Approach	EER%(vein)	EER%(fingerprint)	EER%(vein+fingerprint)
Wang et al. ³²	Gabor wavelet features	2.12%	1.88%	1.24%
Asaari et al. ³⁷	Band limited phase correlation	1.46%	1.53%	0.89%
Qiu et al. ³³	Dual-sliding window and pseudo-elliptical transformer	1.38%	1.43%	0.85%
Qin et al. ³⁸	Deep representation-based feature extraction	0.97%	1.02%	0.72%
The proposed method	Ours(Fusion)	0.42%	0.33%	0.12%

challenges rather than relying on dataset-specific heuristics: by targeting fundamental issues of region misalignment and deformation (instead of optimizing for idiosyncratic data traits), the framework proves to be a practical solution for pose-robust

finger biometric recognition.

Conclusion

This study addresses the performance degradation of finger identification systems caused by longitudinal rotation in contactless image acquisition, which leads to inconsistent imaging regions and non-rigid deformations. To solve these core issues, a novel pose-robust finger identification framework is proposed, integrating mutual region extraction, minimum convolution point feature learning, and weighted score fusion. The mutual region extraction strategy adaptively aligns inconsistent imaging areas via vertical gradient calculation and minimum distance metrics, effectively eliminating longitudinal rotation-induced content discrepancies. The minimum convolution point feature learning module, combined with a deformation-aware matching algorithm, captures discriminative non-linear structural features and enhances robustness against finger deformation. The weighted fusion of structural and texture features fully leverages their complementary advantages, further improving recognition discriminability. Extensive experiments on four public datasets, SDUMLA-HMT, MMCBNU_6000, FV-USM, LFMB-3DFB demonstrate that the proposed framework achieves state-of-the-art methods in terms of recognition accuracy and efficiency, verifying its effectiveness and generalization.

In future work, we aim to extend the framework to achieve robustness against arbitrary finger poses (e.g., lateral tilt, pitch, and compound poses), moving beyond the current focus on longitudinal rotation. Additionally, we will explore the integration of the proposed handcrafted features with deep learning models to further enhance feature representation capability, expanding the framework's applicability to more complex practical scenarios.

References

1. Minaee, S., Abdolrashidi, A., Su, H., Bennisamoun, M. & Zhang, D. Biometrics recognition using deep learning: A survey. *Artif. Intell. Rev.* **56**, 8647–8695 (2023).
2. Shakeel, M. S., Zhang, Y., Wang, X., Kang, W. & Mahmood, A. Multi-scale attention guided network for end-to-end face alignment and recognition. *J. Vis. Commun. Image Represent.* **88**, 103628 (2022).
3. Chen, Y. *et al.* Accurate iris segmentation and recognition using an end-to-end unified framework based on madnet and dsanet. *Neurocomputing* **517**, 264–278 (2023).
4. Qiu, Y. *et al.* Ifvit: Interpretable fixed-length representation for fingerprint matching via vision transformer. *IEEE Transactions on Inf. Forensics Secur.* (2024).
5. Yang, Z., Luo, D., Zeng, M., Wan, H. & Kang, W. An image super-resolution based method for palmprint recognition. In *Chinese Conference on Biometric Recognition*, 100–109 (Springer, 2024).
6. Huang, J., Bhattacharjee, S., Marcel, S. & Kang, W. Study of full-view finger vein biometrics on redundancy analysis and dynamic feature extraction. *IEEE Transactions on Inf. Forensics Secur.* (2025).
7. Qin, H., Gong, C., Li, Y., Gao, X. & El-Yacoubi, M. A. Label enhancement-based multiscale transformer for palm-vein recognition. *IEEE Transactions on Instrumentation Meas.* **72**, 1–17 (2023).
8. Yang, W., Chen, Z., Huang, J. & Kang, W. A novel system and experimental study for 3d finger multibiometrics. *IEEE Transactions on Biom. Behav. Identity Sci.* **4**, 471–485 (2022).
9. Ren, H., Sun, L., Guo, J. & Han, C. A dataset and benchmark for multimodal biometric recognition based on fingerprint and finger vein. *IEEE Transactions on Inf. Forensics Secur.* **17**, 2030–2043 (2022).
10. Priesnitz, J., Huesmann, R., Rathgeb, C., Buchmann, N. & Busch, C. Mobile contactless fingerprint recognition: implementation, performance and usability aspects. *Sensors* **22**, 792 (2022).
11. Duan, Y., Feng, J., Lu, J. & Zhou, J. Estimating fingerprint pose via dense voting. *IEEE Transactions on Inf. Forensics Secur.* **18**, 2493–2507 (2023).
12. Liu, F., Yang, G., Yin, Y. & Wang, S. Singular value decomposition based minutiae matching method for finger vein recognition. *Neurocomputing* **145**, 75–89 (2014).
13. Song, W. *et al.* A finger-vein verification system using mean curvature. *Pattern Recognit. Lett.* **32**, 1541–1547 (2011).
14. Zhu, Y., Yin, X., Jia, X. & Hu, J. Pair-relationship modeling for latent fingerprint recognition. *arXiv preprint arXiv:2207.00587* (2022).
15. Wu, J.-D. & Liu, C.-T. Finger-vein pattern identification using svm and neural network technique. *Expert. Syst. with Appl.* **38**, 14284–14289 (2011).

16. Rojas, A. & Dolecek, G. J. Fingerprint verification system based on dwt, multiple domain feature extraction, and ensemble subspace classifier. *Memoria Investig. en Ingenieria* 32–49 (2022).
17. Zhang, Z. & Wang, M. Multi-feature fusion partitioned local binary pattern method for finger vein recognition. *Signal, Image Video Process.* **16**, 1091–1099 (2022).
18. Momani, A. A. & Kóczy, L. T. A novel fingerprint identification fuzzy system using a center-distance weighted local binary pattern. In *Computational Intelligence and Mathematics for Tackling Complex Problems* 5, 33–40 (Springer, 2024).
19. Kim, H.-G. *et al.* Illumination normalization for sift based finger vein authentication. In *International Symposium on Visual Computing*, 21–30 (Springer, 2012).
20. Cao, K. & Jain, A. K. Automated latent fingerprint recognition. *IEEE transactions on pattern analysis machine intelligence* **41**, 788–800 (2018).
21. Wen, Q. E., Kamaruddin, N. M. & Rosdi, B. A. Raspberry pi-based finger vein recognition system using pcanet. In *Journal of Physics: Conference Series*, vol. 1529, 022068 (IOP Publishing, 2020).
22. Zhang, Y. *et al.* Agcnn: adaptive gabor convolutional neural networks with receptive fields for vein biometric recognition. *Concurr. Comput. Pract. Exp.* **34**, e5697 (2022).
23. Qiu, S., Liu, Y., Zhou, Y., Huang, J. & Nie, Y. Finger-vein recognition based on dual-sliding window localization and pseudo-elliptical transformer. *Expert. Syst. with Appl.* **64**, 618–632 (2016).
24. Tareen, S. A. K. & Raza, R. H. Potential of sift, surf, kaze, akaze, orb, brisk, agast, and 7 more algorithms for matching extremely variant image pairs. In *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 1–6 (IEEE, 2023).
25. Yin, Y., Liu, L. & Sun, X. Sdumla-hmt: A multimodal biometric database. In *Chinese conference on biometric recognition*, 260–268 (Springer, 2011).
26. Lu, Y., Xie, S. J., Yoon, S., Wang, Z. & Park, D. S. An available database for the research of finger vein recognition. In *2013 6th International congress on image and signal processing (CISP)*, vol. 1, 410–415 (IEEE, 2013).
27. Asaari, M. S. M., Suandi, S. A. & Rosdi, B. A. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert. Syst. with Appl.* **41**, 3367–3382 (2014).
28. Yang, W., Chen, Z., Huang, J., Wang, L. & Kang, W. Lfmb-3dfb: A large-scale finger multi-biometric database and benchmark for 3d finger biometrics. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, 1–8 (IEEE, 2021).
29. Chen, J. *et al.* Wld: A robust local image descriptor. *IEEE transactions on pattern analysis machine intelligence* **32**, 1705–1720 (2009).
30. Dalal, N. & Triggs, B. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, 886–893 (Ieee, 2005).
31. Vega, A. P., Travieso, C. M. & Alonso, J. B. Biometric personal identification system based on patterns created by finger veins. In *3rd IEEE International Work-Conference on Bioinspired Intelligence*, 65–70 (IEEE, 2014).
32. Wang, R., Wang, G., Chen, Z., Zeng, Z. & Wang, Y. A palm vein identification system based on gabor wavelet features. *Neural Comput. Appl.* **24**, 161–168 (2014).
33. Qiu, S., Liu, Y., Zhou, Y., Huang, J. & Nie, Y. Finger-vein recognition based on dual-sliding window localization and pseudo-elliptical transformer. *Expert. Syst. with Appl.* **64**, 618–632 (2016).
34. Xie, S. J., Lu, Y., Yoon, S., Yang, J. & Park, D. S. Intensity variation normalization for finger vein recognition using guided filter based singe scale retinex. *Sensors* **15**, 17089–17105 (2015).
35. Meng, X., Yang, G., Yin, Y. & Xiao, R. Finger vein recognition based on local directional code. *Sensors* **12**, 14937–14952 (2012).
36. Lu, Y. *et al.* Efficient descriptor of histogram of salient edge orientation map for finger vein recognition. *Appl. Opt.* **53**, 4585–4593 (2014).
37. Asaari, M. S. M., Suandi, S. A. & Rosdi, B. A. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert. Syst. with Appl.* **41**, 3367–3382 (2014).
38. Qin, H. & El-Yacoubi, M. A. Deep representation-based feature extraction and recovering for finger-vein verification. *IEEE Transactions on Inf. Forensics Secur.* **12**, 1816–1829 (2017).

Funding

This work was supported by the National Natural Science Foundation of China(No. 62566012, No. 62661011), the Science and Technology Plan Projects of Guizhou Province(QKHJCZK[2026]MS312, QKHJCZK[2024]YB493, QKHJCZK[2026]ZD083, QKHJCZK[2026]YB246), the Natural Science Research Project of Education Department of Guizhou Province(No. QJJ2023061), the Guangdong Basic and Applied Basic Research Foundation under Grant(No.2024A1515110009) and the Funded Scientific Research Projects of Guizhou Minzu University(GZMUZK[2023]YB04).

Data Availability Statement

Detailed information and access approaches for the dataset analysed during the current study are described as follows:

1. The SDUMLA-HMT dataset²⁵ is publicly available on the official website of the Machine Learning and Applications Lab at Shandong University: <https://time.sdu.edu.cn/kycg/gksjk.htm>.

2. The MMCBNU_6000 dataset²⁶ was originally released by the Multimedia Lab of Jeonbuk National University. Its official website is no longer accessible due to institutional website adjustment. Complete dataset specifications, acquisition protocols and background information are documented in its official peer-reviewed publication: Lu, Y., Xie, S. J., Yoon, S., Wang, Z. H., & Park, D. S. (2013). An available database for the research of finger vein recognition. In *2013 6th International Congress on Image and Signal Processing (CISP)* (Vol. 1, pp. 410–415). IEEE. <https://doi.org/10.1109/CISP.2013.6744030>.

3. The FV-USM dataset²⁷ is a publicly available finger vein benchmark dataset released by Universiti Sains Malaysia. It can be accessed at the official website: http://drfendi.com/fv_usm_database/.

4. The LFMB-3DFB dataset²⁸ is a publicly available large-scale 3D finger multi-biometric benchmark dataset, released by the Biometric and Intelligent Perception Lab (BIP-Lab) of South China University of Technology. It can be accessed at the official GitHub repository: <https://github.com/SCUT-BIP-Lab/LFMB-3DFB>.

Author contributions statement

Weili Yang: Conceptualization, Writing-original draft, Methodology; Dewen Kong: Writing-original draft; Fujian Feng: Conceptualization, Data curation, Validation; Yonghuai Yang: Conceptualization, Data curation; Li Deng: Validation, Data curation; Liu ZhongChi: Validation, Data curation; Dacan Luo: Supervision, Data curation, Conceptualization