

# Phase-Aggregated Dual-Branch Network for Efficient Fingerprint Dense Registration

Xiongjun Guan<sup>1</sup>, Jianjiang Feng<sup>1</sup>, *Member, IEEE*, and Jie Zhou<sup>1</sup>, *Senior Member, IEEE*

**Abstract**—Fingerprint dense registration aims to finely align fingerprint pairs at the pixel level, thereby reducing intra-class differences caused by distortion. Unfortunately, traditional methods exhibited subpar performance when dealing with low-quality fingerprints while suffering from slow inference speed. Although deep learning based approaches shows significant improvement in these aspects, their registration accuracy is still unsatisfactory. In this paper, we propose a Phase-aggregated Dual-branch Registration Network (PDRNet) to aggregate the advantages of both types of methods. A dual-branch structure with multi-stage interactions is introduced between correlation information at high resolution and texture feature at low resolution, to perceive local fine differences while ensuring global stability. Extensive experiments are conducted on more comprehensive databases compared to previous works. Experimental results demonstrate that our method reaches the state-of-the-art registration performance in terms of accuracy and robustness, while maintaining considerable competitiveness in efficiency.

**Index Terms**—Fingerprint, registration, distortion, deep neural network, phase.

## I. INTRODUCTION

**B**IOMETRIC systems identify individuals through anatomical or behavioral characteristics. Over the past few decades, many biometric traits have been suggested such as fingerprint, face, iris, vein, signature, and voice. Among them, fingerprint is widely used in civil and criminal applications due to its high distinguishability and stability. Although researchers have proposed many fingerprint matching algorithms, it is still a challenging task to deal with intra-class differences resulting from fingerprint distortion [1].

Fingerprint dense registration algorithms aims to establish dense correspondences at the pixel-level between fingerprint pairs. Ridge curves of the registered fingerprint pairs can be strictly aligned, thus significantly reducing the negative impact of skin distortion and improving matching accuracy [2]. Additionally, multiple fingerprints of the same finger can be mosaicked into a complete image after registration, thereby reducing storage requirements and expanding the effective

Manuscript received 20 September 2023; revised 15 January 2024; accepted 14 May 2024. Date of publication 20 May 2024; date of current version 28 May 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 61976121 and Grant 62376132. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Fernando Alonso-Fernandez. (Corresponding author: Jianjiang Feng.)

The authors are with the Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: gxj21@mails.tsinghua.edu.cn; jfeng@tsinghua.edu.cn; jzhou@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TIFS.2024.3403507

1556-6021 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.  
See <https://www.ieee.org/publications/rights/index.html> for more information.

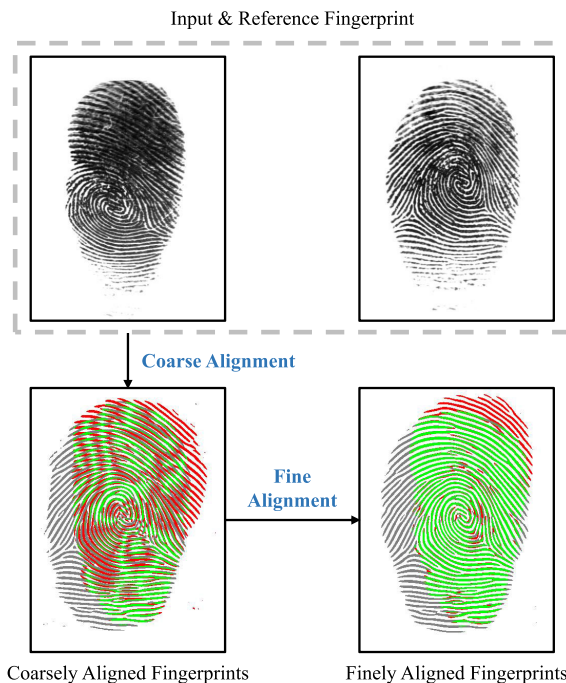


Fig. 1. Flowchart of proposed two-step fingerprint dense registration. Green areas indicate overlapping ridges, gray and red indicate non-overlapping ridges of the two fingerprints respectively.

matching area (especially important for small fingerprints) [3]. Moreover, the registration results can serve as ground truth for certain tasks, such as distortion rectification [4] or modality transformation [5], which can conveniently obtain precise real data instead of synthetic.

Early conventional fingerprint registration algorithms commonly employ spatial transformation models based on matched minutiae pairs [6], [7], [8]. However, these methods cannot achieve precise alignment at ridge level and exhibit instability in regions with sparse minutiae or low quality. In subsequent studies, it is developed and subdivided into two steps: first, a coarse alignment is conducted using rigid transformation or Thin Plate Spline transformation (TPS), then a fine alignment is performed to make local precise adjustments on ridge curves. Although traditional dense registration methods [2], [9], [10] are theoretically able to align local ridges strictly, they are very slow and susceptible to noise in practice. On the other hand, dense registration methods based on convolutional neural networks [3], [11] shows significant advantages in efficiency and robustness, but still needs to be improved in accuracy. Therefore, we are

motivated to design an innovative fingerprint dense registration method to integrate the advantages of both types of fingerprint dense registration solutions. Phase features [9], which can be easily obtained through 2D complex Gabor filters and accurately describe the offset of ridge points, meets our demands conceptually. Inspired by multi branch networks [12], [13], [14], [15], we explicitly introduce the phase feature and make it as a branch of proposed network to better express the dense mapping relationship between fingerprint pairs.

In this paper, we propose a Phase-aggregated Dual-branch Registration Network (PDRNet) to combine the strengths of traditional fingerprint dense registration methods and deep learning. The flowchart of our registration algorithm is shown in Fig. 1. For a fingerprint pair, named input fingerprint and reference fingerprint respectively, we first apply the TPS transformation based on matching minutiae for coarsely alignment (a robust rigid transformation method is used as the guarantee), then estimate the dense deformation field through PDRNet and perform fine alignment. In contrast to previous deep learning methods that simply use pass-through [11] or encoder-decoder structures [3], PDRNet introduces a dual-branch structure with multi-stage interactions between correlation features (specifically referred to phase in this paper) at high resolution and texture features at low resolution, to perceive local fine differences while ensuring global stability. Moreover, the deformation field is estimated based on probability distribution of discretized intervals instead of previous direct regression methods, inspired by [16], [17], and [18].

We conducted extensive experiments on databases containing different types of fingerprint impressions, including different sensing technologies (optical, thermal wiped, latent, non-contact) and different skin conditions (normal, dry, wet, distorted, incomplete, aged). Experimental results demonstrate that the proposed algorithm achieves state-of-the-art registration performance, while also having great competitiveness in terms of model size and inference speed.

The main contributions of this work can be summarized as follows:

- Phase features that perform well in traditional fingerprint registration [9] are introduced into the proposed convolutional neural network, enabling it to perceive more refined correlation information;
- We introduce a multi-stage feature interaction mechanism to handle dual-branch information, which can combine the robustness of low-resolution texture features and the sensitivity of high-resolution correlation features.
- The numerical regression task in previous works is transformed into interval classification, further improving the performance of deformation field estimation.
- Extensive experiments are conducted on more abundant datasets compared to previous studies. Comprehensive evaluation of registration and matching performance is performed across multiple fingerprints impressions.

The paper is organized as follows. Section II reviews the related works. Section III introduces the proposed dense registration algorithm. Section IV describes experiment datasets

we used in training and evaluation, and implementation details. Section V presents the experimental results and discussions. Finally, we make conclusions in Section VI.

## II. RELATED WORK

Briefly, fingerprint registration aims to establish the correspondence between input fingerprints and reference fingerprints, and then transform the fingerprint pairs to align them as closely as possible. According to the transformation model, fingerprint registration can be divided into two categories: rigid transformation and elastic registration. From the distribution density of control points, elastic registration can be further subdivided into sparse or dense registration.

### A. Fingerprint Rigid Transformation

Early fingerprint registration studies mostly utilize rigid transformations, which only consider the relative translation, rotation, or scaling relationships between image pairs. These algorithms commonly estimate transformation parameters by minimizing the projection error of matching minutiae [7], [19], [20], [21], orientation field [22], [23] or image correlation [24], [25]. On the other hand, some researchers recently utilize Spatial Transformer Network (STN) combined with subsequent recognition tasks to estimate affine transformation parameters of each input fingerprint [26], [27], or use more complex networks to directly estimate the fingerprint pose [28], [29]. Although these algorithms are applied to a single fingerprint, they have a certain ability to roughly align fingerprint pairs. Obviously, these models lack the capability to handle the nonlinear distortion of fingerprints.

### B. Fingerprint Sparse Registration

Sparse registration models calculate transformation parameters based on the position of control points. Among them, TPS based models that use minutiae correspondences are widely used [6], [8], [30], [31] because it can align the extracted stable feature points while ensuring smooth mapping relationships. However, these methods cannot accurately register areas without or far from minutiae. Researchers have introduced various extended features, such as the orientation field [32], period map [33], ridge curve [34] and network descriptor [17], [35], to improve the accuracy of feature point extraction and matching. Nevertheless, these advancements do not fully address the issue of unreliable results in low-quality areas. Furthermore, it poses a significant challenge to establish dense correspondence between featureless ridge points solely based on sparsely distributed control points in space.

### C. Fingerprint Dense Registration

Fingerprint dense registration aims to establish pixel by pixel correspondence between fingerprint pairs. Researchers typically use methods introduced in Section II-A or Section II-B to roughly align fingerprints and then make fine adjustments in the following stage.

Si et al. [2] proposed a dense registration method based on block-based image correlation and Markov optimization, which performs better than sparse registration methods.

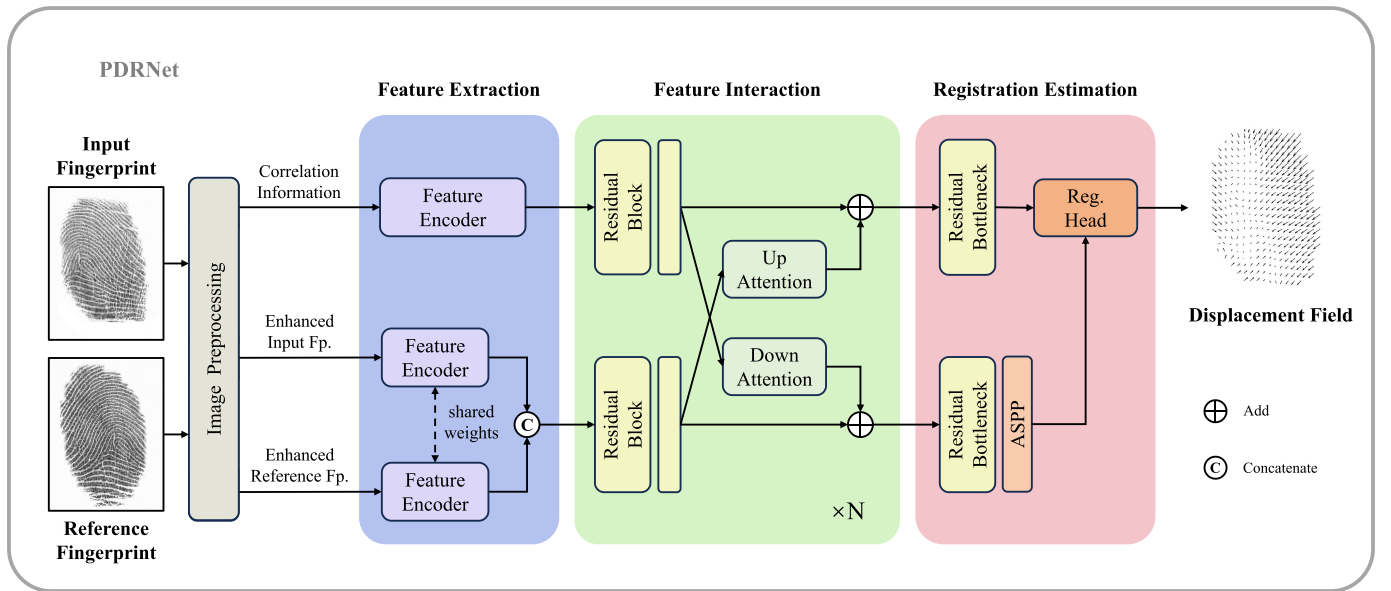


Fig. 2. An overview of our dense deformation estimation network. ‘Fp.’ and ‘Reg.’ are the abbreviations of ‘fingerprint’ and ‘registration’ respectively. The network includes two encoder branches for extracting features of texture and correlation information, a multi-stage interaction module for fusing multi-scale and multi-semantic features, and an registration estimation module to predict the pixel-wise deformation field. The specific details of block architectures and image preprocessing flow are shown in Fig. 3 and Fig. 4 respectively.

Cui et al. [9] introduced the concept of phase demodulation into fingerprint registration, resulting in substantial improvements in both speed and accuracy. Lan et al. [10] also proposed a simple but effective registration algorithm, which implements iterative optimization based on correlation and orientation field, but sometimes be unstable on images with significant grayscale variations. This kind of methods all depend on manually designed image features and are susceptible to areas with low quality or large distortion. In addition, their optimized ways of iteration or traversal are computationally expensive.

Convolutional neural networks are used to regress pixel-wise displacements in methods of Cui et al. [3], [11]. Compared with traditional algorithms, deep learning based methods show desirable advantages in terms of robustness and efficiency. However, their works simply connected a siamese network with straight network [11] or encoder-decoder [3] without incorporating additional constraints on the learned features, which might not be sensitive enough to differences in fingerprint pairs.

### III. METHOD

In this paper, we estimate the pixel-level displacement fields between fingerprint pairs in two stages to perform dense registration. Fig. 1 gives the complete flowchart of the proposed algorithm. For a pair of fingerprints, our algorithm first apply the TPS model based on matching minutiae for coarse alignment. A robust rigid transformation method based on orientation and period map is used as a fallback. The preliminarily aligned fingerprints are then preprocessed as enhancement images and correlation information (determined as phase in final approach) and input into the proposed network, whose structure is shown in Fig. 2, so as to obtain a dense displacement field from the input fingerprint to the

reference fingerprint. Finally, the input image is moved pixel by pixel according to the network prediction result to approximate the reference fingerprint.

#### A. Coarse Alignment

Same as previous works [2], [9], [3], [11], we perform coarse registration based on fingerprint minutiae. VeriFinger [36], a widely used commercial software, is utilized for minutiae extraction and matching. The paired points are subsequently used as control points for TPS transformation to coarsely align the input and reference fingerprints. It should be noted that the TPS model may not be reliable when the number of paired minutiae is small. Therefore, we switch to conduct a global search for the optimal rigid transformation based on orientation and period maps in order to obtain a more accurate alignment result, due to their advantages in registration robustness in these cases [37]. The function for parameter search is defined as

$$\begin{aligned} \operatorname{argmax}_{x,y,\theta} \quad & \| \text{OriDiff} (O_R, O_I(x, y, \theta)) \leq \theta_t \ \& \\ & \text{PedDiff} (P_R, P_I(x, y, \theta)) \leq p_t \|_0, \end{aligned} \quad (1)$$

where  $x$ ,  $y$  and  $\theta$  represent the translation and rotation parameters,  $O$  is the orientation map,  $P$  is the period map, subscripts I and R represent the input and reference fingerprint respectively, functions  $\text{OriDiff}()$  and  $\text{PedDiff}()$  calculate the difference of orientation and period maps at the corresponding location. Fixed thresholds  $\theta_t$  and  $p_t$  are set to  $10^\circ$  and 1 pixel. Both orientation and period map are sampled on blocks of  $8 \times 8$  blocks. We perform the above rigid transformation when paired minutiae are less than 4, otherwise use TPS-based sparse elastic registration.

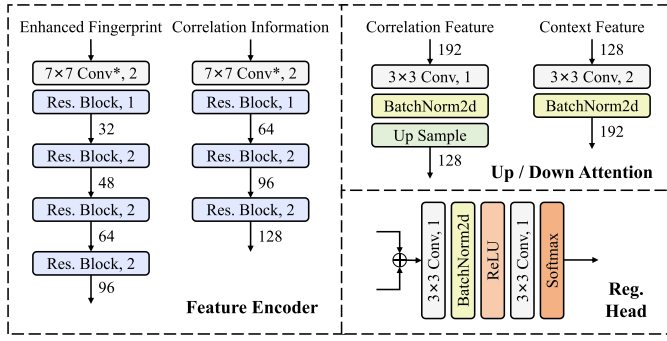


Fig. 3. The specific architecture of blocks utilized in the proposed fingerprint dense registration network. “\*” indicates that this layer connects convolution, batch normalization and ReLU in series. Numbers on the right side of each module and arrows represent the stride and current channel number respectively.

## B. Image Preprocessing

Image quality of fingerprints greatly affects the performance of subsequent algorithms. Researchers have proposed several enhancement methods in the preprocessing stage to improve the clarity of ridge structures. Among them, context filtering in the spatial domain based on Gabor filters [38] can effectively remove undesired noise while retaining the true ridge structure, which is currently one of the most popular technologies [1]. On the other hand, phase features can also be calculated through Gabor filter banks [17]. In other words, we can naturally obtain phase feature from the intermediate results of enhancement without additional complex operations. Therefore, we integrate fingerprint enhancement and phase feature extraction in the same module. For convenience, this module is shown in Fig. 2 as the first module of the proposed network, but it can also be executed independently of PDRNet.

The 2-D Gabor filter combines Gaussian waves and sine waves. Its expression in the complex domain after adjustment for fingerprints is:

$$G(x, y) = \exp\left(-\left(\frac{x'^2}{2\sigma_x^2} + \frac{y'^2}{2\sigma_y^2}\right)\right) \exp(i \cdot 2\pi f_0 x'), \quad (2)$$

where  $\sigma_x$  and  $\sigma_y$  are the standard deviation of corresponding direction,  $x'$  and  $y'$  are the original coordinates  $x$  and  $y$  rotated by angle  $\theta$ ,  $f_0$  is the filtering frequency. The  $\theta$  and  $f_0$  parameters should be adjusted according to the orientation and period of local ridges in practical. Let  $Z(x, y)$  denote the filtered result with  $G(x, y)$ , the enhanced fingerprint  $E$  and phase  $\phi$  of the original image can be calculated as

$$E = \text{Norm}(\text{Re}[Z]),$$

$$\phi(x, y) = \text{atan2}(\text{Re}[Z(x, y)], \text{Im}[Z(x, y)]), \quad (3)$$

where  $\text{Norm}()$  is to perform global normalization,  $\text{Re}[Z]$  and  $\text{Im}[Z]$  are the real and imaginary components of the complex signal  $Z$  respectively.

A partial structure of FingerNet [17] is isolated and adjusted (according to Equation 3) for this task because it can obtain the above features quickly and accurately. As show in Fig. 5, this network converts the Gabor filter bank as a set of convolution kernels with fixed parameters and estimates the orientation

field through another branch to select appropriate filtering results. The mask of the fingerprint is also estimated. Other enhancement methods based on Gabor filtering can also be used to obtain similar results.

In addition to enhancement images  $E$  and phases  $\phi$ , mask  $M$  is also utilized to reduce interference in non-overlapping areas. For the input fingerprint  $I$  and reference fingerprint  $R$ , the final results of image preprocessing is

$$E'_I = E_I \cdot M_I, \quad E'_R = E_R \cdot M_R,$$

$$\psi = (\phi_I - \phi_R) \cdot M_I \cdot M_R, \quad (4)$$

where  $\psi$  represents the correlation information,  $E'_I$  and  $E'_R$  represent the enhanced input and reference fingerprint respectively. An example is given in Fig. 4, which visualizes the intermediate and final results in the image preprocessing stage.

## C. Network for Dense Displacement Field Estimation

For a pair of fingerprints, our proposed network predicts the dense deformation field from the input fingerprint to the reference fingerprint. A two-branch structure is introduced, which extracts features of different semantics at different resolutions respectively and performs information interaction in multiple stages, to understand local ridge differences finely and robustly. In this paper, phase feature are utilized as correlation information because it is easy to obtain and can accurately describe the displacement relationship at pixel level [9]. The complete structure of PDRNet is given in Fig. 2, which can be divided into four parts: image preprocessing, feature extraction, feature interaction and registration estimation. In Section III-B, we have described how to perform image preprocessing on fingerprint pairs. The following mainly introduces the design motivation and specific implementation of remaining architectures and loss functions of PDRNet.

1) *Feature Extraction*: A natural idea is to extract texture features directly from images. In work [3] and [11], a siamese network with shared parameters was applied, which first extracts the respective texture features from each image separately and then merges and analyzes them. These implementations are similar to the control points based elastic registration in traditional methods. As analyzed in Section II-B, even if stable feature points can be extracted and matched, their number is not enough to establish a pixel-wise dense connections. Therefore, we introduce correlation information that can accurately represent the displacement relationship and explicitly use it as network input, while retaining the previous disposal mode of texture features. As shown in Fig. 3, downsampled convolutions are stacked as feature encoders in order to balance registration accuracy and model parameters. For features of correlation information, we only downsample them to 1/8 size to retain richer spatial details. On the other hand, texture features are calculated using a deeper encoder (1/16 of the original size) to extract more abstract semantic information to increase the robustness of registration.

Heatmap of features extracted by the respective encoders are shown in Fig. 6. It can be seen that the texture branch

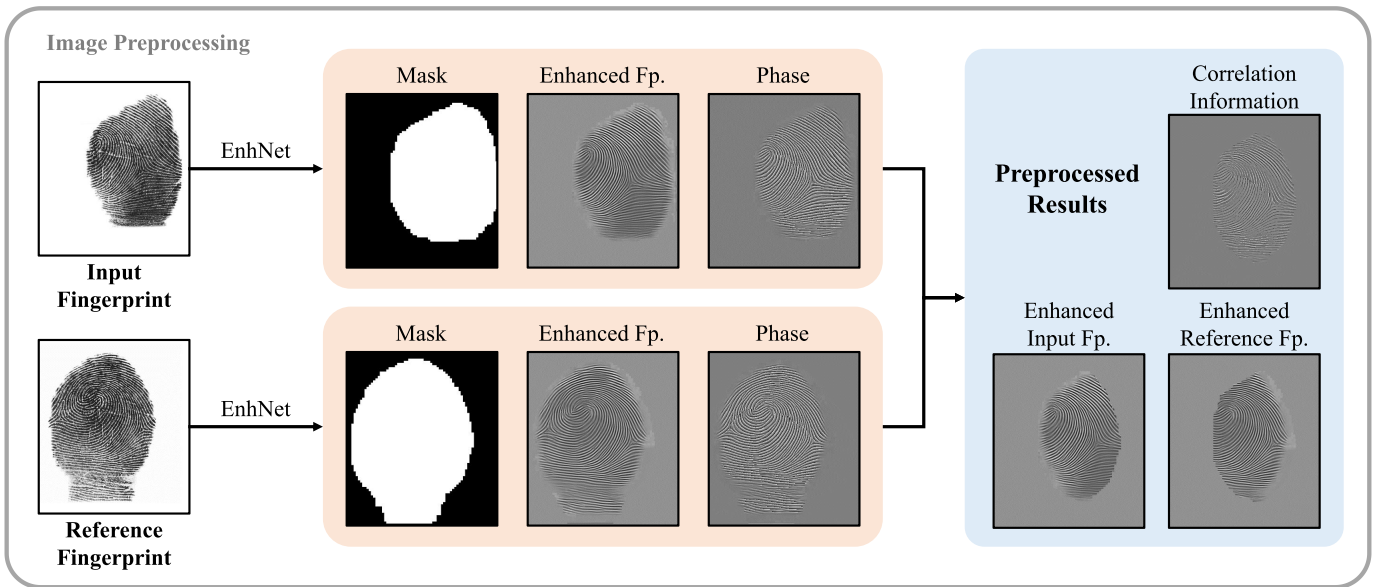


Fig. 4. A visual example of image preprocessing. Red rectangles represent the intermediate results of fingerprint enhancement. Blue rectangle represents the final result of image preprocessing, which corresponds to the output of the same module in Fig. 2. The structure of ‘EnhNet’ is shown in Fig. 5.

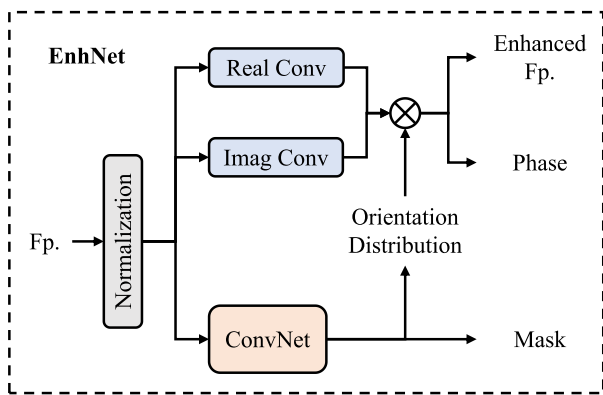


Fig. 5. A simplified schematic of fingerprint enhancement network. The main structure is separated from FingerNet [17], and only the output is adjusted according to Equation 3.

without additional constraints pays more attention to point-level features, and the high corresponding areas are sparsely distributed and relatively rough; while features inferred from correlation information pay more attention to regional properties, and the high-response areas are densely distributed and relatively fine. This case strongly demonstrates that the proposed two branches effectively capture information with distinct semantics. One reasonable explanation is that the input to texture feature branch is the respective enhanced result and there is no information exchange between enhanced fingerprint pairs during feature extraction, which will make the corresponding encoder tend to focus on stable features from a single image in the form of points (such as minutiae and singular points) and ignore low-texture areas that are highly repetitive or easily confused. On the other hand, the input to correlation information branch is the phase difference between fingerprint pairs, whose value of each pixel clearly characterizes the offset of ridge alignment [9], resulting in a

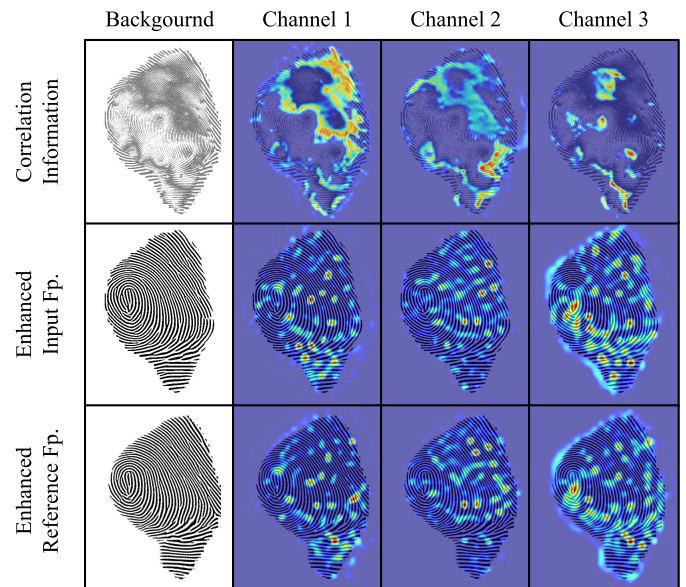


Fig. 6. Examples of heatmap visualized from extracted features, which reflects the interesting areas of encoders from different branches. Each row, from left to right, is the background of subsequent images and attention maps from three channels output by the respective encoders. The background from bottom to top is the binary image of two fingerprints and the XOR result between them.

regional response in the feature distribution due to the spatial continuity of fingerprint deformation.

2) *Feature Interaction*: Previous works [3], [11] have verified that networks can establish the relationship between fingerprint pairs from extracted texture features. Furthermore, the studies conducted by [9] and [39] proved that networks have the ability to unwrap phase and thus resolve the displacement between fingerprints. Therefore, it is theoretically feasible to analyze the displacement through dual-branch features in PDRNet. Inspired by other multi-branch network structure designs [12], [13], [14], [15] and multi-stage feature iteration

strategies [40], [41], we introduce a bilateral structure to analyze and interact the information from two branches in several iterations. Residual blocks [42] are exploited to further express features and avoid gradient problems, while simple convolution and sampling layers are used for information exchange between branches. It should be noted that the shape of each feature is not changed in order to facilitate arbitrary adjustment of iterations.

3) *Registration Estimation*: The iterated results are first fed into a residual bottleneck [42] to further integrate features, where the channels are doubled to carry denser information. The ASPP module [43] (with dilated ratio of 1,2,4 and average pooling) is connected next to capture texture information at multiple scales and fuse features. Here we perform such a transformation flow because phase is used as correlation information in this paper, which can be mapped to displacement more directly and interpretably than texture features [9]. Fig. 3 shows the specific structure of the registration head. Inspired by [16], [17], and [18], we use interval classification to express the deformation field, which can limit the range to avoid unreasonable output and characterize the numerical distribution thus reduce the learning difficulty, instead of previous direct regression [3], [11]. Let  $p^t$  and  $z^t$  represent the numerical value and predicted probability corresponding to  $t$ -th category, the final deformation field  $D$  is calculated as

$$D = \frac{1}{\sum_T p^t} \sum_T p^t z^t, \quad (5)$$

and subsequently interpolated to the full size. Considering that large relative distortions are basically removed through the coarse alignment in Section III-A, we set the value range of displacement in both directions to  $[-30, 30]$  pixels and divide it into 25 equal intervals, that is, 50 channels are output with 1/8 size of original image. The deformation field in non-overlapping regions can be approximated by TPS interpolation based on estimated results in overlapping regions.

4) *Loss Function*: Considering that directly using 0-1 labels cannot accurately express values especially at the boundaries of intervals, we use Gaussian smoothing label strategy to generate the ground truth of class probability  $p_{\text{gt}}$  as

$$p_{\text{gt}}^t(z) = \frac{1}{\sum_t p_{\text{gt}}^t(z)} \exp\left(-\frac{z - z^t}{2\sigma^2}\right), \quad (6)$$

where  $z$  represents the real displacement,  $z^t$  represents the discrete value corresponding to the  $t$ -th category. The variance  $\sigma$  is set to 2.0 empirically. Focal loss [44] is applied to optimize these imbalanced multi-class distributions, which is defined as

$$\begin{aligned} \mathcal{L}_{\text{cla}} &= -\frac{1}{|M|} \sum_M \sum_{t=1}^T \alpha (1 - q^t)^\gamma \log(q^t), \\ q^t &= p_{\text{gt}}^t p^t + (1 - p_{\text{gt}}^t) (1 - p^t), \end{aligned} \quad (7)$$

where  $M$  is the mask of the common area of fingerprint pairs,  $p^t$  is the probability that the deformation is inferred in  $t$ -th category,  $T$  is the total number of intervals. Hyperparameters  $\alpha$  and  $\gamma$  are fixed to 1.0 and 2.0 respectively. Another

objective function  $\mathcal{L}_{\text{smo}}$  is also performed to constrain the spatial smoothness of probability distribution:

$$\mathcal{L}_{\text{smo}} = -\frac{1}{|M|} \sum_M \sum_{t=1}^T |\Delta p^t|, \quad (8)$$

where  $\Delta$  is the standard Laplacian filter. Overall, the complete loss function is

$$\mathcal{L} = \mathcal{L}_{\text{cla}} + \lambda \cdot \mathcal{L}_{\text{smo}}. \quad (9)$$

The weight  $\lambda$  is fixed to 1.0 in this paper.

#### IV. DATASET DESCRIPTION

Extensive experiments are conducted on multiple databases containing fingerprints with different impressions, including widely used public datasets *FVC2004 DB1\_A & DB3\_A* [45], *NIST SD27* [46], *Tsinghua Distorted Fingerprint Database (TDF)* [47] and several private datasets *Hisign Latent*, *THU Old*, *Hisign Multi-pose Plain Fingerprint Database (Hisign MPF)*, *Hisign Contact-based 2D to Multi-pose Contactless 2D Fingerprint Database (Hisign C2CL)*. Table I provides a comprehensive description about the composition and usage of these datasets. In the following we will introduce the details of training and test data used in experiments.

##### A. Training Data Building

Similar to previous deep learning methods [3], [11], we extract the real distortion field from *TDF* and use it to transform fingerprints in *Hisign Latent*, which enables us to obtain a large number of fingerprint pairs in two impressions and the ground truth of corresponding displacements. Specifically, we use VeriFinger [36] to extract and track minutiae of distorted fingerprint videos in *TDF*, and calculate TPS transform as the distortion field through paired points of first and last frames. For a certain fingerprint  $I$  in *Hisign Latent*, a deformation field  $D$  in *TDF* is randomly selected and used for synthesis as:



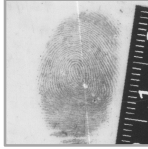



$$\begin{aligned} I'(x + D_x, y + D_y) &= I(x, y), \\ F(x, y) &= D(x, y), \\ F'(x + D_x, y + D_y) &= -D(x, y), \end{aligned} \quad (10)$$

where  $I'$  is the conjugate fingerprint,  $x$  and  $y$  represent the corresponding direction components,  $F$  and  $F'$  represent the displacement field registered from  $I$  to  $I'$  and the opposite scenario respectively. In this way, a total of 20,918 sets similar to  $\{I, I', F\}$  are conveniently generated and used for network training. In order to increase the diversity of fingerprint poses and categories, data augmentation strategies are used in the training stage, which include mirror flipping, rotation (by 90, 180 or 270 degree) and swapping (as  $\{I', I, F'\}$ ).






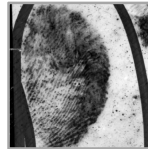
##### B. Matching Protocols in Testing

In order to balance the number of genuine and imposter matches, specific protocols are set for different databases, as shown in Table I. There are also some other settings: (i) considering *FVC2004 DB1\_A* contains some strongly distorted fingerprints that may affect the accuracy of coarse





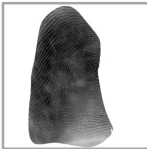

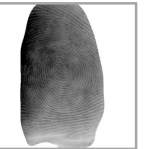
TABLE I  
ALL FINGERPRINT DATASETS USED IN EXPERIMENTS

Database	Hisign Latent			TDF		THU Old
	Rolled / Plain		Latent	Plain		Plain
Image						
Sensor	Inking / Optical		-	Optical		Optical
Description	10459 pairs latent fingerprints from real crime scenes			320 videos large distortion		136 fingers × 2 wrinkled and low quality
Usage	Training			Training		Registration accuracy Matching performance
Genuine Match	\			\		136 <sup>a</sup>
Imposter Match	\			\		9,180 <sup>b</sup>

Database	FVC2004 DB1_A			FVC2004 DB3_A	NIST SD27	
	Plain			Plain	Rolled	Latent
Image						
Sensor	Optical			Thermal sweeping	Inking	-
Description	100 fingers × 8 large distortion and various quality			100 fingers × 8 large distortion	258 pairs latent fingerprints from real crime scenes	
Usage	Registration accuracy Matching performance			Registration accuracy Matching performance	Matching performance	
Genuine Match	2,800 <sup>a</sup>			2,800 <sup>a</sup>	258 <sup>a</sup>	
Imposter Match	4,950 <sup>b</sup>			4,950 <sup>b</sup>	66,306 <sup>c</sup>	

Database	Hisign MPF			Hisign C2CL			
	Plain			Contact-based (C)	Contactless (CL)		
Image							
Sensor	Optical			Optical	Camera		
Description	100 fingers × 3 poses (left / middle / right)			200 fingers × 1	200 fingers × 3 poses (left / middle / right)		
Usage	Registration accuracy			Registration accuracy Matching performance			
Genuine Match	200 <sup>d</sup>			400 (CL-CL) <sup>d</sup> , 600 (C-CL) <sup>e</sup>			
Imposter Match	\			19,900 (CL-CL) <sup>f</sup> , 19,900 (C-CL) <sup>f</sup>			

<sup>a</sup> Each fingerprint is matched with other fingerprints from the same finger. The symmetric matches are avoided.

<sup>b</sup> First fingerprints of each finger are matched each other. The symmetric matches are avoided.

<sup>c</sup> Each latent fingerprint is matched with all rolled fingerprints from other fingers.

<sup>d</sup> All middle fingerprints are matched with the corresponding fingerprints on left and right side poses. The symmetric matches are avoided.

<sup>e</sup> Each contact-based fingerprint is matched with other contactless fingerprints from the same finger. The symmetric matches are avoided.

<sup>f</sup> A contactless fingerprint with specified pose in each group is sequentially selected and matched with other corresponding fingerprints. The symmetric matches are avoided.

alignment, we implemented a fingerprint distortion rectification process [4] marked with \* in experiments to distinguish from the original dataset. (ii) we do not evaluate the matching performance in *Hisign MPF*, because the genuine and imposter scores of original images are already perfectly separated; (iii) considering efficiency or fingerprint quality, subsets of *Hisign C2CL* and *THU Old* are selected and used in

experiments; (iv) in all datasets containing three poses, the genuine matching between the left and right side poses is not calculated because their overlapping area is too small; (v) contactless-contactless (CL-CL) and contact-contactless (C-CL) matching experiments are implemented in *Hisign C2CL* to evaluate the performance of fingerprint registration on multi-modality.

TABLE II  
REGISTRATION ACCURACY OF DIFFERENT FINGERPRINT REGISTRATION ALGORITHMS

Method	FVC2004 DB1_A		FVC2004 DB1_A*		Hisign MPF		THU Old		Hisign C2CL(CL-CL)		Hisign C2CL(C-CL)	
	NCC	VF	NCC	VF	NCC	VF	NCC	VF	NCC	VF	NCC	VF
TPS Based	0.20	366	0.19	362	0.22	320	0.20	322	0.15	361	0.09	321
Phase Based [9]	<b>0.65</b>	404	<b>0.67</b>	400	0.62	350	0.63	350	<b>0.60</b>	391	<b>0.56</b>	354
DRN (local) <sup>†</sup> [11]	0.52	293	0.54	289	0.51	267	0.52	269	0.46	266	0.43	221
DRN (global) <sup>†</sup> [3]	0.55	388	0.58	388	0.56	340	0.56	343	0.51	375	0.45	335
Proposed <sup>†</sup>	<b>0.65</b>	<b>417</b>	<b>0.67</b>	<b>410</b>	<b>0.64</b>	<b>357</b>	<b>0.65</b>	<b>353</b>	0.59	<b>397</b>	<b>0.56</b>	<b>368</b>

<sup>†</sup> Registration method based on neural networks.

V. EXPERIMENTS

In this section, we compare the proposed method with state-of-the-art algorithms. TPS transformation based on matching minutiae is used as a benchmark for comparison due to its simplicity and practicality. In addition, we compared phase registration [9], which is representative of traditional methods, and two typical deep learning methods that estimate displacement locally [11] (called DRN (local)) or globally [3] (called DRN (global)). The performance of fingerprint dense registration schemes are comprehensively evaluated in terms of registration accuracy, matching performance, and efficiency. Moreover, ablation experiments are conducted to demonstrate the effectiveness of modules and strategies in PDRNet.

A. Evaluation Protocols

Image correlator and VeriFinger [36] are used to reflect the similarity between two fingerprints, consistent with previous works [2], [3], [9], [10], and [11]. Let  $I$  and  $M$  represent the image and mask of fingerprints, for any two fingerprints 1 and 2 the correlation score is calculated as

$$NCC = \frac{\sum_M (I_1 - \bar{I}_1) \cdot (I_2 - \bar{I}_2)}{\sqrt{\sum_M (I_1 - \bar{I}_1)^2 \cdot \sum_M (I_2 - \bar{I}_2)^2}}, \quad (11)$$

where  $M$  is the common area of  $M_1$  and  $M_2$ ,  $\bar{I}$  represents the mean value of image  $I$  in  $M$ . This metric can sensitively reflect the degree of ridge overlapping and is easy to implement. Since the experiments mainly focus on the relative improvement brought by registration algorithms rather than the absolute performance, the image correlator can be regarded as a representative of image-based matchers. On the other hand, we choose to employ VeriFinger SDK 12.0 [36], a widely used commercial software, to measure the alignment of minutiae due to its superior performance in this type of matchers. For convenience, the matching score of VeriFinger is referred to as  $VF$  in the following. It should be noted that matching methods based on global descriptors, such as [5] and [26], are not applied in this paper because they may not pay equal attention to every local fine structures, that is to say, the scores obtained in this manner lack sufficient interpretability for the performance of dense fingerprint registration. Furthermore, fixed-length global descriptors have limited capability in distinguishing relative translation differences of ridges compared to image correlation, which is precisely concerned in dense fingerprint registration.

B. Registration Accuracy

In this subsection, genuine matching are conducted using image correlator and VeriFinger matcher to quantitatively evaluate the contribution of dense registration algorithms to the alignment of ridges and minutiae. We utilize databases *FVC2004 DB1\_A*, *Hisign MPF*, *THU Old* and *Hisign C2CL* because they fit the scenarios in practical applications, such as fingerprint mosaicking and cross-modal comparison.

As shown in Table II, the proposed method significantly outperforms other learning based methods in all datasets and maintains advantages over conventional methods. This suggests that our network is capable of establishing dense relationships between fingerprint pairs more precisely while ensuring high stability of relative structures in space.

Several typical examples are given in Fig. 7 in order to show the effect of fingerprint registration methods more intuitively. From the comparisons on the left, it can be seen that phase registration [9] is difficult to deal with scenes of incomplete (low-quality areas in row 1), complex texture structures (singular regions in rows 2 and 3) and interlaced ridges (folds and distortion in row 4). On the other hand, the local displacement registration network [11] is more robust in the above areas, but it only gives a local optimal solution which is burr in space and obviously reflected in registration results. Subsequent improvements proposed by Cui et al. [3] significantly improved the global smoothness, but there are still misalignments in some ridge areas. One reasonable explanation is that the network tends to learn features of point structures without additional guidance (see Fig. 6), which are just lacking in these problem regions. Our method integrates the advantages of traditional algorithm and deep learning, showing higher accuracy and stability in registration.

Meanwhile, the representative failure cases on the right of Fig. 7 show that our algorithm still needs to be improved in some extreme scenarios, such as: (i) image defects caused by low contrast or extensive wrinkles (row 1 and 2); (ii) lack of sufficient reference information due to limited overlap (row 3); (iii) severe spatial structural dislocation caused by large distortion (row 4).

It should be mentioned that the performance of phase based registration [9] in this paper is higher than those reported in previous papers [3], [9], [10], [11]. This is because we use a newer version of VeriFinger (from 6.2 to 12.0) which performs better in extracting binary images. Fig. 8 shows the impact of software version on phase registration.



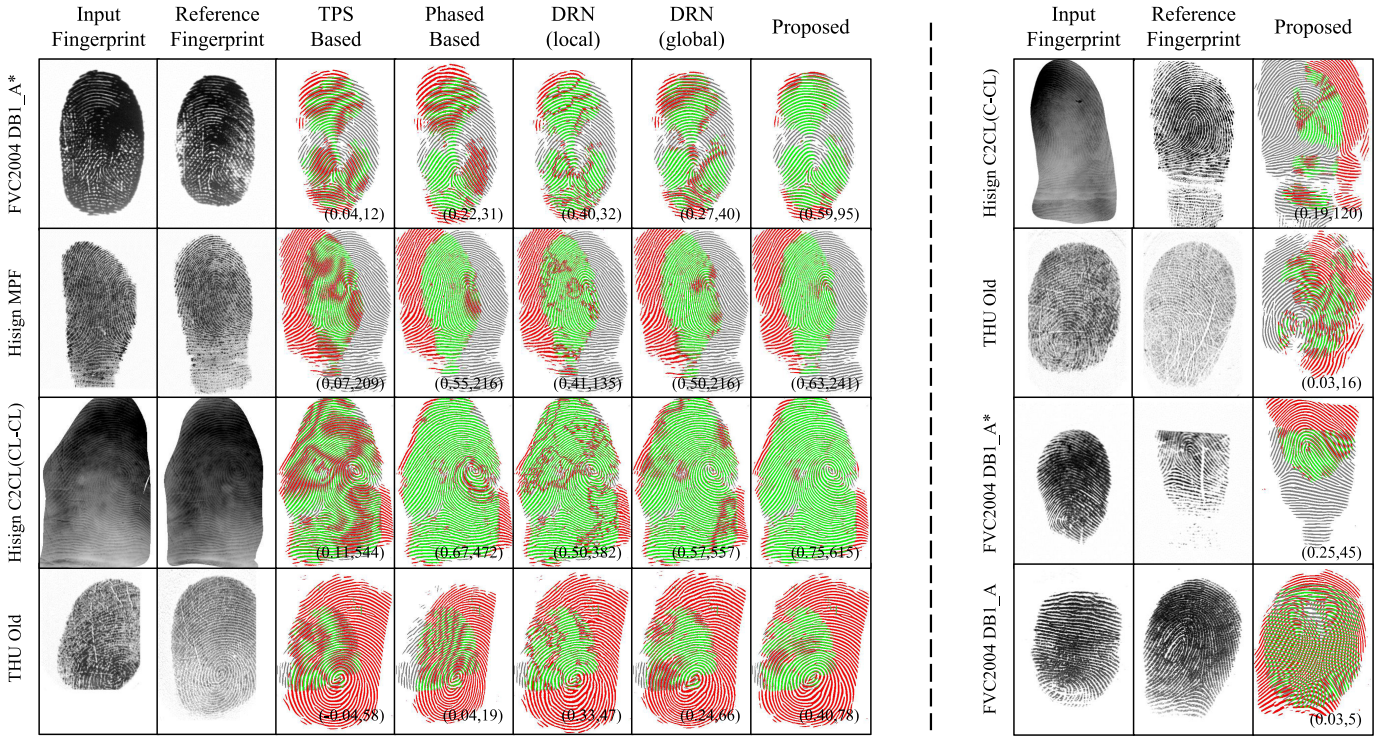


Fig. 7. Examples of fingerprint registration for genuine matching fingerprints. The left side compares the performance of different registration methods in typical scenarios. Some representative cases where our approach fails are shown on the right. The beginning of each row gives the name of corresponding datasets, and numbers in brackets are matching scores by image correlator and VeriFinger. Green indicates overlap, while red and gray indicate non-overlapping areas of respective fingerprints.

TABLE III  
MATCHING PERFORMANCE BY IMAGE CORRELATOR WITH DIFFERENT FINGERPRINT REGISTRATION ALGORITHMS

Method	FVC2004 DB1_A			FVC2004 DB1_A*			FVC2004 DB3_A			THU Old			Hisign C2CL(CL-CL)			Hisign C2CL(C-CL)		
	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER
TPS Based	70.2	33.8	11.5	72.6	46.9	11.6	74.6	65.4	10.4	70.6	37.2	11.7	61.4	78.1	12.5	45.3	99.6	22.3
Phase Based [9]	97.8	3.16	1.12	98.7	2.78	0.64	93.1	9.23	1.79	94.9	9.53	3.68	97.8	<b>13.0</b>	0.96	96.2	48.3	1.17
DRN (local) <sup>†</sup> [11]	91.1	25.5	2.49	87.2	20.8	2.75	72.5	46.4	5.98	76.5	53.5	8.65	71.8	82.7	5.75	67.0	92.3	6.67
DRN (global) <sup>†</sup> [3]	95.4	12.1	2.42	97.5	8.00	1.10	91.8	30.1	2.01	91.9	14.7	2.21	95.8	28.2	1.25	95.3	42.8	1.19
Proposed <sup>†</sup>	<b>98.9</b>	<b>1.54</b>	<b>0.79</b>	<b>99.8</b>	<b>0.41</b>	<b>0.17</b>	<b>98.4</b>	<b>2.78</b>	<b>1.24</b>	<b>98.5</b>	<b>2.21</b>	<b>1.47</b>	<b>99.5</b>	21.7	<b>0.25</b>	<b>99.5</b>	<b>4.16</b>	<b>0.33</b>

<sup>†</sup> Registration method based on neural networks.

TAR represents TAR@FAR = 0.1%, FMR represents ZeroFMR.

TABLE IV  
MATCHING PERFORMANCE BY VERIFINGER MATCHER WITH DIFFERENT FINGERPRINT REGISTRATION ALGORITHMS

Method	FVC2004 DB1_A			FVC2004 DB1_A*			FVC2004 DB3_A			THU Old			Hisign C2CL(CL-CL)			Hisign C2CL(C-CL)		
	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER	TAR	FMR	EER
TPS Based	98.8	1.86	0.75	99.4	1.61	0.45	98.9	<b>1.32</b>	0.70	96.3	5.15	1.60	99.7	0.69	0.26	99.7	<b>0.33</b>	0.32
Phase Based [9]	<b>99.2</b>	<b>1.04</b>	0.61	<b>99.5</b>	<b>0.57</b>	0.41	98.6	1.81	0.80	97.1	3.41	2.20	99.7	<b>0.25</b>	0.25	99.7	0.91	0.32
DRN (local) <sup>†</sup> [11]	97.2	5.32	1.62	97.7	2.68	1.28	95.9	5.64	2.13	93.4	8.09	4.21	98.3	5.25	0.78	98.7	4.33	1.20
DRN (global) <sup>†</sup> [3]	98.6	2.37	0.83	99.4	1.14	0.36	98.8	1.71	0.72	97.8	4.32	2.25	99.5	1.00	0.49	99.7	0.83	0.32
Proposed <sup>†</sup>	<b>99.2</b>	1.16	<b>0.58</b>	<b>99.5</b>	0.71	<b>0.23</b>	<b>99.1</b>	1.52	<b>0.66</b>	<b>98.5</b>	<b>1.47</b>	<b>1.41</b>	<b>100</b>	0.50	<b>0.02</b>	<b>99.8</b>	0.58	<b>0.16</b>

<sup>†</sup> Registration method based on neural networks.

TAR represents TAR@FAR = 0.1%, FMR represents ZeroFMR.

### C. Matching Performance

Further experiments are conducted on several databases with multiple modalities, including three widely used public datasets *FVC2004 DB1\_A*, *FVC2004 DB3\_A*, *NIST SD27* and two additional private datasets *THU Old*, *Hisign C2CL*,

to examine the assistance of fingerprint registration algorithms on matching performance. Similar to previous works, we evaluate the similarity score of mated or non-mated fingerprint pairs using image correlator and VeriFinger matcher.

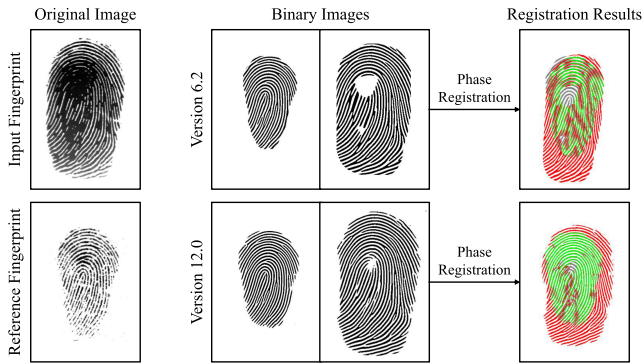


Fig. 8. Registration results based on phase information with binary images extracted by different versions of VeriFinger. Green indicates overlap of two fingerprint ridges, while gray and red indicate no overlap.

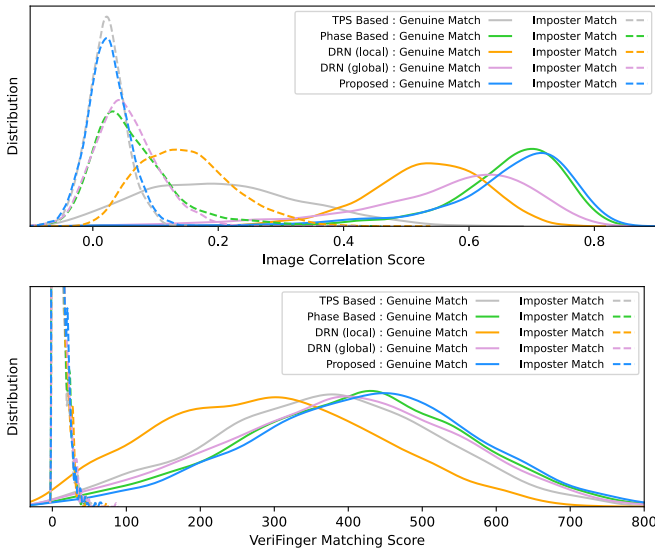


Fig. 9. Distribution of genuine and imposter matching scores on FVC2004 DB1\_A by image correlator (top) and VeriFinger matcher (bottom). The scale on vertical axis is not displayed because we are more concerned with the relative values of probability density.

We first present the score distribution of genuine and imposter matches on *FVC2004 DB1\_A* to qualitatively assess the performance of different registration algorithms. As shown in Fig. 9, our proposed method outperforms the others in improving genuine matching scores. In addition, our method obviously has more concentrated distribution and lower average score on imposter matches of image correlation. DRN (local) [11] exhibits the poorest performance because it lacks global constraints. Other suboptimal methods [3], [9] may align local blocks of non-mated fingerprints although global information exchange are introduced explicitly or implicitly. As analyzed above, our network aggregates the advantages of traditional method and deep learning, which simultaneously focuses on the characteristics of ridge lines and anchor points while conducting global information interaction in multiple stages, thus achieving the highest precision and stability.

Table III and Table IV show the comparison of matching performance after different registration methods. Three representative indicators True Accept Rate (TAR), False Match Rate (FMR) and Equal Error Rate (EER), which are commonly used in biometric recognition systems, are listed to briefly reflect the accuracy in identification scenarios. The corresponding

TABLE V  
ABLATION STUDY OF THE PROPOSED NETWORK WITH DIFFERENT MODULES AND STRATEGIES ON HISIGN MPF

Modules & Strategies					NCC
Correlation branch	Context branch	Information interaction	Fusion block	Reg. head	
-	✓*	-	-	cla	0.51
✓*	-	-	-	cla	0.60
✓	✓	-	✓	cla	0.57
✓	✓	✓	-	cla	0.62
✓§	✓	✓	✓	cla	0.59
✓	✓	✓	✓	reg	0.49
✓	✓	✓	✓	cla	0.64

\* For fairness, channels of corresponding models are adjusted to ensure that the model size is similar to others.

§ The difference calculated by subtracting two fingerprint images is input into this branch instead of phase in other groups.

TABLE VI  
ABLATION STUDY OF THE PROPOSED NETWORK WITH DIFFERENT NUMBERS OF STACK STAGES OF FEATURE INTERACTION MODULE ON HISIGN MPF

Stacking stages	2	3	4	5	6
NCC	0.59	0.62	0.64	0.63	0.63
Param (M)	8.71	10.9	13.0	15.2	17.4

Detection Error Tradeoff (DET) curves are shown in Fig. 10 and Fig. 11 for more complete information. Experimental results demonstrate that the proposed method surpasses other methods in almost all evaluation aspects, and its advantages in TAR and EER are particularly obvious.

Furthermore, we calculate the Cumulative Matching Characteristic (CMC) curves on latent fingerprint database *NIST SD27* to evaluate the performance of registration algorithms on low-quality fingerprints. As shown in Fig. 12, our method is most robust to these complex and difficult samples. It is worth mentioning that deep learning based method [3] perform better than traditional method [9], which is different from the results on other datasets. This phenomenon occurs because the dense presence of incomplete or contaminated areas in latent fingerprints disrupts normal feature extraction, which usually leads to misjudgments by those algorithms. In contrast, deep learning methods can reduce the interference of local misinformation by selectively using more stable features.

#### D. Ablation Study

We perform ablation studies on *Hisign MPF* and use Equation 11 to examine the performance in registration accuracy. Table V presents the experiment results of specific modules and strategies in our proposed network, where “cla” and “reg” denote the construction of registration head in form of classification (introduced in Section III-C3) or regression (same as previous works [3], [11]). The information exchange part between two branches in the feature interaction stage is called “information interaction”, and ASPP used in the registration estimation stage for improving compatibility is represented by “fusion block”. These comparisons first verify the effectiveness of correlation information compared to directly extracting texture feature without constraints, while

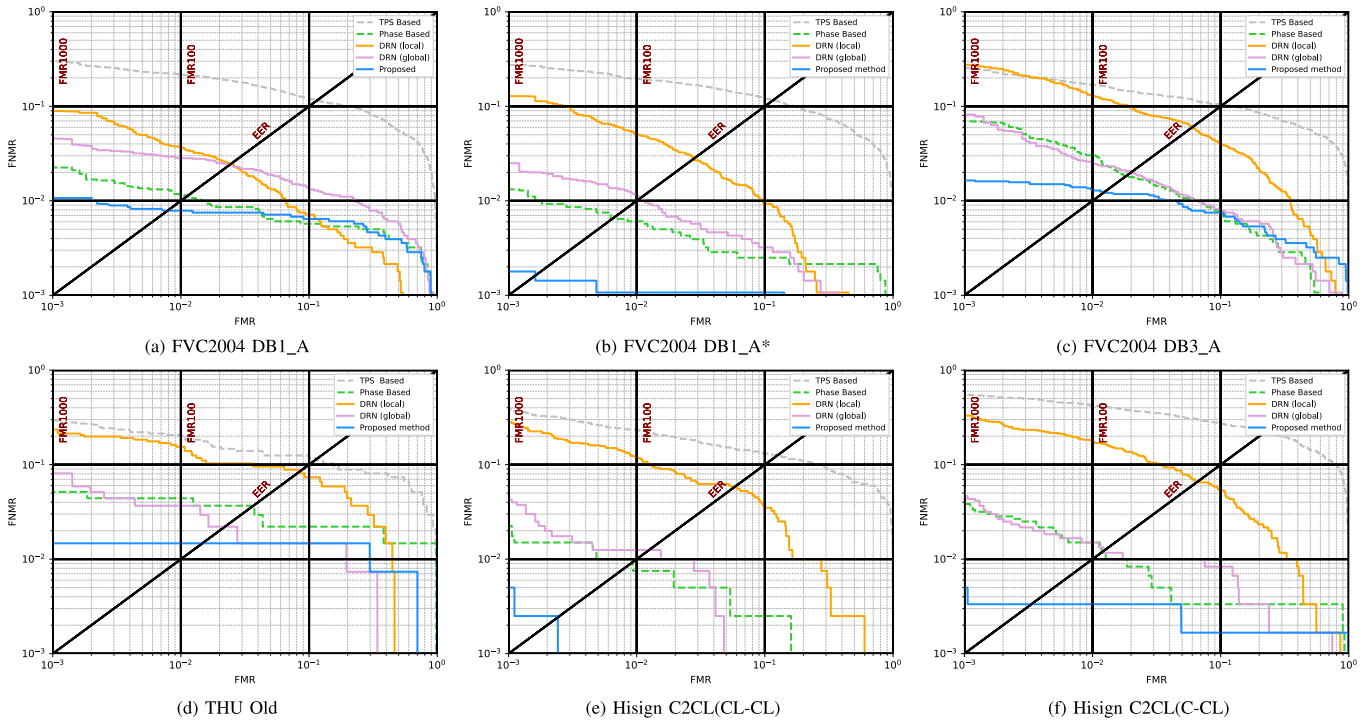


Fig. 10. DET curves by image correlator. Solid and dotted lines represent deep learning methods and traditional methods respectively.

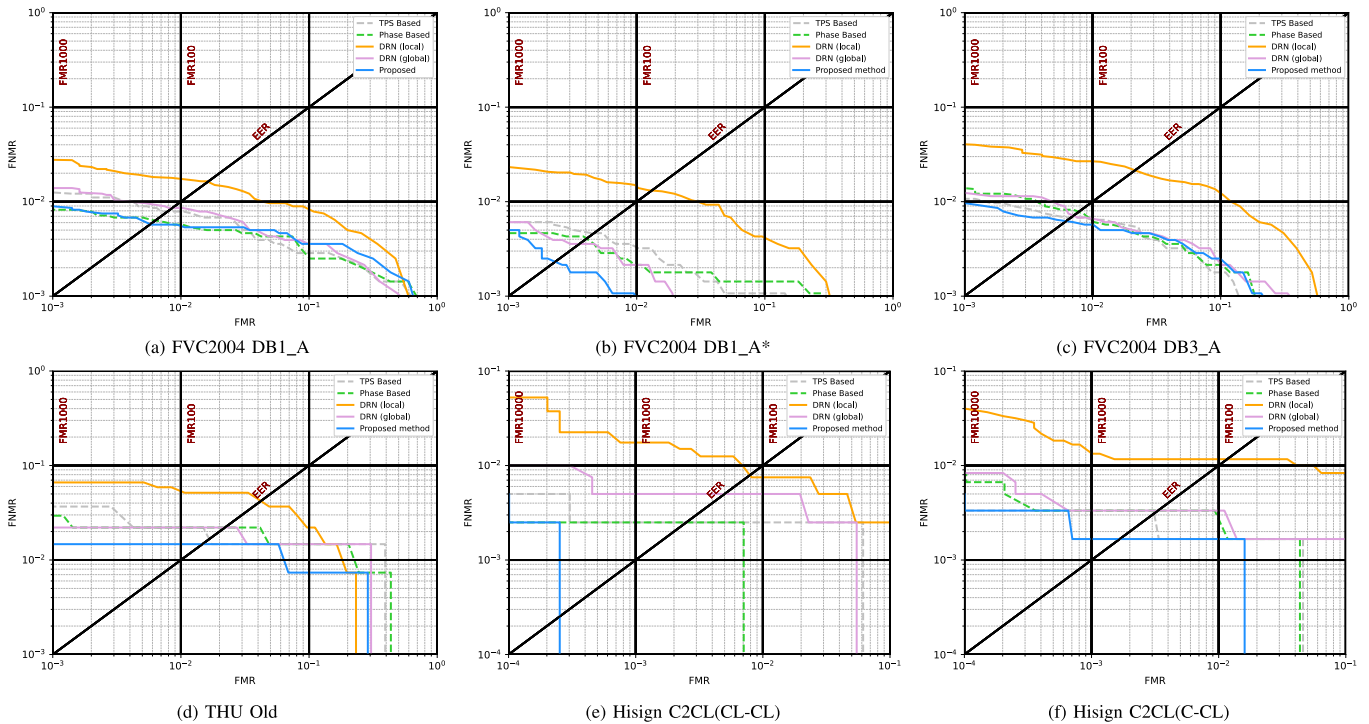


Fig. 11. DET curves by VeriFinger matcher. Solid and dotted lines represent deep learning methods and traditional methods respectively.

proving that integrating the two can achieve better performance. In particular, using phase as correlation information is significantly superior to image difference because phase contains richer information about the direction and value of displacement at each pixel [9]. The results also strongly validate the positive effects of specific designs in our network, including dual semantic feature fusion strategy, corresponding

auxiliary module ASPP and the decision of output format. In our proposed network, the stack number of information interaction module can be freely adjusted according to actual data scale. Table VI shows corresponding ablation results under the data protocol of this paper, where the accuracy keeps improving with more stacking stages until the model overfits after 4 steps.

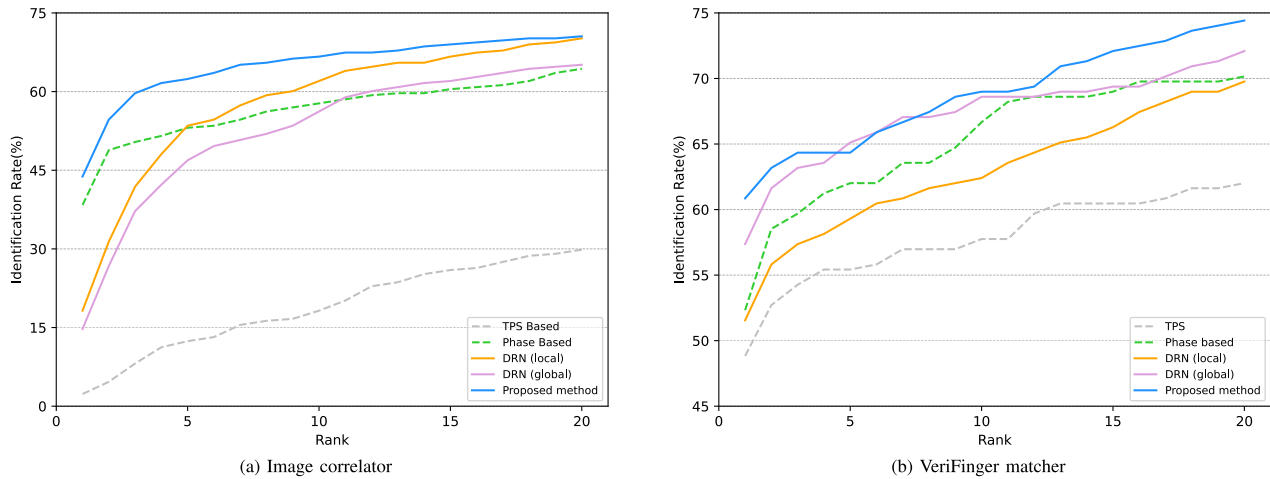


Fig. 12. CMC curves by image correlator and VeriFinger matcher with different fingerprint registration algorithms on NIST SD27. Solid and dotted lines represent deep learning methods and traditional methods respectively.

TABLE VII

MODEL SIZE AND AVERAGE TIME COST OF DIFFERENT FINGERPRINT REGISTRATION ALGORITHMS FOR PROCESSING A  $640 \times 640$  FINGERPRINT PAIR IN HISIGN MPF

Method	Param (M)	Time (s)
TPS Based	-	0.41
Phase Based [9]	-	21.3
DRN (local) <sup>†</sup> [11]	0.20	1.17
DRN (global) <sup>†</sup> [3]	65.0	0.56
Proposed <sup>†</sup>	13.0	0.43

<sup>†</sup> Registration method based on neural networks.

E. Efficiency Analysis

Model size and inference speed of different fingerprint registration algorithms on *Hisign MPF* are listed in Table VII. The time covers the process from inputting a pair of  $640 \times 640$  fingerprints to outputting the corresponding deformation field, which is measured on a single NVIDIA GeForce RTX 3090 GPU by setting the batch size to 1, with a 2.4 GHz CPU. All algorithms are implemented in Python. It can be seen that all deep learning based methods are significantly faster than traditional methods [9]. As can be observed from Table VII and Section V-B, V-C, our proposed fingerprint registration algorithm is quite competitive in efficiency while leading in accuracy and robustness.

VI. CONCLUSION

In this paper, we propose a fingerprint registration algorithm to estimate the dense deformation field between two fingerprints. Both high-resolution phase information extracted by Gabor filters and low-resolution texture information extracted by convolutional layers are utilized in our proposed network. A multi-stage dual-branch information interaction mechanism is introduced to aggregate features of these two semantics at different resolutions. Moreover, we use discrete classification in the output header instead of the previous direct regression, which implicitly introducing the correlation between numerical distributions. Extensive experiments demonstrate that our method surpasses state-of-the-art fingerprint registration algorithms in accuracy and robustness, while also exhibiting

notable efficiency advantages. On the other hand, current solution are still not ideal in some extreme scenarios where fingerprints are severely contaminated or barely overlap. In the future, we will further explore the design of feature extraction and multi-information fusion to overcome these difficulties.

REFERENCES

- [1] D. Maltoni, D. Maio, A. K. Jain, and J. Feng, *Handbook of Fingerprint Recognition*, 3rd ed. Berlin, Germany: Springer, 2022.
- [2] X. Si, J. Feng, B. Yuan, and J. Zhou, "Dense registration of fingerprints," *Pattern Recognit.*, vol. 63, pp. 87–101, Mar. 2017.
- [3] Z. Cui, J. Feng, and J. Zhou, "Dense registration and mosaicking of fingerprints by training an end-to-end network," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 627–642, 2021.
- [4] X. Guan, Y. Duan, J. Feng, and J. Zhou, "Regression of dense distortion field from a single fingerprint image," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 4377–4390, 2023.
- [5] S. A. Grosz, J. J. Engelsma, E. Liu, and A. K. Jain, "C2CL: Contact to contactless fingerprint matching," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 196–210, 2022.
- [6] A. M. Bazen and S. H. Gerez, "Fingerprint matching by thin-plate spline modelling of elastic deformations," *Pattern Recognit.*, vol. 36, no. 8, pp. 1859–1867, Aug. 2003.
- [7] M. Tico and P. Kuosmanen, "Fingerprint matching using an orientation-based minutia descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 1009–1014, Aug. 2003.
- [8] A. Ross, S. Dass, and A. Jain, "A deformable model for fingerprint matching," *Pattern Recognit.*, vol. 38, no. 1, pp. 95–103, Jan. 2005.
- [9] Z. Cui, J. Feng, S. Li, J. Lu, and J. Zhou, "2-D phase demodulation for deformable fingerprint registration," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 3153–3165, Dec. 2018.
- [10] S. Lan, Z. Guo, and J. You, "Pre-registration of translated/distorted fingerprints based on correlation and the orientation field," *Inf. Sci.*, vol. 520, pp. 292–304, May 2020.
- [11] Z. Cui, J. Feng, and J. Zhou, "Dense fingerprint registration via displacement regression network," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2019, pp. 1–8.
- [12] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5693–5703.
- [13] C. Yu, C. Gao, J. Wang, G. Yu, C. Shen, and N. Sang, "BiSeNet v2: Bilateral network with guided aggregation for real-time semantic segmentation," *Int. J. Comput. Vis.*, vol. 129, no. 11, pp. 3051–3068, Nov. 2021.
- [14] S. Kumaar, Y. Lyu, F. Nex, and M. Y. Yang, "CABiNet: Efficient context aggregation network for low-latency semantic segmentation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 13517–13524.
- [15] H. Pan, Y. Hong, W. Sun, and Y. Jia, "Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 3448–3460, Mar. 2023.

- [16] S. Gidaris and N. Komodakis, "LocNet: Improving localization accuracy for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 789–798.
- [17] Y. Tang, F. Gao, J. Feng, and Y. Liu, "FingerNet: An unified deep network for fingerprint minutiae extraction," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 108–116.
- [18] B. Doosti, S. Naha, M. Mirbagheri, and D. J. Crandall, "HOPE-net: A graph-based model for hand-object pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6607–6616.
- [19] N. K. Ratha, K. Karu, S. Chen, and A. K. Jain, "A real-time matching system for large fingerprint databases," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 799–813, Aug. 1996.
- [20] Y. S. Moon, H. W. Yeung, K. C. Chan, and S. O. Chan, "Template synthesis and image mosaicking for fingerprint registration: An experimental study," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Sep. 2004, p. 409.
- [21] E. Zhu, J. Yin, and G. Zhang, "Fingerprint matching based on global alignment of multiple reference minutiae," *Pattern Recognit.*, vol. 38, no. 10, pp. 1685–1694, Oct. 2005.
- [22] N. Yager and A. Amin, "Evaluation of fingerprint orientation field registration algorithms," in *Proc. 17th Int. Conf. Pattern Recognit.*, 2004, pp. 641–644.
- [23] L. Liu, T. Jiang, J. Yang, and C. Zhu, "Fingerprint registration by maximization of mutual information," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1100–1110, May 2006.
- [24] A. M. Bazen, G. T. Verwaaijen, S. H. Gerez, L. P. Veulenturf, and B. J. Van Der Zwaag, "A correlation-based fingerprint verification system," in *Proc. Workshop Circuits, Syst. Signal Process. (CSSP)*, 2000, pp. 205–213.
- [25] T. Hatano et al., "A fingerprint verification algorithm using the differential matching rate," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, vol. 3, 2002, pp. 799–802.
- [26] J. J. Engelsma, K. Cao, and A. K. Jain, "Learning a fixed-length fingerprint representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 6, pp. 1981–1997, Jun. 2021.
- [27] S. A. Grosz and A. K. Jain, "AFR-net: Attention-driven fingerprint recognition network," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 6, no. 1, pp. 30–42, Jan. 2024.
- [28] Q. Yin, J. Feng, J. Lu, and J. Zhou, "Joint estimation of pose and singular points of fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1467–1479, 2021.
- [29] Y. Duan, J. Feng, J. Lu, and J. Zhou, "Estimating fingerprint pose via dense voting," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 2493–2507, 2023.
- [30] A. Almansa and L. Cohen, "Fingerprint image matching by minimization of a thin-plate energy using a two-step algorithm with auxiliary variables," in *Proc. 5th IEEE Workshop Appl. Comput. Vis.*, Dec. 2000, pp. 35–40.
- [31] A. Ross, S. Shah, and J. Shah, "Image versus feature mosaicing: A case study in fingerprints," *Proc. SPIE*, vol. 6202, pp. 70–81, Apr. 2006.
- [32] F. Chen, J. Zhou, and C. Yang, "Reconstructing orientation field from fingerprint minutiae to improve minutiae-matching accuracy," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1665–1670, Jul. 2009.
- [33] D. Wan and J. Zhou, "Fingerprint recognition using model-based density map," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1690–1696, Jun. 2006.
- [34] J. Feng, Z. Ouyang, and A. Cai, "Fingerprint matching using ridges," *Pattern Recognit.*, vol. 39, no. 11, pp. 2131–2140, Nov. 2006.
- [35] K. Cao, D.-L. Nguyen, C. Tymoszek, and A. K. Jain, "End-to-end latent fingerprint search," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 880–894, 2020.
- [36] *VeriFinger SDK 12.0*. Accessed: Sep. 8, 2023. [Online]. Available: <https://www.neurotechnology.com/verifinger.html>
- [37] R. P. Krish, J. Fierrez, D. Ramos, J. Ortega-Garcia, and J. Bigun, "Pre-registration of latent fingerprints based on orientation field," *IET Biometrics*, vol. 4, no. 2, pp. 42–52, Jun. 2015.
- [38] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 2, no. 7, pp. 1160–1169, 1985.
- [39] G. E. Spoorthi, R. K. S. S. Gorthi, and S. Gorthi, "PhaseNet 2.0: phase unwrapping of noisy data based on deep learning approach," *IEEE Trans. Image Process.*, vol. 29, pp. 4862–4872, 2020.
- [40] J. Hur and S. Roth, "Iterative residual refinement for joint optical flow and occlusion estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5747–5756.
- [41] Z. Teed and J. Deng, "RAFT: Recurrent all-pairs field transforms for optical flow," in *Computer Vision—ECCV*, Glasgow, U.K. Cham, Switzerland: Springer, 2020, pp. 402–419.
- [42] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [43] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 801–818.
- [44] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
- [45] *FVC 2004: The Third International Fingerprint Verification Competition*. Accessed: Sep. 8, 2023. [Online]. Available: <http://bias.csr.unibo.it/fvc2004/databases.asp>
- [46] *NIST Special Database 27/27A: Fingerprint Minutiae From Latent and Matching Tenprint Images*. Accessed: Sep. 8, 2023. [Online]. Available: <https://www.nist.gov/itl/iad/image-group/nist-special-database-2727a>
- [47] X. Si, J. Feng, J. Zhou, and Y. Luo, "Detection and rectification of distorted fingerprints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 555–568, Mar. 2015.



**Xiongjun Guan** received the B.S. degree from the Department of Automation, Tsinghua University, Beijing, China, in 2021, where he is currently pursuing the Ph.D. degree under the supervision of Prof. Jianjiang Feng. His research interests include fingerprint recognition, computer vision, and pattern recognition.



**Jianjiang Feng** (Member, IEEE) received the B.Eng. and Ph.D. degrees from the School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, China, in 2000 and 2007, respectively. From 2008 to 2009, he was a Post-Doctoral Researcher with the PRIP Laboratory, Michigan State University. He is currently an Associate Professor with the Department of Automation, Tsinghua University, Beijing. His research interests include fingerprint recognition and computer vision.



**Jie Zhou** (Senior Member, IEEE) received the B.S. and M.S. degrees from the Department of Mathematics, Nankai University, Tianjin, China, in 1990 and 1992, respectively, and the Ph.D. degree from the Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China, in 1995. From 1995 to 1997, he was a Post-Doctoral Fellow with the Department of Automation, Tsinghua University, Beijing, China. Since 2003, he has been a Full Professor with the Department of Automation, Tsinghua University. In recent years, he has authored more than 300 papers in peer-reviewed journals and conferences. Among them, more than 100 papers have been published in top journals and conferences, such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON IMAGE PROCESSING, and CVPR. His research interests include computer vision, pattern recognition, and image processing. He is a fellow of IAPR. He received the National Outstanding Youth Foundation of China Award. He is an Associate Editor of IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and two other journals.