PAPER

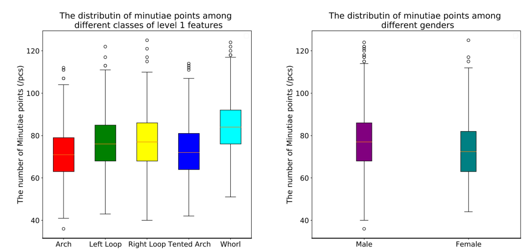
Deep Learning-inspired Automatic Minutiae Extraction From Semi-automated Annotations

Hongtian ZHAO[†a)], ***Member***, **Hua YANG**[††], ***and*** **Shibao ZHENG**[††], ***Nonmembers***

1. Introduction of Fingerprint Datasets

To obtain a suitable fingerprint training dataset with minutiae annotations, we construct a new dataset named Fingerprint Minutiae Dataset (FMD). The FMD includes the location coordinates and orientation values of each minutiae point in a fingerprint. For this dataset, we select 2910 images from the publicly available NIST SD04 dataset [1]. The NIST SD04 dataset is specifically distributed for the fingerprint classification task and contains 4000 8-bit encoded images. The selected images are categorized into five classes (Arch, Left Loop, Right Loop, Tented Arch, and Whorl) based on the pattern near the singularity points. Each image has a size of 512×512, with 32 rows of pixel blanks at the bottom. The labeling algorithm is implemented in MATLAB, and during the labeling process, we manually review and correct any inaccuracies in the minutiae annotations, involving at least two annotators. The aligned minutiae points in minutiae dataset are stable and representative, as they can be used to determine the uniqueness of a fingerprint [2–4].

The dataset comprises 2910 fingerprint images with a total of 223,207 minutiae, averaging 76.7 minutiae points per image. We conducted statistical analysis on the distribution of images and minutiae based on original classes and gender divisions. The results are summarized in Table 1. For visualization and analysis purposes, a boxplot (Figure 1) is generated, where the red line represents the median value. The number of minutiae points in Arch, Left Loop, Right Loop, Tented Arch, and Whorl images fall within the intervals of [39, 103], [42.5, 110.5], [41, 113], [38.5, 106.5], and [60, 116], respectively. For Male and Female images, the intervals are [41, 113] and [34.5, 110.5], respectively. These intervals align well with the fingerprint quality standards [40, 100] [2]. The distribution of minutiae points demonstrates a relatively balanced distribution among different classes and genders, with slightly higher median values for Whorl and Male images. This observation suggests an association between fingerprint level 1 features and gender characteristics. Additionally, a few outliers (e.g., minutiae Num $\geq$ 110.5



(a) Level-1 features (b) Genders

Fig. 1: Distribution of minutiae points across Level-1 Features and Genders.

for Left Loop) were identified outside of the main intervals. However, these minor deviations are not expected to significantly affect the labeling results, as the overall distribution of minutiae points remains fairly consistent. Therefore, the annotated minutiae dataset meets the requirements for subsequent training applications in theory.

Compared with the revoked NIST SD27, we use more fingerprint images for testing, which include 516 fingerprints and more than 258 fingerprints in the NIST SD27 dataset. We compare our minutiae dataset with the FVC 2004 dataset [5], a benchmark for fingerprint recognition. Table 2 shows the comparison results, including statistical information on the FVC 2004 dataset obtained from [6]. Compared to the standard fingerprint distribution of [40, 100], the proposed dataset exhibits a more reasonable distribution of minutiae counts in each fingerprint.

2. Comparative Analysis of Deep Learning Models for Fingerprint Minutiae Extraction

To objectively evaluate the backbone network, we benchmarked ResNet and other prevalent architectures, including VGGNet [8], InceptionNet [9], XceptionNet [12], DenseNet [13], MobileNet [10], and EfficientNet [11], in our experiments. We conducted experiments on publicly available fingerprint datasets, including NIST SD04 and FVC 2004. For each model, fingerprint feature extraction modules were implemented based on their respective core ideas. To ensure fair comparison, consistent preprocessing and augmentation were performed on all models. In this study, we employ the F_1 Score, Mean Location Error (LE), and Mean Orientation Error (OE) as the evaluation metrics. Because the overall model size remains relatively consistent, the difference in inference time can be considered negligible. Therefore, we focus solely on presenting the F_1 Score, LE,

Manuscript received January 1, 2015.
Manuscript revised January 1, 2015.
† The author is with College of Mathematics and System Science, Xinjiang University, 830046, Urumqi, China.
† The author is with SEIEE of SJTU, 200240, Shanghai, China.
a) E-mail: zhaohongtian@xju.edu.cn
DOI: 10.1587/trans.E0.??.1

Table 1: The statistics of minutiae points in different categories and genders

Tuple count	Categories of level-1 feature classification					Genders	
	Arch	Left Loop	Right Loop	Tented Arch	Whorl	Male	Female
Image	534	589	606	572	609	2410	500
Minutiae	38123	45299	46971	41572	51242	186551	36656

Table 2: The detailed attributes comparison of different datasets

Dataset	Image Size	Number of Minutiaes		
		Avg	Max	Min
FVC2004DB1A	640×480	40.96	80	11
FVC2004DB3A	300×480	40.76	76	11
NISTSD0406	512×512	78.01	114	45
NISTSD0407	512×512	79.96	121	49

and OE in our experiments. The F_1 Score, being the harmonic mean of precision and recall, offers a comprehensive representation of the overall performance of the detector. The mean localization error is a metric that quantifies the average Euclidean distance between the predicted and ground truth positions of fingerprint minutiae. The mean error of angle is a metric that assesses the average angular deviation between the predicted and actual orientations of fingerprint minutiae.

Table 3 manifests performance comparison of different models on fingerprint minutiae extraction task. Fig. 2 shows two fingerprint image samples obtained from the NISTSD 04 and FVC 2004 datasets, along with the corresponding minutiae detection results of several state-of-the-art models. In each group of images, the top image represents the deep model detection results and the corresponding ground truth, while the bottom image represents corresponding pure detection results and the ground truth minutiae comparison. We observe that although VGGNet performs well in image recognition tasks, its performance in fingerprint minutiae extraction is slightly inferior to ResNet, possibly due to its deep hierarchical structure not being suitable for capturing subtle detailed features. The Inception model, with its multi-scale convolutional kernel, is capable of capturing details at different levels. The Xception model, utilizing depthwise separable convolution, improves parameter efficiency and helps in learning finer features with limited data, achieving relatively better performance on both datasets compared to InceptionNet. However, their overall performance is not as good as ResNet. DenseNet enhances feature transfer and improves the learning of detail features through feature reuse, but it may suffer from insufficient feature reuse as depth increases, thereby affecting generalization. Moreover, deeper networks are usually harder to train due to issues like noisy gradient updates, which can affect the learning process. Therefore, models that perform well on the NIST SD04 dataset may have poor generalization ability. MobileNet is designed for mobile and embedded devices, and its lightweight structure may be beneficial for deploying fingerprint recognition systems in resource-constrained environments. However, its accuracy on NIST SD04 is relatively low. EfficientNet exhibits excellent capability in extracting complex fingerprint features, which contributes to the generation of well-generalized trained models. The findings of our study reveal that while EfficientNet shows superior performance in generalization across a multitude of models, ResNet has been adopted as the baseline for our investigation. This decision is informed by ResNet's exemplary proficiency in feature extraction, the ease with which it can be implemented and deployed, and its demonstrated robustness in accurately extracting a diverse range of fingerprint minutiae.

References

[1] NIST special database 4, Aug. 27, 2010. [Online]. https://www.nist.gov/srd/nist-special-database-4

[2] R. Bansal *et al.*, "Minutiae extraction from fingerprint images-a review," *arXiv:1201.1422*, 2011.

[3] H. Lin, W. Yifei, and A. Jain, "Fingerprint image enhancement: algorithm and performance evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 777-789, 1998.

[4] A. Chowdhury *et al.*, "Can a CNN Automatically Learn the Significance of Minutiae Points for Fingerprint Matching?," in *Proc. WACV.*, 2020, pp.351-359.

[5] FVC2004: The Third International Fingerprint Verification Competition. http://bias.csr.unibo.it/ fvc2004/.

[6] B. Zhou, C. Han, Y. Liu, T. Guo, and J. Qin, "Fast minutiae extractor using neural network," *Pattern Recognit.*, vol. 103, p. 107273, 2020.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. CVPR.*, 2016, pp. 770-778.

[8] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409.1556*, 2014.

[9] C. Szegedy, W. Liu, Y. Jia, *et al.* "Going deeper with convolutions," in *Proc. CVPR.*, 2015, pp. 1-9.

[10] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv:1704.04861*, 2017.

[11] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. ICML.*, 2019, pp. 6105-6114.

[12] F. Chollet, "Xception: Deep Learning with Depthwise

Table 3: Comparative Ablation Study of Backbone Networks for Fingerprint Feature Estimation: Evaluating the Impact of VGG, Inception, Xception, DenseNet, MobileNet, EfficientNet, and ResNet on F_1 Score, Localization Error (LE), and Orientation Error (OE).

F_1/LE/OE	NISTSD		FVC2004	
Model	0406	0407	DB1	DB2
VGGNet	0.8917/**1.7958**/0.0326	0.8907/**1.7711**/0.0331	0.3199/5.4717/0.1181	0.4620/3.8472/0.0746
Inception	0.8868/1.9996/0.0356	0.8884/1.9816/0.0354	0.1207/3.9802/0.0792	0.5403/3.7707/0.0674
Xception	0.8913/1.9612/0.0340	0.8966/1.9338/0.0330	**0.5650**/4.4865/0.0886	0.6054/3.7494/0.0672
DenseNet	0.8946/1.9749/0.0327	0.8979/1.9706/0.0332	0.1642/**3.8622**/**0.0763**	0.5035/3.8680/**0.0630**
MobileNet	0.8881/2.0644/0.0342	0.8873/2.0422/0.0341	0.4674/4.6422/0.0937	**0.7799**/3.6671/0.0665
EfficientNet	0.8821/2.3007/0.0355	0.8835/2.2956/0.0361	0.5002/5.1321/0.1011	0.7547/3.7846/0.0700
ResNet	**0.9007**/1.8708/**0.0303**	**0.9021**/1.8430/**0.0298**	0.4977/4.2298/0.0858	0.7239/**3.6379**/0.0714

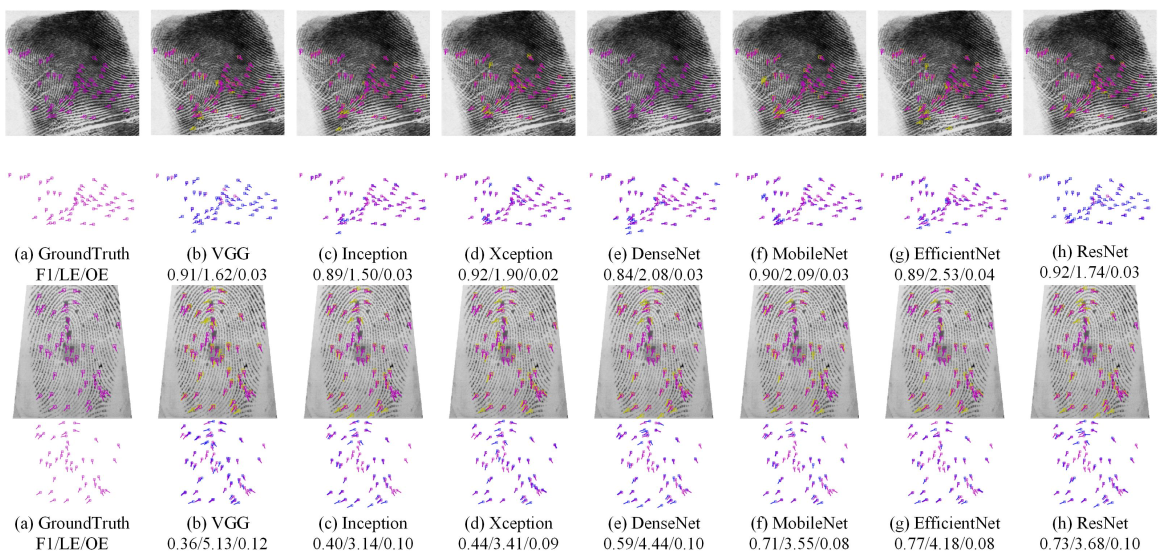


Fig. 2: Evaluation of fingerprint minutiae detection on sample images from NIST SD4 and FVC2004 datasets. In the conducted experiments, the employment of ResNet as the backbone network demonstrates superior robustness in fingerprint minutiae detection across varied image inputs.

Separable Convolutions," in *Proc. CVPR*, Jul. 2017, pp. 1800-1807.

[13] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proc. CVPR*, Jul. 2017, pp. 2261-2269.