PAPER

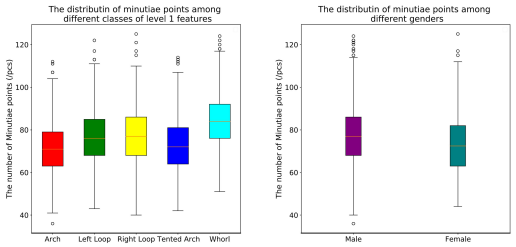
Deep Learning-inspired Automatic Minutiae Extraction From Human-in-the-loop Guided Annotations

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

1. Appendix

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.



(a) Level-1 features (b) Genders

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

Additionally, a few outliers (e.g., minutiae Num ≥ 110.5 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], our dataset exhibits a more reasonable distribution of minutiae counts in each fingerprint.

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/.

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

[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.