IVPAI 2026

The 8th International Conference on Image, Video Processing and Artificial Intelligence

Contactless fingerprint feature extraction integrating minutiae and texture features

Paper ID：26

Linqi Zhang[1], Jiaqi Gao[1], Yongzhong Hu[1]

1. Qingshuihe Campus of UESTC, University of Electronic Science and Technology of China (UESTC), Chengdu, China

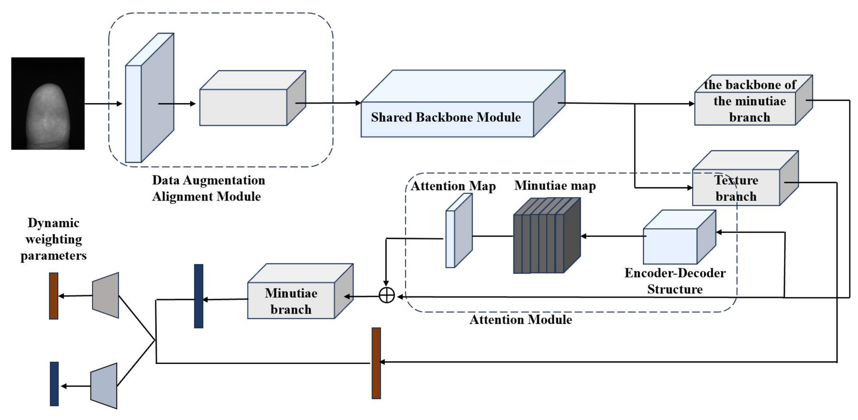
E-mail: china.hyz@263.net

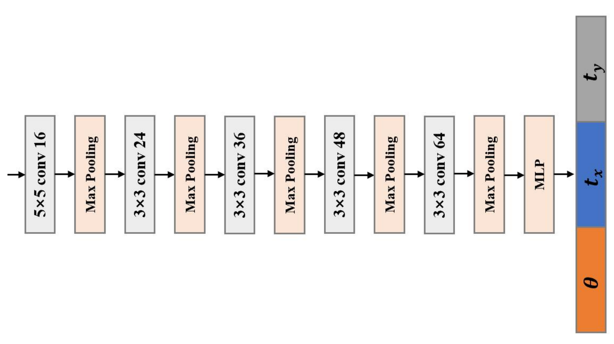
Abstract

Contactless fingerprint images offer improved hygiene and safety compared to traditional methods due to their contactless acquisition process, leading to higher user acceptance. However, images captured by mobile or high-resolution cameras exhibit significant variations from conventional contact-based fingerprints, which both promotes research in contactless fingerprint recognition and introduces substantial challenges. To address these issues, this paper proposes a convolutional neural network for fixed-length feature extraction and matching. The network integrates minutiae and texture information, and incorporates an alignment module along with data augmentation to handle geometric variations. A minutiae map is introduced to assist feature learning in the minutiae branch and to guide attention map generation within the attention mechanism. Additionally, supervision is imposed on the attention branch to facilitate end-to-end training. To address the multi-task learning setting, a dynamic weighting strategy is adopted to automatically learn the relative importance of different tasks. Experimental results demonstrate that the proposed components jointly improve recognition performance. Compared with traditional methods and minutiae-based approaches, the proposed method achieves significant performance gains, validating its effectiveness.

Methods

- Overall architecture design



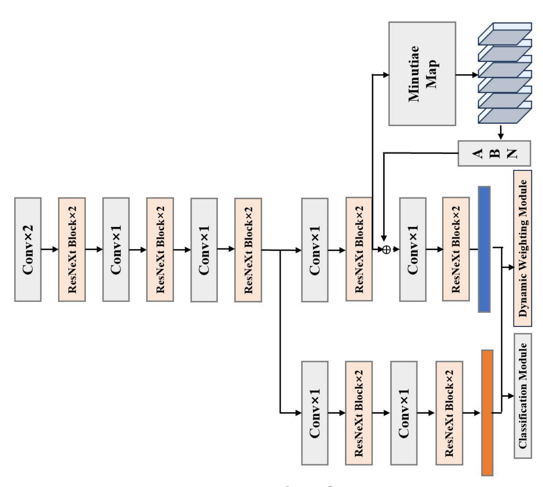


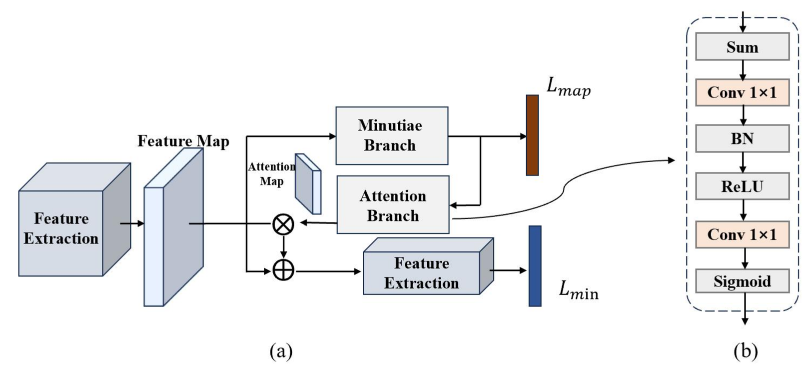
- Alignment module and data augmentation

An STN-based alignment module enables end-to-end image alignment, while dynamic data augmentation simulates real-world pose and environmental variations.

- Texture branch and minutiae branch

The backbone alternates Conv and ResNeXt modules to robustly extract semantically rich features while preventing degradation. These features then feed into two complementary branches: a texture branch for fixed-length feature extraction, and a minutiae branch that utilizes heatmaps (encoding minutiae position and direction) as a regression supervision signal to guide effective training.





- Attention module

A supervised attention mechanism leverages minutiae maps to highlight discriminative local textures through three stages: feature extraction, attention generation, and feature refinement.

- Loss function and dynamic weight design

The overall loss function employs a dynamic weighting mechanism to adaptively balance multi-task learning, combining cross-entropy and center loss for feature classification with Mean Squared Error (MSE) for minutiae regression.

Results

1. Comparative experiments

The proposed method is evaluated against three baseline pipelines on a benchmark dataset: mindtct+MCC, MinutiaeNet+Bozorth3, and MinutiaeNet+MCC. Comparative results are detailed in Table 1.

Table 1. Comparative experimental results on the benchmark contactless fingerprint dataset.

Method	EER	Top1 Accuracy rate
MCC	16.5%	83.3%
MinutiaeNet	25.9%	82%
MCC+MinutiaeNet	14.5%	81%
Proposed	3.5%	89.8%

By jointly modeling global texture and minutiae features with center loss, the proposed method achieves a 3.5% EER and an 89.8% Top-1 accuracy, significantly outperforming baselines like mindtct and MinutiaeNet.

2. Ablation study

Table 2. Performance metrics from the ablation study on the benchmark dataset.

Number	Attention Module	Dynamic Weighting Design	Alignment and Data Augmentation	EER	Top1 Accuracy rate	Top5 Accuracy rate
1	√	√	√	3.5%	89.8%	97.5%
2	×	√	√	4.5%	89.3%	96.3%
3	×	×	√	6.6%	83.3%	92%
4	√	√	×	6%	85.3%	94.3%

- Alignment and Augmentation: Reduces EER by 2.5% and improves Top-1 accuracy by 4.5%.
- Attention Mechanism: Further reduces EER by 1.0%.
- Dynamic Weighting: Decreases EER by an additional 2.1%.
- Full Model: Achieves peak performance with 89.8% Top-1 and 97.5% Top-5 accuracy in 1:N identification.

Table 3 Performance metrics from the ablation study on the Hong Kong Polytechnic University dataset.

Number	Attention Module	Dynamic Weighting Design	Alignment and Data Augmentation	EER	Top1 Accuracy rate	Top5 Accuracy rate
1	√	√	√	0.8%	98.84%	99.91%
2	×	√	√	1.4%	98.09%	99.87%
3	×	×	√	2.1%	95.73%	99.6%
4	√	√	×	1.5%	97.33%	99.96%

- Attention Mechanism: Reduces EER by 0.6%.
- Dynamic Weighting: Further reduces EER by 0.7%.
- Full Model: Achieves peak performance with 98.84% Top-1 and 99.91% Top-5 accuracy, validating the proposed architecture's superiority.

Conclusions

This paper proposes a multi-task CNN that leverages minutiae maps to guide attention generation. The network consists of a shared backbone and two branches (minutiae and texture), built upon stacked Conv-ResNeXt modules for multi-level feature learning. Minutiae maps serve as both auxiliary supervision and guidance for attention generation, emphasizing regions around minutiae while suppressing irrelevant features. A spatial alignment module is introduced to handle pose variations via lightweight geometric correction. Data augmentation is applied to enhance generalization on limited datasets. Additionally, a dynamic weighting strategy is employed to adaptively balance multi-task losses, improving training efficiency and performance.