

基于神经网络的手指三模态多级特征编码融合方法

温梦娜, 叶子云, 赵子豪, 石滨萌

深圳职业技术学院粤港澳大湾区人工智能应用技术研究院, 广东 深圳

收稿日期: 2023年6月14日; 录用日期: 2023年7月5日; 发布日期: 2023年7月18日

摘要

本文提出了一种基于神经网络的手指三模态多级特征编码融合方法。首先, 针对手指三模态原始图像分别构建轻量级卷积神经网络实现多级特征提取。其次, 利用聚合向量编码方式对手指三模态多级特征进行编码, 分别得到指静脉编码特征、指纹编码特征和指节纹编码特征。然后, 利用构建的全连接神经网络模型分别将手指三模态浅层特征和深层特征进行融合。最后, 将手指三模态浅层融合特征和深层融合特征进行串联融合, 得到手指三模态融合特征。实验结果表明通过该方法得到的融合特征的识别准确率为99.76%, 说明该方法能够得到具有良好个性特征表达能力的融合特征, 可以有效提高个体身份识别精度。

关键词

生物特征识别, 手指三模态, 特征编码, 多特征融合

Fusion Method of Finger Trimodal Multilevel Coded Features Based on NN

Mengna Wen, Ziyun Ye, Zihao Zhao, Binmeng Shi

Institute of Applied Artificial Intelligence of the Guangdong-HongKong-Macao Greater Bay Area, Shenzhen Polytechnic, Shenzhen Guangdong

Received: Jun. 14th, 2023; accepted: Jul. 5th, 2023; published: Jul. 18th, 2023

Abstract

In this paper, a neural network based fusion method of finger trimodal multi-level coded features is proposed. Firstly, lightweight convolutional neural networks are constructed respectively for

文章引用: 温梦娜, 叶子云, 赵子豪, 石滨萌. 基于神经网络的手指三模态多级特征编码融合方法[J]. 图像与信号处理, 2023, 12(3): 270-278. DOI: 10.12677/jisp.2023.123027

original images of finger trimodal to extract multi-level features. Secondly, coding the finger trimodal multi-level features based on local aggregate vector theory, finger-vein coding features, fingerprint coding features and finger-knuckle coding features are obtained respectively. Then, the finger trimodal shallow coded features and deep coded features are respectively fused using a fully connected neural network. Finally, the finger trimodal shallow fusion features and deep fusion features are fused in series. The experimental results show that the recognition accuracy of fusion features obtained by this method is 99.76%, which indicates that this method can obtain fusion features with good expression ability of personality features, and can effectively improve the accuracy of individual identity recognition.

Keywords

Biometric Recognition, Finger Trimodal, Feature Coding, Multi-Feature Fusion

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着科学技术的不断发展，人们对生物特征识别技术的要求越来越高。其中，手部特征具有较高的配合度和友好性，基于手部的单模态生物特征识别技术已被广泛应用于社会生活。相比于单模态生物特征识别技术，因为具有识别精度更高、安全性更高和应用范围更广等优点，多模态生物特征识别技术受到越来越多的关注[1]。在实际应用中，基于手部特征的生物特征识别技术能够有效克服对其他特征采集影响较大的普遍化个性因素。另外，同一根手指的不同位置具有指纹、指节纹和指静脉三个关键模态的特征信息[2] [3]，在手指三模态图像采集工作中，这种紧凑的生理结构便于单个采集装置一次完成三种模态特征信息的采集工作[4] [5]。这种工作模式不仅大大减少了数据采集的成本，更重要的是排除了大部分采集环境对手指三模态数据的影响。

多模态生物特征融合可分为像素层、特征层、匹配层和决策层等四个层次，它们相互联系，各有特点，但这四个层次的基本方法都不能科学地刻画手指的整体特征[6]。由于指纹、指静脉、指节纹三者的成像结构不同，手指粗细不一，导致 ROI 的长宽比不同，另外三者图像内容、噪声等方面也存在较大差异，像素层的融合并不能有效整体刻画手指特征[7]。匹配层和决策层是以单模态生物特征的匹配结果和识别结果为基础的，能够实现多模态信息的相互融合[8]。但是存在算法依赖性强、规模大时效率较低、识别性能提升有限等缺点。而且，在匹配层和决策层的融合过程中忽略了多模态生物特征的整体表达所具有的辨析能力，这样的融合就无法从物理上表达手指特征的整体性[9]。另外，这两种融合策略对单模态单元的性能要求较高，数据存储空间的消耗较大，非常不便于单机实现和应用推广。相比于以上融合策略，特征层的融合能够最大限度地发挥特征的辨析能力，较好地平衡数据处理的难易程度和数据所包含的信息量大小之间的关系[10]。由于生物特征属性的不同和成像水平的限制，目前手指三模态图像的成像原理各不相同，采集到的图像质量较低。利用传统滤波器提取特征信息时，需要对手指三模态图像进行预处理，费时费力，而且，传统滤波器是由人工设计完成，采集到的特征有时并不具有很好的区分度[11]。近年来，神经网络技术发展迅速，在计算机视觉与模式识别领域中卷积神经网络的应用最为广泛[12] [13] [14]。卷积神经网络具有强大的特征提取能力，能够从图像自身特性出发，提取出更符合图像本身的特征[15] [16] [17]。另外，特征编码通过改变特征的表达方式，可以有效消除不同滤波器带来的差异性，

能够完整保留特征自身的属性,使得多模态特征融合更具科学性和区分性[18]。

综上,本文提出了一种基于神经网络的手指三模态特征编码融合方法。针对手指三模态图像分别构建轻量级卷积神经网络进行多级特征提取。利用聚合向量编码方式对手指三模态特征进行编码,分别得到指静脉编码特征、指纹编码特征和指节纹编码特征。通过构建的神经网络模型将手指三模态浅层编码特征进行融合,得到手指三模态浅层融合特征。同时,利用相同的融合方案对手指三模态深层编码特征进行融合,得到手指三模态深层编码特征。最后,将手指三模态浅层融合特征和手指三模态深层融合特征进行串联融合,得到手指三模态多级融合特征。

2. 局部聚合向量

局部聚合向量(Vector of Local Aggregated Descriptors, VLAD)是一种编码方式,主要内容是基于特征空间的局部标准来聚合特征描述符。通过用向量的方式描述图像特征来实现对图像的精简表达,表征了向量相对中心的分布。局部聚合向量编码的流程如下,首先提取图像特征。然后利用提取的特征训练码书,原则上,码书越大,编码精度越高,但是要综合考虑计算量与精度的平衡。量化图像特征,利用最近邻算法将所有的描述子分配到 k 个聚类中心,对每个聚类中心做残差和,对残差和做 L_2 归一化,拼接成一个 $k \times d$ 维的向量, d 是特征维度。上述过程如公式(1)所示。

$$v_{i,j} = \sum x_j - c_{i,j} \quad (1)$$

其中, x_j 是图像的第 j 个特征点, $c_{i,j}$ 是相对应的最近的聚类中心, $x_j - c_{i,j}$ 表示特征点与距离特征点最近的聚类中心之间的求差运算, $v_{i,j}$ 是指以 c_i 为聚类中心的特征点 x 的每一维的值与聚类中心的每一维的差的和。

3. 手指三模态特征编码融合

生物特征识别技术的核心问题是如何将生物特征转换为可以进行相似性度量的表示,传统的方法是抽取图像的多个局部特征然后再将其聚合为单个紧致向量。最后,基于各种距离对向量进行相似性度量,从而实现个体身份的鉴别。

3.1. 手指三模态特征提取

传统的特征提取算法如 SIFT、SURF 等是基于人们的经验知识,通过人为设计的滤波器进行特征提取,不可避免地会掺杂人为因素,而且人为设计的滤波器有时并不能很好地适应图像特性。卷积神经网络拥有强大的特征提取能力,虽然提取的特征有时不太适合人脑理解,但无数实验证明利用卷积神经网络提取的特征能够更好地反映图像自身特性,而且可以提取人脑无法发现的特征。

为了从指纹、指静脉和指节纹的原始图像中提取有效特征,我们构建了基于卷积神经网络的特征提取模型。我们知道在训练样本数量理想的前提下,随着网络复杂度的提高,卷积神经网络的性能逐渐提升。但是,当训练样本数量有限时,复杂的网络结构会带来灾难性的问题。为此,我们通过研究已成熟应用的网络结构的特性,构建了轻量级的卷积神经网络模型,用来完成指纹、指静脉和指节纹三种模态的特征提取任务。本文构建的特征提取模型主要包括卷积操作(公式(2))、均值池化操作(公式(3))。同时,为了提高网络的泛化能力和收敛速度,每个池化层后面紧跟一个批归一化层(公式(4)),后面连接一个全连接层,最后选用 softmax 函数(公式(5))实现分类。

$$Z(W, X) = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (2)$$

其中,向量 W 代表卷积核的权重,向量 X 表示感受野内图像的像素值, n 是感受野内像素的个数。

$$p = \frac{1}{M_1 \times M_2} \sum_{i=1}^{M_1 \times M_2} x_i \quad (3)$$

其中, P 代表下采样操作后特征图中的一个元素值, M_1 和 M_2 为下采样区域长和宽, x_i 为下采样区域内的元素。

$$\hat{Z}_j = \frac{Z_j - u_j}{\sqrt{\sigma_j^2 + \varepsilon}} \quad (4)$$

其中, j 代表第 j 个特征图, 引入 ε 是为了防止分母为 0 而产生无效计算。 u_j 为批数据的平均值, σ_j^2 为批数据的方差。

$$s_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (5)$$

其中, $X = \{x_1, x_2, \dots, x_n\}$ 为前一层的输出, S_i 为 Softmax 层第 i 个神经元的输出, softmax 层每个节点的输出值代表的是当前节点类别是图像真实类别的概率。

3.2. 基于 VLAD 的特征编码

假设通过特征提取器获取的特征为 N 个 $S_1 \times S_2$ 的特征图, 首先, 将特征图“展平”为长度为 $S_1 \times S_2 = D$ 的一维向量, 此时, 特征可用描述符 $X_{N \times D} = [x_1, x_2, \dots, x_n]^T$ 表示, 其中, x_i 为一个 D 维的特征向量。然后, 构建包含 K 个元素的码本 $C = \{c_1, c_2, \dots, c_k\}$, 其中, c_k 是一个 D 维的特征向量。手指三模态特征编码的过程如公式(6)~(7)所示。

$$x_i = \text{Connect}(p_1, \dots, p_{S_1 \times S_2}) \quad (6)$$

其中, P_j 为一个特征图中的元素, $\text{Connect}()$ 将输入连接成一维向量。

$$V(j, k) = \sum_{i=1}^N x_i(j) - c_k(j) \quad (7)$$

其中, V 是聚合编码的输出, 是一个 $K \times D$ 维的特征向量。 $x_i(j)$ 是第 i 个视觉描述符的第 j 维, $c_k(j)$ 是第 k 个聚类中心的第 j 维。

3.3. 手指三模态特征编码融合模型

基于神经网络的手指三模态特征编码融合模型的整体结构如图 1 所示, 该模型由手指三模态特征提取模块, 手指三模态浅层特征编码模块, 手指三模态浅层编码特征融合模块和手指三模态深层特征编码模块, 手指三模态深层编码特征融合模块, 以及手指三模态深浅层融合特征串联融合层等组成。卷积神经网络的结构中, 浅层结构接近输入层, 此处的感受野较小, 感受野的重叠度较低, 有利于提取更多的细节, 浅层输出的特征包含更多的位置和细节信息, 语义性较低, 而深层结构中增加了感受野的重叠区域, 对图像信息进行压缩, 输出的特征包含的是图像整体性信息, 具有很强的语义性, 但细节感知力较差。神经网络的浅层特征和深层特征具有优势互补的特点, 综合利用深浅层特征能够更好地刻画个体属性, 更有利于个体身份信息的鉴别。首先, 将手指三模态原始图像数据输入基于卷积神经网络的特征提取模型, 通过不同的网络层分别输出手指三模态浅层特征和手指三模态深层特征。其次, 根据局部聚合向量编码原理分别对手指三模态浅层特征和深层特征进行编码, 得到手指三种模态浅层编码特征和深层编码特征, 共六种编码特征。再次, 手指三模态浅层编码特征输入手指三模态浅层编码特征融合模块实

现手指三模态浅层编码特征的融合，同理，手指三模态深层编码特征输入手指三模态深层编码特征融合模块进行融合。最后，由于编码后的特征描述子代表的是特征相对于聚类中心的距离，消除了不同模态特征的差异性，在特征融合时不用考虑特征对齐等问题，直接将手指三模态浅层融合特征与深层融合特征进行串联，得到手指三模态融合特征。

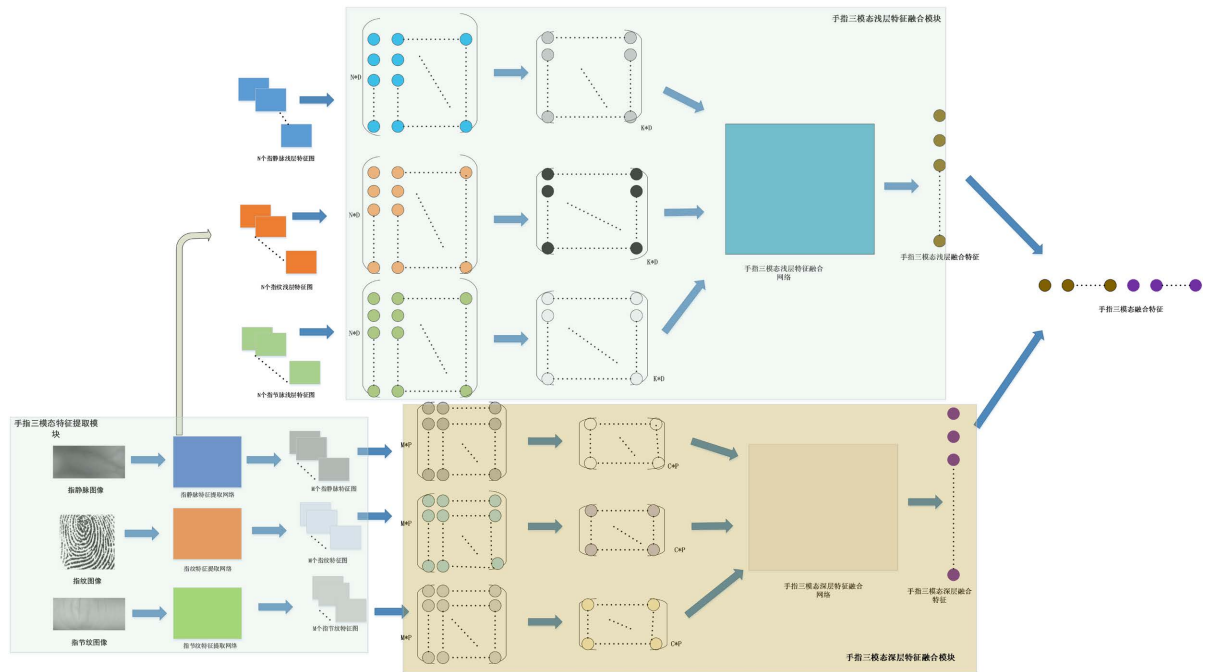


Figure 1. Finger trimodal feature coding fusion model

图 1. 手指三模态特征编码融合模型

4. 实验

4.1. 实验数据

本文实验用的数据为利用自主研发的手指三模态一体化图像采集设备采集的指静脉图像、指纹图像和指节纹图像组成，如图 2 所示。共采集 585 个手指的三模态图像，每个手指通过变化不同姿态分别采集十次，即每根手指的每个模态别采集到十张图像，即 $3 \times (585 \times 10)$ 张图像。



Figure 2. Raw images of finger trimodal
图 2. 手指三模态原始图像

4.2. 实验结果与分析

基于神经网络的手指三模态特征编码融合模型中用于特征提取的部分为手指三模态特征提取模块，该模块由三个并行的卷积神经网络组成，如图 3 所示。其中，用于提取指静脉特征的网络由六个卷积层

组成，用于提取指纹特征的网络由四个卷积层组成，用于提取指节纹特征的网络由六个卷积层组成，这里的卷积层包括卷积操作和均值池化操作。每个池化层后面跟一个批归一化(Batch Normalization, BN)层，用来提高网络泛化能力和收敛速度。对网络进行训练时，单模态特征提取网络的最后面连接一个全连接层，选用 softmax 函数完成分类。单模态特征提取网络训练完成后，保存网络中所有参数的数值。用于特征提取时，单模态特征提取网络丢弃最后的全连接层，输入分别为指静脉、指纹和指节纹的原始图像，输出为特征图。

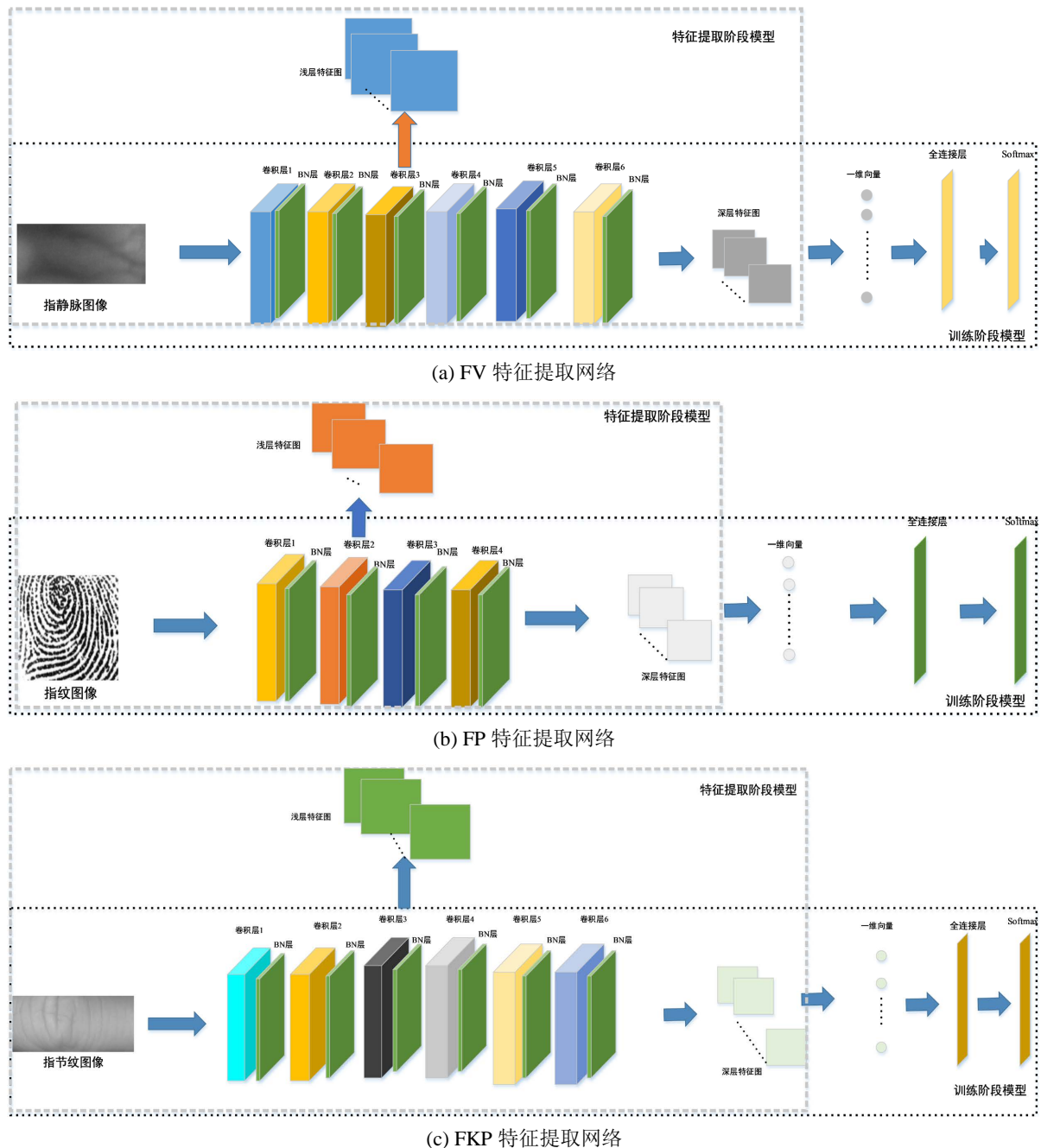


Figure 3. Finger trimodal feature extraction model
图 3. 手指三模态特征提取模型

为了验证我们构建的手指三模态特征提取模块是否具有科学性以及手指三模态特征提取网络获取的特使是否具有较好的类别区分度。分别以指静脉、指纹和指节纹的原始图像作为输入，图像所属类别作为标签，选用交叉熵函数作为损失函数。另外，为了使特征提取网络具有较好的泛化能力，首先，将实验数据的次序进行打乱。然后，将所有实验数据划分为训练数据集和测试数据集，其中，训练数据集占所有数据的 80%，测试数据集占所有数据的 20%。图 4 分别为三个特征提取网络的训练过程，表 1 为训练完成的手指三模态特征提取网络在测试数据集上的识别精度。

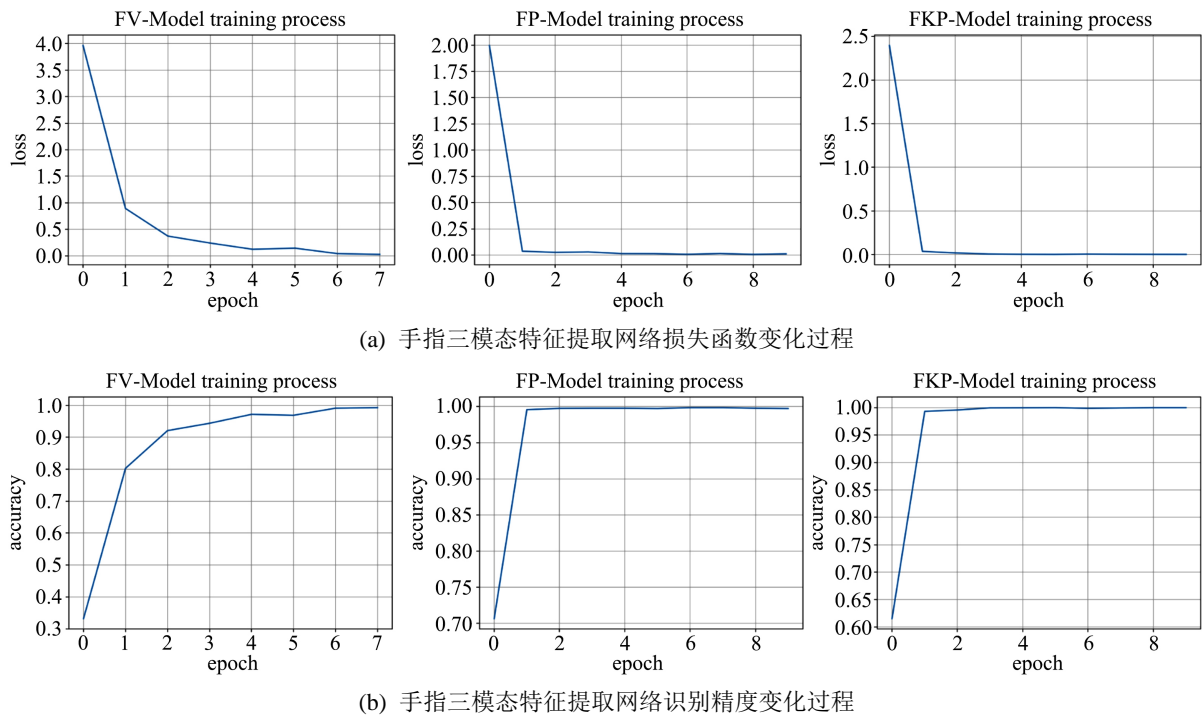


Figure 4. Training process of finger trimodal feature extraction model

图 4. 手指三模态特征提取网络训练过程

Table 1. Identification accuracy of finger trimodal feature extraction model

表 1. 手指三模态特征提取模型识别精度

模态 - 特征	FV-特征	FP-特征	FKP-特征
精度(%)	97.43	93.24	98.37

综合分析上述实验结果，可以看出本文构建的手指三模态特征提取网络是可训练的，训练过程中损失函数逐渐变小，识别精度逐渐变大，最终稳定在理想值周围，且训练完成的网络能够得到较高的识别精度。说明本文构建的手指三模态特征提取网络能够提取到有效的手指三模态特征。

为了验证局部聚合编码的有效性，首先将特征根据上文提到的局部聚合向量编码方法进行编码，然后通过余弦距离测量相似度，最后将待识别的图像归为相似度最大的类别中。表 2 为手指三模态编码特征的识别精度。

分析表 2 可以看出，基于编码后的特征能够得到较高的分类正确率，说明利用局部聚合向量编码算法对通过神经网络提取到的手指三模态特征进行编码有效的，可以得到识别精度较高的编码特征。

Table 2. Identification accuracy of finger trimodal coded feature
表 2. 手指三模态编码特征的识别精度

模态特征	FV 浅层特征	FP 浅层特征	FKP 浅层特征	FV 深层特征	FP 深层特征	FKP 深层特征
精度(%)	94.86	94.09	93.24	95.66	95.73	95.78

为了验证本文提出的手指三模态编码特征融合方法的有效性, 本文进行了以下实验, 首先将编码完成的手指三模态浅层特征作为构建的神经网络的输入层, 然后通过余弦距离测量相似度, 最后将待识别的图像归为相似度最大的类别中, 表 3 为上述实验结果。

Table 3. Identification accuracy of finger trimodal fusion feature
表 3. 手指三模态融合特征的识别精度

模态 - 特征	手指三模态浅层融合特征	手指三模态深层融合特征	深浅层特征串联融合特征
精度(%)	93.24	97.43	99.76

对比分析表 2 和表 3 中的结果, 可以得出通过本文提出的手指三模态编码特征融合算法得到的融合特征能够描述能力较强的特征, 可以有效提高个体识别精度。

参考文献

- [1] 王俊生, 王波, 韩庆芝. 一种多模态生物特征融合算法探究[J]. 智能计算机与应用, 2018, 8(6): 87-89.
- [2] 刘伟业, 鲁慧民, 李玉鹏, 等. 指静脉识别技术研究综述[J]. 计算机科学, 2022, 49(Z1): 1-11.
- [3] Hou, C. and Jia, W. (2022) Combining Band-Limited OTSDF Filter and Directional Representation for Palmprint Recognition. Springer, Cham. https://doi.org/10.1007/978-3-031-20233-9_5
- [4] 张帆. 基于指节纹的身份识别方法研究[D]: [硕士学位论文]. 西安: 西安理工大学, 2019.
- [5] Khodadoust, J., Medina-Perez, M.A., Monroy, R., et al. (2021) A Multibiometric System Based on the Fusion of Fingerprint, Finger-Vein and Finger-Knuckle-Print. *Expert Systems with Applications*, **176**, Article ID: 114687. <https://doi.org/10.1016/j.eswa.2021.114687>
- [6] 苗迪. 基于特征层与分数层的多模态生物特征融合[D]: [博士学位论文]. 合肥: 中国科学技术大学, 2017.
- [7] 戴丽丽. 指关节图像的特征抽取方法研究[D]: [硕士学位论文]. 南京: 南京理工大学, 2015.
- [8] Gowda, S. and Kumar, A. (2019) Dimension Reduction of Hand and Face Feature Level Fusion in Multimodal Biometric Authentication. *International Journal of Innovative Technology and Exploring Engineering*, **8**, 2930-2937. <https://doi.org/10.35940/ijitee.I8924.078919>
- [9] Asaari, M., Suandi, S.A. and Rosdi, B.A. (2014) Fusion of Band Limited Phase Only Correlation and Width Centroid Contour Distance for Finger Based Biometrics. *Expert Systems with Applications*, **41**, 3367-3382. <https://doi.org/10.1016/j.eswa.2013.11.033>
- [10] Yang, W., Wang, S., Hu, J., et al. (2018) A Fingerprint and Finger-Vein Based Cancelable Multi-Biometric System. *Pattern Recognition*, **78**, 242-251. <https://doi.org/10.1016/j.patcog.2018.01.026>
- [11] 秦芳, 顾广华. 基于多尺度局部特征编码与多通道特征融合的图像场景分类[J]. 燕山大学学报, 2019, 43(4): 357-363.
- [12] 吴超, 邵曦. 基于深度学习的指静脉识别研究[J]. 计算机技术与发展, 2018, 28(2): 200-204.
- [13] 何鑫, 陈迅. 基于改进卷积神经网络的指静脉识别[J]. 计算机工程与设计, 2019, 40(2): 562-566.
- [14] Sobhan, S., Ali, D., Hadi, K., et al. (2018) Multi-Level Feature Abstraction from Convolutional Neural Networks for Multimodal Biometric Identification. *Proceedings of 2018 24th International Conference on Pattern Recognition (ICPR)*, Beijing, 20-24 August 2018, 3469-3476.
- [15] Tang, S., Zhou, S., Kang, W., et al. (2019) Finger Vein Verification Using a Siamese CNN. *IET Biometrics*, **8**, 306-315. <https://doi.org/10.1049/iet-bmt.2018.5245>

- [16] 介婧, 陈羽川, 郑慧, 等. 一种基于生成对抗网络和卷积神经网络的指静脉识别方法[P]. 中国专利, 114743278A, 2022-07-12.
- [17] Fang, Y., Wu, Q. and Kang, W. (2018) A Novel Finger Vein Verification System Based on Two-Stream Convolutional Network Learning. *Neurocomputing*, **290**, 100-107. <https://doi.org/10.1016/j.neucom.2018.02.042>
- [18] 杨晓生. 虹膜图像预处理及特征编码方法的研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2016.