

# 基于CNN的手指三模态多级特征融合与识别

温梦娜<sup>\*</sup>, 叶子云<sup>#</sup>

深圳职业技术学院, 广东 深圳

收稿日期: 2021年11月25日; 录用日期: 2021年12月21日; 发布日期: 2021年12月28日

## 摘要

随着人们对生物特征识别的要求不断提高, 手指因为包含有丰富的生物特征信息受到越来越多的关注。手指同时携带了指纹、指节纹和指静脉三种模态信息, 作为一个整体来表征个体的生物特性, 具有得天独厚的优势。本文中, 我们提出了一种基于CNN (Convolutional Neural Network)的手指三模态多级特征融合与识别算法, 以融合指纹、指静脉和指节纹三种模态的特征信息来进行身份识别。该融合框架由联合优化的三部分组成, 手指三模态特征提取模块, 手指三模态浅层特征融合模块和手指三模态多级特征融合模块。手指三模态特征提取模块由三个并行的特征提取框架组成, 分别用于提取手指单模态多级特征。浅层特征融合模块以CNN作为融合框架, 以手指三模态特征提取网络的浅层输出特征作为输入, 来获取三模态浅层融合特征。多级特征融合模块以全连接网络作为融合框架, 手指三模态浅层融合特征和手指三模态的深层特征作为全连接网络的输入, 以实现手指三模态多级特征融合。最后在手指三模态数据库上进行了大量实验, 本文提出的算法识别精度可以达到99.91%。实验结果表明本文提出的手指三模态多级特征融合与识别算法性能优于其他现有的算法。

## 关键词

指纹, 指静脉, 指节纹, 多级特征融合, 生物特征识别, 卷积神经网络

# CNN-Based Finger Trimodal Multi-Level Features Fusion and Recognition

Mengna Wen<sup>\*</sup>, Ziyun Ye<sup>#</sup>

Shenzhen Polytechnic, Shenzhen Guangdong

Received: Nov. 25<sup>th</sup>, 2021; accepted: Dec. 21<sup>st</sup>, 2021; published: Dec. 28<sup>th</sup>, 2021

<sup>\*</sup>第一作者。

<sup>#</sup>通讯作者。

## Abstract

With the increasing demand of biometric recognition, finger, which contains abundant biometric information, has played an important role in the identification recognition field. There are finger-print (FP), finger-knuckle-print (FKP) and finger-vein (FV) on different parts of a finger. As a whole, finger has unique superiority to explore multimodal fusion recognition technology. In this paper, we propose a finger trimodal multi-level features fusion framework based on CNN, which integrates feature information of finger-print, finger-vein and finger-knuckle-print for identification. The proposed trimodal multi-level features fusion framework consists of finger three-modal features extraction module, three-modal shallow features fusion module and three-modal multi-level features fusion module, which are jointly optimized. The finger three-modal features extraction module is composed of three parallel and independent CNN frameworks, which are used to extract multilevel features of FP, FV and FKP separately. The three-modal shallow features fusion module uses CNN as fusion framework, and takes the shallow features of the three biometric patterns as inputs to obtain shallow fusion feature. The multilevel features fusion module takes the fully connected network as the fusion framework. The trimodal biometric shallow fusion feature and the deep features serve as the inputs of the full connection network to realize trimodal multi-level features fusion. Experimental results show that the recognition accuracy of the proposed algorithm can be as high as 99.91%. A comprehensive analysis of many experimental results shows that the proposed trimodal multi-level features fusion framework performs better than other state-of-the-arts.

## Keywords

**Fingerprint, Finger-Vein, Finger-Knuckle-Print, Multi-Level Features Fusion, Biometrics Recognition, Convolutional Neural Network**

Copyright © 2021 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

在当今社会中，生物特征识别技术拥有着广阔的需求，被广泛应用于各个领域。随着科学技术不断发展，人们对身份鉴别技术的要求也在不断提高，相比于传统的身份识别媒介如卡片、密码等，人体生物特征具有便携、不会遗失等多种优势[1]。人体生物特征大致可分为生理特征(如人脸、虹膜、指纹、掌纹、静脉等)和行为特征(如步态、签名、击键习惯等)。根据其载体位置不同，生理特征又可以分为头部特征、手部特征和形体特征等[2]。随着身份鉴别技术的应用场景逐渐多元化，单模态生物特征识别技术有时无法满足人们对安全性的要求。为了提高生物特征识别技术的准确性、稳定性和可靠性，通过综合利用人体多个模态的生物特征信息来共同刻画和表达个体的身份属性，被称为“多模态生物特征识别”[3][4]。在手部多种模态特征载体中，手指的表现最为突出，手指同时携带了指纹(fingerprint, FP)、指节纹(finger-knuckle-print, FKP)和指静脉(finger vein, FV)三个关键模态的特征信息，天然综合了手指生物组织内外层的生理特征，并且具有充分的互补性[5]。例如，指纹无法判别活体性，指静脉天然地具有活体性；指纹处于受力面，很容易磨损，而指节纹(背侧)则处于非受力面，表面比较完整；指静脉成像清晰度较低，指节纹图像成像清晰稳定。手指三模态特征的互补性充分说明了综合利用三者信息能使手指在身

份识别上更具有稳定性和普遍性，也更能发挥手指作为特征载体的科学价值与应用价值[6]。

从融合的角度，多模态生物特征的融合可分为四个层次：像素层、特征层、匹配层和决策层。这四个层次的融合相互联系，各有特点，构成了当前多模态生物特征识别的基础框架[7]。相比于其他层的融合策略，特征层融合能最大程度地发挥融合特征的辨析能力，较好地平衡数据处理的难易程度以及数据所包含的信息量大小[8]。在过去的研究中，人们提出了大量的基于特征层的多模态融合算法，但是，这些融合算法，常采用特征串联、并联和加权叠加等方式进行特征融合，这类方法并没有考虑特征之间的相互关系，是一种“强融合”，效果较差[9]。近几年，一些学者通过扩展典型相关分析方法(简称 CCA)实现特征层融合，考虑了不同特征之间的线性关系，但这种方法也只能实现两类特征的互相融合，并且在小样本和非线性条件下效果并不理想[10]。因此，在特征层融合上，泛化、合理的理论依然是欠缺的，这也成了制约手指多模态生物识别技术发展的一个主要瓶颈。

近年来，深度学习技术在计算机视觉领域的飞速发展，为多模态生物特征融合识别带来了新的发展机遇[11]。在多个公开数据集上的测试结果表明，基于深度学习的方法在生物特征识别任务中取得了非常优秀的效果，大大超过了传统识别方法[12]。研究人员已经构建了多种深度神经网络模型，如 AlexNet、VggNet、GoogleNet、ResNet、DenseNet 和 SENET 等，当前的优秀模型在 ImageNet 数据集上已经取得了超越人类的辨识能力，性能卓越。卷积神经网络(Conventional Neural Network, CNN)模型在手指单模态生物特征识别上得到了广泛的应用，从图像增强[13]，特征提取[12][14]，到提高算法的鲁棒性[15][16]，CNN 算法都获得了令人满意的效果。然而，就手指多模态融合识别而言，基于深度学习的融合算法并未获得良好应用[17]。由于采集环境及手指三模态各自特性的影响，手指三模态图像存在尺度不一致现象，强制性的归一化方式容易破坏图像结构信息，这对于融合识别是不利的。另外，已有的特征层融合算法中，往往只考虑了 CNN 深层的输出特征，却忽略了网络浅层的输出特征[18]。为了有效解决以上问题，本文在 CNN 模型框架下，深入探索了手指三模态的融合识别网络，研究开发出合适的并行网络特征提取方式，建立了手指三模态特征融合机制，解决了由尺度不一致性带来的融合制约问题。

本文主要贡献如下：1) 提出了手指三模态多级特征融合的思想。CNN 每层输出的特征图都保留有一定的空间特征信息，本文提出的多级特征融合方法不仅考虑了单模态网络最后一层的输出特征，也添加了手指三模态的浅层特征，能够更加合理地表征个体的生物特性。2) 构建了基于 CNN 的手指三模态多级特征融合框架。该框架由手指三模态特征提取模块、手指三模态浅层特征融合模块和手指三模态多级特征融合模块三部分组成。该框架可以通过端到端的训练方式，实现并行特征提取网络和融合网络的联合训练，得到联合最优参数。3) 针对浅层特征融合模块和多级特征融合模块，分别提出了两种特征融合策略。融合模块能够完整地嵌入原始网络，实现端到端的训练。提出的两种融合策略能够在不改变三模态图像数据原始输入尺寸的前提下，实现不同尺寸的卷积特征融合。本文第 2 节为相关工作，简要介绍了已有的多模态特征融合算法；第 3 节重点叙述了本文提出的基于 CNN 的手指三模态多级特征融合识别算法；第 4 节是对比实验以及对实验结果的描述和分析；第 5 节是本文研究工作的总结和对未来工作的展望。

## 2. 相关工作

由于多模态生物特征识别技术的应用越来越广泛，在过去的十几年中，多模态融合算法的相关研究得到了蓬勃发展。1995 年，Brunelli 等人提出了多模态融合识别的思想，并通过融合声音和人脸特征建立了一个多模态生物识别系统[19]。200 年，Ross 和 Jain 将多模态融合分为三个层次，指明了多模态融合的发展方向[2]。此后，不同层次的多模态融合策略迅速发展。2011 年，Park Y. H. 等人提出了一种基于决策层的指纹和指静脉融合识别算法[9]。2016 年，Yang 等人提出了一种基于多模态手指特征的空间圆

形造粒方法[10]。Bai G 等人提出了一种基于像素的粒化融合手指识别方法,但是由于这种算法是对像素进行运算的,运算及存储耗费的资源过大[20]。Zhang H 等人提出了一种基于图的特征提取方法并且设计了两种特征融合框架[21]。虽然传统的融合算法在很多邻域均取得了不错的成绩,但是,因为有过多的人为因素参与其中,仍然具有一定的缺点。而且,传统的融合算法往往通过简单地剪裁和比例归一化等方法来调整输入数据的大小,这种简单粗暴的方法容易破坏图像的结构,造成有用信息的缺失。

近年来,随着深度学习在图像处理领域的迅速发展,生物特征识别也从手工提取特征转变为自动学习的方式。又由于卷积神经网络优秀的图像处理能力,研究人员提出了许多基于卷积神经网络的多模态融合算法。Zhang H 等人提出了一种基于 RGB-D 图像的人脸识别方法,该方法能够通过深度学习网络学习多模态互补的人脸特征,进而提高识别精度[22]。Wang 等人提出了一个基 CNN 的多模态特征学习框架,用于识别 RGB-D 目标,随后他们设计了一个深度学习框架来学习多模态的特定相关特征[23]。Sobhan Soleymani 等人提出了一种基于卷积神经网络的多级特征提取和融合表达的方法,用以提取不同的模态的特征信息,来完成多级特征的联合表达[24]。

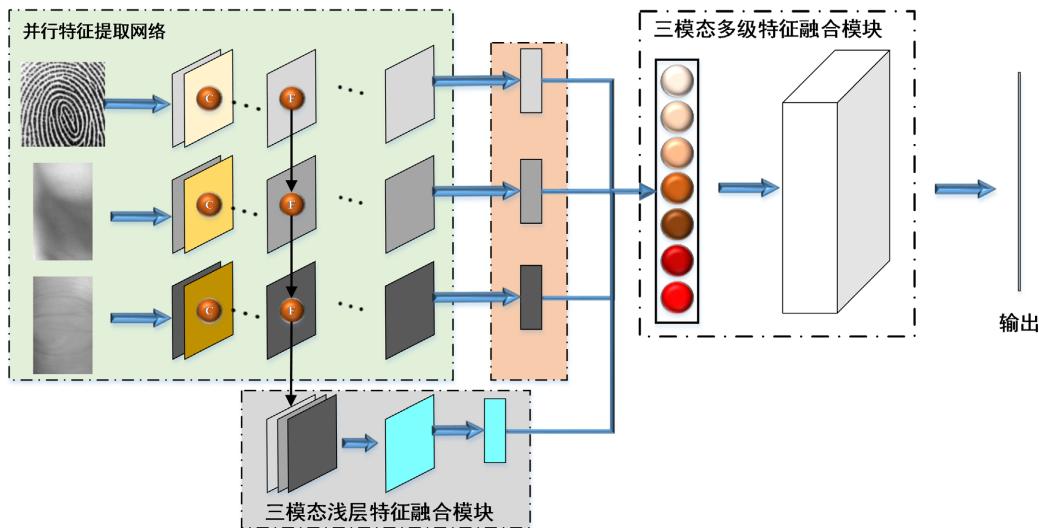
### 3. 基于 CNN 的手指三模态多级特征融合框架

近年来,CNN 在图像处理领域的表现优异,被认为是学习图像内容的最佳技术之一。卷积神经网络的本质是模拟实现抽象特征(高层)与目标数据(输入)之间的非线性映射。参数数量(卷积核)和深度(层数)是体现网络容量(Capacity)的重要指标,它们结合激励层(Active Layer)、池化层(Pooling Layer)和全连接层,共同形成一个网络结构来表达数据集与特征集之间的非线性关系。卷积神经网络以分层的方式堆叠多个线性和非线性处理单元,因而具有了学习不同抽象级别上复杂表示的能力。卷积神经网络每层输出的特征图代表了输入图像的特征在不同抽象级别上的表示,网络浅层输出的特征图保留了更多的空间信息,网络深层输出的特征图则提供了更加复杂和抽象的特征信息。为了能够同时利用浅层和深层的特征来实现更高的识别精度,我们构建了手指三模态多级特征融合框架。该框架由联合优化的三部分组成,手指三模态特征提取模块,手指三模态浅层特征融合模块和手指三模态多级特征融合模块。手指三模态特征提取模块由三个并行且独立的 CNN 特征提取网络构成,分别对应于指静脉图像、指纹图像和指节纹图像,旨在提取手指单模态特征信息;在手指三模态浅层特征融合模块中,引入了矩阵型融合策略。由于手指三模态特征提取网络输出的浅层特征图尺寸并不一致,我们通过设置行高不相等的非对称卷积核,可以有效实现输出特征图的尺寸转化。然后,利用卷积神经网络的内部结构实现三模态浅层特征图的有机融合,得到三模态浅层融合特征;在手指三模态多级特征融合结构中,采用向量型融合策略,采用全局池化方式将卷积输出的特征图转化为一维向量特征。最后,将得到的浅层融合特征与深层特征进行串联,并将串联后的特征输入为特征映射函数  $f(x)$  中,在实际训练中可以以多层感知器(MLP)进行表达,目的是进行输出向量维度的规范化。基于 CNN 的手指三模态多级特征融合框架示意图如图 1 所示。

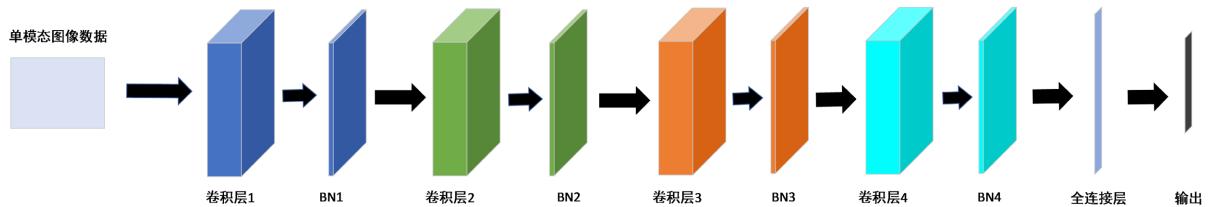
#### 3.1. 手指三模态特征提取模块

为了从指纹、指静脉和指节纹三种模态的图像数据中提取多级特征,我们在三模态多级特征融合框架中嵌入了并行的手指三模态特征提取网络,由三个独立的 CNN 特征提取框架组成,旨在提取手指单模态特征信息。网络输入不需改变手指三模态图像尺度,网络结构和参数也无需进行一致化要求,可依据成熟的单模态识别网络进行结构配置。如图 2 所示,基于 CNN 的单模态特征提取网络由四个卷积层组成,每个卷积层后面跟了一个归一化(Batch Normalization, BN)层,用来提高网络泛化能力和提高网络的收敛速度,网络的最后添加一个全连接层。在嵌入三模态多级特征融合框架之前,对单模态特征提取网络单独进行训练,此时,单模态专用网络的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数值。构建三模态融合网络模型时，三个并行单模态专用网络丢弃最后的全连接层作为融合网络的组成部分，预训练时保存的参数作为初始化参数。



**Figure 1.** Architecture of proposed network  
**图 1.** 网络结构



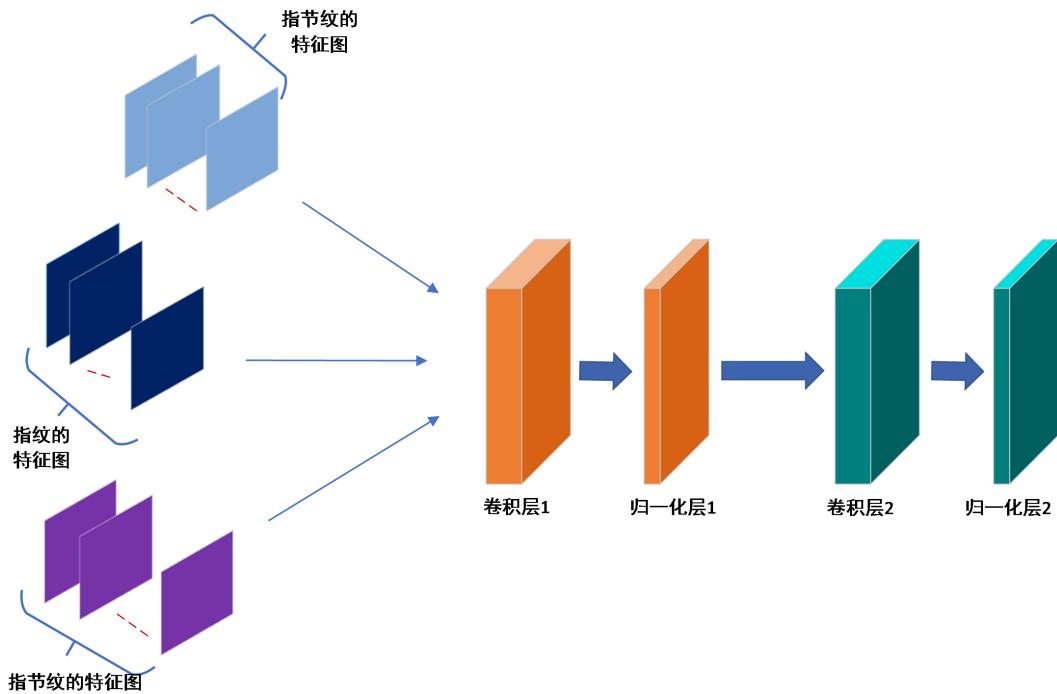
**Figure 2.** Single mode feature extraction network  
**图 2.** 单模态特征提取网络

### 3.2. 手指三模态浅层特征融合模块

卷积神经网络拥有优秀的特征提取能力，而且，通过设计网络输入的通道数，卷积神经网络可以同时对具有相同尺寸的多个矩阵数据进行处理。将三个并行单模态特征提取网络浅层输出的特征图作为卷积神经网络的输入，不仅能够自动地实现三模态浅层特征的融合，而且因为卷积神经网络具有强大的学习能力，这样获取的三模态融合特征更符合三个模态的内在机制。

虽然 CNN 结构在多种任务中取得了优秀的成绩，但是却有一个天然缺点，影响了更广泛的应用。CNN 模型对输入图片尺寸有一致化的要求，无法处理其他尺寸的数据。为能够利用卷积神经网络实现指纹、指静脉和指节纹三模态的浅层特征融合任务，我们在构建单模态特征提取网络时，联合考虑输入图像的空间尺寸和单模态特征提取网络各层的超参数，通过设置行高不相等的非对称卷积核，再配合滑窗和池化等方式，可以确定某一层输出特征图的尺寸。

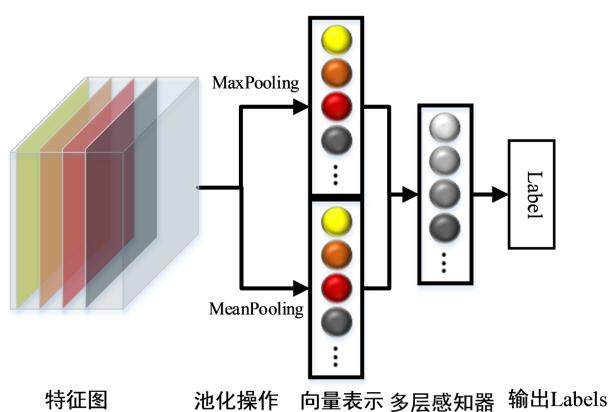
本文构建的三模态浅层特征融合网络如图 3 所示，该网络结构由两个卷积层组成。每个卷积层后面跟一个 BN 层，用于加速网络的收敛。网络的最后添加一个全局池化层，用于将获取的三模态融合特征图映射为一维特征向量。该三模态浅层特征融合网络作为重要的一部分被嵌入到三模态多级特征融合框架中，如图 1 的下半部分。



**Figure 3.** Shallow features fusion network  
**图 3.** 浅层特征图融合网络

### 3.3. 多级特征融合模块

在 CNN 框架中, 每一层的输出代表输入的不同抽象特征表示, 其中更深层的输出提供了更复杂和更抽象的特征。而浅层的输出则保留了更多的空间特征信息。为了综合利用单模态特征提取网络不同层的输出特征, 我们设计了三模态多级特征融合策略。首先, 将浅层融合特征和单模态特征提取网络的深层特征利用全局池化方式转化为一维向量特征。然后, 将一维特征向量输入简单的全连接网络, 实现特征匹配。具体过程如图 4 所示。



**Figure 4.** Vector-type features fusion  
**图 4.** 向量型特征融合

本文拟采用全局最大值池化(MaxPooling)和全局均值池化(MeanPooling)方式将卷积输出映射为特征向量。其中, 全局最大值池化方式就是取每一个感受野内的最大值, 全局均值池化方式提取每个感受野

内的平均值。一方面，向量型特征融合方式可以有效解决输入特征尺度不一致问题，MaxPooling 与 MeanPooling 的计算与输入特征的尺寸无关，最终形成一维向量特征。另一方面，该方式可以有效避免单模态局部缺陷对于整体识别模型的影响，全局池化方式将感受野扩展到整个像素空间，局部变化对于向量型输出而言影响较小。在实际训练中后面可以接多层感知器(MLP)进行表达，从而进行输出向量维度的规范化。

## 4. 实验与分析

### 4.1. 实验数据

文章中的实验数据为使用本研究团队自制的手指三模态图像采集设备采集的指纹、指静脉和指节纹图像。指纹图像的尺寸为  $152 \times 152$  像素，指静脉图像的尺寸为  $91 \times 200$  像素，指节纹图像的尺寸为  $90 \times 200$  像素。共采集了 585 个手指对象，每个手指的每种模态采集 10 张图像。该数据库由三个文件夹组成，FP 文件夹包含 585 个手指的 5850 张指纹图像；FV 文件夹包含 585 个手指的 5850 张指静脉图像；FKP 文件夹包含 585 个手指的 5850 张指节纹图像。

### 4.2. 实验平台

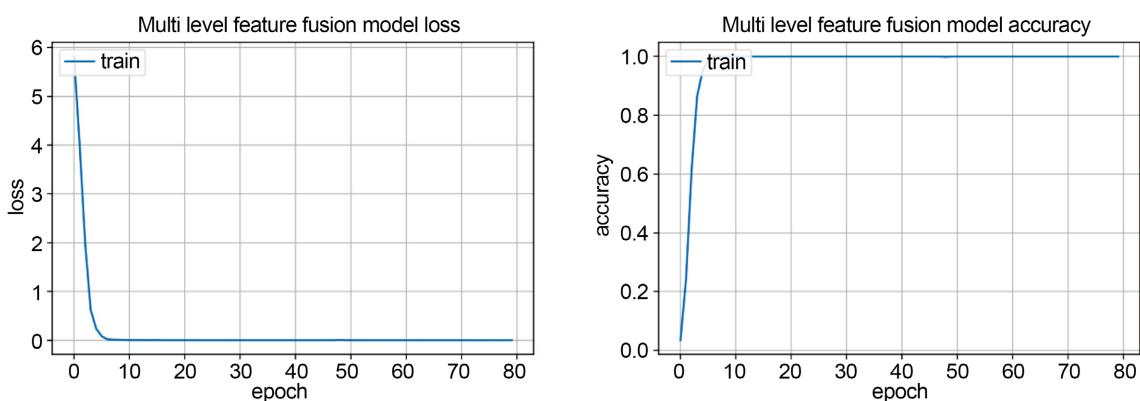
本文实验在单机平台下完成，系统为 win10、8 核 Intel Core i7-9700 CPU、16G 内存、NVIDIA GeForce RTX2070，实验所需环境为 Python3.5.4、tensorflow1.7.0。

### 4.3. 实验结果与分析

本文中实验的目的是验证提出的手指三模态多级特征融合框架的有效性。因此可以通过验证以下内容，对提出的融合模型的有效性进行评估：1) 验证该三模态多级特征融合框架的特性；2) 验证不同全局池化方式对识别精度的影响；3) 验证多级特征融合带来的识别精度的提升；4) 验证联合训练带来的识别精度的提升；

#### 4.3.1. 三模态多级特征融合网络的特性

为了验证三模态多级特征融合网络的特性，我们对网络的训练效率、识别精度进行实验验证，同时与针对单模态的卷积神经网络模型做了对比实验。进行以上实验时，为了使实验数据的特性能够均匀地分布在数据集内，将所有数据及对应标签进行了无序打乱。然后，使用“留出法”将数据集随机划分为训练集和验证集，其中训练集数据占整个数据集的 80%，20% 的数据为验证集。实验结果如图 5 和表 1 所示。



**Figure 5.** Training process of three-mode multilevel feature fusion network

**图 5.** 三模态多级特征融合网络的训练过程

从图 5 中可以看出, 在训练过程中, 随着 epoch 值不断增加, 网络的损失函数(loss)和识别精度(accuracy)不断变化。当 epoch 值达到 25 时, 损失函数值和精度值趋于稳定, 说明此时网络模型达到收敛状态。从图中可以看出, 当网络模型达到收敛状态后, 网络的损失值和精度值都处于稳定状态, 并未出现剧烈变化, 而且当网络达到稳定状态后, 精度值较高, 损失函数值较低。说明本文提出网络模型不仅能够实现较高的识别精度, 而且具有较高的稳定性。

**Table 1.** Recognition results of single mode frame and fusion frame  
**表 1. 单模态框架与融合框架的识别结果**

模型类别	识别精度(%)	训练用时(s)
CNN_FV	99.36	908.35
CNN_FP	99.66	880.37
CNN_FKP	99.74	676.01
融合框架	99.91	771.99

观察表 1 中的实验结果, 可以发现本文提出的网络模型的识别精度均高于基于卷积神经网络的单模态识别网络的识别精度。对比表 1 中的第四列数据, 我们可以看到本文提出的融合网络达到收敛状态的训练用时要低于单模态识别网络的训练用时。综合分析以上结果, 本文提出的手指三模态多级特征融合识别网络不仅拥有较高的识别精度, 而且能够较早地达到收敛状态, 计算资源消耗较低, 充分证明了本文提出的算法的可用性。

#### 4.3.2. 不同全局池化方式的效果

在多级特征融合模块中, 为了解决由于特征图尺寸不一致无法进行融合的问题, 我们引入了向量型特征融合策略。该融合策略中重要的一步是利用全局池化方式将卷积输出映射为一维特征向量。全局池化又可以分为全局最大值池化(GlobalMaxPooling)和全局均值池化(GlobalAvgPooling)。为了验证这两种全局池化方式在本文提出的融合框架中的有效性, 我们分别构建了应用这两种全局池化方式的融合框架进行实验。实验结果如表 2 所示。

从表 2 可以看出, 嵌入全局池化层的网络模型均能取得较高的识别精度, 这是因为全局池化方式将感受野扩展到整个像素空间, 局部变化对于向量型输出而言影响较小, 可以有效避免单模态局部缺陷对于整体识别模型的影响。对比表 2 中的识别结果可知, 嵌入全局均值池化层的网络识别精度要稍高于嵌入全局最大值池化层的网络, 说明全局均值池化方式更适合本文提出的手指三模态多级特征融合网络。

**Table 2.** Recognition results of different global pooling methods  
**表 2. 不同全局池化方式识别效果**

全局池化类别	最大值池化	均值池化
识别精度(%)	99.89	99.91

#### 4.3.3. 不同融合策略对识别精度的影响

为了验证手指三模态多级特征融合策略对识别效果的影响, 本文对手指三模态多级特征融合策略的识别精度进行了实验验证, 同时对比了仅用三模态深度特征的融合策略的识别精度。实验结果如表 3 所示。

观察表 3 可以看出, 相比于单独使用深层特征图进行融合的算法, 本文提出的手指三模态多级特征融合框架拥有更高的识别精度。这是因为多级特征融合框架不仅仅保留了三模态深层特征图的信息, 而

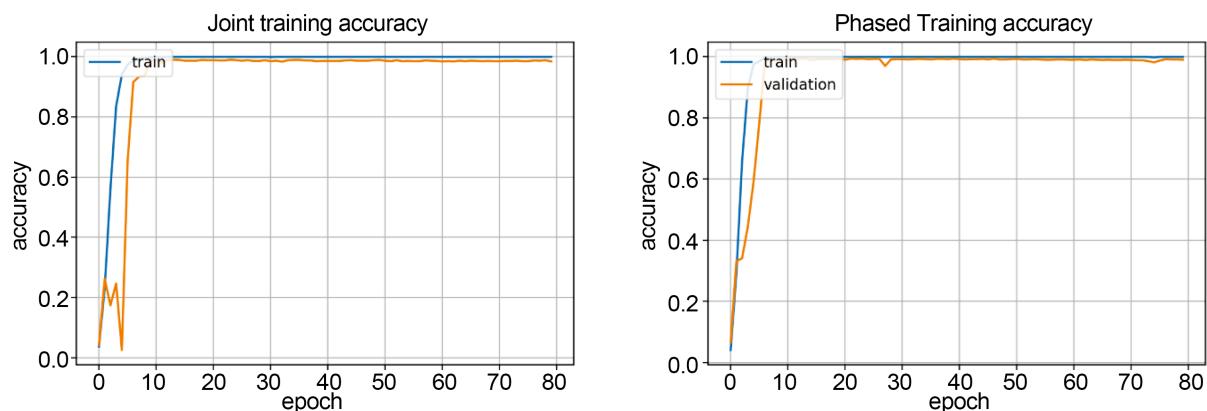
且加入了手指三模态浅层的特征。由于手指单模态特征提取网络的浅层输出保留了更多的空间信息, 综合了手指三模态浅层与深层特征的融合特征能更加有效地表征个体特性。

**Table 3.** Recognition results of different global pooling methods**表 3. 不同全局池化方式识别效果**

融合类别	单级特征融合	均值池化
识别精度(%)	99.89	99.91

#### 4.3.4. 融合框架联合训练的有效性

本文首次提出了融合框架联合优化的概念, 在对网络进行训练时, 将三个并行单模态特征提取网络通过预训练获得的参数作为初始化值, 利用反向传播算法对整个融合框架进行联合训练, 确保整个网络框架获得联合最优值。为了验证联合训练的有效性, 我们对比了联合训练方式与分阶段训练方式的训练过程, 并且比较了两种训练方式获得的网络的识别效果。实验结果如图 6 和表 4 所示。

**Figure 6.** Training process of different training strategies  
**图 6.** 不同训练策略训练过程**Table 4.** Recognition accuracy of different training strategies**表 4. 不同训练策略下模型的识别精度**

训练方式	分阶段训练	联合训练
识别精度(%)	98.84	99.91

#### 4.3.5. 对比实验

为了验证本文提出的融合框架的有效性, 我们对比了不同融合算法, 比较了本方法与其他方法的 ROC 曲线(Receiver Operating Characteristic), 如图 7 所示。在特征比对算法中, 设定不同的阈值 T 会得到不同的错误拒绝率 FRR (False Reject Rate)和错误接受率 FAR (False Accept Rate), 以 FRR 为 X 轴, FAR 为 Y 轴绘制到直角坐标系上就得到了 ROC 曲线。ROC 曲线反映了算法 FRR 和 FAR 间的权衡关系。ROC 曲线与  $y = x$  的交点越靠近原点说明算法性能越好。

观察图 7 可以清楚地发现, 相比于其他的融合算法本文提出的融合策略拥有更高的识别精度。对于基于以手指作为载体的生物特征识别, 本文提出的三模态多级特征融合方法可以准确地提取典型特征。并且, 获取的融合特征具有良好的表征能力, 增加模态之间的区分性, 可以有效提高识别精度。

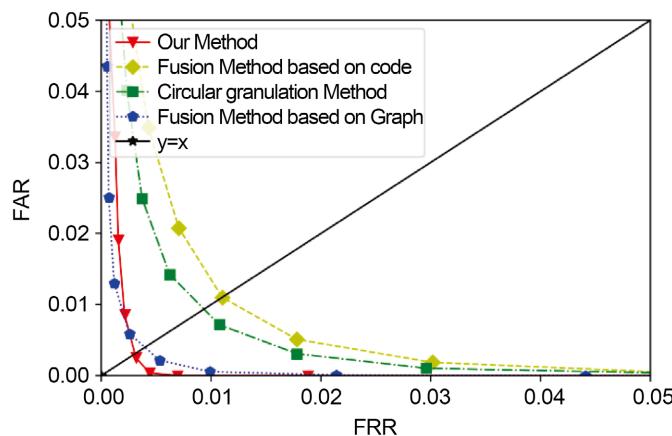
**Figure 7.** Comparison results of ROC for different fusion methods

图 7. 不同融合算法 ROC 结果比较

## 5. 总结

本文提出了一种基于卷积神经网络的手指三模态多级特征融合识别算法, 利用不同的融合策略实现了三模态多级特征融合的目的, 通过不同的尺寸归一化方法解决了不同尺寸特征图无法融合的问题。通过多种对比实验说明了本文提出的三模态多级特征融合框架通过综合利用手指三模态的生物特征信息来共同刻画和表达个体的身份属性, 能够有效提高生物特征识别技术的准确性、普遍性和可靠性。

本文的研究重点是将手指看作多模态生物特征载体, 建立手指特征的整体表达, 可实质性、最大程度地挖掘手指在身份识别方面的使用价值。安全性是生物特征识别的根本, 手指特征整体表达以融合特征为基础, 本身具备特征加密性。未来我们将构建深度学习框架下的手指三模态深度表达模型, 探索全新的网络训练机制, 实现自学习的特征聚合方式。

## 参考文献

- [1] Nguyen, K., Fookes, C., Sridharan, S., Tistarelli, M. and Nixon, M. (2018) Super-Resolution for Biometrics: A Comprehensive Survey. *Pattern Recognition*, **78**, 23-42. <https://doi.org/10.1016/j.patcog.2018.01.002>
- [2] Ross, A. and Jain, A. (2003) Information Fusion in Biometrics. *Pattern Recognition Letters*, **24**, 2115-2125. [https://doi.org/10.1016/S0167-8655\(03\)00079-5](https://doi.org/10.1016/S0167-8655(03)00079-5)
- [3] Ross, A. and Jain, A.K. (2004) Multimodal Biometrics: An Overview. 2004 12th European Signal Processing Conference, Vienna, 6-10 September 2004, 1221-1224.
- [4] Nageshkumar, M., Mahesh, P. and Swamy, M.S. (2009) An Efficient Secure Multimodal Biometric Fusion Using Palmprint and Face Image. *International Journal of Computer Science Issues (IJCSI)*, **1**, 49-53.
- [5] Mittal, N. (2014) Hand Based Biometric Authentication. Banasthali University, Rajasthan.
- [6] Shaikh, J. and Uttam, D. (2016) Review of Hand Feature of Unimodal and Multimodal Biometric System. *International Journal of Computer Applications*, **133**, 19-24. <https://doi.org/10.5120/ijca2016907853>
- [7] Asaari, M.S.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 Application*, **41**, 3367-3382. <https://doi.org/10.1016/j.eswa.2013.11.033>
- [8] Khellat-Kihel, S., Abrishambaf, R., Monteiro, J., et al. (2016) Multimodal Fusion of the Finger Vein, Fingerprint and the Finger-Knuckle-Print Using Kernel Fisher Analysis. *Applied Soft Computing*, **42**, 439-447. <https://doi.org/10.1016/j.asoc.2016.02.008>
- [9] Gudavalli, M., Raju, S.V., Babu, A.V. and Kumar, D.S. (2012) Multimodalbiometrics—Sources, Architecture and Fusion Techniques: An Overview. 2012 International Symposium on Biometrics and Security Technologies, Taipei, 26-29 March 2012, 27-34. <https://doi.org/10.1109/ISBAST.2012.24>
- [10] 孙权森, 曾生根, 王平安, 夏德深. 典型相关分析的理论及其在特征融合中的应用[J]. 计算机学报, 2005, 28(9):

1524-1533.

- [11] Fayyaz, M., Hajizadeh-Saffar, M., Sabokrou, M., Hoseini, M. and Fathy, M. (2015) A Novel Approach for Finger Vein Verification Based on Self-Taught Learning. *9th Iranian Conference on Machine Vision and Image Processing (MVIP)*, Tehran, 18-19 November 2015, 88-91. <https://doi.org/10.1109/IranianMVIP.2015.7397511>
- [12] 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>
- [13] Radzi, S., Hani, M. and Bakhteri, R. (2016) Finger-Vein Biometric Identification Using Convolutional Neural Network. *Turkish Journal of Electrical Engineering & Computer Sciences*, **24**, 1863-1878.
- [14] Qin, H. and El-Yacoubi, M. (2017) Deep Representation-Based Feature Extraction and Recovering for Finger-Vein Verification. *IEEE Transactions on Information Forensics and Security*, **12**, 1816-1829. <https://doi.org/10.1109/TIFS.2017.2689724>
- [15] Tang, S., Zhou, S., Kang, W., Wu, Q. and Deng, F. (2019) Finger Vein Verification Using a Siamese CNN. *IET Biometrics*, **8**, 306-315. <https://doi.org/10.1049/iet-bmt.2018.5245>
- [16] Hou, B. and Yan, R. (2018) Convolutional Auto-Encoder Model for Finger-Vein Verification. 2018 *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome, 11-13 June 2018, 1-5. <https://doi.org/10.1109/MeMeA.2018.8438719>
- [17] Liu, W., Li, W., Sun, L., Zhang, L. and Peng, C. (2018) Finger Vein Recognition Based on Deep Learning. 2017 12th *IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Siem Reap, 18-20 June 2017, 205-210. <https://doi.org/10.1109/ICIEA.2017.8282842>
- [18] Raghavendra, R., Venkatesh, S., Raja, K.B. and Busch, C. (2017) Transferable Deep Convolutional Neural Network Features for Finger Vein Presentation Attack Detection. 2017 5th *International Workshop on Biometrics and Forensics (IWBF)*, Coventry, 4-5 April 2017, 1-5. <https://doi.org/10.1109/IWBF.2017.7935108>
- [19] Brunelli, R. and Falavigna, D. (1995) Person Identification Using Multiple Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**, 955-966. <https://doi.org/10.1109/34.464560>
- [20] Bai, G. and Yang, J. (2016) A New Pixel-Based Granular Fusion Method for Finger Recognition. 8th *International Conference on Digital Image Processing*, Chengdu, 20-22 May 2016, Article ID: 100330. <https://doi.org/10.1117/12.2245270>
- [21] Zhang, H., Li, S., Shi, Y. and Yang, J. (2019) Graph Fusion for Finger Multimodal Biometrics. *IEEE Access*, **7**, 28607-28615. <https://doi.org/10.1109/ACCESS.2019.2902133>
- [22] Zhang, H., Han, H., Cui, J., Shan, S. and Chen, X. (2018) RGB-D Face Recognition via Deep Complementary and Common Feature Learning. 2018 13th *IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, Xi'an, 15-19 May 2018, 8-15. <https://doi.org/10.1109/FG.2018.00012>
- [23] Wang, A., Cai, J., Lu, J. and Cham, T. (2015) Mmss: Multi-Modal Sharable and Specific Feature Learning for RGB-D Object Recognition. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 1125-1133. <https://doi.org/10.1109/ICCV.2015.134>
- [24] Soleymani, S., Dabouei, A., Kazemi, H., Dawson, J. and Nasrabadi, N.M. (2018) Multi-Level Feature Abstraction from Convolutional Neural Networks for Multimodal Biometric Identification. 24th *International Conference on Pattern Recognition (ICPR 2018)*, Beijing, 20-24 August 2018, 3469-3476. <https://doi.org/10.1109/ICPR.2018.8545061>