

基于多模态融合的残缺掌纹识别

刘晨龙, 杨璐, 郝凡昌

(山东建筑大学 计算机科学与技术学院, 山东 济南 250101)

摘要: 针对掌纹图像在实际应用中会因为各种因素出现成像不完整导致识别准确率下降的问题, 提出一种融合指静脉特征的残缺掌纹识别方法, 以增强鲁棒性并提升识别性能。该方法创新性地引入自适应注意力机制, 以动态调整掌纹和指静脉特征的融合权重, 从而有效弥补掌纹信息缺失的影响。该方法采用基于VGG16的特征提取网络分别提取掌纹和指静脉特征, 将特征图压缩为固定维度的特征向量, 消除了输入图像的尺寸差异, 增强了对局部特征的鲁棒性。实验结果表明, 相较于仅使用残缺掌纹, 自适应融合策略在两个数据集上的识别准确率分别提升了7.61%与6.82%, 验证了该方法的有效性。此外, 为进一步验证特征提取网络对识别性能的影响, 还进行了不同特征提取网络的对比实验, 证明了合理的特征提取和融合策略对于残缺掌纹识别至关重要。

关键词: 残缺掌纹识别; 多模态融合; 自适应注意力机制; 生物特征识别; 卷积神经网络

DOI: 10.11907/rjtk.251182

扫描二维码阅读全文:

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2026)006-0188-07



Incomplete Palmprint Recognition Based on Multimodal Fusion

LIU Chenlong, YANG Lu, HAO Fanchang

(School of Computer Science and Technology, Shandong Jianzhu University, Ji'nan 250101, China)

Abstract: To address the issue of incomplete palmprint images in practical applications, which can lead to decreased recognition accuracy, a robust incomplete palmprint recognition method integrating finger vein features is proposed. This method innovatively introduces an adaptive attention mechanism to dynamically adjust the fusion weights of palmprint and finger vein features, effectively compensating for the missing palmprint information. A feature extraction network based on VGG16 is employed to separately extract palmprint and finger vein features, compressing feature maps into fixed-dimensional feature vectors. This process eliminates variations in input image sizes and enhances robustness to local features. Experimental results demonstrate that the adaptive fusion strategy improves recognition accuracy by 7.61 and 6.82 percentage points on two different datasets compared to using incomplete palmprints alone, validating the effectiveness of the proposed method. Furthermore, to assess the impact of feature extraction networks on recognition performance, comparative experiments with different networks were conducted. The results indicate that an appropriate feature extraction and fusion strategy is crucial for incomplete palmprint recognition.

Key Words: incomplete palmprint recognition; multimodal fusion; adaptive attention mechanism; biometric recognition; convolutional neural network

0 引言

生物特征识别技术凭借高安全性与便捷性的双重优势, 已在身份认证、金融支付、边境安检等诸多领域实现规

模化应用^[1-2]。掌纹识别作为生物特征识别技术的重要分支, 凭借自身独有的特征稳定性与个体区分度, 在身份验证场景中具有较高的应用价值。相较于人脸识别, 掌纹特征不受面部表情变化的干扰, 识别稳定性更强; 与指纹识别相比, 掌纹的磨损程度更低, 且包含丰富的纹理与几何

收稿日期: 2025-04-09

基金项目: 国家自然科学基金项目(62076151); 山东省泰山学者资助项目(tsqn202211182); 山东省自然科学基金项目(ZR2022MF272); 山东省重点技术研发计划项目(2019GGX101068)

作者简介: 刘晨龙(2001-), 男, 山东建筑大学计算机科学与技术学院硕士研究生, 研究方向为生物特征识别; 杨璐(1988-), 女, 博士, 山东建筑大学计算机科学与技术学院教授、硕士生导师, 研究方向为生物特征识别、机器学习; 郝凡昌(1981-), 男, 博士, 山东建筑大学计算机科学与技术学院副教授、硕士生导师, 研究方向为机器学习、图像处理、算法分析与设计。本文通讯作者: 郝凡昌。

特征, 能为身份识别提供更充分的判别依据^[3-5]。

深度学习技术的快速发展推动掌纹识别研究取得了诸多突破, 在采集条件理想、图像数据完整的场景下, 基于深度学习的掌纹识别模型已能实现较高的识别精度。但在实际应用场景中, 掌纹图像采集过程易受拍摄角度、光照条件、手部遮挡、成像模糊等多种因素影响, 极易出现图像信息残缺的问题。这会直接导致特征提取的有效性下降, 进而影响后续的特征匹配与识别结果。

残缺掌纹识别的难点在于图像缺失区域往往包含关键的个体差异化特征, 这类信息的丢失会大幅降低识别模型的判别能力。传统掌纹识别方法以手工提取的纹理、几何特征为核心, 对掌纹图像的完整性具有较强依赖性, 在面对残缺掌纹时鲁棒性不足。深度学习虽能自动挖掘图像深层特征, 但其特征学习过程仍以完整的输入信息为基础, 当掌纹图像关键区域缺失时, 模型的识别性能也会显著下滑。因此, 如何提升模型对残缺掌纹图像的处理能力, 成为掌纹识别领域亟待解决的问题。

针对这一问题, 现有研究从特征增强、数据预处理、多模态信息融合等多个方向展开探索。其中, 多模态信息融合因能利用不同生物特征的互补性弥补单一模态的信息缺失, 成为提升生物识别系统鲁棒性的重要研究方向。将指静脉特征与掌纹特征相结合, 借助合理的融合策略实现信息互补, 能够有效改善掌纹信息残缺带来的识别精度下降问题, 这也成为本文解决残缺掌纹识别问题的核心思路。

1 相关工作

掌纹识别领域的传统方法以纹理与几何特征为核心开展身份验证, 通过提取掌纹纹理模式、局部弯曲度、拇指区域等细节特征完成分类^[6-8]。文献[9]设计一种具有鲁棒性的算法实现掌纹局部脊方向与频率的精准估计, 为脊线和细节特征提取提供支撑, 并通过对齐式细节匹配算法完成掌纹匹配。文献[10]提出掌纹特征线族概念, 将纹线特征划分为主次两类, 依托双向匹配策略输出识别结果。文献[11]则将脊线距离作为辅助特征设计新型细节匹配算法, 在提升掌纹对齐速度与匹配精度的同时有效降低了时间成本。但这类传统方法均高度依赖完整的掌纹图像, 在掌纹存在局部缺失时, 特征提取的鲁棒性大幅下降, 难以保障识别结果的稳定性。

近年来, 深度学习技术推动掌纹识别研究取得了显著突破, 卷积神经网络(Convolutional Neural Network, CNN)的应用大幅提升了掌纹识别精度^[12-14]。文献[15]基于MobileNet设计轻量级掌纹识别网络, 通过压缩网络层数、扩大卷积核, 实现了模型在移动设备上的高效运行。文献[16]针对网络泛化性不足的问题, 提出半预训练的非接触式掌纹识别网络, 融合ResNet-152底层结构与掌纹识别专

用辅助层, 让模型在低质量图像下仍能稳定提取有效的掌纹特征。但上述深度学习方法仍以单一掌纹特征为研究对象, 未能利用其他生物特征的互补信息, 在掌纹残缺场景下的识别精度与鲁棒性仍有提升空间。

为弥补单一模态的缺陷, 多模态融合成为掌纹识别领域的研究热点。通过融合掌纹、指纹、指静脉等不同生物特征开展联合训练, 可有效弥补单一模态在信息遮挡或缺失时的不足, 显著提升识别系统的性能^[17-19]。例如, 文献[20]融合指关节纹与掌纹特征, 通过深度学习提取各模态特征后, 利用多类SVM完成分数级融合。文献[21]提出非平稳特征融合技术, 通过交错矩阵融合人脸与掌纹局部特征, 结合DCT算法与高斯混合模型实现融合特征的类别概率估计。此类方法多通过早、晚期融合策略实现多模态信息整合。

自适应注意力机制作为动态加权融合策略, 在多模态融合中展现出显著优势。其可根据数据质量对不同模态特征分配差异化权重, 动态调整各模态对识别结果的贡献度。在处理残缺或遮挡数据时, 该机制能自动聚焦于信息完整的可靠模态, 大幅提升识别的鲁棒性。基于此, 本文针对残缺掌纹识别的难点, 提出融合指静脉特征的自适应多模态识别方法。以VGG16为基础网络分别提取掌纹与指静脉的深度特征, 同时引入自适应注意力机制, 根据掌纹信息的完整程度动态调整双模态特征的融合权重, 以充分利用可靠的特征信息, 实现残缺掌纹识别性能与鲁棒性的双重提升。

2 基于多模态融合的残缺掌纹识别方法

本文提出一种基于自适应注意力机制的多模态生物特征融合识别网络, 通过结合掌纹和指静脉两种模态的数据, 辅助提高对残缺掌纹图像的识别性能。网络框架如图1所示。该框架以掌纹图像和指静脉图像作为双模态输入, 通过特征提取、动态融合与分类决策3阶段实现残缺掌纹识别。

2.1 共享特征提取模块

特征提取模块负责从输入的掌纹和指静脉图像中提取深度特征。为保证两个模态的特征在同一空间内具有更好的融合效果, 采用共享参数的特征提取器, 即使用同一个卷积神经网络对掌纹和指静脉数据进行特征提取。掌纹和指静脉共享同一特征提取器, 可以通过反向传播的梯度约束, 迫使网络学习对两种模态均敏感的通用滤波器, 例如特定的卷积核可能同时响应掌纹的脊线走向与指静脉的血管曲率。同时, 共享参数可以使模型复杂度降低一半, 改善小样本场景下的过拟合问题。

本文的共享特征提取器基于VGG16网络, VGG16是一个经典的卷积神经网络, 具有强大的特征提取能力。该网络包含13个卷积层、5个最大池化层以及后续的全连接

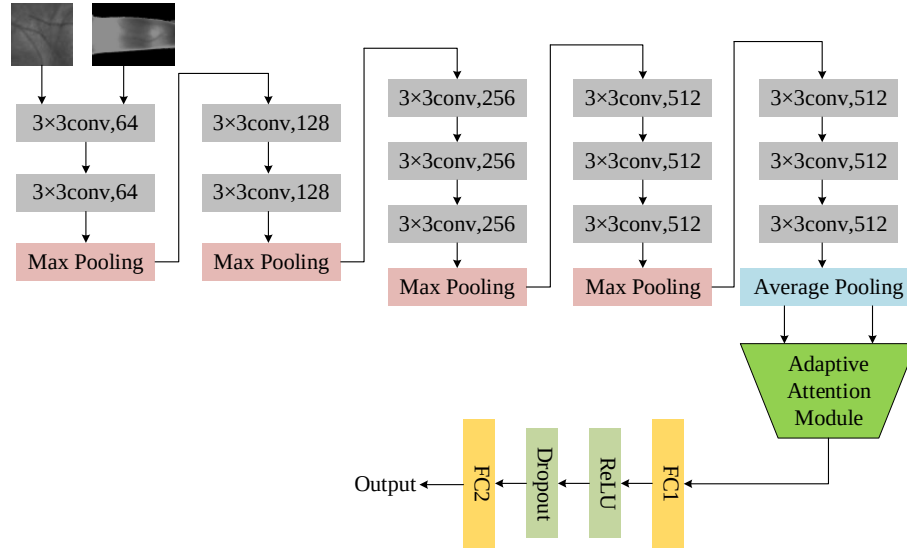


Fig. 1 Method framework

图1 方法网络框架

层。通过使用预训练的 VGG16 网络,可充分利用其在大规模数据集上训练得到的通用特征。为了使其更适用于所做的任务,对其进行了一定修改。

具体而言,首先,移除了原网络末端的全连接层,保留前 13 层卷积层,使得输出仍保持空间结构信息,以适应掌纹和指静脉数据的几何表示。浅层卷积核如 conv1_2 的 3×3 滤波器能有效捕捉掌纹的纹理细节与指静脉的血管交叉点,同时深层卷积层如 conv5_3 能够编码全局空间结构信息,避免高层语义信息的过度抽象,适用于掌纹和指静脉的细粒特征表达。其次,在卷积层后添加全局平均池化层,将特征图压缩为固定维度的特征向量,从而消除输入图像的尺寸差异,并增强对局部特征的鲁棒性。表 1 给出了详细的网络配置细节,包括每一层的输出、内核、步幅和填充情况等。

Table 1 Feature extraction network configuration

表1 特征提取网络配置

层名称	结构	输出尺寸
卷积层 1	$3 \times 3 \times 64$, 步长 1, ReLU	$224 \times 224 \times 64$
卷积层 2	$3 \times 3 \times 64$, 步长 1, ReLU	$224 \times 224 \times 64$
最大池化	2×2 , 步长 2	$112 \times 112 \times 64$
卷积层 3	$3 \times 3 \times 128$, 步长 1, ReLU	$112 \times 112 \times 128$
卷积层 4	$3 \times 3 \times 128$, 步长 1, ReLU	$112 \times 112 \times 128$
最大池化	2×2 , 步长 2	$56 \times 56 \times 128$
卷积层 5	$3 \times 3 \times 256$, 步长 1, ReLU	$56 \times 56 \times 256$
卷积层 6	$3 \times 3 \times 256$, 步长 1, ReLU	$56 \times 56 \times 256$
卷积层 7	$3 \times 3 \times 256$, 步长 1, ReLU	$56 \times 56 \times 256$
最大池化	2×2 , 步长 2	$28 \times 28 \times 256$
卷积层 8	$3 \times 3 \times 512$, 步长 1, ReLU	$28 \times 28 \times 512$
卷积层 9	$3 \times 3 \times 512$, 步长 1, ReLU	$28 \times 28 \times 512$
卷积层 10	$3 \times 3 \times 512$, 步长 1, ReLU	$28 \times 28 \times 512$
最大池化	2×2 , 步长 2	$14 \times 14 \times 512$
卷积层 11	$3 \times 3 \times 512$, 步长 1, ReLU	$14 \times 14 \times 512$
卷积层 12	$3 \times 3 \times 512$, 步长 1, ReLU	$14 \times 14 \times 512$
卷积层 13	$3 \times 3 \times 512$, 步长 1, ReLU	$14 \times 14 \times 512$
全局平均池化	—	$1 \times 1 \times 512$

2.2 自适应注意力模块

自适应注意力模块是本文方法的核心部分,其作用是根据每个模块的特征质量,自动计算每个模态的注意力权重,并将不同模态的特征加权融合。自适应注意力机制可以有效应对多模态数据中的信息不平衡问题,尤其是在处理残缺数据时,能够动态调整各模态权重。具体处理流程如图 2 所示。

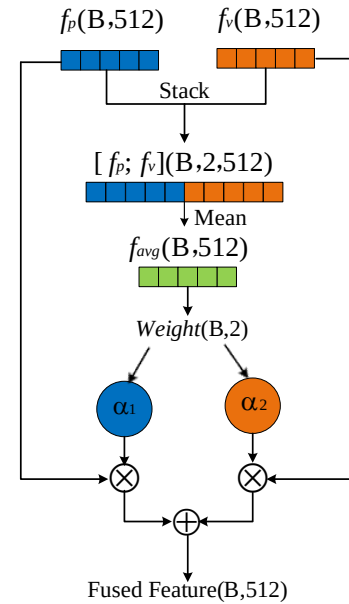


Fig. 2 Adaptive attention mechanism feature processing flow

图2 自适应注意力机制特征处理流程

在该模块中,首先,对每个模态的特征进行堆叠。设掌纹特征向量为 f_p ,指静脉特征向量为 f_v ,二者沿模态维度堆叠为张量 $F = [f_p; f_v] \in \mathbb{R}^{2 \times 512}$ 。其次,计算每个样本的平均特征,沿模态维度取均值,得到全局平均特征 $f_{avg} = \frac{1}{2} (f_p + f_v)$ 。该操作捕捉双模态的共性信息,为权重生成

提供统计依据。相较于直接拼接特征, 均值池化可减少后续全连接层输入维度的50%, 显著降低了计算量。

下一步通过权重生成网络为每个特征计算权重。权重生成网络由两层全连接层构成, 第一层的作用是维度压缩, 其将512维的输入映射至128维隐空间, 激活函数采用ReLU, 表达式如公式(1)所示:

$$h = \text{ReLU}(W_1 f_{\text{avg}} + b_1) \quad (1)$$

其中, $W_1 \in \mathbb{R}^{128 \times 512}$ 。第二层是权重计算层, 将上一步的输出进一步映射至二维输出, 经Softmax归一化后得到模态权重 α_p, α_v , 满足 $\alpha_p + \alpha_v = 1$ 。

通过加权融合后, 得到的融合特征表示如公式(2)所示:

$$F_{\text{fused}} = \alpha_p f_p + \alpha_v f_v \quad (2)$$

2.3 分类模块

分类模块负责将融合后的特征送入全连接层进行分类。具体地, 经过自适应注意力模块加权融合后的特征向量首先通过第一个全连接层进行降维, 减小特征维度, 并通过ReLU函数增强非线性表达能力。之后, 为了防止过拟合, 使用Dropout层进行正则化, 强制网络学习冗余特征表示。具体表达式如公式(3)所示:

$$h_{\text{cls}} = \text{Dropout}(\text{ReLU}(W_{\text{cls}}^1 F_{\text{fused}} + b_{\text{cls}}^1)) \quad (3)$$

最终通过另一个全连接层映射至类别空间, 输出维度为身份类别数C, 如公式(4)所示:

$$y = W_{\text{cls}}^2 h_{\text{cls}} + b_{\text{cls}}^2 \in \mathbb{R}^C \quad (4)$$

为了优化网络参数, 使得网络能根据融合特征有效区分不同身份, 采用交叉熵损失函数进行训练。交叉熵损失在分类任务中得到了广泛应用, 其能够衡量模型预测分布与真实分布之间的差异。其数学表达如下:

$$L_{\text{CE}} = -\sum_{i=1}^C y_i \log(\hat{y}_i) \quad (5)$$

其中, C为身份类别数, y_i 为真实身份类别的one-hot编码, $\hat{y}_i$ 为模型预测的类别概率。通过softmax运算得到:

$$\hat{y}_i = \frac{\exp(z_i)}{\sum_{j=1}^C \exp(z_j)} \quad (6)$$

其中, z_i 是网络在类别i处的输出logit值。

为兼顾模型的收敛速度与泛化能力, 本文在训练过程中采用多种优化策略: ①选用Adam优化器进行网络参数更新, 将初始学习率设置为0.001, 使梯度更新过程更稳定, 同时引入L2正则项对模型参数进行约束, 抑制模型过拟合; ②在数据预处理阶段引入水平翻转等数据增强技术, 扩充训练样本的多样性, 提升模型对输入图像变化的适应性。基于上述优化策略, 本文将模型的总损失函数定义为交叉熵损失与L2正则项之和。其数学表达式如公式(7)所示:

$$L_{\text{total}} = L_{\text{CE}} + \lambda \|W\|^2 \quad (7)$$

其中, $\lambda \|W\|^2$ 是L2正则项, λ 是正则化系数。通过该损失函数能够有效优化融合特征, 提高分类的准确率。

3 实验与结果分析

为验证本文提出的基于自适应注意力机制的残缺掌纹识别方法的有效性与鲁棒性, 本文设计了一系列对比实验与消融实验, 从融合有效性、跨数据集泛化性、特征提取网络适配性3个维度对模型性能进行全面评估。实验所用数据均来自公开的掌纹与指静脉数据集, 并针对残缺掌纹的实际应用场景进行了针对性的预处理, 以确保实验结果的真实性与可参考性。

3.1 实验数据集与预处理

在实际的生物特征识别应用中, 掌纹与指静脉特征均采集自同一手部, 同源的双模态数据能有效提升识别系统的性能, 但在目前公开的生物特征识别数据集中, 尚未提供掌纹与指静脉的同源联合数据集。因此, 本文选用独立的公开掌纹数据集与指静脉数据集, 通过编号匹配的方式人为定义同源关系, 以此验证所提出的多模态融合方法的有效性。

Tongji Contactless Palmprint Dataset是由同济大学建设的非接触式掌纹数据库, 该数据库包含来自300名志愿者共600个手掌的掌纹图像^[22]。数据采集过程分为两个阶段, 每个手掌每阶段采集10幅图像, 共12 000幅图像, 平均采集间隔为61天。同济掌纹数据库示例如图3所示。

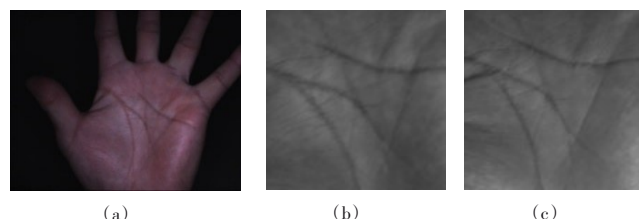


Fig. 3 Examples of Tongji palmprint dataset

图3 同济掌纹数据库示例

Tsinghua Palmprint Database是由清华大学建设的接触式掌纹数据库^[23]。该数据库包含使用Hisign商用掌纹扫描仪捕获的来自80名受试者的1 280幅掌纹图像, 每人每只手采集8幅图像, 所有图像大小均为2 040×2 040像素。清华掌纹数据库示例如图4所示。

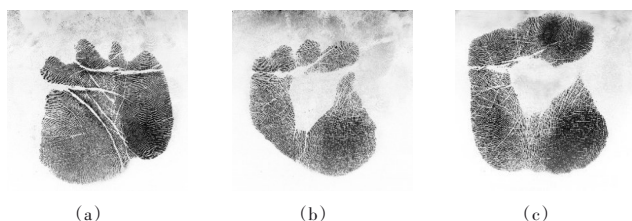


Fig. 4 Examples of Tsinghua palmprint database

图4 清华掌纹数据库示例

SDUMLA-HMT是由山东大学建立的同源多模态特征数据库, 包含来自106个人的真实多模态数据^[24]。本文主要使用其指静脉数据, 在指静脉数据采集过程中, 每个受

试者被要求提供双手食指、中指和无名指的图像,对每个手指分别采集6次以获得6幅指静脉图像,共获得3 816幅指静脉图像,图像大小为320×240像素。山东大学指静脉数据库示例如图5所示。

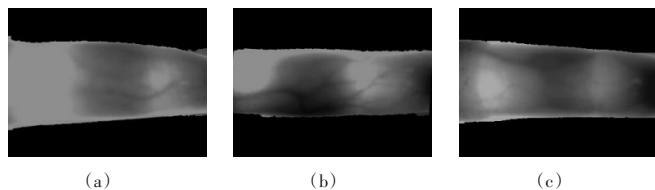


Fig. 5 Examples of Shandong University finger vein database

图5 山东大学指静脉数据库示例

由于掌纹数据集和指静脉数据集来源不同,本文通过编号匹配的方式进行身份绑定,即掌纹数据集和指静脉数据集中相同编号的个体被认为是同一身份。例如,Tongji数据集中的第1号个体与SDUMLA-HMT数据集中的第1号个体视为同一身份,并用于模型训练和测试。尽管该匹配策略并非真实的同源数据,但在实验中仍可以验证多模态融合方法在身份识别任务中的有效性,并为后续研究提供参考。

在真实场景中,掌纹识别系统可能会遇到因手部遮挡、光照变化或采集不完整等因素导致的残缺掌纹问题。因此,为了模拟这种情况,本文在数据预处理中随机为掌纹图像添加遮挡,以增强模型的鲁棒性,并验证其对不完整掌纹的识别能力。具体而言,在Tongji数据集中的图像中随机选择一个区域进行遮挡,遮挡比例控制在10%~40%之间,并采用白色填充来模拟信息缺失。遮挡区域的位置和大小均随机变化,以避免学习固定模式。Tsinghua数据集因其本身即存在采集时捺印不完整的情况,所以不作处理。处理后的同济掌纹数据库示例如图6所示。

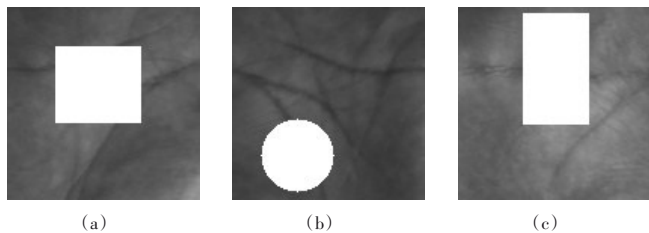


Fig. 6 Examples of processed Tongji palmprint database

图6 处理后的同济掌纹数据库示例

3.2 融合有效性验证

为了验证不同模态特征在身份识别任务中的作用,使用处理后的同济掌纹数据集和山东大学指静脉数据集,分别测试了仅使用残缺掌纹特征、仅使用指静脉特征、掌纹与指静脉特征直接拼接、掌纹与指静脉特征基于自适应注意力融合4种特征输入方式下的识别准确率。实验结果如表2所示。

从表2可以看出,仅使用单一模态的识别效果较差,其中残缺掌纹的准确率为89.17%,而指静脉的准确率仅为

Table 2 Results of different fusion methods

表2 不同融合方式结果

特征输入方式	识别准确率/%
仅使用残缺掌纹	89.17
仅使用指静脉	80.83
特征直接拼接	94.22
自适应注意力融合	96.78

80.83%,表明单一生物特征的判别能力存在局限性。相比之下,直接拼接掌纹与指静脉特征的方式显著提高了准确率,达到94.22%,说明双模态融合可以有效提升识别性能。进一步地,采用自适应注意力机制对掌纹和指静脉特征进行融合,在测试集上的准确率达到96.78%,相比直接拼接的方式提高了2.56%,说明自适应注意力机制能够有效增强多模态特征的表达能力,提高生物特征融合的鲁棒性。

实验结果表明,相较于单一模态输入,多模态融合可以显著提升身份识别的准确率。自适应注意力机制能够更有效地权衡不同模态的重要性,进一步提高识别性能,同时能获得较高质量的数据,具有良好的鲁棒性。针对掌纹不完整这种低质量数据,该方法能够动态调整掌纹和指静脉模态的贡献,弥补掌纹特征的缺失,使得模型在面对掌纹残缺、部分遮挡等情况时,仍能保持较强的识别能力,提高对低质量数据的适应性。

3.3 不同数据集对比

为了进一步验证本文方法的泛化能力,将同济掌纹数据集更换为清华掌纹数据集,并在相同的实验设置下进行了测试。清华掌纹数据集的图像质量、采集环境和同济掌纹数据集存在一定差异,从而使掌纹特征的整体可区分性可能有所变化。实验仍采用4种不同的特征输入方式进行测试,结果如表3所示。

Table 3 Results of different datasets

表3 不同数据集结果

特征输入方式	识别准确率/%
仅使用残缺掌纹	86.04
仅使用指静脉	80.83
特征直接拼接	90.13
自适应注意力融合	92.86

从实验结果可以看出,以清华掌纹数据集为实验对象时,仅使用残缺掌纹特征的识别准确率降至86.04%,相较同济掌纹数据集下降明显。出现这一现象的主要原因是清华掌纹数据集中部分样本存在图像模糊、天然残缺严重等问题(见图7),这类低质量样本会影响掌纹特征提取的有效性,导致模型无法充分挖掘掌纹的判别特征,进而降低了识别准确率。

但即便在低质量的清华掌纹数据集场景下,双模态融合仍能显著提升模型的识别性能。其中,特征直接拼接方式的识别准确率为90.13%,相较于仅使用残缺掌纹特征方式提升了4.09%。本文提出的自适应注意力融合方法识别准确率为92.86%,相较于仅使用残缺掌纹特征方式提升了

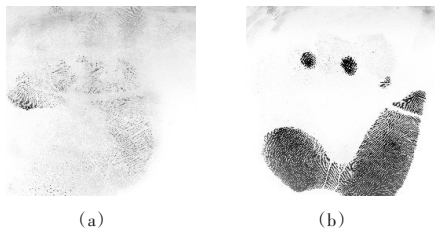


Fig. 7 Examples of low-quality data in Tsinghua palmprint database
图7 清华掌纹数据库低质量数据示例

约7%,相较于特征直接拼接方式提升了2.73%。这一结果表明,自适应注意力机制能更有效地挖掘双模态特征之间的互补信息,即使在掌纹特征质量较差的场景下,也能通过动态调整权重实现高效的特征融合,提升识别性能。

同时,同济掌纹数据集与清华掌纹数据集在成像方式、图像分辨率、采集环境等方面存在显著差异,而本文方法在两个数据集上均能保持稳定的识别性能,证明了该方法具有良好的跨数据集的鲁棒性与泛化能力,能适配不同采集场景下的掌纹识别任务。

3.4 不同特征提取网络对比

为了验证不同特征提取网络对识别效果的影响,本文将特征提取网络由VGG16替换为ResNet50,并在同济掌纹数据集和山东大学指静脉数据集上采取与上文相同的实验策略进行测试,结果如表4所示。

Table 4 Results of different feature extraction networks
表4 不同特征提取网络结果 (%)

特征输入方式	VGG16 识别准确率	ResNet50 识别准确率
仅使用残缺掌纹	89.17	82.44
仅使用指静脉	80.83	86.61
特征直接拼接	94.22	91.35
自适应注意力融合	96.78	94.89

从实验结果可以看出,使用VGG16作为特征提取网络时,大部分情况下的识别准确率均优于使用ResNet50的情况。但在仅使用指静脉特征的情况下,ResNet50取得了86.61%的识别准确率,相比VGG16反而有所提升。主要原因是VGG16采用3×3卷积核,具有较小的感受野,深度较浅,所以特征提取过程中保留了更多局部纹理信息,能够较好地捕捉掌纹的细节特征。而ResNet50由于网络更深,并且在早期层级进行了较大步长的池化,可能导致部分局部纹理信息丢失,使得掌纹识别性能下降。指静脉识别主要依靠其血管结构,通常具有较强的全局连贯性。ResNet50通过残差连接和更深的网络结构,能够学习到更具全局性的特征表达,因此针对指静脉数据的识别效果优于VGG16。

上述实验结果表明,不同生物特征类型存在差异,对特征提取网络的适配性也各不相同。在多模态融合识别任务中,VGG16能兼顾掌纹与指静脉的特征提取需求,综合识别性能更优,因此成为本文的基础特征提取网络。这也说明在多模态生物特征识别研究中,根据不同模态的特

征特点选择合适的特征提取网络是提升模型识别性能的关键。

4 结论

本文针对掌纹在被遮挡或损坏情况下识别精度降低的问题,提出一种基于自适应注意力的多模态融合识别方法,并进行了系统性的实验和分析。主要结论如下:

(1)多模态融合是提升残缺掌纹识别性能的核心策略。掌纹与指静脉特征存在显著的互补性,掌纹特征的优势在于局部纹理细节,指静脉特征的优势在于全局结构信息,二者融合能有效弥补残缺掌纹的信息缺失问题。相较于单一模态识别方法,双模态融合策略能大幅提升残缺掌纹的识别准确率。

(2)自适应注意力机制为多模态融合提供了更优的实现方式。相较于特征直接拼接的静态融合策略,自适应注意力机制能根据双模态特征的质量动态调整权重系数,从而有效处理多模态数据的信息不平衡问题,在掌纹残缺时自动突出指静脉特征的判别作用,进一步提升了识别系统的鲁棒性。

(3)特征提取网络与生物特征的适配性对识别效果具有显著影响。不同的网络结构在特征提取上各有侧重,VGG16的小卷积核与浅网络结构更适配掌纹局部纹理特征的提取,ResNet50的残差连接与深网络结构更适合指静脉全局结构特征的提取。这提示在多模态融合任务中,需根据不同模态的特征特点选择适配的基础网络。

(4)本文方法具备良好的泛化能力与鲁棒性。该方法在同济、清华两个不同成像方式与图像质量的掌纹数据集上均保持了稳定的识别性能,能有效适配实际应用中复杂的掌纹采集场景,可应对图像遮挡、模糊、残缺等多种低质量图像情况。

整体而言,本文的实验结果表明,合理的特征提取网络选择与多模态融合策略,是提升多模态生物特征识别性能的关键因素。后续研究将进一步探索针对不同生物特征的专用特征提取网络,构建多网络协同的多模态融合方法,实现掌纹和指静脉特征的精细化提取与融合。同时,将收集并构建掌纹与指静脉的同源联合数据集,在更贴近实际应用的场景中验证方法的有效性,并探索该方法在复杂采集环境下的应用优化策略,推动其在实际生物特征识别系统中的落地与应用。

参考文献:

[1] SHARMA S, SAINI A, CHAUDHURY S. A survey on biometric cryptosystems and their applications [J]. Computers Security, 2023, 134: 103458.
[2] MALLOULI F, KHELIFI N, HELLAL A, et al. Biometric authentication comparison: toward secure human recognition[C]//2023 International Conference on Computational Science and Computational Intelligence, 2023: 1264-1270.

- [3] ZHONG D X, ZHU J S, DU X F. Review of palmprint recognition[J]. Pattern Recognition and Artificial Intelligence, 2019, 32(5):436-445.
钟德星,朱劲松,杜学峰. 掌纹识别研究进展综述[J]. 模式识别与人工智能, 2019, 32(5):436-445.
- [4] ZHAO S P, FEI L K, WEN J. Multiview-learning-based generic palmprint recognition: a literature review[J]. Mathematics, 2023, 11(5): 1261.
- [5] ALAUSA D, ADETIBA E, BADEJO J, et al. Contactless palmprint recognition system: a survey[J]. IEEE Access, 2022, 10: 132483-132505.
- [6] WIJESIRIWARDENE T I, KARUNAWALLABHA J R, WANNIARACHCHI V U, et al. Optimization of partial palmprint identification process using palmprint segmentation[C]// 22nd International Business Information Management Association Conference, 2013: 641-648.
- [7] PRASAD S K, ROUTRAY S, KHURANA R, et al. Classification of palmprint using principal line[C]//3rd International Conference on Information Systems, Technology and Management, 2009: 208-219.
- [8] BRUNO A, CARMINETTI P, GENTILE V, et al. Palmprint principal lines extraction[C]//2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications Proceedings, 2014: 50-56.
- [9] JAIN A K, FENG J J. Latent palmprint matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(6):1032-1047.
- [10] LI W X, XIA S X, ZHANG D P, et al. A novel palmprint recognition method based on bidirectional matching of mainline features[J]. Computer Research and Development, 2004(6):996-1002.
李文新,夏胜雄,张大鹏,等. 基于主线特征的双向匹配的掌纹识别新方法[J]. 计算机研究与发展, 2004(6):996-1002.
- [11] CHEN J L, GUO Z H. Palmprint matching by minutiae and ridge distance[C]//2nd International Conference on Cloud Computing and Security, 2016: 371-382.
- [12] GENOVESE A, PIURI V, SCOTTI F, et al. Touchless palmprint and finger texture recognition: a deep learning fusion approach[C]//2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, 2019: 1-6.
- [13] FAN R, HAN X H. Deep palmprint recognition algorithm based on self-supervised learning and uncertainty loss[J]. Signal Image and Video Processing, 2024, 18(5): 4661-4673.
- [14] TRABELSI S, SAMAI D, DORNAIKA F, et al. Efficient palmprint biometric identification systems using deep learning and feature selection methods[J]. Neural Computing & Applications, 2022, 34: 12119-12141.
- [15] JIA W, REN Q, ZHAO Y, et al. EEPNet: an efficient and effective convolutional neural network for palmprint recognition[J]. Pattern Recognition Letters, 2022, 159: 140-149.
- [16] CHAI T T, PRASAD S, YAN J, et al. Contactless palmprint biometrics using deepNet with dedicated assistant layers[J]. Visual Computer, 2023, 39(9): 4029-4047.
- [17] TAOUCHE C, BERKANE M, BATOUCHE M, et al. Multimodal biometric systems[C]// 2014 International Conference on Multimedia Computing and Systems, 2014: 301-308.
- [18] CHEN G, LUO D C, LIAN F Z, et al. A multimodal biometric recognition method based on federated learning[J]. IET Biometrics, 2024(1): 5873909.
- [19] PAHUJA S, GOEL N. Multimodal biometric authentication: a review[J]. AI Communications, 2024, 37(4): 525-547.
- [20] ATTIA A, MAZAA S, AKHTAR Z, et al. Deep learning-driven palmprint and finger knuckle pattern-based multimodal person recognition system[J]. Multimedia Tools and Applications, 2022, 81(8): 10961-10980.
- [21] AHMAD M I, WOO W L, DLAY S. Non-stationary feature fusion of face and palmprint multimodal biometrics[J]. Neurocomputing, 2016, 177: 49-61.
- [22] ZHANG L, LI L D, YANG A Q, et al. Towards contactless palmprint recognition: a novel device, a new benchmark, and a collaborative representation based identification approach[J]. Pattern Recognition, 2017, 69: 199-212.
- [23] DAI J F, FENG J J, ZHOU J. Robust and efficient ridge-based palmprint matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34: 1618-1632.
- [24] YIN Y L, LIU L L, SUN X W. SDUMLA-HMT: a multimodal biometric database[J]. Biometric Recognition, 2011, 7098: 260-268.

(责任编辑:黄健)