

基于注意力机制与残差网络的掌纹脊线距离估计

张浩,陈亚南,杨璐,许丽娜,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:掌纹脊线距离是掌纹脊线的重要纹理属性之一,在掌纹识别算法与图像质量评估方法中常作为一个重要参数被进行参考。目前常用的掌纹脊线距离估计方法为频域方法,但已有的频域方法对掌纹图像质量与完整性要求高,导致适用范围受限,且测量精度仍需进一步提高。针对上述问题,提出一种基于注意力机制与残差网络的掌纹脊线距离估计方法,将手工标注获取的脊线距离值作为输入掌纹图像块的类别进行识别。该方法以VGG16网络为基础,将CBAM注意力模块有机引入,增强网络对脊线结构信息的关注;并对损失函数进行改进,添加残差机制,来避免梯度消失问题。所用数据集由手工标注完成,针对数据集中存在的样本分布不均衡问题,采用数据增强的方法削弱类别分布间的不平衡,并设计样本不均衡损失函数对模型进行优化。实验结果表明,该方法在该数据集上相比频域方法在识别准确率上提升了14.2%,且相比部分经典深度学习方法在准确率上至少提升了5.1%。

关键词:掌纹脊线距离;VGG16;残差网络;注意力机制;损失函数

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2023)07-0061-07

doi:10.3969/j.issn.1673-629X.2023.07.009

Palmprint Ridge Distance Estimation Based on Attention Mechanism and Residual Network

ZHANG Hao, CHEN Ya-nan, YANG Lu, XU Li-na, HAO Fan-chang
(School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250101, China)

Abstract: Palmprint ridge distance is one of the important texture attributes of palmprint ridge. It is often referred as an important parameter in palmprint recognition algorithm and image quality evaluation method. At present, the commonly used method to estimate palmprint ridge distance is frequency domain method, but the existing frequency domain method has high requirements on the quality and integrity of palmprint image, resulting in limited application scope and further improvement of measurement accuracy. To solve these problems, a palmprint ridge distance estimation method based on attention mechanism and residual network is proposed. The ridge distance value obtained by manual annotation is recognized as the category of the input palmprint image block. Based on VGG16 network, CBAM attention module is introduced to enhance the network's attention to ridge structure information. The loss function is improved and the residual mechanism is added to avoid the gradient disappearance. The data set is manually annotated. For the unbalanced distribution of samples in the data set, we employ the method of data enhancement to weaken the imbalance between the class distributions, and design the sample unbalanced loss function to optimize the model. Experimental results show that the proposed method has a 14.2% improvement in recognition accuracy compared with the frequency domain methods on this dataset, and at least a 5.1% improvement in recognition accuracy compared with some classical deep learning methods.

Key words: palmprint ridge distance; VGG16; residual network; attention mechanism; loss function

0 引言

掌纹身份鉴别是生物特征识别领域中一个重要的研究问题,在网络空间安全、刑事侦查、司法鉴定等领域有重要的应用价值。相比其他生物特征,掌纹拥有更大的面积以及更为丰富的纹理信息,具有采集成本

低廉、用户易于接受、安全稳定性高等优点^[1]。掌纹识别的现有研究主要基于手掌上三类纹线:乳突纹、屈肌线和皱褶。如图1所示,乳突纹又被称为脊谷纹线或脊线,通常是平行结构的曲线,脊线的终端点、分叉点等特殊细节结构被称为细节点,掌纹刑侦专家手工进

收稿日期:2022-09-18

修回日期:2023-01-31

基金项目:国家自然科学基金(62076151);山东省自然科学基金(ZR2022MF272);山东建筑大学特聘教授支持计划项目(X22049Z)

作者简介:张浩(1997-),男,硕士,研究方向为图像处理、生物特征识别;通讯作者:郝凡昌(1981-),博士,副教授,研究方向为机器学习、图像处理。

行掌纹识别时,主要把掌纹脊线作为特征的基础生理信息^[2-3]。掌纹脊线距离是掌纹脊线的重要纹理属性之一^[4],其定义为沿着垂直于脊线的方向连接相邻两条平行脊线中心的线段长度。掌纹平均脊线距离(下文简称脊线距离)是指一定区域内多对脊线间距离的平均值,在某些掌纹识别算法^[4-6]与图像质量评估方法^[7]中常作为一个重要参数被进行参考。因此,精准地测量掌纹脊线距离对掌纹身份鉴别领域的研究具有重要意义^[7]。

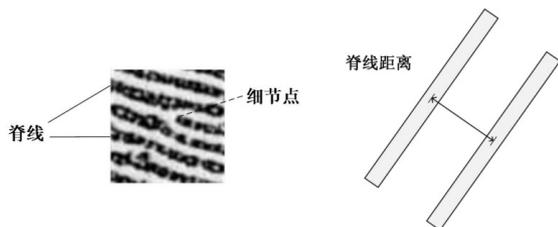


图 1 掌纹图像块及相关概念示例

已有的脊线距离的测量方法主要分为两类:频域方法与空域方法。在空域方法中,Zhang 等人^[8]提出了一种基于脊线搜索与统计窗口的估计方法,在空间域上进行指纹脊线距离的计算。空域方法常应用于指纹脊线距离估计,由于掌纹面积大、皱褶多等特点很少应用于掌纹脊线距离估计,而当前掌纹识别相关工作中多以频域方法进行脊线距离估计。在频域方法中,Jain 等人^[2]通过离散傅里叶变换与区域生长算法选取频域中合适的强正弦波,通过该正弦波的参数计算掌纹块中的脊线周期,脊线周期的倒数即为脊线距离,并将其作为掌纹匹配中的重要参数。在此基础上,任春晓等人^[9]设计了一种指纹图像脊线距离估计算法,通过使用傅里叶变换、平均信息熵和加权欧氏距离等方法计算脊线距离。Hao 等人^[7]采用频域方法对掌纹图像块进行脊线距离估计,并设计环形掩膜来确定频域内高峰搜索范围,减少计算与干扰信息,同时将脊线距离作为掌纹质量评估中的重要指标进行参考。但是,常用的频域方法中的区域生长算法需要相对完整的掌纹图像来进行局部领域间脊线距离信息的校正,对掌纹图像的质量要求高,且用于单独较小掌纹块的测量效果并不佳。而频域方法中对掌纹图像块脊线距离估计,并未充分使用图像频域中的相位角等信息,脊线距离估计的精度还需进一步提高。

近些年来,随着深度学习技术的发展,在掌纹识别中越来越多的研究关注到基于深度学习方法从数据中学习掌纹特征。Zhong 等人^[10]基于深度哈希网络与生物特征图匹配方法,提出了一种融合掌纹和手背静脉的多生物特征算法。2021 年,吴碧巧等人^[11]提出基于迁移学习的网络模型,实现端到端的高分辨率掌纹图像识别。Liu 等人^[12]基于注意力的生成对抗网

络提出了掌纹方向场恢复模型,使用注意力模块引导网络关注修复区域。在已有研究中,深度学习技术的应用对解决上述掌纹识别与特征提取问题起到了很好的效果。

但现有相关研究并未关注到使用深度学习技术进行脊线距离估计。VGG (Visual Geometry Group)^[13]网络作为深度学习中一种经典卷积神经网络,其模型结构简单,可应用性强,网络识别效果明显,综合性能优于一般的典型网络,初步研究发现该网络对于输入尺寸较小的掌纹图像块的处理仍存在网络退化、梯度消失及网络所学习特征不佳等问题。因此,该文引入残差结构与注意力机制,有效解决了上述问题,并对掌纹图像纹理特征中的关键信息进行提取,在提高模型学习能力的同时避免梯度消失。将手工标注获取的脊线距离值按一定值域划分后作为输入掌纹图像块的类别进行识别,并以 VGG16 网络为基础,提出了一种基于注意力机制与残差结构的脊线距离估计网络。另外,针对手工标注数据集中掌纹图像块脊线距离类别存在的分布不均衡问题,采用数据增强方法削弱类别分布间的不平衡,并设计样本不均衡损失函数,来增强网络对样本数目较少类别的关注,提高模型的泛化能力与识别精度。

1 卷积神经网络与注意力机制

1.1 卷积神经网络

卷积神经网络是深度学习算法中一个重要组成部分,一般由若干个卷积层、池化层和全连接层组成。其中,卷积层用于图像特征的提取,池化层用于对卷积层提取的特征进行抽样,全连接层用于将卷积网络提取的特征连接起来,最后通过多分类器输出对图像的分类预测。VGG 是一种经典卷积神经网络,相比 AlexNet^[14]模型在图像分类任务上取得了更好的效果,VGG 模型使用 3×3 大小的小尺寸卷积核和 2×2 大小的池化层构造深度更深的卷积神经网络,其网络结构设计方法为构建深度卷积神经网络提供了方向。一方面,通过 3×3 小卷积核的堆叠使用来获取更大的感受野,减少网络参数,且在文中工作中较大感受野的选取,可以考察到多对相邻脊线间的距离信息,从而使模型关注的信息更加全面。另一方面,卷积层中 ReLU 激活函数与全连接层中 Dropout 函数的使用增加了网络中的非线性映射,增强网络的泛化能力,从而能够关注解决脊线分叉、涡形变化等多种局部细微变化带来的差异问题。

VGG16 模型由 13 层卷积和 3 层全连接层组成,模型结构简单、可应用性强,但是对于输入尺寸较小且结构简单的掌纹图像块,网络退化现象仍比较明显,考

考虑到残差网络对于解决网络退化与梯度消失问题具有不错的效果,对此以 VGG16 网络为基础引入恰当的残差结构来改善上述问题,同时亦改进保持了网络的轻量化。

1.2 网络注意力机制

为使模型更加有效地关注到脊线结构信息,并对掌纹脊线距离进行更加精准的测量估计,该文在模型设计中引入注意力机制。计算机视觉中的注意力机制借鉴了人类视觉的注意力思维方式,学习获取重点关注的目标,从而在大量无关背景区域中筛选出具有重要价值信息的目标区域,帮助人类更加高效地处理视觉信息^[15]。注意力机制的核心思想为基于原有数据中内容的关联性来突出数据中的重要特征,常见的注意力机制有通道注意力、空间注意力、多阶注意力等。在实现上,注意力机制通过神经网络的学习生成与输入特征通道或空间维度大小相同的掩码矩阵,以此表示输入特征所需网络关注的得分,两者对应维度相乘后输入特征中更重要的信息将获得更大权重,来提高网络对输入特征中重要区域关注程度。CBAM(Convolutional Block Attention Module)注意力机制^[16]由通道注意力机制与空间注意力机制串联而成,总体结构如图2所示。

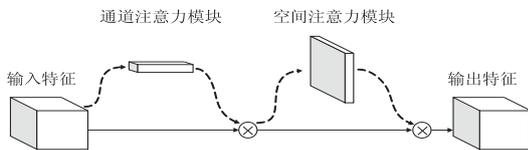


图2 CBAM注意力机制总体结构

通道注意力机制模块如图3所示,具体实现过程为:将 $C \times H \times W$ 大小的输入特征图经全局平均池化和全局最大池化分别生成两个 $C \times 1 \times 1$ 大小的特征图 GAP 与 GMP,经过两层共享权值的神经网络,为了减少网络参数,第一层神经元个数为 C/r , r 为缩减比率,在文中方法中 r 设置为 16,激活函数为 ReLU,第二层神经元个数为 C ,由此得到两组 $C \times 1 \times 1$ 大小的中间特征图 GAP' 与 GMP' ;两组特征图在对应维度相加后,经过 Sigmoid 激活函数得到通道注意力特征图;而后,将通道注意力特征图和输入特征图进行乘法操作,由此对输入特征图进行注意力叠加,生成空间注意力机制模块需要的输入特征。

空间注意力机制模块如图4所示,具体实现过程为:将上一模块输出的特征图作为输入特征图,并在通道层次上进行全局平均池化和全局最大池化生成两个 $1 \times H \times W$ 大小的特征图 GAP 与 GMP,并将其拼接为 $2 \times H \times W$ 大小的特征图,以此共享同一网络;而后,经过一层 7×7 卷积操作,降维为 $1 \times H \times W$ 大小,再经过 Sigmoid 激活函数得到空间注意力特征图;最后,将空

间注意力特征图和输入特征图进行乘法操作,得到最终经先后叠加通道注意力与空间注意力的输出特征图。

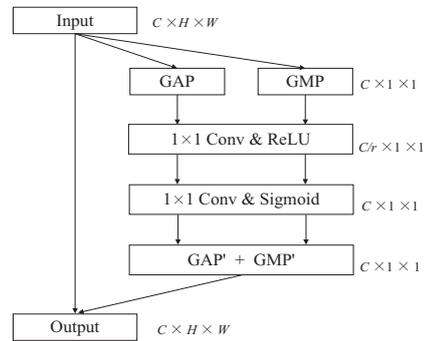


图3 通道注意力模块结构

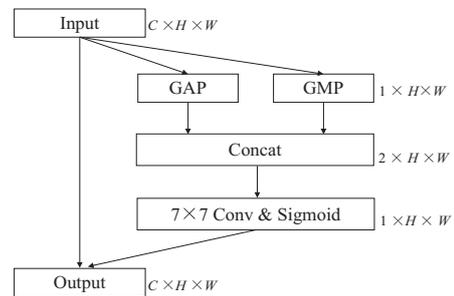


图4 空间注意力模块结构

与其他注意力机制相比,CBAM注意力机制同时叠加了通道注意力与空间注意力两种不同机制,所生成的注意力特征内容更加丰富,使得原有通道间的信息交互与空间内的信息交互联系更加密切。同时,两种注意力机制均通过使用全局平均池化和全局最大池化操作来分别生成输入特征的全局和局部信息,以此在各注意力机制内同时关注输入特征的局部与全局特征生成注意力网络掩码,使得网络参考信息的范围更全面。对于主要由线条构成的掌纹图像块,CBAM注意力机制可使模型更加有效地关注到其结构信息,从而对掌纹脊线距离进行更加精准的测量估计,由此选取该注意力机制对模型进行改进。

2 基于样本不均衡损失函数的残差网络

2.1 嵌入 CBAM 的残差网络

该文将裁剪后的 64×64 像素大小的掌纹图像块作为输入,可更精细化计算原掌纹图像的脊线距离分布,且对该尺寸的掌纹图像块进行标注,标注准确性高。由于输入图像尺寸较小,选用 VGG16 网络作为基础网络具有更好的适用性。对网络有机加入 CBAM 注意力模块,可提高网络对特征空间中重要纹理特征的关注程度,且能够增强网络中间层的特征表达,提高网络的学习能力。同时,为有效避免模型存在的梯度消失问题,对网络改进添加残差结构与批量归一化(BN)层。设计实现两种残差子模块,如图5所示。由

于 CBAM 模块中空间注意力机制使用 7×7 大小的卷积核来提取空间内多条脊线间的局部相关性,因此嵌入 CBAM 的残差子模块应用于网络前端,使模型着重关注输入图像块中的关键纹理特征,而普通残差子模块应用于网络后端,且在同一子模块内卷积层的输出通道数设置相同。

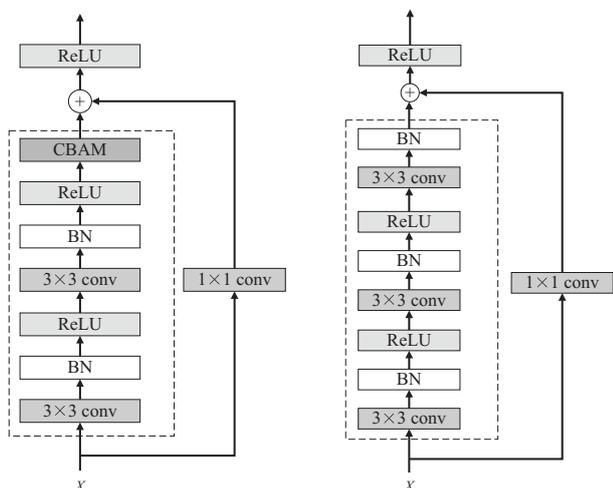


图 5 Residual-CBAM-Block 注意力残差模块与普通 Residual-Block 残差模块

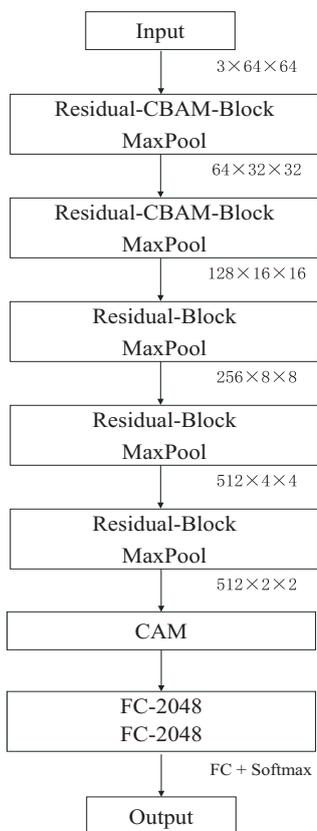


图 6 引入残差机制与 CBAM 注意力模块的 Vgg16-CBAM-Residual 网络

引入残差机制与注意力模块的 VGG16 网络结构如图 6 所示,网络流程如下:首先,为拓展网络输入通道信息及增加初始卷积计算,将输入的掌纹灰度图像

块预处理为 3 通道,且经实验验证图像通道设置为 3,网络性能更佳。而后,输入图像经过 2 次嵌入 CBAM 的残差模块,并且每次输出进行最大池化操作,再经过 3 次普通残差模块,同样每次输出进行最大池化操作,此时输出特征图大小为 512×2×2。而后,将上述输出特征图输入通道注意力模块(CAM)计算叠加其通道注意力,特征图展平后输入全连接层模块,经过 2 层全连接层 FC-2048 输出特征。最后,经过类别输出全连接层 FC 与 Softmax 输出对各个类别的预测概率。

2.2 样本不均衡损失函数

在深度学习中,若训练数据不均衡,会使模型更偏向于数目较多的样本,而数量较少的样本类别得不到有效的关注,从而影响模型的泛化能力与识别精度。因此,只有训练数据分布均衡,才能避免训练模型产生严重的偏向性,以此加强模型泛化能力与识别精度。所用的数据集是随机选取手工标注,而手掌各局部区域的脊线距离值多分布于 9 至 12 像素间^[7],如图 7 所示,标注数据集由此存在类别分布不均衡问题。

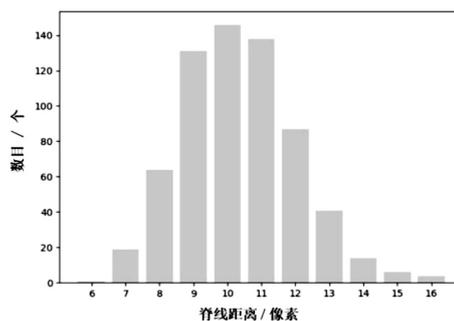


图 7 标注数据集脊线距离值分布

通过数据增强与设置样本不均衡下的分类损失函数来解决标注数据集中存在的分布不均衡问题。在后文中,将具体介绍通过增广类别与数据增广来增强及均衡化训练数据。在损失函数设置上,图像识别中常用多分类交叉熵损失函数,公式如下:

$$L = \frac{1}{N} \sum_{i=1}^N L_i = - \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (1)$$

其中, N 为样本数目, M 为类别数目, y_{ic} 为符号函数,若样本 i 的真实类别为 c ,取 1,否则取 0, p_{ic} 为观测样本 i 属于 c 的预测概率。

类别分布的不平衡可归结为难易样本分布的不平衡^[17],为使模型更加关注难分样本,相对减少关注易分样本,对此将交叉熵损失函数修改设置为:

$$L_{Focal_i} = - \sum_{c=1}^M y_{ic} (1 - p_{ic}) \log(p_{ic}) \quad (2)$$

但是由于标注误差与样本中离群点的存在,导致上述目标函数会使模型过于关注极难分样本,影响模型泛化能力。而难易样本分布的不平衡可进一步归结为梯度分布的不平衡^[18],梯度分布计算公式如下:

$$GD(g) = \frac{1}{l} \sum_{i=1}^N \delta(g_i, g) \quad (3)$$

其中,

$$\delta(x, y) = \begin{cases} 1, & \text{if } y - \frac{\varepsilon}{2} \leq x < y + \frac{\varepsilon}{2} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$l = \min(g + \frac{\varepsilon}{2}, 1) - \max(g - \frac{\varepsilon}{2}, 0) \quad (5)$$

$$g_i = \sum_{c=1}^M y_{ic}(1 - p_{ic}) \quad (6)$$

通俗来讲, $GD(g)$ 即是以梯度 g 为中心, 宽度为 ε 区域内的样本密度。

设置梯度划分范围后, 由统计的初始样本预测概率值所在梯度分布, 得知易分样本与难分样本所在梯度模长较大, 而使模型更加关注两者之间的普通样本, 可更有效地提升模型的泛化能力, 对此该文改进的目标函数公式如下:

$$L_Focal_G_i = \frac{L_Focal_i}{GD(g_i)} = - \sum_{c=1}^M \frac{y_{ic}(1 - p_{ic}) \log(p_{ic})}{GD(g_i)} \quad (7)$$

该目标函数融合了两种样本不均衡损失函数思想, 使得样本所在的梯度密度在模型训练前期起主导作用, 使模型更加关注普通样本, 并避免对过难样本的关注。而在训练后期, 随着模型精度的提高, 样本梯度 g_i 普遍减小, 样本梯度密度更加集中, 梯度密度无法发挥有效的作用, 而此时 L_Focal_i 中设置的难易权重可相对更有效地发挥作用。在文中工作中, 目标函数的设置可使模型在训练过程中更加关注样本数目较少的类别, 同时更加关注普通样本的识别正确率, 减少样本类别不均衡对模型训练的影响。

3 实验

3.1 数据集

该文使用的掌纹图像来自两个包含全掌纹和现场掌纹的开放数据库, 即现场掌纹识别数据库 (LPIDB v1.0)^[19] 和清华大学高分辨率掌纹数据库 (THU-HRPD)^[4-5]。LPIDB v1.0 包含两种模式的掌纹图像。一种模式包括在受控环境中直接从物体扫描出的完整掌印, 而另一种模式包括用黑色粉末和胶带从模拟的无法控制的犯罪现场所提取的现场掌纹。该数据库图像来自 51 个人的 102 个不同手掌的 102 个全掌纹和 380 个现场掌纹数据。LPIDB v1.0 图像均为 8 位灰度图像, 分辨率为 500 DPI (Dots Per Inch), 其中, 全掌纹图像的大小约为 2 000×2 000 像素, 而现场掌纹图像大小的范围在大约 150×200 像素到 2 000×800 像素之间。THU-HRPD 图像数据由商用掌纹扫描仪拍摄采集, 包含来自 80 个人的 160 个不同手掌, 每个手掌对

应有 8 幅全掌纹图像, 实验数据仅从中取一幅图像, 所有图像均来自非重复手掌。THU-HRPD 的 160 幅图像中, 均为 500 DPI, 大小为 2 040×2 040 像素, 其中, 有 24 幅是 JPG 格式的 RGB 彩色图像, 其余是位深度为 8 的 JPG 格式的灰度图像。

该文选取上述数据库中的图像来标注创建掌纹脊线距离数据集, 具体实现过程为: 将掌纹图像裁剪为 64×64 大小的非重叠图像块, 作为待标注数据集, 并从中随机选取图像清晰且来自尽量多不同手掌、不同区域的图像块进行手工标注。如图 8 所示, 标注采用选取平行线条标点并计算两点间距离方式, 以像素为单位计算脊线距离。在掌纹图像块内取不同区域的多对脊线进行标注计算, 取平均值的整数表示该图像块的平均脊线距离, 并将其写入图像块的文件名中作为标签使用。

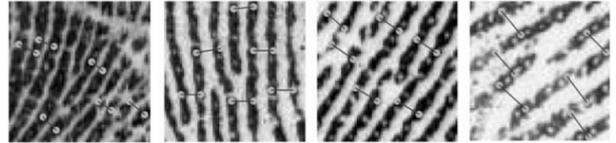


图 8 掌纹脊线距离标注过程效果

该文标注图像数量为 651 幅, 并将上述手工标注计算的脊线距离值作为图像的样本类别。由于部分类别图像数目太少、类别分布不平衡、标注存在误差等情况, 原标注类别间存在一定的重叠部分且即便是专家手工标注也不易设定严格的界限, 以此采用增广类别的方式改善类别分布, 将脊线距离为 6、7 像素合并为“7 像素邻域”类别, 脊线距离为 9、10、11 像素合并为“10 像素邻域”类别, 脊线距离为 12、13 像素合并为“12 像素邻域”类别, 脊线距离为 14、15、16 像素合并为“14 像素邻域”类别, 类别数目缩小为 5 个。增广类别后, 为确保样本类别分布均衡, 依据类别内所含的样本数目, 通过高斯滤波、中值滤波及添加噪声等方式进行不同程度的数据增广, 对样本数目少的类别进行最大程度的数据增广, 样本数目相对多的类别只选取添加噪声的方式进行数据增广或不进行数据增广。数据增广后的数据集包含 1 441 幅图像, 且极大改善了原始标注数据中样本分布不均衡问题, 如图 9 所示。

3.2 实验设置

实验均在 Python 3.6.9 环境下进行编写, 实现框架采用 Pytorch。实验将数据增强后的 64×64 像素大小的图像作为输入, 共 1 441 幅图像。其中, 1 020 幅图像作为训练集, 431 幅图像作为验证集。网络共迭代 200 轮, 选用自适应矩估计 (Adam) 优化器, 初始学习率为 0.002, 迭代 40 轮后, 网络每迭代 40 轮, 学习率衰减为现学习率的 0.6 倍, 批次大小设为 128。上述常数均按照经验值进行设定。采用准确率 Accuracy

作为实验的评价指标,公式为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{P + N} \quad (8)$$

其中,TP 为将正样本预测正确的数目,TN 为将负样本预测正确的数目,P 为正样本数目,N 为负样本数目。

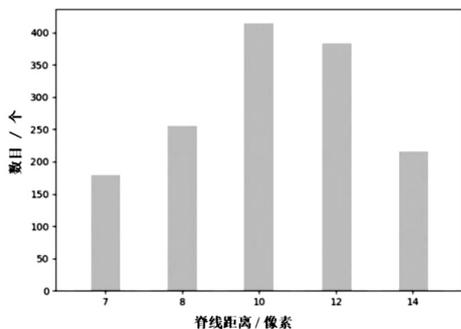


图 9 数据增广与类别合并后 1 441 幅图像脊线距离值分布

3.3 实验结果与分析

为了验证文中方法的效果,进行了三组实验,分别为:传统频域方法对比实验、消融对比实验、深度学习方法对比实验。

传统掌纹脊线距离估计方法主要是通过选择在频域中选取合适的强正弦波参数来计算局部掌纹图像块中的脊线距离,且传统方法常作为辅助工作进行介绍,未进行系统性详细研究。由此,在与传统频域方法对比实验中,选取文献[7]中最新改进的对掌纹图像块的脊线距离估计频域方法进行实验对比,结果如表 1 所示。结果上看,文中方法(Vgg16-CBAM-Residual)相比传统频域方法准确率有较大提升,表明通过文中方法能够有效学习获取掌纹图像块脊线距离特征,同时传统频域方法对掌纹图像块脊线距离的估计精确度存在一定局限性,深度学习对掌纹纹理特征的学习具有一定的应用价值。

表 1 频域方法实验对比结果

方法	准确率/%
频域方法 ^[7]	62.1
Vgg16-CBAM-Residual	76.3

为了验证文中注意力机制、残差结构以及样本不均衡损失函数设置的有效性,进行了消融实验。在实验设置相同的条件下,设置三类方法:VGG16 网络加入 CBAM 注意力机制(Vgg16-CBAM)、VGG16 网络加入残差结构(Vgg16-Residual)以及文中方法(Vgg16-CBAM-Residual),分别在四种不同损失函数下进行实验对比,结果如表 2 所示。

消融实验结果显示,文中方法将 CBAM 注意力机制与残差结构同时加入网络能够起到更好的效果。一方面,这是因为 CBAM 注意力机制在网络前端的加入

增强了模型对掌纹图像块在通道与空间层次中关键纹理信息与细粒度信息的学习,同时在网络后端的通道注意力机制进一步增强了对当前特征图中重要通道信息的关注。另一方面,由于输入掌纹图像块较小,残差结构的加入,在增加模型网络深度的同时,解决了梯度消失问题,同时增强了模型对掌纹图像中脊线结构信息的学习。在损失函数设置方面,所提出的损失函数效果优于现有其他损失函数,表明提出的基于样本难分程度与梯度密度的样本不均衡损失函数,在模型训练学习的不同阶段,能够更加有效地关注到少样本类别与难分样本,减少样本不均衡对模型所带来的影响。

表 2 消融实验对比结果

方法	损失函数	准确率/%
Vgg16-CBAM		69.8
Vgg16-Residual	CE Loss	71.7
Vgg16-CBAM-Residual		71.9
Vgg16-CBAM		71.2
Vgg16-Residual	Focal Loss	73.2
Vgg16-CBAM-Residual		74.2
Vgg16-CBAM		70.1
Vgg16-Residual	GHM Loss	72.2
Vgg16-CBAM-Residual		73.8
Vgg16-CBAM		72.2
Vgg16-Residual	GHM-Focal Loss	75.4
Vgg16-CBAM-Residual		76.3

在与不同深度学习网络的对比实验中,选取 VGG16、Resnet-34(Residual Network)^[20]、Densenet-161(Densely Connected Convolutional Networks)^[21] 三种经典深度学习网络模型在相同实验设置下进行对比,并使用以上三种网络模型在参考文献中的默认参数设置进行实验。如表 3 所示,所提出的方法效果最佳,这进一步表明,以 VGG16 网络为基础模型进行改进添加的注意力机制与残差结构,能够使模型更有效地学习掌纹图像块中脊线距离特征信息。同时,模型所设置的样本不均衡损失函数能更加关注到训练过程中的难分样本,以此提高模型的泛化力。相较于其他残差网络,改进的网络模型虽参数量较多,但模型整体性能表现更好。

表 3 深度学习方法实验对比结果

方法	准确率/%	Paras/M
VGG16 ^[13]	67.5	134.3
Resnet-34 ^[20]	70.1	21.3
Densenet-161 ^[21]	71.2	26.5
Vgg16-CBAM-Residual	76.3	23.6

4 结束语

该文提出了一种基于注意力机制与残差网络的掌纹脊线距离估计方法。该方法基于 VGG16 网络进行改进,为网络添加残差结构与 CBAM 注意力机制,增强网络对图像纹理特征中关键信息的提取,提高模型学习能力,且避免梯度消失。同时,设计基于样本难分程度与梯度密度的样本不均衡损失函数,增强网络对样本数目较少类别的关注,提高模型的泛化能力与识别精度。所用数据集由手工标注完成,并采用数据增强方式削弱减少类别分布不均衡问题。在该数据集上,文中方法与其他深度学习网络相比脊线距离估计效果更佳,与传统频域方法相比脊线距离估计准确率提升明显,且对深度学习在纹理图像特征提取方面的应用具有一定借鉴意义。文中方法在脊线距离估计精度上有提升,但数据集仍存在一定标注误差且文中方法对脊线距离估计的细分程度还仍需提高,未来研究可更加关注掌纹图像脊线距离标注信息的完善以及脊线距离估计精确程度的进一步提高。

参考文献:

- [1] ZHANG D D. Palmprint authentication [M]. New York: Springer Science & Business Media, 2004.
- [2] JAIN A K, FENG J. Latent palmprint matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(6): 1032–1047.
- [3] LIU E, JAIN A K, TIAN J. A coarse to fine minutiae-based latent palmprint matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2307–2322.
- [4] DAI J, ZHOU J. Multifeature-based high-resolution palmprint recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(5): 945–957.
- [5] DAI J, FENG J, ZHOU J. Robust and efficient ridge-based palmprint matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(8): 1618–1632.
- [6] CHEN J, GUO Z. Palmprint matching by minutiae and ridge distance [C]//International conference on cloud computing and security. [s.l.]: Springer, 2016: 371–382.
- [7] HAO F, YANG L, YANG G, et al. RFPIQM: ridge-based forensic palmprint image quality measurement [J]. IEEE Access, 2018, 6: 62076–62088.
- [8] ZHANG L, YIN Y, YANG D. Fingerprint ridge distance estimation based on ridge search [C]//2009 international joint conference on bioinformatics, systems biology and intelligent computing. Shanghai: IEEE, 2009: 604–607.
- [9] 任春晓, 尹义龙, 马军, 等. 一种指纹纹线距离估计算法 [J]. 模式识别与人工智能, 2009, 22(4): 554–559.
- [10] ZHONG D, SHAO H, DU X. A hand-based multi-biometrics via deep hashing network and biometric graph matching [J]. IEEE Transactions on Information Forensics and Security, 2019, 14(12): 3140–3150.
- [11] 吴碧巧, 邢永鑫, 王天一. 基于 VGG16 和迁移学习的高分辨率掌纹图像识别 [J]. 智能计算机与应用, 2021, 11(5): 37–42.
- [12] LIU B, FENG J. Palmprint orientation field recovery via attention-based generative adversarial network [J]. Neurocomputing, 2021, 438: 1–13.
- [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv:1409.1556, 2014.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84–90.
- [15] MNIH V, HEES N, GRAVES A. Recurrent models of visual attention [J]. Advances in Neural Information Processing Systems, 2014, 27: 2204–2212.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the European conference on computer vision (ECCV). Munich: Springer, 2018: 3–19.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//Proceedings of the IEEE international conference on computer vision. Venice: IEEE, 2017: 2980–2988.
- [18] LI B, LIU Y, WANG X. Gradient harmonized single-stage detector [C]//Proceedings of the AAAI conference on artificial intelligence. Hawaii: AAAI, 2019: 8577–8584.
- [19] MORALES A, MEDINA-PÉREZ M A, FERRER M A, et al. LPIDB v1.0 – Latent palmprint identification database [C]//IEEE international joint conference on biometrics. Florida: IEEE, 2014: 1–6.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 770–778.
- [21] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Hawaii: IEEE, 2017: 4700–4708.