

文章编号: 2095-2163(2021)05-0037-06

中图分类号: TP391

文献标志码: A

## 基于 VGG16 和迁移学习的高分辨率掌纹图像识别

吴碧巧, 邢永鑫, 王天一

(贵州大学 大数据与信息工程学院, 贵阳 550025)

**摘要:** 由于高分辨率掌纹图像尺寸大、掌纹图像数量少等特点, 目前主流的方法采用细节特征点匹配, 其算法设计复杂, 识别精度不高。针对以上问题, 本文提出基于迁移学习的高分辨率掌纹图像识别方法, 该方法以 VGG16 为基础网络, 将在 imagenet 数据集上训练好的权重参数, 用于初始化所有的卷积层; 使用图像增强技术将高分辨率掌纹图像分别 4、9、16、25 等分, 采用投票的方法得到整个掌纹图像的准确率, 最高可达到 99.69%。经实验证明, 该方法可以实现端到端的高精度高分辨掌纹图像识别, 识别率优于以往的基于细节特征点匹配方法。

**关键词:** 生物特征识别; 高分辨率掌纹图像; 迁移学习; 图像增强; VGG16

### High-resolution palmprint image recognition based on VGG16 and transfer learning

WU Biquiao, XING Yongxin, WANG Tianyi

(College of Big Data and Information Engineering, GuiZhou University, Guiyang 550025, China)

**[Abstract]** Due to the large size of high-resolution palmprint images and the small number of palmprint images, the current mainstream method adopts detail feature point matching, which has complicated algorithm design and low recognition accuracy. To solve the above problems, a high resolution palmprint image recognition method based on transfer learning is proposed in this paper. This method takes VGG16 as the base network and initializes all convolutional layers with weight parameters trained on ImageNet data set. The high resolution palmprint images are divided into 4, 9, 16 and 25 equal parts by image enhancement technology, and the accuracy of the whole palmprint image is obtained by voting method, which can reach up to 99.69%. Experimental results show that this method can realize end-to-end high precision and high resolution palmprint image recognition, and the recognition rate is better than the previous method based on detail feature point matching.

**[Key words]** biometric recognition; high-resolution palmprint images; transfer learning; image enhancement; VGG16

## 0 引言

掌纹作为一种重要的生物特征, 和指纹一样, 可用于身份验证、身份识别、公共安全等多个领域。低分辨率掌纹图像的分辨率介于 75~150 dpi 之间, 以照相机、摄像机等采集为主<sup>[1]</sup>, 多为非接触式采集的掌纹图像, 其研究成果相对较多<sup>[2-4]</sup>。高分辨率掌纹图像的分辨率在 300 dpi 以上, 采用扫描仪扫描掌纹, 属于接触式采集的掌纹图像。高分辨率掌纹图像由于其分辨率高, 所以尺寸会较大, 对其进行处理的计算成本也更高, 但细节特征点很多<sup>[5]</sup>, 包含丰富的信息, 在公安刑侦领域和法律层面都具有重要的应用。在犯罪现场不仅会留下指纹, 还会留下残缺或完整的掌纹信息<sup>[6]</sup>。

目前, 对高分辨率掌纹图像的识别主要采用对

掌纹轮廓、掌纹主线、褶皱线、细节特征点、三角点等特征点进行特征提取、匹配的方法。Jain 等人<sup>[7]</sup>以细节为特征, 用固定长度的细节描述符, 捕获每个细节周围的独特信息。基于对齐的匹配算法用于匹配掌纹, 对实时扫描手掌图样和潜在掌纹的识别率分别为 78.7% 和 69%。Feng 等人<sup>[8]</sup>提出了掌纹表示的 Gabor 幅相模型, 将 Adaboost 算法引入模型训练中, 将所选弱分类器的加权线性组合的响应值用于细节可靠性测量和不可靠的去除。Fei 等人<sup>[9]</sup>提出了圆边界一致性, 使用加长的 Gabor 滤波器设计来获取更可靠的细节特征点。Liu 等人<sup>[10]</sup>对掌纹匹配设计了基于细节簇和细节匹配传播的粗略匹配策略, 识别准确率为 79.4%。由于受到 3 条主线和噪声的强烈影响以及图像采集时的不规范行为, 高分辨率掌纹图像的部分区域图像质量较差, 传统的掌

**基金项目:** 贵州省科技厅与贵州大学科技合作计划项目(黔科合 LH 字[2016]7431 号)。

**作者简介:** 吴碧巧(1994-), 女, 硕士研究生, 主要研究方向: 深度学习、图像识别; 邢永鑫(1993-), 男, 硕士研究生, 主要研究方向: 深度学习、目标检测; 王天一(1989-), 男, 博士, 副教授, 主要研究方向: 量子通信、图像处理、计算机视觉。

**通讯作者:** 王天一 Email: tywang@gzu.edu.cn

**收稿日期:** 2021-01-03

哈尔滨工业大学主办 ◆ 学术研究与应用

纹识别方法在方法设计上复杂,识别时间长且识别效果不理想。

深度学习能够通过自动学习从大数据中获得有效的特征表示并进行识别,免除了繁杂的特征提取算法和掌纹匹配过程,但对训练数据的数量和质量要求较高。目前,高分辨率掌纹数据库大多数数据量小且单个的掌纹图像大,直接在深度学习网络使用难以达到好的识别效果。

针对以上问题,本文以 VGG16 为基础网络,提出基于迁移学习的掌纹识别技术。利用迁移学习,通过共享特征把从源域学习到的信息迁移转换到目标域中,以解决目标域带标签样本数据量较少的识别问题。本文使用由自然图像组成的 ImageNet 预训练网络模型得到初始化权重,由于 ImageNet 训练的神经网络偏向于纹理,因而适用于对高分辨率掌纹图像的识别。针对数据量小,单个数据大的问题,可利用图像增强技术将单张的高分辨率掌纹图像进行等大小的裁剪分块,得到同一个掌纹的不同块掌纹作为新的数据集,增大了数据集进行图像的识别,对同一张掌纹图像的不同块掌纹块进行投票得到一个更好的识别效果。

1 卷积神经网络和迁移学习

1.1 卷积神经网络

卷积神经网络源自于 Lecun 等人提出的 LeNet<sup>[11]</sup>,包含一系列的卷积层、池化层和全连接层。卷积层使用一组可训练的卷积核对输入图像进行特征提取,卷积层的参数共享机制大大减少了参数的数量<sup>[12]</sup>,使得参数的数量与输入图片的大小无关,增加了网络的泛化能力。池化层一般在卷积层后,主要有平均池化和最大池化,池化层在保留特征图主要特征的情况下,可以通过减少特征图的大小来减少参数量,防止网络过拟合,提高模型的泛化能力。全连接层的核心操作是矩阵向量乘法,前一层所有神经元与后一层的所有神经元相连。因为全连接层所有神经元都相连的特性,一般全连接层的参数是最多的,存在着一定的参数冗余。全连接层一般位于网络模型的最后几层,处理前面卷积层提取到的高级特征,在整个卷积神经网络中扮演着“分类器”的角色。

1.2 VGG16 模型

VGG 模型是由 Simonyan 等人于 2014 年提出的<sup>[13]</sup>,模型非常简单,只有卷积层、最大池化层和全连接层。VGG 模型首次提出了小卷积核的优势,在

卷积层使用非常小的(3×3)的卷积核,添加更多的卷积层使得网络变得更深。使用 2 个 3×3 的卷积核可以代替 5×5 的卷积核,使用 3 个 3×3 的卷积核可以代替 7×7 的卷积核<sup>[14]</sup>。使用更小的卷积核可以减少网络的参数,提升网络的深度,增强了网络的表达能力;同时更多的使用非线性激活函数,可以提高网络的判断能力。VGG 网络在图像识别领域的成功运用,代表着更深层的网络,可显著提升图像识别的效果。

VGG 网络中所有卷积层的卷积核大小都是(3×3),几个卷积层后面接一个最大池化层为一个 block。VGG16 的网络结构如图 1 所示。VGG16 共有 5 个 block,每个 block 的通道数一致,最大池化层减少特征图的尺寸。随着卷积层一层层的运算,卷积核输出的内容越来越抽象,保留的信息也越来越少,特征图的尺寸也越来越小,通道数越来越多,以提取更多的抽象特征。Block5 完成后将特征图平铺成一维数据作为全连接层的输入。最后 3 层为全连接层,前 2 层每层具有 4 096 个神经元,使用 Relu 非线性激活函数,第 3 层有 1 000 个神经元(1 000 个类别)使用 softmax 函数。

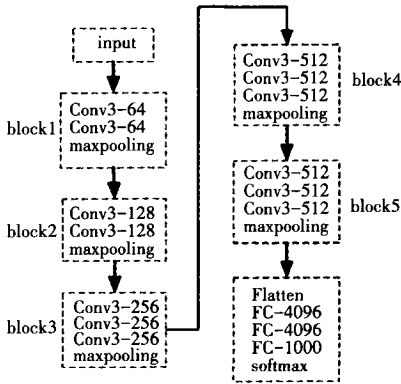


图 1 VGG16 网络结构

Fig. 1 Network structure of VGG16

1.3 迁移学习

迁移学习的目的是将源于学习到的知识应用推广到目标域中<sup>[15]</sup>。在深度学习中,训练数据的获取至关重要,当训练样本不够时,往往识别效果也不尽如人意。有时因为客观原因无法获得足够的数据样本,或者大量没有进行标注的数据样本,无法直接使用,需要耗费大量的人力去标注,网络无法得到足够的训练。迁移学习则可以将从其它相似数据集上训练得到的网络权重迁移到目标网络,能够更快更好地进行参数的训练,而不必从头训练。

深度迁移学习主要有 3 种方式:使用其它数据集训练好的权重系数,作为初始化权重参数,进行接



下来的训练;使用其它数据集的权重参数,冻结预测层之前的所有权重参数,进行接下来的训练;使用其它数据集训练好的权重参数,有选择地冻结一部分层,对其余层进行参数的更新和训练。

2 基于迁移学习的 VGG 模型

2.1 数据集以及图像预处理

目前存在的高分辨率掌纹图像数据库主要有 2 类:一种是公安部门采集的标准掌纹档案库(不公开),其次是实验条件控制下采集的公开用于科研的高分辨率掌纹图像数据库。本文使用的公开数据库为 THU 高分辨率掌纹数据库<sup>[16]</sup>,该数据库中采集了 80 人的左右手掌,共有 160 个不同手掌的 1 280 张高分辨率掌纹图像。其中,单个手掌有 8 张掌纹图像。数据库中的每一张高分辨率掌纹图像的分辨率均为 500 ppi,大小为 2 040x2 040 像素,大部分掌纹图像是灰度图像,但有少数掌纹图像是彩色图像。处理数据时,将所有的高分辨率掌纹图像读取为灰度图像。图 2 中显示了 2 张 THU 数据库中的高分辨率掌纹图像。数据库中的有的掌纹采集的很完整,整个掌纹的脊线、褶皱、谷线、主线都被清楚的采集到,如图 2 中的(a)图所示。但也有少数的掌纹只采集到一部分掌纹信息,如图 2 中的(b)图,三条主线信息全部丢失。

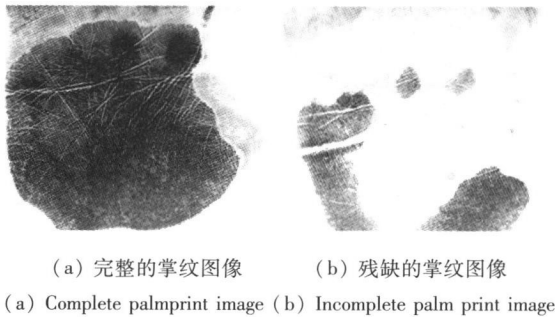


图 2 高分辨率掌纹图像

Fig. 2 High resolution palmprint image

在本文实验中,将采集到的每个手掌掌纹图像中的六张图像,即  $160 \times 6 = 960$  张掌纹图像作为训练样本;其余的  $160 \times 2 = 320$  张掌纹图像作为测试样本。数据集里原本高分辨率掌纹图像大小为  $2\,040 \times 2\,040$  像素,为了减小运算量,提高运算速度,对数据集进行降采样预处理,处理后的掌纹图片大小为  $510 \times 510$  像素。

2.2 VGG16 深度迁移模型

本文使用在 imagenet 数据集上训练得到的权重参数作为初始化参数,训练识别高分辨率掌纹图像。

虽然这些参数并不是针对高分辨率掌纹图像的,但这些参数对纹理有很好的分辨能力。高分辨率掌纹图像主要依靠纹理进行识别而非形状,理论上此迁移学习会有很好的结果。

将 VGG16 在 imagenet 数据集上训练得到的参数信息中卷积层部分的参数,用来初始化本实验的 VGG16 模型的卷积层参数。将原模型中的 flatten 操作换成全局最大池化操作,可以有效的较少参数数量,防止过拟合。将第 3 个全连接层的 softmax 激活函数换成 relu 激活函数,并在其后接一个全连接层 160 个神经元,代表着 160 个类别。softmax 激活,得到预训练 VGG16 模型,如图 3 所示。增加一个全连接层能提升模型的非线性表达能力,综合前面卷积层提取到的特征,更好的完成分类任务。

预训练 VGG16 模型的最后 4 层全连接层使用随机参数初始化方法,前 3 层的激活函数为 relu 非线性激活函数,最后一层是 softmax 激活函数。数据库中的掌纹图像基本为灰度图像,即图像的通道数为 1,采用预训练模型的输入通道为 3,故将高分辨率掌纹图像的通道复制 3 次,变成 3 通道图像。

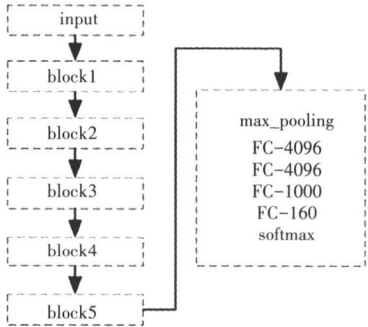


图 3 VGG16 迁移模型

Fig. 3 Transfer model of VGG16

2.3 在 VGG16 深度迁移模型上进行图片增强

预处理后的高分辨率掌纹图像大小为  $510 \times 510$  像素,对预处理后的高分辨率掌纹图像进行数据增强的等分裁剪。

图 4 为按照不同的尺寸裁剪后的左上角的第一个掌纹图像块,可以看到,图片裁剪的越小,其包含的有效信息就越少。

表 1 为数据增强的数据,裁剪份数为将一个完整的高分辨率掌纹图像有重叠的等分后得到的掌纹块数量。分别有 4 组实验数据:4 等分、9 等分、16 等分和 25 等分。其中,掌纹块为裁剪后单个的掌纹图像块大小,重叠部分为裁剪后 2 个相邻的掌纹图像块在一个维度上重叠的部分。

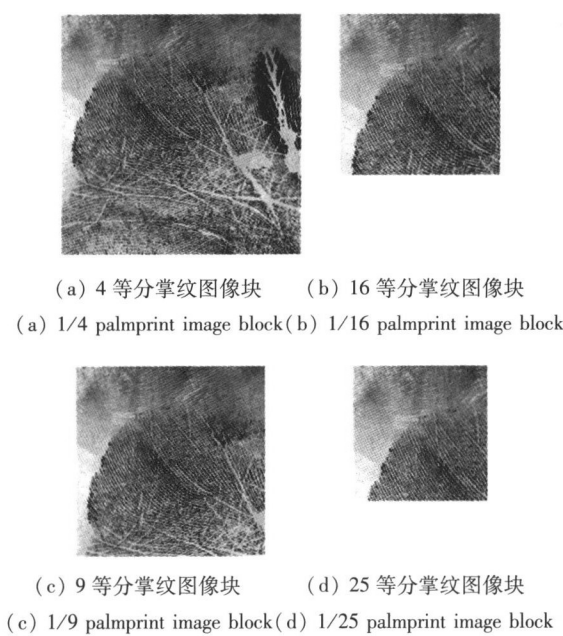


图 4 裁剪后的掌纹图像块  
Fig. 4 Palmprint image blocks

表 1 数据增强数据

Tab. 1 Data of data enhancement

裁剪份数	掌纹块大小(像素)	重叠部分(像素)
4	300×300	90
9	230×230	90
16	180×180	70
25	150×50	60

在一组实验中,如掌纹图像 9 等分实验,将训练集和测试集都进行同样的 9 等分裁剪的操作,会得到  $960 \times 9 = 8\,640$  个训练掌纹图像块,  $320 \times 9 = 2\,880$  个测试掌纹图像块。训练集会将裁剪后得到的掌纹图像块顺序打乱,送入 VGG16 深度迁移模型进行训练,得到训练好的参数模型。测试集则不打乱顺序,一张掌纹图像的 9 个掌纹图像块都是按顺序依次排列。将裁剪得到的掌纹图像块放入训练好的 VGG16 深度迁移模型进行测试,将得到每一张掌纹图像块的预测值。

将得到的预测结果进行重新排列。例如,对 9 等分得到的掌纹图像块的预测值处理为:将所有的预测值 reshape 对应为  $(320, 9)$  的队列。其中 320 代表测试集的 320 张掌纹图像,一行有 9 个掌纹图像块的预测值,这 9 个掌纹图像块来自于同一张掌纹图像的不同部分。采用硬投票的方式,即直接用掌纹图像块的类别值,对每一行的 9 个掌纹图像块的预测值进行投票。基于少数服从多数的原则,票数多的预测值即为这张掌纹图像的预测值,若有 2 个类别投票结果一致,则随机选择一个预测值。320

行每行都会得到一个预测值即为测试集的测试结果,再将该值与每个掌纹图像的标签进行比对,若相等结果为 1,否则为 0。求取所有结果的平均值,即为最后的测试准确率。

3 实验结果分析

实验均在 Ubuntu18.04.4LTS 操作系统上进行,采用 tensorflow 深度学习框架,i7-9700 处理器、RTX 2080Ti 显卡、32GB 内存。所有实验都采用 VGG16 模型,损失函数默认使用 Cross Entropy Loss、Adam 优化器,Batch Size 为 8,学习率为  $1 \times 10^{-5}$ 。

3.1 VGG16 上的识别效果

将预处理后的数据集分别在随机初始化参数的 VGG16 与使用 imagenet 比赛中预训练模型参数的 VGG16 上运行。

图 5 为随机初始化参数的 VGG16 与使用 imagenet 比赛中预训练模型参数的 VGG16 的训练结果对比图。横坐标为训练次数,纵坐标分别为准确率和损失值。从图中可以看出,无论准确率还是损失值,相比于随机初始化的网络,采用预训练权重网络收敛的更快且更加平稳。

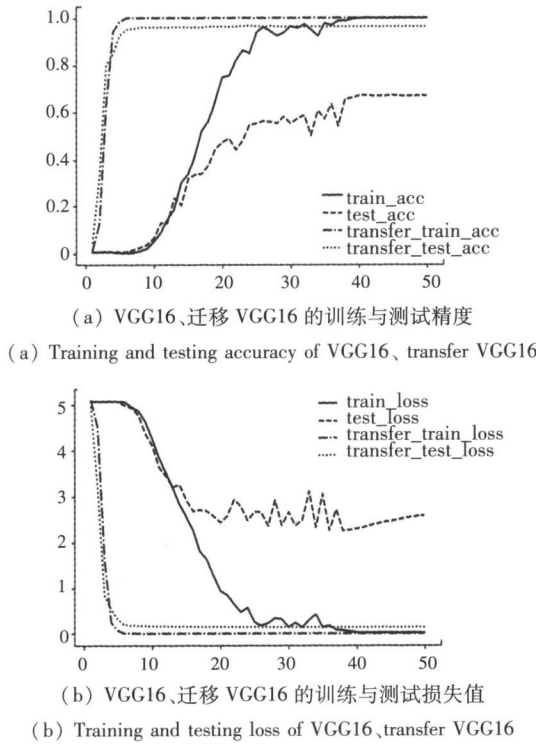


图 5 VGG16 运行结果

Fig. 5 VGG16 running results

由图 5(a) 可以看到,随机初始化参数的 VGG16 在第 38 个 epoch 时,训练达到最好效果。此时训练准确率为 99.89%,测试准确率为 66.87%。此后,训

训练准确率和测试准确率没有明显的变化,图 5(b) 训练损失值不再变化,而测试损失却开始增加,此时网络已经过拟合。而使用训练权重的网络在第 5 个 epoch 时训练准确率就已达到 100%;测试准确率为 95%,测试准确率在第 17 个 epoch 时达到了 96.56%,之后不再变动。

### 3.2 基于迁移的 VGG16 上识别效果

图 6 为 4 组实验的准确率曲线图。从左至右分别是 4、9、16、25 等分掌纹图像块。曲线图的横坐标为迭代次数,纵坐标为准确率。由图可知,掌纹图像块的尺寸越大,包含的有效信息越多,测试准确率越高,而且曲线越平稳,波动越小。

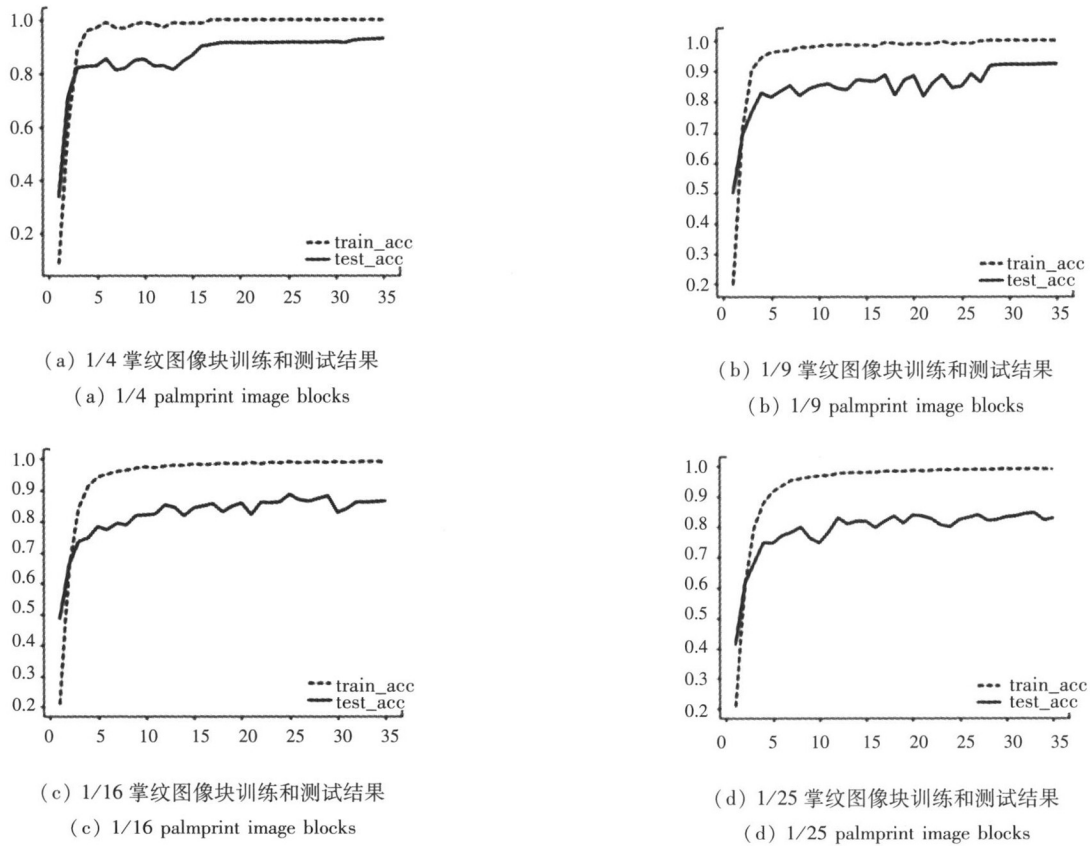


图 6 不同掌纹图像块的准确率

Fig. 6 The accuracy of different palmprint image blocks

图 7 为裁剪成不同份数的掌纹图像的测试准确率。从图中可以看到,各个准确率曲线都存在一定的波动性,掌纹图像分的越小最后投票得到的掌纹图像准确率的波动性越小。因此尽可能地将图片裁剪,最后投票得到的准确率越高,且结果更加的稳定。25 等分的掌纹图像的识别率最好,16 等分略好于 9 等分,最差的是 4 等分。

掌纹图片分的越小,就有越多的掌纹图像块参与投票,投票机制的容错率越高。另一方面,掌纹图像分的越多,总的重叠部分越多,4 等分的重叠部分为 90 像素,9 等分为 180 像素,16 等分为 210 像素,25 等分为 240 像素,对纹理的特征提取有更多的冗余,虽然有更好的识别精度,但也增加了计算成本。

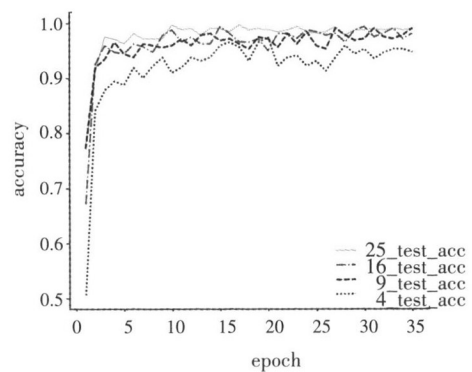


图 7 不同裁剪份数掌纹图像测试准确率

Fig. 7 Test accuracy of palmprint images with different cropped copies

表 2 不同裁剪份数的测试结果  
Tab. 2 Test results of with different cropped copies

裁剪份数	掌纹图像块 准确率%	掌纹图像 准确率%	单张训练 时间/ms	训练时间/s
1	96.56	96.56	43	41
4	93.47	97.5	16	61
9	92.36	99.06	10	86
16	87.20	99.38	7	110
25	84.76	99.69	6	138

表 2 为一张掌纹图像裁剪为不同的大小后掌纹图像块的测试识别率、掌纹图像的测试识别率与训练时间对比。由表可知,掌纹图片裁剪的越小,掌纹图像块的识别准确率也越低;掌纹图片裁剪的越小,掌纹图像块越小,单张掌纹图像块训练的时间也越少,但总的训练时间有所增加,其中,25 等分的掌纹图像块的训练时间是 4 等分训练时间的两倍多。综合准确率与训练时间,最好的裁剪份数为 9。9 等分的图像准确率为 99.06%,相比于没有裁剪的掌纹图像准确率增加了 2.5%,训练时间增加了一倍。16 等分、25 等分掌纹图像准确率相比与 9 等分只增加了百分之零点几,训练时间却增加了 28%和 49%。

表 3 不同算法识别率对比  
Tab. 3 Comparison of recognition rates of different algorithms

算法	识别率%
VGG16	66.87
VGG16_transfer	96.56
Inception_v3	89.69
Inception_v3_transfer	99.06
文献[7]	78.70
文献[8]	67.83
文献[10]	79.40
文献[17]	71.50
proposed	99.69

表 3 为不同算法的识别精度对比,可以看到,迁移学习可以极大的提升网络的识别精度。表中列举了 4 个其它方法得到的高分辨率掌纹图像识别率,都是采用的传统的机器学习,可以看到,本文提出的方法识别精度远远高于机器学习基于细节特征匹配的方法。

4 结束语

目前深度学习的网络模型都需要大量的数据进行训练,数据量不够会出现网络无法训练或者过拟合,本文采用 VGG16 为基础网络,结合迁移学习的

方法,对高分辨率掌纹识别的准确率能达到 96.56%,采用了迁移学习的 VGG16 网络其在准确率、收敛速度和稳定性都优于随机初始化的模型。使用图像增强技术对高分辨率掌纹图像进行处理,以掌纹图像 4 等分、9 等分、16 等分、25 等分为例,验证了掌纹图像分块识别投票算法的可行性。在实际应用中,考虑到计算成本和计算时间,选择将高分辨率掌纹图像 9 等分将会是更好的选择。

参考文献

[1] 王瀚,刘重晋,付翔,等. 基于学习的高分辨率掌纹细节点质量评价方法[J]. 软件学报,2014,25(9):2180-2186.

[2] 郑仕伟,韩俊刚. 基于深度学习的非接触掌纹识别方法[J]. 微电子学与计算机,2018,35(4):98-102.

[3] 王杰. 动态交叉模板与方向滤波融合的掌纹褶皱检测方法[D]. 西安:西安电子科技大学,2014.

[4] 王海纶,李书杰,贾伟,等. 卷积神经网络在掌纹识别中的性能评估[J]. 中国图象图形报,2019,24(8):1231-1248.

[5] 刘重晋,王瀚,封举富. 基于 Gabor 相位和图像质量评价的高分辨率掌纹细节点提取算法[J]. 北京大学学报(自然科学版),2015,51(3):384-390.

[6] 吴春生,冯才刚,迟学斌. 刑侦领域高分辨率掌纹识别技术及快速匹配方法[J]. 中国科学院大学学报,2014,31(4):555-563.

[7] JAIN A K, FENG J. Latent palmprint matching[J]. IEEE Trans Pattern Anal Mach Intell, 2009, 31(6):1032-1047.

[8] FENG J F, LIU C J, WANG H, et al. High-resolution palmprint minutiae extraction based on Gabor feature[J]. enceChina,2014,57(11):112106.

[9] FEI Lunke, TENG Shaohua, WU Jigang, et al. Enhanced minutiae extraction for high - resolution palmprint recognition [ J ]. International Journal of Image and Graphics, 2017, 17 ( 4 ): 1750020.

[10] LIU Eryun, Jain, et al. A coarse to fine minutiae-based latent palmprint matching[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(10):2307-2322.

[11] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient - based learning applied to document recognition[J]. Proceedings of the IEEE,1998,86( 11 ): 2278-2324.

[12] 张功国,吴建,易亿,等. 基于集成卷积神经网络的交通标志识别[J]. 重庆邮电大学学报(自然科学版),2019,31(4):571-577.

[13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10) [2018-04-12].

[14] 伍思雨,冯骥. 基于改进 VGGNet 卷积神经网络的鲜花识别[J]. 重庆师范大学学报(自然科学版),2020,37(4):124-131.

[15] 孙伟师. 基于迁移学习的天气图像识别[J/OL]. 西华大学学报(自然科学版):1-5[2020-10-11].

[16] DAI J, FENG J, ZHOU J. Robust and efficient ridge-based palmprint matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012,34(8):1618-1632.

[17] 费伦科. 掌纹图像多级特征提取与识别算法研究[D]. 哈尔滨:哈尔滨工业大学,2017.