

基于改进 DenseNet 网络的书法字体识别算法^①



麦良廷, 梁 艳, 潘家辉, 黄嘉琳, 陈禧琳, 余依聪

(华南师范大学 软件学院, 佛山 528225)

通信作者: 梁 艳, E-mail: liangyan@m.scnu.edu.cn

摘 要: 汉字书法是中华传统文化的代表, 但是, 由于书法字体具有风格迥异、结构复杂、变形繁多等特点, 给大众学习和欣赏书法带来了极大障碍. 为了解决普通老百姓解读书法作品的困难, 提出一种基于改进 DenseNet 网络的书法字体识别算法, 设计区域权值比例池化规则替换传统 DenseNet 网络的最大池化和平均池化规则, 采用 Nadam 算法优化模型训练效果, 进行自适应学习率调整, 此外, 提出基于剪枝技术的模型裁剪策略, 在保证识别性能的同时, 提高了模型的训练效率. 实验结果表明, 在由楷书、行书、隶书和篆书 4 类字体组成的混合字体数据集中, 本文算法获得了 96.13% 的识别率, 优于另外 5 种深度学习模型.

关键词: 深度学习; DenseNet; 书法字体识别; 池化规则; 模型裁剪

引用格式: 麦良廷, 梁艳, 潘家辉, 黄嘉琳, 陈禧琳, 余依聪. 基于改进 DenseNet 网络的书法字体识别算法. 计算机系统应用, 2022, 31(2): 253-259. <http://www.c-s-a.org.cn/1003-3254/8326.html>

Calligraphic Font Recognition Algorithm Based on Improved DenseNet Network

MAI Gen-Ting, LIANG Yan, PAN Jia-Hui, HUANG Jia-Lin, CHEN Xi-Lin, SHE Yi-Cong

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: Chinese calligraphy is one of the representatives of Chinese traditional culture. However, the different styles, complex structures, and various distortions of calligraphic fonts have brought great obstacles to learning and appreciating calligraphy for the public. A calligraphic font recognition algorithm based on an improved DenseNet network is proposed to solve the difficulty of ordinary people in interpreting calligraphy works. A regional weight ratio pooling rule is designed to replace the maximum pooling and average pooling rules of the traditional DenseNet network. The Nadam algorithm is used to adjust the adaptive learning rate and optimize the model training effect. In addition, a model pruning strategy based on the pruning technology is proposed, which ensures a strong recognition performance and improves the training efficiency of the model. The experimental results show that in a mixed font data set composed of four types of fonts, namely the standard script, the running script, the clerical script, and the seal script, the proposed algorithm obtains a recognition rate of 96.13%, which is better than those of the other five deep learning models.

Key words: deep learning; DenseNet; calligraphic font recognition; pooling rules; model pruning

中国书法艺术不但是中华民族的瑰宝, 更是人类的共同财富. 随着信息化的不断发展, 书法作为中国传统艺术急需被继承与弘扬. 针对手写体字符识别的研究由来已久, 书法字作为一种特殊的手写体, 由于存在不同字体风格以及不同人所写书法存在个体差异等原

因, 对书法字体识别研究带来极大的挑战. 手写汉字书法字体识别的研究, 不仅为智能字符识别提供一种新的解决方案, 而且对于弘扬中华传统文化, 为广大书法爱好者学习、欣赏和传承书法艺术提供极大的帮助, 具有重要的理论价值和社会意义.

^① 基金项目: 广州市科技计划重点领域研发计划 (202007030005)

收稿时间: 2021-04-27; 修改时间: 2021-05-19; 采用时间: 2021-06-02; csa 在线出版时间: 2022-01-17

早期书法字体识别主要采用基于传统机器学习的方法。段银雷^[1]通过 K-means 方法进行分类,但在聚类过程中需要人工参与调试参数,未能完全实现自动分类。肖斌等人^[2]在特征提取方面采用了边缘法、骨架法、笔画法的思想,并且利用 SVM 方法对特征样本进行识别,但该方法在不同网格区域中并未通过加权处理来加强关键区域位置特征所占比重。Zhang 等人^[3]提出一种多尺度特征提取方法,在不同尺度参数下构造一系列最优特征空间,并采用嵌套子集马氏距离分类器实现字符识别,在常规数据集和草书数据集测试,分别获得 99.3% 和 88.4% 的准确率。但是基于传统机器学习的算法泛化性不强,无法对存在多种风格的大规模书法字体数据进行有效分类。Gao 等人^[4]提出一种基于检索的快速识别方案,设计出一种新颖的形状描述符,用于表示书法字符图像,以进行高效有效的检索,但其对噪声等干扰比较敏感。

近年来随着深度学习相关理论的发展和完善,深度学习模型由于具有较强的特征学习能力,受到众多学者关注,应用于汉字书法识别领域。曲延直^[5]采用改进后的深度残差网络 ResNet-50 对汉字书法进行识别,在自建的数据集进行实验,证明该模型能通过增加网络深度提高模型性能。但由于训练所用的数据集规模较小,对模型的训练产生一定影响,在测试集仅获得 90.6% 的准确率。Gao 等人^[6]通过实验证明了利用深度卷积神经网络提取的深度特征比手工设计的特征能更好地捕获书法风格信息。Melnik 等人^[7]提出一种用于离线 HCCR 的高性能 CNN-Melnik 网络,仅对一种风格的手写数据进行训练,其精度可达到 97.61%,但并未说明对于多种汉字风格的数据的识别效果如何。Zou 等人^[8]研究 CNN 模型的损失函数,提出字符分类和相似度排序监督信号是互补的,最终准确率为 95.58%,但该方法仅针对硬笔手写汉字进行识别。许嘉谕等人^[9]采用 GoogLeNet Inception-v3 模型判断书法作品图像的书法风格,然后利用 ResNet-50 残差网络进行单字字形识别,但该方法仅针对楷书和篆书两种字体进行识别,适用性较低。

稠密卷积神经网络 (dense convolutional network, DenseNet)^[10]是 2017 年由 Huang 等人提出的一种经典的卷积神经网络 (convolutional neural network, CNN)^[11]架构。DenseNet 通过任意两层之间的直接连接加强特征的传递,更有效地利用了特征,同时其核心模块

dense block 的设计使得网络不容易过拟合。与其他 CNN 模型相比, DenseNet 具有明显优势。然而,传统的 DenseNet 网络在池化规则、训练策略等方面仍需要进一步优化。同时,该网络中的参数存在大量冗余,逐层连接的模式使得网络中的每一层均需依赖于它上面所有层的特征,影响其计算效率。

因此,针对中国书法字体数量庞大、结构复杂、变形繁多的特点,本文提出一种改进的 DenseNet 模型用于汉字书法字体识别,从池化规则、训练策略、模型裁剪 3 个方面进行优化,降低计算复杂度和模型复杂度。结合数据增强方法,我们对改进后的 DenseNet 网络模型进行训练与测试。实验结果表明,与传统的 DenseNet 网络相比,改进的 DenseNet 网络获得了更高的识别准确率,同时,在计算效率方面也明显优于传统的 DenseNet 网络。本文所提方法具有更强的鲁棒性与泛化能力,为现实场景中的计算机辅助书法字体识别提供了高效的解决方案。

1 模型结构

1.1 DenseNet 网络

DenseNet 是密集型网络模型,一般具有 L 层的传统卷积网络具有 L 个连接(每一层与其后一层之间有一个连接),而 DenseNet 具有 $L(L+1)/2$ 个直接连接。这一结构使得参数的重用性增大,在一定程度上防止了过拟合现象的发生。DenseNet 由 5 个 Dense block 组成,而 Dense block 采用了激活函数在前,卷积层在后的顺序,即 BN-ReLU-Conv 的顺序,这种方式也被称为预激活, DenseNet 的结构如图 1 所示。

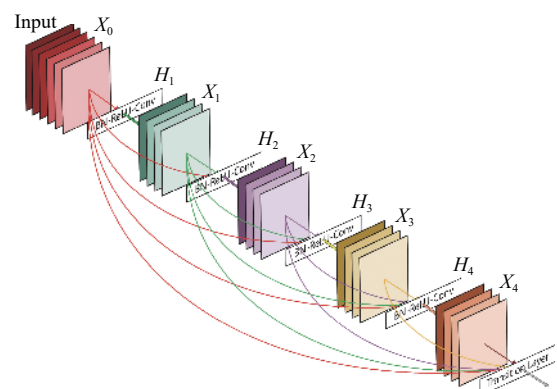


图 1 DenseNet 的 5 层结构图

DenseNet 的 L 层都实现了一个非线性变换 $K_i(\cdot)$, 其中, i 表示第 i 层, 假设输入图像为 h_0 , 那么, 第 i 层的输

出 h_i 可根据下式计算得到:

$$h_i = K_i([h_0, h_1, \dots, h_{i-1}]) \quad (1)$$

从式(1)可见,第 i 层以之前所有层的特征映射 h_0, h_1, \dots, h_{i-1} 作为输入,提升了网络层间的信息流,使其更易于训练^[12].

1.2 改进的 DenseNet 网络

为了适应中国书法字体数量庞大、结构复杂、变形繁多的特点,本文从池化规则、训练策略、模型裁剪3个方面对传统的 DenseNet 网络进行改进,具体如下:

(1) 将传统 DenseNet 网络的最大池化规则和平均池化规则进行融合,设计了区域权值比例池化规则,从而减少特征提取的误差;

(2) 采用 Nadam 算法对模型训练效果进行优化,让模型能够进行自适应学习率调整,提高模型收敛速度和模型的性能,加强对不同风格书法字体的特征提取;

(3) 针对存在的连接冗余问题,提出一种可行的基于剪枝算法^[13]的模型裁剪策略,对 DenseNet 网络中贡献不大的 BN 层进行裁剪,得到更轻量的汉字书法识别模型,从而提高模型的计算效率。

1.2.1 网络结构的确定

DenseNet 算法有4种网络结构,分别为 DenseNet-121、DenseNet-169、DenseNet-201、DenseNet-264。书法汉字识别与一般的图像识别任务有所不同,书法汉字图片一般为黑白色,比彩色图片所具有的特征少了许多。同时,相同汉字不同风格的字体形状相差不大,特征不明显,因此需要更深的网络来获取特征^[14]。经过实验,DenseNet-201 与 DenseNet-264 所能达到的识别率是最高的,但 DenseNet-264 与 DenseNet-201 相比,识别率并无明显的提升,因此本文选取 DenseNet-201 作为基础模型,并以此进行改进,最终得到的 DenseNet 网络结构如图2所示。

1.2.2 设计区域权值比例池化规则进行特征提取

卷积神经网络特征提取的误差主要来自两个方面^[15],一方面是邻域大小受限造成的估计值方差增大;另一方面是卷积层参数误差造成估计均值的偏移^[16]。DenseNet 中的平均池化(average-pooling)规则^[17]能够有效地减少第1种误差,但是容易造成正负激活值互相抵消的情况出现,从而导致丢失重要信息。而最大池化(max-pooling)规则^[18]可以减少第2种误差,但不能解决第1种误差,因此本文设计了区域权值比例池化(regional-weight-pooling, RWP)规则,将最大池化与平均池化规则进行融合,最大程度地减少这两种误差。

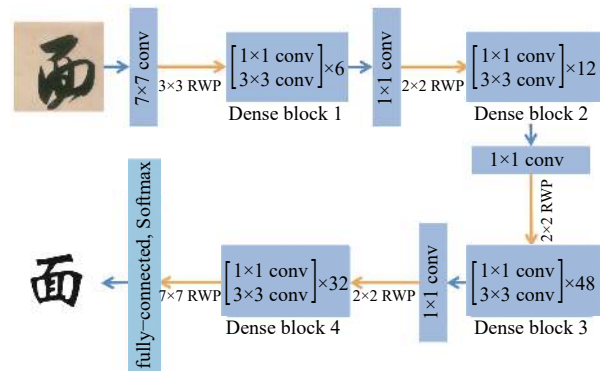


图2 改进的 DenseNet 网络结构

设输入数据为 $M \times N$ 矩阵,采样核大小为 n ,使用区域权值比例池化规则进行池化,则池化窗口区域数据为 a_1, a_2, \dots, a_n ,每个数据的权值比例为 w_1, w_2, \dots, w_n ,则该池化窗口区域的池化结果 r 可根据式(2)计算得到:

$$r = \sum_{i=1}^n w_i a_i \quad (2)$$

其中,

$$w_i = \frac{a_i + |a_{\min}|}{\sum_{j=1}^n (a_j + |a_{\min}|)} \quad (3)$$

在计算权值 w_i 时,为保证其均为正数,对池化窗口区域数据加上一个最小值 a_{\min} 的绝对值,最小值 a_{\min} 根据式(4)计算得到:

$$a_{\min} = \min(a_1, a_2, \dots, a_n, 0) \quad (4)$$

1.2.3 采用 Nadam 算法优化模型训练效果

传统的 DenseNet 模型一般采用 SGD (stochastic gradient descent) 优化算法,但 SGD 下降速度慢,更新参数较为频繁,使得代价函数容易发生严重的震荡,而且容易收敛到局部最优。考虑到 Nadam 优化器^[19]的参数更新不受梯度伸缩变换影响^[20],能实现学习率的自动调整,且能够更好更快地找准梯度下降的方向,更容易达到一个全局最优,同时, Nadam 优化器实现简单,计算高效,因此,本文采用 Nadam 算法对模型进行优化,让模型能够更好地进行自适应学习率调整。

设待优化参数为 ϵ ,目标函数为 $f(\epsilon)$,学习率为 α ,用 k 表示每个 epoch,则 k 时刻的梯度 g_k 可通过下式计算获得:

$$g_k = \nabla f(\epsilon_{k-1}) \quad (5)$$

根据历史梯度计算一阶动量 F_k 和二阶动量 S_k ,其中,前者是各个时刻梯度方向的指数移动平均值,后者用于度量历史时刻的梯度更新频率^[21],两者的初始值

均为 0, 具体计算公式如下:

$$F_k = \beta_1 F_{k-1} + (1 - \beta_1) g_k \quad (6)$$

$$S_k = \beta_2 S_{k-1} + (1 - \beta_2) g_k^2 \quad (7)$$

其中, β_1 和 β_2 表示两个超参数, 前者控制一阶动量, 后者控制二阶动量. β_1 的经验值为 0.9, 这就意味着下降方向主要是此前累积的下降方向, 并略微偏向当前时刻的下降方向. β_2 的经验值为 0.999, 说明二阶动量基本取决于此前累积的梯度更新频率. 在迭代初期, F_k 和 S_k 容易趋于 0, 因此引入偏置纠正, 如式 (8) 和式 (9) 所示:

$$\hat{F}_k = \frac{F_k}{1 - \beta_1^k} \quad (8)$$

$$\hat{S}_k = \frac{S_k}{1 - \beta_2^k} \quad (9)$$

根据式 (10) 自适应调整学习率:

$$\alpha_k = \alpha_{k-1} / \sqrt{\hat{S}_k} \quad (10)$$

接着根据式 (8) 与 (10) 计算当前时刻的下降梯度:

$$\eta_k = \alpha_k \hat{F}_k \quad (11)$$

最后根据下降梯度更新参数 ε :

$$\varepsilon_k = \varepsilon_{k-1} - \eta_k \quad (12)$$

1.2.4 采用自适应裁剪策略对模型进行裁剪

传统的 DenseNet 网络连接冗余^[22], 导致计算效率较低. Liu 等人^[23]采用 BN 层归一化激活值的方法, 实现高效鉴别和剪枝不重要的层. 受该方法的启发, 本文提出自适应裁剪策略, 设计基于剪枝算法的模型裁剪策略, 利用 L1 正则化对通道进行剪枝以实现网络瘦身.

首先, 令 DenseNet 网络中的 BN 层的 γ 参数作为网络瘦身的缩放因子, 以减少剪枝给网络带来的额外开销. 接着, 利用 L1 正则化使得网络中的一些缩放因子的大小趋向于 0, 缩放因子的大小越接近于 0, 表明该缩放因子对应的通道越不重要. 而具体要剪掉哪些通道, 需要通过一个阈值进行判断. 假设裁剪阈值为 T , 先对所有缩放因子的绝对值进行升序排序, 将该序列中前 T 位置的缩放因子所对应的通道剪掉, 最后得到一个较少参数、运行时内存占用小、计算量较低的紧凑网络.

2 数据与实验

2.1 数据源及预处理

2.1.1 收集书法作品集并建立字库

由于目前公开的书法字数据集较少, 并且数据集

规模无法满足本实验的需求, 故采用爬虫技术收集汉字, 从国学大师网站 (www.guoxuedashi.com) 分别爬取楷书、行书、隶书和篆书 4 类字体的常用 3 500 个汉字的单字图像. 进行预处理后, 对每一张图片进行人工审核, 确保每张图片都是合格的, 最后得到 4 类字体的常用汉字单字字库. 分类字体数据集规模如表 1 所示.

表 1 分类字体数据集

书体	单字类别数量 (个)	总图片数 (张)
楷书	3 488	52 724
篆书	3 286	71 396
行书	3 499	143 368
隶书	3 207	48 171

2.1.2 图像预处理

为了使神经网络能够更好地提取图像的数据特征, 本文先对图像进行预处理. 首先对图像进行尺寸归一化, 处理后的图像尺寸均为 224×224. 为了降低噪声对汉字识别的准确性和鲁棒性的影响, 采用高斯滤波对图像进行平滑去噪^[24], 指定高斯核的大小为 3×3. 通过加权平均对图像进行灰度化并采用大类间方差法 (OTSU) 对图像进行全局二值化处理^[25], 使字体和背景分别为黑白两色. 最后对处理后的图像进行数据增强, 对同一张字体图像采用 15% 范围内随机水平、垂直平移和随机放大的方案, 最终得到原基础上 10 倍的数据量.

考虑到汉字在图像中的位置不一致会对识别的结果造成一定的影响, 本文采用连通域提取的方法, 保留连通域内的部分, 即图像中的汉字部分. 由于汉字的笔划不一定完全粘连, 例如“彳”部首可能会被分成 3 个连通域, 因此对提取的连通域计算并集, 组成最大连通域. 根据该区域的顶点坐标值对图像进行裁剪, 得到最终的目标汉字区域. 图像预处理过程如图 3 所示.

2.2 实验配置及参数设置

本文提出了一种改进的 DenseNet 卷积网络的手写书法字体识别算法, 其软硬件平台: CPU: Intel(R) CORE(TM) i5-8250U 3.40 GHz; GPU: NVIDIA GeForce RTX 2060; 操作系统: Windows 10 专业版; 深度学习框架: TensorFlow 2.2. 为了验证本文提出的书法字体识别算法的有效性, 本文采取两种训练策略, 一种为将每种风格字体的识别模型分开进行训练, 最终得到 4 个不同风格的字体内容识别模型; 另一种为将 4 种风格字体数据集进行融合, 形成一个字体风格混合数据集进行训练, 最终得到一个可以识别楷书、隶书、篆书、行书 4 种风格字体的字体内容识别模型. 经过最终调整, 本文提出的网络模型参数设置如下: depth=40,

growth_rate=10, bottleneck=True, reduction=0.5, 初始学习率设置为 0.001, batch 为 64, 其余对比算法则采用 GitHub 项目源码进行训练测试. 为确定本文算法中网络训练时 epoch 参数的大小, 我们采用第 2 种训练策略进行训练和验证, 实验结果如图 4 所示.

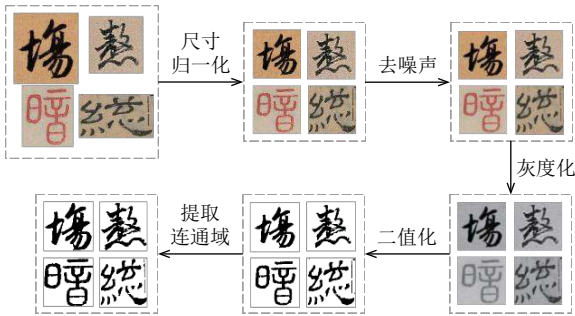


图3 图像预处理过程

从图 4 可发现, 在 epoch 为 25 时, 曲线已趋于平缓, 模型的识别准确率和损失值基本不再变化, 说明模型已经收敛, 此时识别准确率已趋于稳定且已达到最优值. 因此在后续的实验中, epoch 设置为 25.

2.3 模型裁剪策略中阈值的确定

为了确定模型裁剪的阈值, 本文使用传统 DenseNet-201 网络模型, 选取行书数据集进行测试, 实验结果如表 2 所示.

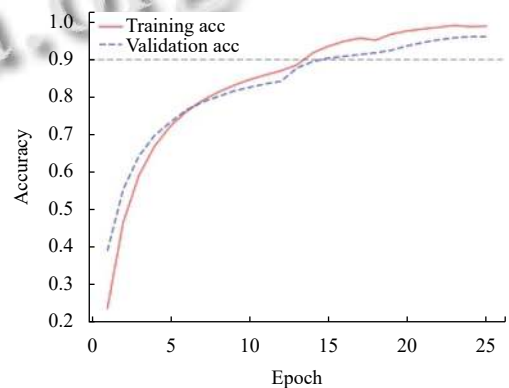
由表 2 可知, 当模型裁剪阈值小于 30% 时, 识别率变化不大, 均能达到 91% 以上, 这说明被裁剪部分的模型连接存在冗余. 当模型裁剪阈值逐渐增大时, 识别率也随之下降. 当阈值为 70% 时, 识别率下降到 31.69%. 因此, 我们最终把模型裁剪阈值设置为 30%, 在保证识别性能的前提下, 降低了模型的计算量.

2.4 不同风格字体单独进行模型训练的实验结果与分析

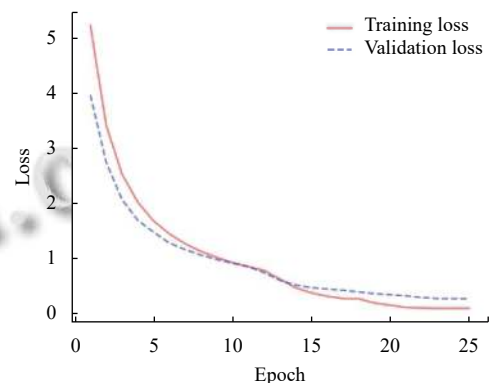
为了验证本文算法的有效性和可靠性, 本文分别对楷书、篆书、隶书和行书 4 种风格字体的识别模型进行训练. 将数据集按照 8:2 的比例划分训练集和测试集, 使用改进的 DenseNet-201 算法对其进行训练. 我们在测试集测试了本文算法, 同时与传统 DenseNet-201、ResNet-50^[26]、AlexNet^[15]、GoogLeNet V4^[27]、DPN-92^[28] 5 种算法进行对比, 识别结果如表 3 所示.

从表 3 可以看到, 本文方法针对楷书、篆书、行书、隶书 4 种风格字体的平均识别率为 94.69%, 比传统 DenseNet-201 算法高 1.78%, 能够准确识别出绝大部分单字书法图像内容. 虽然 DPN 结合了 ResNet 和 DenseNet 的优点, 参数更少, 但书法图像所包含的特征信息太多, DPN 无法提取到更多有用的分类特征, 因此

在书法字体识别任务中表现较差, 平均识别率仅为 82.94%. AlexNet 的网络结构较为简单, 无法很好地提取到书法图像的特征, 平均识别率也仅为 83.92%. 传统 DenseNet-201、ResNet-50、GoogLeNet V4 由于其网络结构较深, 因此表现不错, 平均识别准确率均达到 90% 以上. 虽然传统 DenseNet-201 算法所训练的隶书识别模型准确率高于本文改进的 DenseNet 算法所训练得到的模型, 但从总体来分析, 本文改进的 DenseNet 算法平均识别率明显高于其他 5 种算法模型, 这说明经过改进后的 DenseNet-201 算法在书法字体识别任务中可以获得更高的性能.



(a) 识别准确率变化趋势图



(b) 损失函数变化趋势图

图4 本文算法的识别准确率及损失函数变化趋势图

表2 使用不同阈值进行模型裁剪的识别结果 (%)

阈值	识别率
0	91.08
10	91.05
20	91.07
30	91.08
40	90.76
50	86.13
60	74.36
70	31.69

表3 在分类字体数据集的识别结果比较

算法模型	单字风格	测试集数量	正确识别数量	错误识别数量	识别率 (%)	平均识别率 (%)
传统DenseNet-201	楷书	10 544	9 950	594	94.37	92.91
	篆书	14 279	13 449	830	94.19	
	行书	28 673	26 115	2 558	91.08	
	隶书	9 634	9 140	494	94.87	
ResNet-50	楷书	10 544	9 314	1 230	88.33	91.50
	篆书	14 279	13 245	2 098	92.76	
	行书	28 673	26 250	2 423	91.55	
	隶书	9 634	8 992	603	93.34	
AlexNet	楷书	10 544	9 160	1 384	86.87	83.92
	篆书	14 279	11 760	2 519	82.36	
	行书	28 673	24 183	4 490	84.34	
	隶书	9 634	7 873	1 761	81.72	
GoogLeNet V4	楷书	10 544	9 231	1 313	87.55	90.72
	篆书	14 279	13 612	667	95.33	
	行书	28 673	26 041	2 632	90.82	
	隶书	9 634	8 388	1 246	87.07	
DPN-92	楷书	10 544	8 920	1 624	84.6	82.94
	篆书	14 279	12 487	1 792	87.45	
	行书	28 673	23 667	5 006	82.54	
	隶书	9 634	7 284	2 350	75.61	
本文方法	楷书	10 544	10 084	460	95.64	94.69
	篆书	14 279	13 676	603	95.78	
	行书	28 673	26 703	1 970	93.13	
	隶书	9 634	9 075	559	94.20	

2.5 所有风格字体混合进行模型训练的实验结果与分析

为了进一步提高对各种风格单字书法图片的识别率, 本文将4种风格的书法字体图片进行混合, 形成一个字体风格混合的数据集, 我们称之为混合字体数据集, 图片总数为315 659张. 数据集也是按照8:2的比例划分训练集和测试集. 对模型进行训练后的实验结果如表4所示. 同时, 表4也列出了传统DenseNet-201、ResNet-50、AlexNet、GoogLeNet V4、DPN-92算法在混合字体数据集的测试结果.

表4 在混合字体数据集的识别结果比较

算法模型	测试集数量	正确识别数量	错误识别数量	识别率 (%)
传统 DenseNet-201	63 131	59 261	3 870	93.87
ResNet-50	63 131	57 430	5 701	92.97
AlexNet	63 131	52 146	10 985	82.6
GoogLeNet V4	63 131	58 194	4 937	92.18
DPN-92	63 131	51 187	11 944	81.08
本文方法	63 131	60 688	2 443	96.13

从表4的实验结果可以看到, 与另外5种算法相比, 本文方法获得了最高的平均识别率96.13%, 比传统DenseNet-201算法高出2.26%. 此外, 与在分类字体数

据集的实验结果相比, 传统DenseNet-201、ResNet-50、GoogLeNet V4以及本文方法均取得了更高的平均识别率, 说明随着数据集的增大, 模型能够学习到更多的特征. 但值得注意的是, AlexNet和DPN-92模型在混合字体数据集取得了更低的平均识别率, 这可能是由于所选模型没有能力学习这么大的训练数据集, 或者发生过拟合现象而导致的.

3 结束语

本文提出了一种改进的DenseNet网络用于汉字书法字体识别, 通过设计区域权重比例池化规则进行特征提取, 合理地减少了卷积神经网络进行特征提取时带来的误差; 采用Nadam算法对DenseNet网络进行了优化, 让模型能够进行自适应学习率调整, 加强对不同风格书法字体的特征提取; 针对传统DenseNet模型存在的冗余问题, 本文提出了一种可行的自适应模型裁剪方案, 以得到更紧凑的汉字书法字体识别模型. 本文算法在分类字体数据集和混合字体数据集分别取得了94.69%和96.13%的识别率, 明显优于另外几种常见的卷积神经网络模型. 在未来的工作中, 我们将引入更多风格字体进行研究, 进一步优化算法模型, 提高算法的泛化性能.

参考文献

- 1 段银雷. 基于 K-means 的图像文字识别与提取研究. 电子技术与软件工程, 2015, (9): 88.
- 2 肖斌, 黄襄念. 基于 SVM 的几种汉字特征提取法比较研究. 西华大学学报 (自然科学版), 2009, 28(5): 70–74. [doi: 10.3969/j.issn.1673-159X.2009.05.020]
- 3 Zhang JY, Ding XQ, Liu CS. Multi-scale feature extraction and nested-subset classifier design for high accuracy handwritten character recognition. Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000. Barcelona: IEEE, 2000. 581–584.
- 4 Gao PC, Wu JQ, Lin Y, *et al.* Fast Chinese calligraphic character recognition with large-scale data. Multimedia Tools and Applications, 2015, 74(17): 7221–7238. [doi: 10.1007/s11042-014-1969-3]
- 5 曲延直. 基于深度学习的汉字书法识别研究. 电子测试, 2019, (24): 44–46, 61. [doi: 10.3969/j.issn.1000-8519.2019.24.017]
- 6 Gao PC, Gu G, Wu JQ, *et al.* Chinese calligraphic style representation for recognition. International Journal on Document Analysis and Recognition (IJDAR), 2017, 20(1): 59–68. [doi: 10.1007/s10032-016-0277-z]
- 7 Melnyk P, You ZQ, Li KQ. A high-performance CNN method for offline handwritten Chinese character recognition and visualization. Soft Computing, 2020, 24(11): 7977–7987. [doi: 10.1007/s00500-019-04083-3]
- 8 Zou JY, Zhang JL, Wang LD. Handwritten Chinese character recognition by convolutional neural network and similarity ranking. arXiv: 1908.11550, 2019.
- 9 许嘉谕, 林楚焯, 陈志涛, 等. 基于深度学习的手写书法字体识别算法. 计算机系统应用, 2021, 30(2): 213–218. [doi: 10.15888/j.cnki.csa.007819]
- 10 Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2261–2269.
- 11 LeCun Y, Boser B, Denker JS, *et al.* Backpropagation applied to handwritten zip code recognition. Neural Computation, 1989, 1(4): 541–551. [doi: 10.1162/neco.1989.1.4.541]
- 12 杨其睿. 基于改进的 DenseNet 深度网络火灾图像识别算法. 计算机应用与软件, 2019, 36(2): 258–263. [doi: 10.3969/j.issn.1000-386x.2019.02.046]
- 13 He YH, Zhang XY, Sun J. Channel pruning for accelerating very deep neural networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 1398–1406.
- 14 张福成. 基于卷积神经网络的书法风格识别的研究 [硕士学位论文]. 西安: 西安理工大学, 2018.
- 15 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, *et al.* eds. Advances in Neural Information Processing Systems. Cambridge: The MIT Press, 2012. 1097–1105.
- 16 Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. arXiv: 1603.07285, 2018.
- 17 Lin M, Chen Q, Yan SC. Network in network. arXiv: 1312.4400, 2014.
- 18 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1–9.
- 19 Dozat T. Incorporating nesterov momentum into Adam. Proceedings of International Conference on Learning Representations. San Juan: ICLR, 2016. 1–4.
- 20 Li LT, Xu WZ, Yu H. Character-level neural network model based on Nadam optimization and its application in clinical concept extraction. Neurocomputing, 2020, 414: 182–190. [doi: 10.1016/j.neucom.2020.07.027]
- 21 Ruder S. An overview of gradient descent optimization algorithms. arXiv: 1609.04747, 2017.
- 22 Huang G, Liu SC, van der Maaten L, *et al.* CondenseNet: An efficient densenet using learned group convolutions. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 2752–2761.
- 23 Liu Z, Li JG, Shen ZQ, *et al.* Learning efficient convolutional networks through network slimming. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2755–2763.
- 24 任晓文, 王涛, 李健宇, 等. 基于深度学习的异噪声下手写汉字识别的研究. 计算机应用研究, 2019, 36(12): 3878–3881.
- 25 张俊松. 书法碑帖图像去噪、轮廓拟合及纹理建模研究 [博士学位论文]. 杭州: 浙江大学, 2007.
- 26 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778.
- 27 Szegedy C, Ioffe S, Vanhoucke V, *et al.* Inception-v4, inception-ResNet and the impact of residual connections on learning. Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco: AAAI Press, 2017. 4278–4284.
- 28 Chen YP, Li JN, Xiao HX, *et al.* Dual path networks. arXiv: 1707.01629, 2017.