



DOI:10.12404/j.issn.1671-1815.2403015

引用格式:胡瑞朋,何春燕,张伟明,等.基于卷积循环神经网络的手写汉字文本识别[J].科学技术与工程,2025,25(4):1547-1554.

Hu Ruipeng, He Chunyan, Zhang Weiming, et al. Handwritten chinese character text recognition based on convolutional recurrent neural network[J]. Science Technology and Engineering, 2025, 25(4): 1547-1554.

基于卷积循环神经网络的手写汉字文本识别

胡瑞朋^{1,2}, 何春燕^{1,2*}, 张伟明^{1,2}, 赵立新^{1,2}, 李明博³

(1. 河北工程大学机械与装备工程学院, 邯郸 056038; 2. 河北工程大学河北省智能工业装备技术重点实验室, 邯郸 056038;
3. 集智康(北京)科技有限公司, 北京 102600)

摘要 为了解决卷积循环神经网络(convolutional recurrent neural networks, CRNN)手写汉字文本识别网络模型的训练参数大、文本识别率低等问题,提出一种基于注意力双向长短期记忆网络(based on attention bi-directional long short-term memory network, AT-BLSTM)和知识蒸馏(knowledge distillation, KD)技术的手写汉字识别方法。通过对 AT-BLSTM 网络的输入向量特征赋予不同的权重,使模型训练数据集更加高效、准确;通过 KD 技术将一个高性能的大模型获取的知识传输到一个小模型中,在确保模型准确性的同时,减少训练参数和内存占比,得到一个性能更优的轻量级训练模型。该方法通过多组实验对比,汉字识别准确率提高了 6.7%,训练参数减少 15.94 M。该网络模型识别准确率达到 97.9%,汉字识别效果更好。

关键词 卷积循环神经网络(CRNN); 手写汉字文本识别; 注意力机制; 知识蒸馏(KD)

中图分类号 TP391.1;

文献标志码 A

Handwritten Chinese Character Text Recognition Based on Convolutional Recurrent Neural Network

HU Rui-peng^{1,2}, HE Chun-yan^{1,2*}, ZHANG Wei-ming^{1,2}, ZHAO Li-xin^{1,2}, LI Ming-bo³

(1. Institute of Machinery and Equipment Engineering, Hebei University of Engineering, Handan 056038, China;
2. Key Laboratory of Intelligent Industrial Equipment Technology of Hebei Province, Hebei University of Engineering, Handan 056038, China; 3. Ji Zhi Kang (Beijing) Technology Co., Ltd., Beijing 102600, China)

[Abstract] In order to solve the problems of large training parameters and low text recognition rate of convolutional recurrent neural networks (CRNN) handwritten Chinese character recognition network model, a novel method for handwritten Chinese character recognition based on attention bi-directional long short-term memory network(AT-BLSTM) and knowledge distillation (KD) technology was proposed. By assigning different weights to the input vector features of AT-BLSTM network, the model training data set was more efficient and accurate. Through KD technology, the knowledge acquired from a large high-performance model was transferred to a small model, which ensured the accuracy of the model, reduced the training parameters and internal storage ratio, and obtained a lightweight training model with better performance. Through the comparison of multiple groups of experiments, the accuracy of Chinese character recognition is increased by 6.7%, and the training parameters are reduced by 15.94 M. The recognition accuracy of this network model reaches 97.9%, and the recognition effect of Chinese characters is better.

[Keywords] convolutional recurrent neural networks (CRNN); handwritten Chinese character text recognition; attention mechanism; knowledge distillation(KD)

随着科学技术的高速发展和社会的生活需要,汉字识别技术得到了更多的关注,将纸质文档信息识别转换成电子信息存储能够给人们生活带来极大的便利,也为其他国家学习汉字提供更好、更便利的帮助。因此,通过计算机技术快速获取文本信息并准确识别成为汉字识别的一个重要研究方向。

汉字识别可以分为印刷体汉字识别和手写体汉字识别^[1]。而手写体汉字识别又可分为在线手写体汉字识别和离线手写体汉字识别。离线手写体汉字识别的难点是通过识别人们在纸张上书写好的汉字信息,由于书写者有自己独特的书写方式和习惯,并且提取不到笔画顺序特征,所以识别的

收稿日期:2024-04-24; 修订日期:2024-11-27

基金项目:河北省教育厅科学研究项目(CXY2024046)

第一作者:胡瑞朋(1999—),男,汉族,河北邯郸人,硕士研究生。研究方向:机器视觉、信息处理。E-mail:hurpeng1106@163.com。

*通信作者:何春燕(1979—),女,汉族,河北邯郸人,博士,副教授。研究方向:机器人及自动化。E-mail:hcy0509@163.com。

难度和复杂度都很高。因此,离线手写汉字识别研究成为接下来汉字识别领域的热点问题之一。黄洋^[2]、袁柱^[3]、王永强^[4]在深度学习方面对离线手写汉字识别进行了研究。

与其他语言不同的是,汉字有其自己的特点,主要表现在:①汉字结构复杂、数量庞大,《信息交换用汉字编码字符集》(GB/T 2312—1980)中一级字库就高达 3755 个;②汉字存在很多形近字,具有几乎相同的空间结构和笔画形式;③手写体是人为书写,不同的人有不同的书写方式和习惯;④人为书写还可能存在字体笔画粘连现象。基于以上问题,离线手写汉字文本识别是一个非常困难且具有挑战性的问题^[5-6]。

离线手写汉字识别通常通过依赖序列模型识别技术^[7],一般可以分为过分割方法和无分割方法。过分割方法对于字符重叠和粘连现象很难得到处理。因此,研究出一种不需要将文本分割成单字符的无分割方法得到更多的关注。于是,一种名为隐马尔可夫模型(hidden Markov model, HMM)被关注并融合到了离线手写汉字识别技术中,并且取得了很好的效果^[8]。隐马尔可夫模型是一种时间序列的概率模型,它通过对训练样本进行分析和学习,然后学习状态之间的转移概率和观测之间的发射概率,从而实现离线手写汉字文本序列的识别。Du 等^[9]和 Wang 等^[10]又进一步研究和扩展了 HMM 网络模型在离线手写体中文文本识别中的应用。但是隐马尔可夫模型没有记忆性,随着文本字符长度的增加,导致识别性能下降。

神经网络的发展使得该问题得到解决,循环神经网络(recurrent neural network, RNN)的发展既避免了字符分割的问题,又解决了带记忆性联系上下文的问题。Shi 等^[11]提出了卷积循环神经网络(convolutional recurrent neural networks, CRNN),并将该网络模型运用到了场景文本识别。周旋^[12]在卷积循环网络方面也有着一定进展,通过改进算法对场景文字识别有着很大成果。Hinton 等^[13]提出了一种用于提高轻量化模型性能的知识蒸馏方法,知识蒸馏技术的发展有效解决了神经网络训练参数大和内存占比大等问题。Shi 等^[14]提出了 RARE (Robust text recognizer with Automatic Rectification)模型,将注意力机制引入文本识别领域。杨琼^[15]、王天伟^[16]在注意力机制方面也有着研究,通过在神经网络中加入注意力机制网络对手写汉字识别技术研究。

随着卷积神经网络在汉字识别方向的研究和发展,提出一种基于注意力双向长短期记忆(based

on attention bi-directional long short-term memory network, AT-BLSTM)网络和知识蒸馏(knowledge distillation, KD)技术的 AK-CRNN 手写汉字文本识别方法。该方法融合了 AT-BLSTM 使神经网络能对某一部分有用特征得到更多的关注,而忽略掉那些无用的特征;通过融合知识蒸馏技术减少训练参数量来压缩模型体积获得更高效和轻量级的识别模型。现提出一种 AT-BLSTM 注意力模块,它是一种通过结合双向长短期记忆网络(bi-directional long short-term memory network, BLSTM)融入注意力的模块,使神经网络能对某一部分有用特征更加敏感,而忽略掉那些无关的特征,优化模型的特征提取能力,提高神经网络训练效率和文本识别率。引入知识蒸馏技术,在确保模型精度的前提下尽可能地压缩模型体积,得到更小、更快、轻量级的训练模型。在 Pytorch 框架下,结合 OpenCV(开源计算机视觉库)对数字图像处理技术对文本图片进行信息和背景噪声区分开,对倾斜文本图片进行透视变换使得图片文本校正,便于完成手写汉字识别系统的研究。

1 模型设计与方法

1.1 CRNN 模型

CRNN 网络模型的总体结构由卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)和联结时间分类(connectionist temporal classification, CTC)组成,如图 1 所示。该模型的特点是不需要传统的逐个字符拆分再进行识别的方法,就可以进行不定长的序列文本识别,而是文本序列依次通过卷积层、循环层和转录层依赖时序进行序列识别。首先通过 CNN 对文本图像的特征进行提取,然后通过 BLSTM 将提取字符上下文特征序列的特征向量进行整合,再通过 Softmax 函数将每列的特征进行概率分布,最后通过 CTC 损失函数对文本序列进行预测分析。CRNN 网络模型通过利用 BLSTM 和 CTC 损失函数对文本图像中的上下文特征序列进行学习和预测,文本识别准确率得到提升。

CRNN 网络模型解决了无分割字符识别的不定长序列问题,但其存在一定的局限性,这是因为 CRNN 是结合 CNN 和 RNN 两个复杂的模型,导致模型体积很大,在进行较大数据集的训练时,较多的训练参数也会导致训练内存占比增大和识别准确率下降。

1.2 基于注意力双向长短期记忆网络

深度学习中的注意力机制类似一种模仿人脑的机制,通过模仿人类视觉和学习认知能力使神经

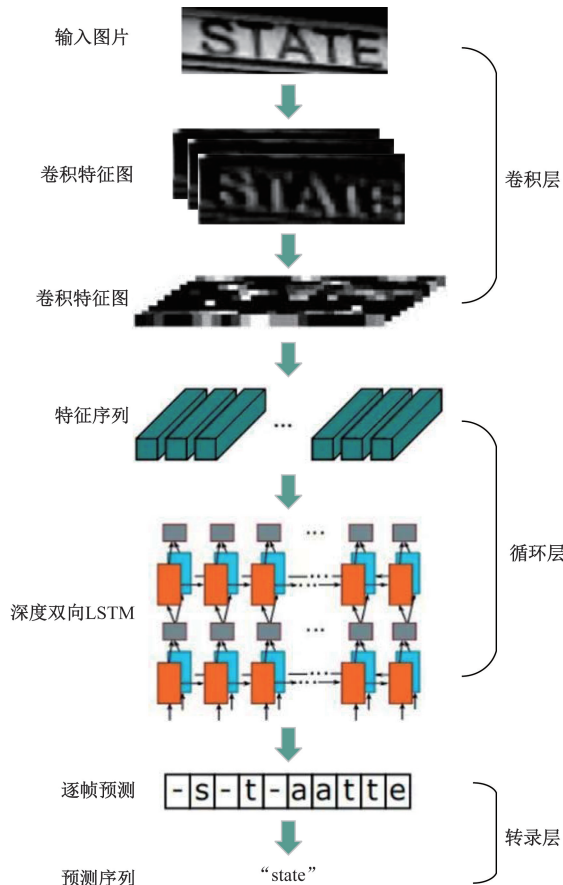


图1 CRNN 模型结构图

Fig. 1 CRNN model structure diagram

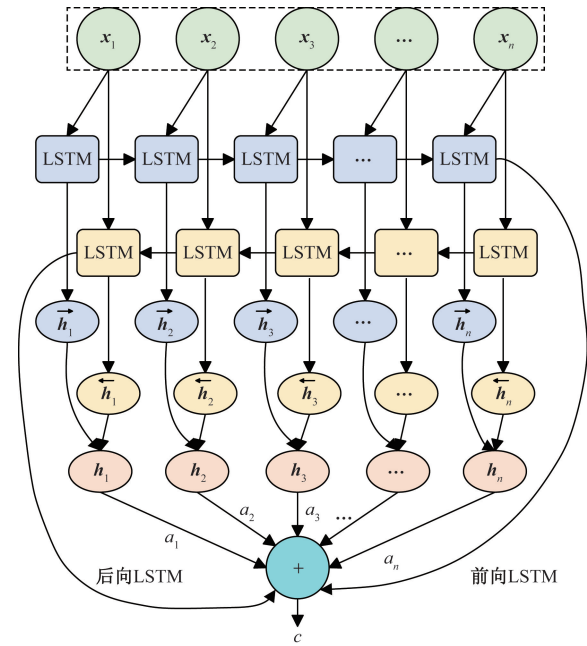
网络能对某一部分有用特征得到更多的关注,而忽略掉那些无用的特征。注意力机制被广泛应用于图像识别、机器翻译、语音识别、文本识别等任务中^[17]。注意力机制可以对输入序列的不同位置分配不同的权重,有利于模型处理每个序列元素时更专注于最相关的部分,从而优化模型的特征提取能力,使得神经网络的训练效率和文本识别率得到提高。

为了让 BLSTM 网络对有用特征更加敏感,更高效的学习相关特征,融入注意力机制对 BLSTM 网络的输入向量特征赋予不同的权重。对相关性大的特征信息赋予较大的权重,对相关性小的信息赋予较小的权重,从而使模型训练数据集可以更加高效、准确。提出一种基于注意力的双向长短期记忆(AT-BLSTM)网络模型,AT-BLSTM 模型结构如图2所示。

$$m_i = f(\mathbf{h}_i, \mathbf{h}_n) \quad (1)$$

$$a_i = \frac{e^{m_i}}{\sum_{i=1}^n e^{m_i}} \quad (2)$$

$$\mathbf{c} = \sum_{i=1}^n a_i \mathbf{h}_i \quad (3)$$



x_1, x_2, \dots, x_n 为 CNN 网络输出的特征向量; h_1, h_2, \dots, h_n 为 LSTM 神经网络每个时间步输出的隐藏层向量; a_1, a_2, \dots, a_n 为注意力机制的注意力权重; \rightarrow 为前向 LSTM 传播;

\leftarrow 为后向 LSTM 传播

图2 AT-BLSTM 模型结构

Fig. 2 AT-BLSTM model structure

式中:函数 f 为隐藏层 h_i 的学习函数; m_i 为每个时间步的隐藏层 h_i 的相似度得分; a_i 为每个隐藏层 h_i 的注意力权重; \mathbf{c} 为上下文信息向量。

1.3 知识蒸馏

知识蒸馏是由 Hinton 等^[13]提出的一种用于提高轻量化模型性能的方法,知识蒸馏模型结构如图3所示。它是将一个高性能的大模型(称为教师模型)获取的知识传输到一个小模型(称为学生模型)中,在确保模型准确性的同时,减小模型的大小,得到一个更小、更快的轻量级训练模型。Hinton 等^[13]在 Softmax 函数中引入蒸馏温度 T ,来增大各类别之间的区别。加入温度 T 的 Softmax 函数概率分布 q_i 如式(4)所示。温度参数的值越大,得到的概率分布越平滑。当 $T = 1$ 时,该函数为原始 Softmax 函数;当接近无穷大时,所有数据类都具有相同的概率分布。

$$q_i = \frac{\exp\left(\frac{z_i}{T}\right)}{\sum_j \exp\left(\frac{z_j}{T}\right)} \quad (4)$$

式(4)中: T 为温度; z_i, z_j 分别为教师模型和学生模型的 Softmax 输出。

知识蒸馏其实是一种模型压缩方法,通过使用学习的方法借助大型网络指导训练一个与其性能

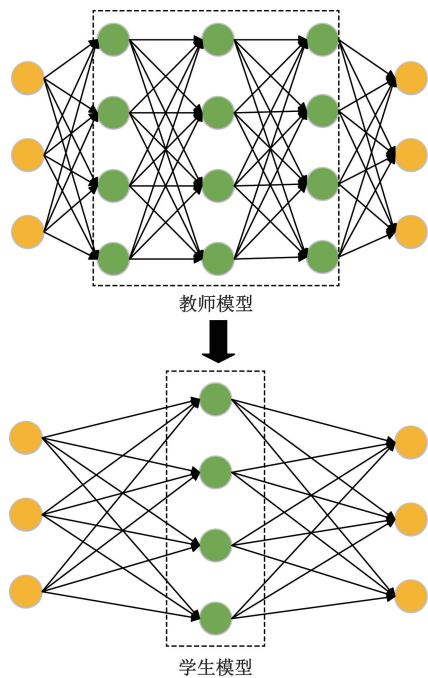


图3 知识蒸馏模型结构

Fig.3 Knowledge distillation model structure

相当的小型网络,从而间接达到压缩的目的^[18]。知识蒸馏具有很好的整体性特点,可以在不改变神经网络结构的前提下,使模型具有较好的可扩展性。

2 模型构建

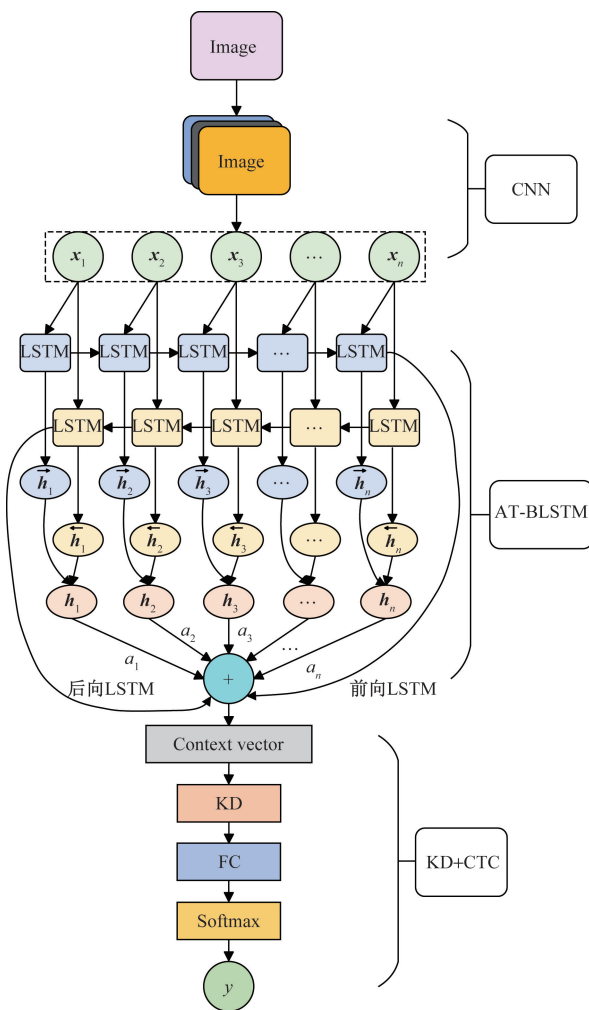
结合 CRNN 模型在进行大数据集训练时的模型体积和训练参数,提出融合注意力机制和知识蒸馏技术,从而减小识别模型体积和减少训练参数,提高文本识别准确率,得到一个更快、更准确地轻量级训练模型 AK-CRNN。

2.1 AK-CRNN 模型

AK-CRNN 模型主要由 CNN、AT-BLSTM、KD + CTC 组成,如图 4 所示。对于不定长文本序列,AK-CRNN模型先通过 CNN 模型进行特征提取,然后将提取的特征向量输入 AT-BLSTM 模型得到上下文信息,通过 KD 技术得到训练好的教师模型,然后将提取到的特征信息传输给学生模型,得到一个轻量级模型,通过全连接层和 Softmax 函数对每列特征进行概率分布,最后通过 CTC 损失函数对文本序列进行预测。

2.2 OpenCV 图像处理

OpenCV 是一种用于图像处理和计算机视觉的开源库,提供了许多强大的功能,其中最重要的功能是图像识别。OpenCV 可以实现许多不同类型的图像识别,这些图像识别技术可以用来识别人脸、文本、场景和其他图像特征。OpenCV 还具有强大的图像处理技术,它可以对图像进行透视变换、仿射变换、边缘检



Context vector 为上下文向量; FC 为全连接层; Softmax 为分类层; y 为输出结果; KD + CTC 为知识蒸馏和文本序列预测

图4 AK-CRNN 模型结构

Fig.4 AK-CRNN model structure

测、滤波去噪和形态学操作等功能^[19-20]。

由于数码设备拍摄、扫描仪扫描、网络截取等原因可能存在文本图片出现倾斜现象,可能会有稍微向左或向右侧倾斜的现象发生;拍摄光线的变化、采集设备的不足和光照不均匀等原因而出现噪声等问题,会对文本图片识别产生一定的干扰。因此,OpenCV 在文字识别中体现着重要作用。OpenCV 可以通过数字图像处理技术对文本图片进行去噪、检测、定位和透视变换等工作,把文字图像中的有用信息和背景噪声区分开,对倾斜文本图片进行透视变换使得图片文本校正,以便于后面的识别工作。

2.3 模型训练

采用 CASIA-HWDB2.0 数据集训练 AK-CRNN 教师模型,首先通过 CNN 对文本图像的特征进行提取,然后通过 AT-BLSTM 注意力模型将提取字符上下文特征序列的特征向量进行整合,通过知识蒸馏

技术得到训练好的教师模型,将教师模型的 Softmax 层输出作为软目标,通过改变温度参数得到教师模型更好的软目标,通过蒸馏将预训练好的教师网络中有用的知识传输给学生模型,得到一个轻量级学生模型,再通过全连接层和 Softmax 函数对每列特征进行概率分布,最后通过 CTC 损失函数对文本序列进行预测。通过以上步骤得到一个性能更优的轻量级网络模型,模型构建完成。

待识别汉字图片通过 OpenCV 进行预处理,预处理工作主要步骤包括灰度图处理、二值化、滤波去噪、图片缩放和透视变换等操作,通过预处理消除文字图片中无关的背景信息,增强文字信息的可检测性,使得识别模型更好地读取和识别图片中的文字信息,使得模型识别效果更好。

通过构建好的模型进行手写汉字识别工作。首先,通过 OpenCV 对待识别图片进行图像处理,经过文本定位、透视变换和去噪工作;然后进行 AK-CRNN 模型进行汉字识别;最后得到识别结果。如图 5 所示。

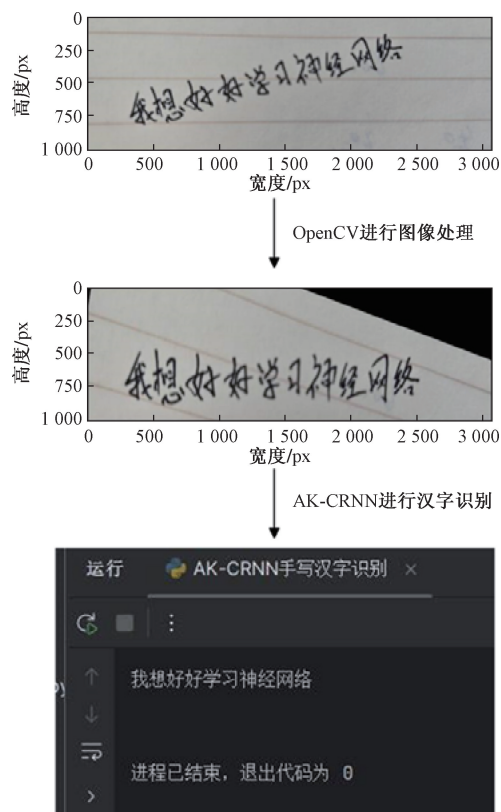


图5 AK-CRNN 网络模型的手写汉字识别

Fig. 5 Handwritten Chinese character recognition based on AK-CRNN network model

3 实验结果与分析

3.1 实验环境与数据集

本次实验仿真都是在以下操作环境进行的:软

件条件为 Python 3.8.5, OpenCV-Python 4.8.1.78, 深度学习框架为 Pytorch 框架,编辑器为 PyCharm 2023 版本;硬件条件为 CPU Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz 2.59, RAM 2 × 8 G, GPU NVIDIA GeForce GTX 1650, 系统为 Windows 10。

本实验采用 CASIA-HWDB2.0 数据集进行一系列实验研究。该数据集由 420 人参与并书写连续文本,总共 5 091 页图像,分割为 52 230 个文本行和 1 349 414 个文字,由飞浆平台公开提供。由于原始数据集文本行宽度像素点不统一,如图 6 所示。为了方便后续数据集的训练工作,将数据集进行整理,得到宽度统一的行文本图片,数据集所有图片都被设置为宽度为 32 个像素点的不定长文本,并按照训练集 80%、验证集 20% 的占比进行划分,处理后部分文本图片如图 7 所示。

3.2 超参数优化

在神经网络学习中,超参数的选择会对实验结果产生一定影响。本实验通过单一变量因素法,选取不同超参数进行优化。经过多次对比试验,学习率设置为 0.001, dropout 率为 0.5, 采用 Adam 优化器,动量初始化为 0.9, 权重衰减系数设为 0.005, batch size 设置为 64, epoch 设置为 100, 神经网络学习效果最佳。

3.3 对比实验

为了验证本方法的有效性,需要通过设置对比实验对模型进行对比分析,实验结果一般采用准确率 (precision, 记为 P)、召回率 (recall, 记为 R) 及测度值 F (F -measure, 记为 F) 3 个评估指标对模型性能进行评估^[21]。3 个指标的表达式分别为

$$P = \frac{TP}{TP + FP} \times 100\% \quad (5)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (6)$$

$$F = \frac{2PR}{P + R} \times 100\% \quad (7)$$

式中: TP 为正确文本预测为正确类别数目; FP 为错误文本预测为正确类数目; FN 为正确文本预测为错误类数目; F 为测度值。

通过上述超参数优化对 AK-CRNN 网络模型进行训练,得到训练结果如下, AK-CRNN 网络模型在 100 次迭代实验中,第 30 次迭代之后识别准确率趋于稳定,最后识别准确率在 97.9%。通过对比 CRNN 原始模型和分别加入 AT-BLSTM 模型、KD 模型以及 AK-CRNN 模型后的数据,分析得出模型性能。各个模型的识别准确率对比实验如图 8 所示,各个模型性能评估对比结果如表 1 所示,各个模型的识别准确率和训练参数量如表 2 所示。

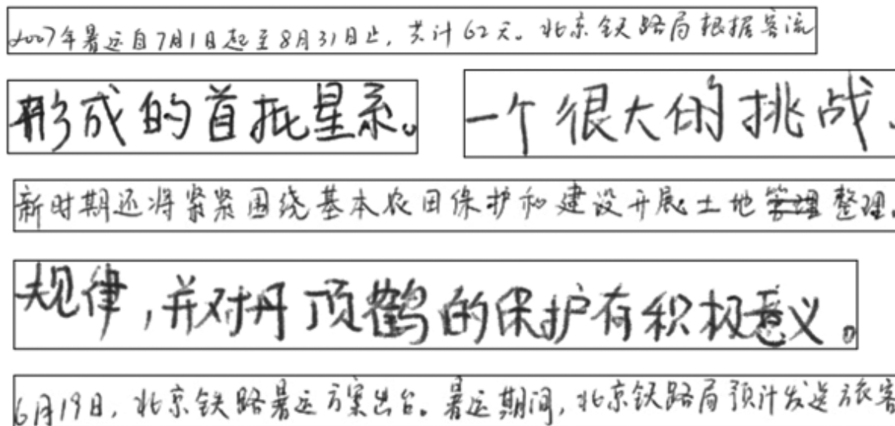


图6 CASIA-HWDB2.0 原始数据集部分文本图片
Fig.6 CASIA-HWDB2.0 raw Data set part text picture

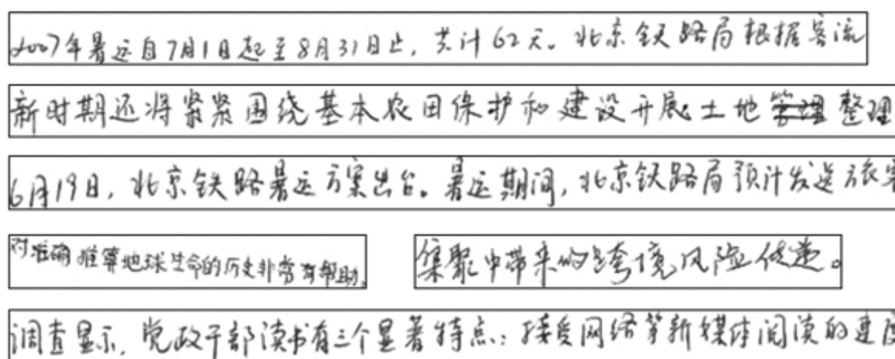


图7 CASIA-HWDB2.0 处理后数据集部分文本图片
Fig.7 CASIA-HWDB2.0 post-processing Data set part text picture

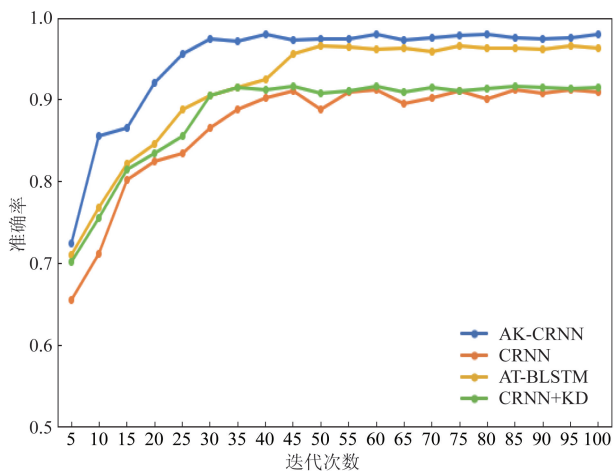


图8 各个模型的识别准确率变化趋势
Fig.8 Change trend of recognition accuracy of each model

表1 各个模型性能评估对比结果
Table 1 Comparison results of performance evaluation of each model

方法	准确率 P/%	召回率 R/%	测度值 F/%
CRNN	91.2	90.3	90.7
AT-BLSTM	96.5	96.8	96.6
CRNN + KD	91.6	90.7	91.1
AK-CRNN	97.9	98.3	98.1

表2 各个模型的识别准确率和训练参数量变化
Table 2 The recognition accuracy of each model and the change of the number of training parameters

方法	准确率 P/%	训练参数/M
CRNN	91.2	47.29
AT-BLSTM	96.5	49.79
CRNN + KD	91.6	30.47
AK-CRNN	97.9	31.35

通过上述实验分析,所提出的模型在评估指标中取得了优异的效果,相较于原始模型的准确率、召回率和测度值均有所提高;加入 AT-BLSTM 网络比原始 CRNN 网络模型的识别准确率提高了 5.3%,但是训练参数增加了 2.5 M;通过加入 KD 技术比原始 CRNN 网络模型的训练参数减少了 16.82 M,但是识别准确率仅仅提高了 0.4%;AK-CRNN 网络模型对比原始 CRNN 网络模型,在训练参数减少 15.94 M 的前提下,识别准确率也提高了 6.7%,比 Peng 等^[22]提出的一种新的基于分割的方法来识别手写中文文本的准确率 97.7% 高出 0.2%。

4 结论

提出一种 AK-CRNN 网络模型的手写汉字识别

方法。该方法基于 AT-BLSTM 网络和 KD 技术有效解决了 CRNN 模型训练参数大和内存占比大的问题,提高了模型训练效率和文本识别率,得到一个体积更小、训练更快的轻量级网络模型。该方法通过实验表现出的优异性能清楚地证明该方法的有效性。通过引入 AT-BLSTM 网络使神经网络能对某一部分有用特征更加敏感,从汉字图片中提取重要特征,有效提高模型最终的训练效率和文本识别率;利用 KD 将教师模型中学到的知识迁移到小网络中,大大减少了模型训练参数的数量,压缩了模型体积,得到一个性能更优的轻量级网络模型。最终,汉字识别准确率提高了 6.7%,训练参数减少了 15.94 M, AK-CRNN 网络模型识别准确率达到 97.9%。

综上所述,注意力机制有效提高了模型的训练效率和文本识别率,虽然训练参数略有增加,但通过引入知识蒸馏技术得到了参数更少的小模型网络,减少了模型训练参数,压缩了模型体积,得到了一个准确率更高、模型体积更小、性能更优的轻量级网络模型。论证了该方法的可行性和正确性。

参 考 文 献

- [1] Chen L, Wang S, Fan W, et al. Beyond human recognition: a CNN-based framework for handwritten character recognition [C]//2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). Kuala Lumpur: IEEE, 2015: 695-699.
- [2] 黄洋. 基于深度学习的脱机手写汉字识别技术研究[D]. 重庆: 重庆邮电大学, 2019.
Hang Yang. The research of offline handwritten Chinese character recognition based on deep learning [D]. Chongqing: Chongqing University of Posts and Telecommunications, 2019.
- [3] 袁柱. 基于深度学习的脱机手写汉字识别的研究与应用[D]. 广州: 广东工业大学, 2020.
Yuan Zhu. Research and application of offline handwritten Chinese character recognition based on deep learning [D]. Guangzhou: Guangdong University of Technology, 2020.
- [4] 王永强. 基于深度学习的脱机手写体汉字识别研究[D]. 昆明: 云南大学, 2019.
Wang Yongqiang. Research on offline handwritten Chinese character recognition based on deep learning [D]. Kunming: Yunnan University, 2019.
- [5] Wang Z R, Du J, Wang J M. Writer-aware CNN for parsimonious HMM-based offline handwritten Chinese text recognition [J]. Pattern Recognition, 2020, 100: 107102.
- [6] 王芳元, 王建平. 仿反馈机制的智能识别方法研究与应用[J]. 合肥工业大学学报(自然科学版), 2018, 41(3): 333-341.
Wang Fangyuan, Wang Jianping. Research and application of the method of intelligent identification based on simulated feedback [J]. Journal of Hefei University of Technology (Natural Science), 2018, 41(3): 333-341.
- [7] Wang Y, Xiao W, Li S. Offline handwritten text recognition using deep learning: a review[J]. Journal of Physics: Conference Series, 2021, 1848(1): 012015.
- [8] Su T H, Zhang T W, Guan D J, et al. Off-line recognition of realistic Chinese handwriting using segmentation-free strategy [J]. Pattern Recognition, 2009, 42(1): 167-182.
- [9] Du J, Wang Z R, Zhai J F, et al. Deep neural network based hidden Markov model for offline handwritten Chinese text recognition [C]//2016 23rd International Conference on Pattern Recognition (ICPR). Cancun: IEEE, 2016: 3428-3433.
- [10] Wang Z R, Du J, Wang W C, et al. A comprehensive study of hybrid neural network hidden Markov model for offline handwritten Chinese text recognition [J]. International Journal on Document Analysis and Recognition (IJ DAR), 2018, 21(4): 241-251.
- [11] Shi B, Bai X, Yao C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(11): 2298-2304.
- [12] 周旋. 卷积循环神经网络在场景文字识别上的应用的研究[D]. 上海: 上海第二工业大学, 2021.
Zhou Xuan. Research on the Application of convolution recurrent neural network in scene text recognition [D]. Shanghai: Shanghai Polytechnic University, 2021.
- [13] Hinton E G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. Computer Science, 2015, 14(7): 38-39.
- [14] Shi B, Wang X, Lü P, et al. Robust scene text recognition with automatic rectification [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 4168-4176.
- [15] 杨琼. 基于注意力学习网络的古籍汉字图像检索模型[D]. 保定: 河北大学, 2021.
Yang Qiong. Ancient Chinese character image retrieval model based on attention learning network [D]. Baoding: Hebei University, 2021.
- [16] 王天玮. 基于注意力机制的文字识别技术研究[D]. 广州: 华南理工大学, 2021.
Wang Tianwei. Research on attention-based character recognition [D]. Guangzhou: South China University of Technology, 2021.
- [17] 张洋, 江铭虎. 基于注意力机制的文本作者识别[J]. 计算机应用, 2021, 41(7): 1897-1901.
Zhang Yang, Jiang Minghu. Authorship identification of text based on attention mechanism [J]. Journal of Computer Applications, 2021, 41(7): 1897-1901.
- [18] 司兆峰, 齐洪钢. 知识蒸馏方法研究与应用综述[J]. 中国图象图形学报, 2023, 28(9): 2817-2832.
Si Zhaofeng, Qi Honggang. Survey on knowledge distillation and its application [J]. Journal of Image and Graphics, 2023, 28(9): 2817-2832.
- [19] 管焱然, 管有庆. 基于 OpenCV 的仿射变换研究与应用[J]. 计算机技术与发展, 2016, 26(12): 58-63.
Guan Yanran, Guan Youqing. Research and application of affine transformation based on OpenCV [J]. Computer Technology and Development, 2016, 26(12): 58-63.

- [20] 孙丹阳, 邵坤艳, 孙春志, 等. 基于图像处理的文档图像矫正系统研究与设计[J]. 电脑与信息技术, 2023, 31(3): 20-23.
Sun Danyang, Shao Kunyan, Sun Chunzhi, et al. Research and design of document image rectification system based on image processing[J]. Computer and Information Technology, 2023, 31(3): 20-23.
- [21] 李超凡, 马凯. 基于注意力机制结合 CNN-BiLSTM 模型的电子病历文本分类[J]. 科学技术与工程, 2022, 22(6): 2363-2370.
Li Chaofan, Ma Kai. Electronic medical record text classification based on attention mechanism combined with CNN-BiLSTM[J]. Science Technology and Engineering, 2022, 22(6): 2363-2370.
- [22] Peng D, Jin L, Ma W, et al. Recognition of handwritten Chinese text by segmentation; a segment-annotation-free approach [J]. IEEE Transactions on Multimedia, 2022, 44(9): DOI:10.1002/(SICI)1097-4571.