

引用格式: YAN Ruyu, WANG Xiaoxia, XI Jiangtao, et al. Handwritten Font Classification Method Based on Ghost Imaging [J]. Acta Photonica Sinica, 2022, 51(11):1111001

闫茹钰,王肖霞,习江涛,等.基于鬼成像的手写字体分类方法[J].光子学报,2022,51(11):1111001

基于鬼成像的手写字体分类方法

闫茹钰¹,王肖霞¹,习江涛^{1,2},杨风暴¹,包达尔罕³

(1 中北大学 信息与通信工程学院,太原 030051)

(2 伍伦贡大学 电气计算机与通信工程学院,澳大利亚 伍伦贡 NSW2522)

(3 西安微电子技术研究所,西安 710054)

摘 要:针对手写字体特征提取难、识别效率低甚至错误的问题,提出一种结合鬼成像原理和分类网络的思想用于手写字体的识别。首先,使用激光器、数字微镜阵列、单像素探测器等搭建鬼成像探测系统,利用系列哈达玛散斑场分别实现对 17 239 幅手写字体的照射;其次,利用单像素探测器对手写字体透射的总光强值进行采集,并将其转换为二维向量作为网络训练的输入;最后基于卷积神经网络在图像分类的优势搭建网络架构且为解决训练过程中的网络退化问题加入残差块。实验结果表明:对于手写数字来说,全连接神经网络与卷积神经网络的精密度、召回率、F1 值分别提高 86.50%/97.25%、86.40%/98.03%、86.31%/97.60%;对于手写字母来说,卷积神经网络在全采样下的精密度、召回率、F1 值分别为 91.87%、90%、90.23%。该方法利用鬼成像原理,仅通过手写字体透射的总光强值而无需对手写字体的特征进行提取与识别便可进行快速分类,大幅提升了手写字体的识别效率。

关键词:鬼成像;手写字体识别;卷积神经网络;准确率;残差

中图分类号:O431.2

文献标识码:A

doi:10.3788/gzxb20225111.1111001

0 引言

手写字体自动识别作为计算机视觉领域中一种高效的判别方法,可在大幅减少人力的情况下提升字体识别效率,在试卷判别、财务报表等涉及大规模手写字体的自动识别中具有重要的实用价值。实际上,由于个人字体风格迥异、书写习惯不同等常常导致不同手写字体存在较大差异,而不像印刷字体有统一的判别标准。因此,在利用自动识别技术进行字体识别时,常常出现判别效率低甚至失效的问题,严重制约了其实用化进程。

近年来,通常利用机器学习、深度学习等对手写字体的字体特征进行提取与匹配,进而基于分类的思想对字体进行识别。例如:文献[1]针对手写数字在人为判别过程中存在识别效果差的问题,提出一种K-最近邻算法(K-Nearest Neighbor, KNN),用于根据投票法则训练手写数字特征,该方法将未知样本和 k 个手写相邻样本中所属类别占比较多的归为一类,从而对MNIST数字图像进行分类,由于采用惰性算法,计算简单;文献[2]开发了一种CNN和SVM相结合的混合模型,该方法将CNN作为特征自动提取器,SVM用于二元分类,进而识别MNIST数据集中的数字,由于CNN的感受野易自动提取到最可区分的特征,因此该混合模型有着较高的准确率;文献[3]针对传统机器学习中模型泛化能力差的问题,基于深度神经网络,依据BP算法在训练过程中传递误差值进行权重值和偏置项的值的更新,最终使得输出手写数字的分类结果越来越接近预期值。该类方法虽然一定程度上解决了人工判读所造成的识别效率低、甚至误判的问题,但由

基金项目:国家自然科学基金(No. 61972363),内燃机可靠性国家重点实验室开放基金资助(No. skler-202011),山西省应用基础研究计划青年科技研究基金(No. 201901D211233),山西省回国留学人员科研资助项目(No. 2021-110)

第一作者:闫茹钰(1996—),女,硕士研究生,主要研究方向为鬼成像。Email: 345313872@qq.com

导师(通讯作者):王肖霞(1980—),女,副教授,博士,主要研究方向为关联成像技术、不确定性信息处理等。Email: wangxiaoxia@nuc.edu.cn

习江涛(1963—),男,教授,博士,主要研究方向为信号处理、光电信号处理、无线通信等。Email: 20180143@nuc.edu.cn

收稿日期:2022-02-06;录用日期:2022-03-14

<http://www.photon.ac.cn>

于手写字体在提取特征时常存在边缘、纹理等细节信息的缺失,导致识别中易出现分类结果的错误。此外,在字体识别时均需要对获取的字体图像进行如平滑、去噪等预处理。然而在进行这些处理时易造成图像细节信息丢失,影响识别效率。

鬼成像^[4-8]作为一种反直觉间接成像方式,无需事先获取物体的细节信息,可在不直视目标物体的情况下实现对目标物体的快速成像,将其与手写字体自动识别技术相结合,可在不提取特征的前提下,实现对手写字体的识别,有效避免了由于特征细节信息的缺失导致的分类结果不准确或错误的出现。鬼成像在该领域的优势已引起了研究学者的关注,如文献[9-10]分别以MNIST、Fashion数据集为例,将鬼成像与深度神经网络相结合提出了一种不需要提取目标物体特征信息的方法从而提升了识别效率。由于该类方法输入量为一维向量,相比二维图像作为神经网络的输入可以减少对目标物体的探测数量从而缩短训练时间。这为特征细节信息缺失导致的低识别效率提供了有效的解决途径。

同时,基于鬼成像的分类方法输入均为系列桶探测器的值转换成的一维向量。该方法的优势在于:桶探测器的值易探测且探测速度快;以少于手写字体像素值(32×32)的一维向量代替二维图像作为输入时,可简化设计的网络结构,大大降低识别的复杂性;直接利用一维向量对手写字体进行分类识别可跳出识别物体就是识别图像的既存思路,为无成像分类的实现提供可能。然而,这些方法都是基于全连接网络结构所提出的,易导致参数膨胀使得在训练过程中出现过拟合现象,进而直接导致手写字体分类错误。可见,如何避免网络的过拟合现象、提升模型的泛化能力是提升手写字体识别效率的关键。与全连接神经网络相比,卷积神经网络(Convolutional Neural Network, CNN)由于卷积核可参数共享,大大减少网络权重的参数大小,从而有效解决网络在训练过程中的过拟合现象,并且该网络具有更强的自适应性,在识别手写字体时准确率更高。因此,将鬼成像与CNN结合来提升手写字体的识别效率、避免分类错误具有现实意义。

鉴于此,本文将鬼成像原理与CNN结合提出一种解决网络过拟合现象的手写字体自动识别方法。该方法以桶探测器值为输入量,通过数据向量化将残差块^[11]加入网络中,利用跳跃连接实现输入的一维向量绕道传输到输出,通过保护输入信息的完整性提高识别效率。

1 结合鬼成像和CNN的手写字体识别

1.1 本文方法

手写字体自动识别方法实现流程如图1所示,其中图1(a)为实现原理流程,图1(b)为实验室开发的探测实验装置。其中,探测系统主要由10uj-5 KHz-532 nm脉冲激光器、像元尺寸为 $13.67 \mu\text{m}$ 且分辨率为 1027×768 的数字微镜阵列(Digital Mirror Device, DMD)、频带宽度直流(Direct Current, DC):DC-11.0 MHz的单

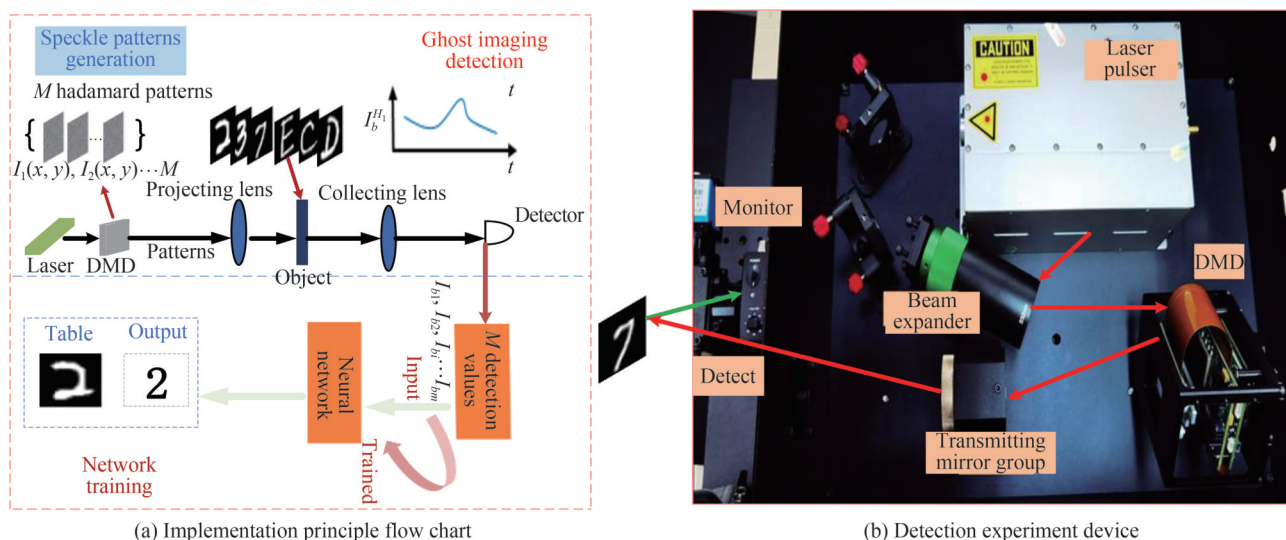


图1 手写字体自动识别方法实现流程

Fig. 1 The realization process of automatic recognition method of handwritten digit

像素探测器、光学镜头等组成。DMD作为鬼成像系统的光源调制模块,是产生散斑的核心部件。经DMD调制后产生的光场为哈达玛散斑场。鬼成像作为一种特殊的成像方式,不需要提前知道目标物体的信息,通过照射散斑就可以对目标物体进行成像。而目标物体是被散斑照射的物体,并不是生成的。本文的目标物体是手写图像,选择公开的MNIST、EMNIST数据集。以 32×32 的手写字体作为被照物体,在全采样时(探测次数为1 024次),单像素探测时间约为0.75 s,速度很快并不会对后续的分类识别造成时间上的大量消耗。

该方法主要由鬼成像探测和网络训练两部分组成。基本实现过程:利用532 nm的激光器照射像元尺寸为 $13.67 \mu\text{m}$ 且分辨率为 $1\ 027 \times 768$ 的DMD设备来产生系列正交的哈达玛(Hadamard)散斑^[12-14],目的是为了减少散斑间的冗余,并利用频带宽度DC-11.0 MHz的单像素探测器来获得物体透射的总光强值。设计的CNN网络由输入层、卷积块、残差块、全连接层、Dropout层和输出层6部分组成。将桶探测器的测量值作为输入并基于CNN网络进行学习,进而给出手写字体的分类结果。具体步骤为:

步骤1:利用532 nm的激光器产生赝热光源,经过DMD调制器来获得不同时刻下照射目标物体且分辨率为 32×32 的Hadamard散斑图案,即 t 时刻时Hadamard散斑图案可表示为 $I_a(x, y; t_i)$,其中 t_i 为不同的时刻, $i = 1, 2, \dots, M$, i 表示测量次数。

步骤2:选用公开手写数字数据集MNIST中包含的0~9十类手写数字和公开手写字母数据集EMNIST^[15]中包含的字母为被照物体,将生成的系列Hadamard散斑图案分别照射某一手写图像并以桶探测器的值来测量被照物体所透射的总光强值。重复上述过程 m 次以获得 m 幅手写图像所对应的桶探测器的值。

步骤3:将探测后的系列桶探测器的值转换为一维向量 I_b 该一维向量可表示为 $I_b = [I_{b1}, I_{b2}, I_{b3}, \dots, I_{bm}]$, $m = 1, 2, \dots, M$ 。并对转换后的一维向量进行归一化处理,采用最小最大规范的方法(min-max normalization)可表示为

$$I'_b = \frac{I_b - \min \{I_b\}}{\max \{I_b\} - \min \{I_b\}} \quad (1)$$

使其映射到0~1范围之内从而加快网络的收敛性。

步骤4:将步骤3获得的一维向量作为CNN网络的输入端进行训练。本方法的卷积神经网络用来连接输入与输出端,其主要由卷积块、残差块、和全连接层3部分组成。

步骤5:利用训练好的CNN分类模型对步骤2中提到的两个数据集中部分手写字体进行测试,最终输出10个神经元对应的10类手写字体。

1.2 基于CNN神经网络架构设计

CNN一般由卷积层、池化层、全连接层构成,可解决全连接网络结构的参数膨胀导致的过拟合问题,并广泛应用于图像分类领域。为解决网络层数增加导致训练时出现梯度消失、梯度爆炸及过拟合现象,将残差块^[16]、Dropout层^[17]加入其中。

设计的CNN网络模型架构如图2所示,其分类模型由输入层、卷积块、残差块、全连接层、Dropout层和输出层6部分组成。其中,输入部分为长度为 M 的归一化桶探测器的值;输出为模型预测的手写字体分类结果。连接输入和输出端的中间模块,其主要由卷积块、残差块、全连接层及Dropout层4部分组成。

卷积块由卷积层、批量归一化层(Batch Normalization, BN)^[18]和修正线性单元(Rectified Linear, Relu)^[19]激活层组成。为了保证卷积前后尺寸不变、padding能够平均分配到卷积张量两边,卷积核的个数一般都设置为奇数,因此,在一维、二维卷积层分别选择卷积核大小为 5×5 、 3×3 的滤波器。另外,本文CNN网络架构的输入为系列光强值,为能更好地处理该系列光强值选择一维卷积层。同时,BN层能优化对权重的调整从而加快网络的训练和收敛速度、控制梯度爆炸防止梯度消失。

基于卷积层的残差学习,其中的skip机制可将输入信息由浅层传递到网络的更深层,避免网络层数增加造成大量手写字体信息丢失及梯度消失的问题,也可提高网络的泛化能力从而提升手写字体的识别准确性。

为实现手写字体的多分类任务,在输出部分连接了三层全连接层,作为CNN结构中的多分类器。

Dropout层也可有效防止网络在训练过程中出现的过拟合问题,提高网络的泛化能力。经过多次实验

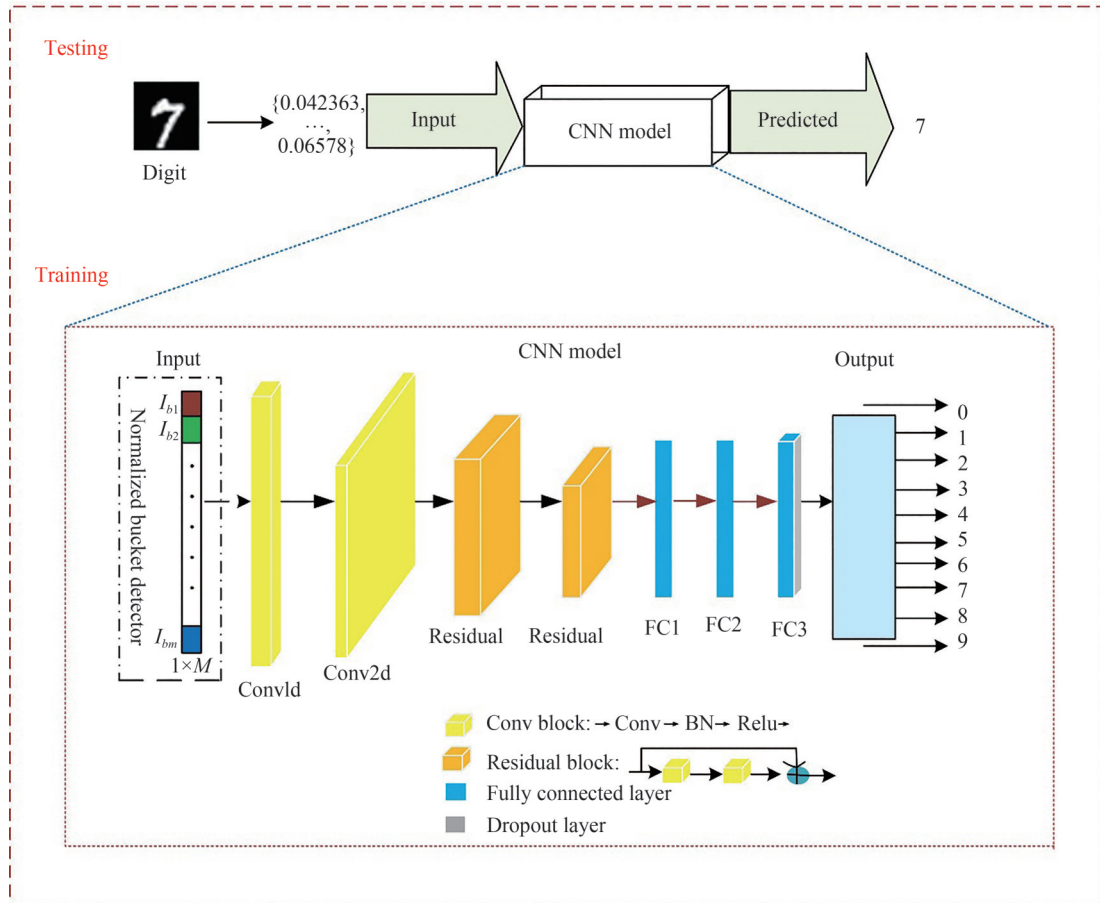


图2 CNN神经网络架构模型
Fig. 2 CNN neural network architecture model

验证,本文选择的Dropout的比例为0.6。

设计的框架使用操作系统为Win10(64位),编程语言python版本3.8(64位),基于PyTorch框架实现了CNN神经网络分类模型。

1.3 数据集准备

鬼成像无需事先获取目标物体的信息,通过散斑照射到目标物体进行成像。被散斑照射的物体就是目标物体。

本文的数据集是从公开的MNIST、EMNIST数据集中分别随机选取其中的17 239幅、1 442幅手写字体图片经过数据扩充 28×28 到 32×32 ,然后再经过1.1节的探测方法得到。MNIST是一个公开的手写数字的图片数据集,内容范围为0~9十类手写数字,由60 000个训练样本和10 000个测试样本组成,每个样本都是一张 28×28 像素的灰度手写数字图片,且EMNIST为MNIST的扩展数据集,其由26类字母组成。

1.4 评价指标的选取

为了验证和分析所提方法的有效性和合理性,以设计的CNN模型对手写字体部分数据集进行测试。并用准确率 A 、召回率 R 、精密度 P 、 F_1 值作为评价指标,即

$$\begin{cases} A = \frac{TP + TN}{TP + FN + TN + FP} \\ R = \frac{TP}{TP + FN} \\ P = \frac{TP}{TP + FP} \\ F_1 = \frac{2 \times P \times R}{P + R} \end{cases} \quad (2)$$

根据分类器在测试数据集上预测的正确或不正确可以分为4种情况^[20],其中,真正(True Positive, TP)表示被模型预测为正的正样本的数量;假正(False Positive, FP)表示被模型预测为正的负样本的数量;假负(False Negative, FN)表示被模型预测为负的正样本的数量;真负(True Negative, TN)表示被模型预测为负的负样本的数量。

2 实例验证与结果分析

2.1 基于 CNN 框架下的实验

实验部分是以图 1(b)所给出的鬼成像探测系统来获得测量值。分别从 MNIST、EMNIST 数据集中随机选取共计 17 239 幅手写字体图像作为该系统中的被照物体,并利用图 1(b)探测装置中的 DMD 生成 Hadamard 散斑,分别照射目标物体 1 024 次(非欠采样)、256 次(欠采样)。为了说明该方法的实用性,分别对非欠采样、欠采样条件下的模型进行训练。由于本文方法采用监督学习,因此,将桶探测器的值及对应的真实手写字体图片作为标签输入到网络中进行训练。

实验结果选取的有效训练周期为 100,以交叉熵损失函数来计算实际输出值与真实值之间的误差。计算原理表示为

$$L = \text{Loss}(\omega, b, x, y) = - \sum y_j \ln(\text{softmax}(j)) \quad (2)$$

式中, Loss 表示损失函数,参数 ω 、 b 分别表示各层线性关系中的权重和偏差, x 是维数为 M 的输入一维向量, y 对应手写字体的分类输出结果, softmax 为输出层的函数, j 是神经元的个数。为了对损失函数进行优化,利用自适应矩(Adaptive Moment Estimation, ADAM)优化器^[21],通过使用参数的梯度值对权重和偏置进行更新。同时,学习率作为训练过程中重要的超参数,过小或过大都会对收敛速度和效果造成影响。在经过多次实验后,此模型选择的初始学习率为 10^{-4} 。

为了验证所构建模型的合理性和有效性,以 1 442 幅手写字体为例分别在全采样和欠采样的情况下将对应的桶探测器的值作为测试集并对训练结束的模型进行验证。

2.2 实验结果及对比分析

本文中的全采样为 1 024 (32×32) 次探测,同样地,0.25 的采样率代表 256 ($32 \times 32 \times 0.25$) 次探测。为验证该方法的有效性,以 MNIST 数据集(随机选取该数据集的 12 039 个训练集,1 242 个测试集)的 0~9 十类数字为目标物体进行实验来说明不同采样率下 CNN、DNN 两种模型准确率分析的结果。图 3(a)和(b)分别为采样率 1.0、采样率 0.25 下的 0~9 每类数字的准确率。

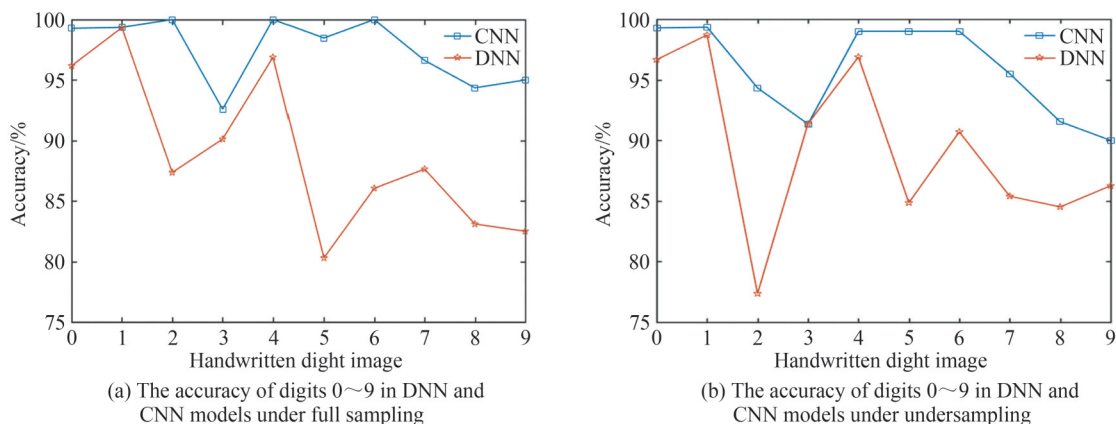


图 3 不同采样率 CNN、DNN 模型下数字 0~9 的准确率

Fig. 3 The accuracy of digit from 0 to 9 under different sampling rates of CNN and DNN models

观察图 3 可知,在相同采样率下,CNN 模型对 0~9 手写数字识别的准确率均高于 DNN 模型。分别测试了全采样和 0.25 采样率下 DNN、CNN 模型中 10 类数字的准确率,其中全采样时 DNN、CNN 模型中 0~9

每类数字的准确率为 96.18%/99.28%、99.34%/99.36%、87.36%/99.99%、90.12%/92.59%、96.88%/99.99%、80.30%/98.48%、86.05%/99.98%、87.64%/96.62%、83.10%/94.36%、82.50%/95%；采样率为 0.25 时两种模型下每类数字的准确率为 96.66%/99.28%、98.7%/99.34%、77.36%/94.33%、91.36%/91.35%、96.88%/99%、84.85%/99%、90.70%/99%、85.39%/95.50%、84.51%/91.54%、86.25%/90%。其中,数字 2、5、6、7、8、9 在全采样和欠采样两种情况下 CNN 模型比 DNN 模型有着显著的提升,全采样时分别提升 12.63%、18.18%、13.93%、8.98%、11.26%、12.5%；欠采样时分别提升 16.97%、14.15%、8.3%、10.11%、7.03%、3.75%。由此可见随着采样率增加识别的准确率也随之提升。因此,定量分析结果与定性分析结果一致。

同时,为了减小实验误差,进一步说明 CNN 模型在所提方法下的优势,记录了两个模型的最低、最高、平均准确率如表 1 所示。

表 1 基于 CNN、DNN 模型准确率比较
Table 1 Comparison of accuracy based on CNN and DNN models

Model	Lowest accuracy	Highest accuracy	Average accuracy
CNN	95.65%	98.47%	97.65%
DNN	87.56%	92.99%	92.34%

由表 1 可见,CNN 模型的最低、最高、平均准确率均高于 DNN 模型下的,且分别提升了 8.09%、5.48%、5.31%,验证了 CNN 模型在该方法下的合理性和有效性。

为了更好地比较两种模型的性能,根据混淆矩阵,在测试部分测试了两种模型的精密度、召回率、 F_1 值,如表 2 所示。可见,相比 DNN 模型,CNN 模型的精密度、召回率、 F_1 值分别提升了 10.75%、11.63%、11.09%,进一步验证了利用桶探测器的值作为网络输入的方法搭建 CNN 模型的合理性。

表 2 DNN、CNN 模型下的精密度、召回率、 F_1 值对比分析
Table 2 Comparative analysis of precision, recall rate and F_1 score under DNN and CNN models

Parameter	DNN	CNN	Lift value
Precision	86.50%	97.25%	10.75%
Recall rate	86.40%	98.03%	11.63%
F_1 score	86.31%	97.60%	11.09%

从表 2 的实验结果可以看出,基于 CNN 模型的手写数字分类模型具有更好的测试稳定性,表明构建的模型显著提高了分类性能。

为了验证所搭建 CNN 模型对不同数据集都有一定的学习效果,以 EMNIST 数据集为例随机选取其中 l、v、y、z、m、n、o、r、s、h 十类(本文所搭建模型为十分类网络)字母图像作为目标物体分别在采样率为 1.0、0.25 的情况下对其所对应的桶探测器的值进行训练和测试。采样率 1.0/0.25 下每类字母对应的准确率的测试结果如图 4(a)、(b)所示。CNN 模型在不同采样率下的精密度、召回率、 F_1 值如图 4(c)、(d)所示。

由图 4(a)、(b)可知,全采样和欠采样情况时 l、v、y、z、m、n、o、r、s、h 十类手写字母在本文提出的 CNN 分类模型下的准确率分别为 90%/60%、90%/85%、100%/70%、90%/60%、95%/95%、70%/75%、85%/90%、95%/85%、100%/70%、85%/80%。除了字母 n、o 以外,其他字母在全采样时均高于欠采样时的准确率,对比分析说明随着采样率的增加准确率也随之提升。实验结果图 4(c)表明在全采样时精密度、召回率、 F_1 值分别为 91.87%、90%、90.23%。说明 CNN 在以字母为目标物体时有较好的分类结果,进一步提高了网络的通用性。

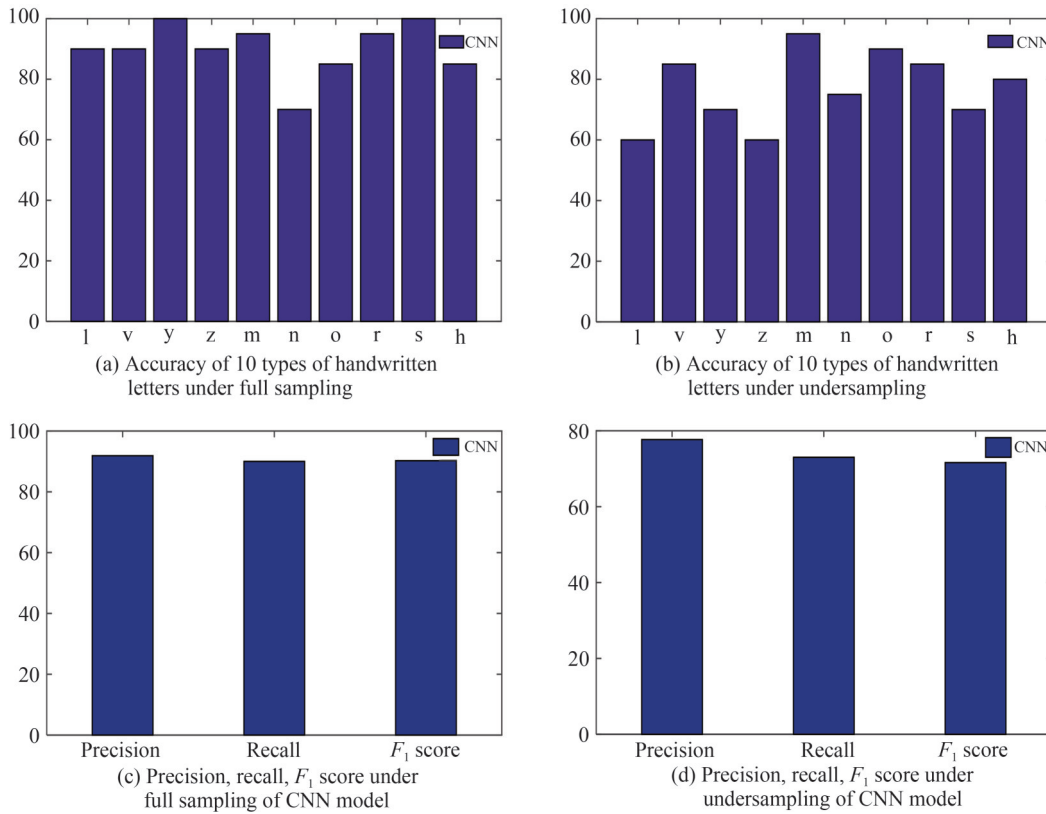


图4 不同采样率下10类字母的准确率及CNN模型精密度、召回率、 F_1 值

Fig. 4 Accuracy rate of 10 letters under different sampling rates and CNN model precision, recall rate, F_1 score

2.4 损失函数分析

研究中发现,在统一的损失函数下,CNN比DNN有更强的收敛性。为了说明这一问题,以交叉熵损失函数为例对CNN、DNN模型下的收敛速度进行对比分析,训练误差曲线如图5所示。

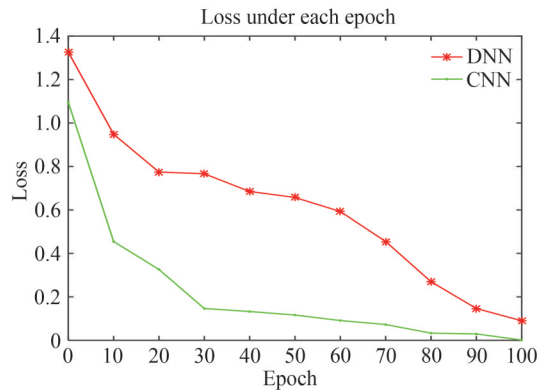


图5 不同模型采用交叉熵损失函数下的损失误差曲线

Fig. 5 The loss error curve of different models using the cross-entropy loss function

图5表明,训练迭代步数为100时两种模型都可以收敛到一定的误差范围内,CNN的收敛效果明显比DNN的更好,且曲线的收敛速度更快,在训练集训练所获得的损失函数最小值约为0.0014。

3 结论

本文提出一种基于卷积神经网络的鬼成像手写字体分类识别方法,能够解决网络训练中的过拟合问题。与全连接神经网络相比,该方法在低采样率的情况下,仍有着较高的准确率,并且卷积神经网络模型的精密度、召回率、 F_1 值均有明显提高。通过实验,进一步验证了本方法的有效可行性,为提升手写字体识别

效率提供了新的解决方案,可推动手写字体识别在实际生活中的应用。在后续的研究中,将进一步优化卷积神经网络结构,使得当目标物体来自其他类型的数据集时仍能达到很好的分类结果。

参考文献

- [1] WANG Yuxiang, WANG Ruijin, LI Dongfen, et al. Improved handwritten digit recognition using quantum k-nearest neighbor algorithm[J]. International Journal of Theoretical Physics, 2019, 58(4) :1-10.
- [2] AHLAWAT S, CHOUDHARY A. Hybrid CNN-SVM classifier for handwritten digit recognition[J]. Procedia Computer Science, 2020, 167(3) :2554-2560.
- [3] SENTHIL T, RAJAN C, DEEPIKA J. An improved optimization technique using deep neural networks for digit recognition[J]. Soft Computing, 2020, 25(2) :1647-1658.
- [4] STREKALOV D V, SERGIENKO AV, KLYSHKO D N, et al. Observation of two-photon “ghost” interference and diffraction[J]. Physical Review Letters, 1995, 74(18) : 3600-3603.
- [5] ASUNDI A K, LING Tianai, YIN Cheng, et al. Optical image encryption via high-quality computational ghost imaging using iterative phase retrieval[J]. Laser Physics Letters, 2018, 15(7) :34-39.
- [6] TAO Yong, WANG Xiaoxia, YANG Fengbao. Edge detection method based on high-pass filter ghost imaging[J]. Laser & Optoelectronics Progress, 2020, 57(2) : 021101.
陶勇,王肖霞,杨风暴.基于高通滤波鬼成像的边缘检测方法[J].激光与光电子学进展,2020,57(2) : 021101.
- [7] LE Mingnan, WANG Gao, ZHENG Huabin, et al. Underwater computational ghost imaging[J]. Optics Express, 2017, 25(19) :22859-22868.
- [8] CHEN Yi, FAN Xiang, CHENG Yubao, et al. Single-arm compressed sensing ghost imaging of light intensity diffusion function analysis method[J]. Acta Photonica Sinica, 2016, 45(9) : 0911002.
陈熠,樊祥,程玉宝,等.单臂压缩感知鬼成像的光强扩散函数分析法[J].光子学报,2016,45(9) : 0911002.
- [9] HE Xing, ZHAO Shengmei, WANG Le. Handwritten digit recognition based on ghost imaging with deep learning[J]. Chinese Physics B, 2021, 30(5) :249-254.
- [10] CAO Jianing, ZUO Yuhui, WANG Huahua, et al. Single-pixel neural network object classification of sub-Nyquist ghost imaging[J]. Applied Optics, 2021, 60(29) : 9180-9187.
- [11] LI Baoqi, HE Yuyao. An improved resnet based on the adjustable shortcut connections[J]. IEEE Access, 2018, 6(6) : 18967-18974.
- [12] YAN Guoqing, YANG Fengbao, WANG Xiaoxia, et al. Computational ghost imaging based on orthogonalized sinusoidal speckle[J]. Laser & Optoelectronics Progress, 2020, 57(4) : 041019.
闫国庆,杨风暴,王肖霞,等.基于正交化正弦散斑的计算鬼成像[J].激光与光电子学进展,2020,57(4) : 041019.
- [13] GAO Zhujun, YIN Jianhua, FU Xiquan, et al. Imaging quality improvement of ghost imaging in scattering medium based on Hadamard modulated light field[J]. Applied Optics, 2020, 59(27) :8472-8477.
- [14] AN Xiaofeng, LI Yanqiu, MA Haiyu, et al. Research on threshold processing of Hadamard code modulation correlation imaging[J]. Infrared and Laser Engineering, 2018, 47(10) : 148-153.
安晓峰,李艳秋,马海钰,等.Hadamard编码调制关联成像的阈值处理研究[J].红外与激光工程,2018,47(10) :148-153.
- [15] XUE Yu, TONG Yiling, Yuan Ziming, et al. Handwritten character recognition based on improved convolutional neural network[J]. Intelligent Automation & Soft Computing, 2021, 29(2) : 497-509.
- [16] WANG Fei, WANG Hao, WANG Haichao, et al. Learning from simulation: an end-to-end deep-learning approach for computational ghost imaging[J]. Optics Express, 2019, 27(18) :25560-25572.
- [17] WU Heng, WANG Ruizhou, ZHAO Genping, et al. Sub-Nyquist computational ghost imaging with deep learning[J]. Optics Express, 2020, 28(3) :3846-3853.
- [18] FENG Wei, SUN Xingyu, LI Xiuhua, et al. High-speed computational ghost imaging based on an auto-encoder network under low sampling rate[J]. Applied Optics, 2021, 60(16) :4591-4598.
- [19] XUAN Liu, TAI Linhan, CHENG Zhou, et al. Computational ghost imaging based on array sampling[J]. Optics Express 2021, 29(26) :42772-42786.
- [20] XIONG Feng, HE Di, LIU Yujie, et al. Pneumonia image classification based on improved VGG19 convolutional neural network (invited)[J]. Acta Photonica Sinica, 2021, 50(10) : 1010001.
熊丰,何迪,刘玉杰,等.基于改进VGG19卷积神经网络的肺炎图像分类(特邀)[J].光子学报,2021,50(10) : 1010001.
- [21] ZHAO Xiaoqiang, SONG Zhaoyang. Adam-optimized CNN super-resolution reconstruction[J]. Journal of Frontiers of Computer Science and Technology, 2019, 13(5) :858-865.
赵小强,宋昭漾.Adam优化的CNN超分辨率重建[J].计算机科学与探索,2019,13(5) :858-865.

Handwritten Font Classification Method Based on Ghost Imaging

YAN Ruyu¹, WANG Xiaoxia¹, XI Jiangtao^{1,2}, YANG Fengbao¹, BAO Daerhan³

(1 School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

(2 School of Electrical Computer and Telecommunications Engineering, University of Wollongong,
Wollongong NSW2522, Australia)

(3 Xi'an Microelectronics Technology Institute, Xi'an 710054, China)

Abstract: With the rapid development of the economy, financial services such as bills are also increasing day by day. Among them, the important information in the bill business, such as personal vouchers, checks, and other bills, requires manual reading and input of a large amount of digital information. In order to avoid the waste of human and financial resources, related researchers classify and recognize handwritten fonts based on the neural network classification idea of deep learning. When the method based on deep learning is used for feature extraction of handwritten fonts, there is often a lack of detailed information such as edges and textures, which leads to the problem of low recognition accuracy.

Aiming at the problem that the features of handwritten digits or letters are difficult to effectively extract, the recognition efficiency was not high and even caused recognition errors, a new automatic recognition method of handwritten digits or letters was proposed by combining the principle of ghost imaging and the classification network based on deep learning. This method utilizes the principle of ghost imaging. It can save the imaging process in the traditional image recognition method, and jumping out of the inherent thinking that identifying objects is identifying images, and can quickly classify the image of handwritten digits and letters only by the total light intensity value transmitted by handwritten digits or letters without extracting and identifying features of handwritten digits or letters. The automatic recognition of handwritten digits or letters based on ghost imaging solves the critical problem of needing to extract digits or letter images features in traditional handwritten font recognition methods, and can greatly improve the recognition efficiency of handwritten digits or letters. Firstly, a ghost imaging detection system is built using commonly used optical instruments such as lasers, digital micromirror arrays, and single-pixel detectors. The laser in the built detection system is used to generate a pseudothermal light source, and the digital micromirror array is used to obtain the Hadamard speckle sequence with a resolution of 32×32 irradiating the target object at different times. And realizing the irradiation of 17 239 handwritten images of handwritten digits and letters. Secondly, the single-pixel detector is used to collect data on the total light intensity value transmitted by the handwritten digits and letters. The data collection process is very fast and does not cause huge time costs. The value of the bucket detector after the collection is converted into a one-dimensional vector, and use the one-dimensional vector corresponds to the handwritten font as the input of network training. Finally, the network framework is built based on the advantages of the convolutional neural network in image classification and is used to solve the problems in the training process. The network degradation problem is added to the residual block structure, which can directly pass shallow information to deeper layers by skipping one or several layers through skip connections. In order to solve the problem of overfitting, the Dropout layer is added to it, and the robustness of the network to the loss of specific neuron connections is improved by reducing the weight.

The experimental results show that: for handwritten digits, compared with the fully connected network, the precision, recall rate and F1 value of the convolutional neural network model are increased by 86.50%/97.25%, 86.40%/98.03%, 86.31%/97.60%; for handwritten letters, the precision, recall, and F1 value of the convolutional neural network under full sampling are 91.87%, 90%, and 90.23%, respectively. At the same time, in the case of undersampling and non-undersampling, the ten types of digits from 0 to 9 under the two models of convolutional neural network and fully connected neural network and randomly selected l, v, y, z, m, n, o, r, s, and h ten types of letters are compared and analyzed. The experimental results show that the accuracy rate of each type of digit and letter of the convolutional neural network is higher than that of the fully connected network under the same conditions. The accuracy of each type of digit and letter under the two models further verifies that as the sampling rate increases, the recognition accuracy also increases. By comparing the evaluation indicators of the convolutional neural network and the fully connected network architecture, the effectiveness and rationality of the proposed method are further illustrated. The classification and recognition results of handwritten letters verified by

experiments further illustrate the versatility of the constructed convolutional neural network. It provides the possibility for the wide application of handwritten fonts in real life. The research on the classification and the recognition of handwritten fonts based on ghost imaging can effectively solve the bottleneck problem of low recognition efficiency of existing font recognition methods.

Key words: Ghost imaging; Handwriting recognition; Convolutional neural network; Accuracy rate; Residual

OCIS Codes: 100.3010; 100.4996; 200.4260

Foundation item: National Natural Science Foundation of China (No. 61972363), Open Fund of the State Key Laboratory of Internal Combustion Engine Reliability (No. skler-202011), Youth Science and Technology Research Fund of the Applied Basic Research Program of Shanxi Province (No. 201901D211233), Scientific Research Funding Project for Returned Overseas Chinese Scholars in Shanxi Province (No. 2021-110)