

基于 Fire Module 卷积神经网络的手写变造数字检测

陈颖, 高树辉*

中国人民公安大学刑事科学技术学院, 北京 100038

摘要 提出基于卷积神经网络(CNN)的手写变造数字检测方法,为变造文件鉴定提供智能化解决方案。实验召集 50 名志愿者,采集 50 种不同品牌型号中性笔,形成 6 类变造笔迹和正常笔迹图像样本,建立了共计 7200 余份的样本数据。在 AlexNet 基础上引入 Fire Module 结构,提出基于变造数字检测的卷积神经网络(FNNet),以 1×1 卷积核代替部分 3×3 卷积核,实行卷积层组装来检测变造样本。实验结果表明,FNNet 在 6 类手写变造数字中的平均测试准确率达 98.36%,比 AlexNet 高 3.01 个百分点。所提方法优于传统特征分类器,为变造笔迹鉴定提供了一种新的方法。

关键词 图像处理; 变造文件鉴定; 变造数字检测; 卷积神经网络

中图分类号 TP391 **文献标志码** A

doi: 10.3788/LOP57.221019

Forgery Numeral Handwriting Detection Based on Fire Module Convolutional Neural Network

Chen Ying, Gao Shuhui*

School of Criminal Investigation and Forensic Science, People's Public Security University of China, Beijing 100038, China

Abstract In this paper, we propose a method of forgery numeral handwriting detection based on convolution neural network (CNN). It provides an intelligent solution for forgery document detection. The experiment convened 50 volunteers and collected image samples of six types of forged handwritings and normal handwriting with 50 different brand pens, and established a total of more than 7200 sample data. Then, we designed a new CNN for forgery numeral handwriting detection called FNNet by introducing Fire Module structure based on AlexNet. We replaced the partial 3×3 convolution kernel with 1×1 convolution kernel and performed convolution layer assembly to detect forged samples. The experimental results show that the average test accuracy of FNNet in the six types of handwritten forgery numbers is 98.36%, which is 3.01 percentage higher than that of AlexNet. The proposed method is superior to traditional feature classifiers; it provides a new method for forged handwriting detection.

Key words image processing; forgery document examination; forgery numeral detection; convolutional neural network

OCIS codes 100.4996; 100.5010; 100.2000

1 引言

随着经济的发展,经济事务往来导致纠纷案件日益增多,一些人为达到经济诈骗的目的,对交易人

之间的借条、发票、票据等文件上的金额、时间等数字进行变造的案例屡见不鲜^[1]。文件司法鉴定的主要思路是通过检测篡改的文本内容来确定文件是原始的还是变造的。添加是最常见的文件变造方法,

收稿日期: 2020-02-27; **修回日期:** 2020-04-08; **录用日期:** 2020-04-27

基金项目: 公安部刑事技术双十计划重点攻关项目(2019SSGG0102)、2018 年上海市现场物证重点实验室开放课题基金(2018XCWZK24)

* **E-mail:** gaoshuhui@ppsuc.edu.cn

由于种类、成分、形成时间及改写方式的不同,添加笔画的笔墨与原文的笔墨存在差异,但添改内容通常是由相同颜色的笔墨改动而成的,一般来说人们不能在第一时间判断是否为添改笔画。目前,文件司法鉴定的方法有物理和化学两种。化学检验主要有薄层色谱、高效液相色谱等。例如,Harris^[2]利用薄层分析法选择溶剂对笔墨进行分类,进而对变造文件进行检验。物理检验主要有特种光源识别、光谱成像技术等。如Reed等^[3]使用VSC6000高光谱成像系统检测分析了不同型号凝胶笔形成的字迹在可见红外光谱范围内形成的光谱差异,对文件字迹笔墨进行鉴别,从而进行笔墨错配检测。但这两种传统方法硬件成本高,且化学方法不能保存文件原貌。

人工智能的兴起与图像处理技术的提高,模式识别广泛应用于各领域中,在变造文档鉴定中的应用也逐渐宽泛,其主要优势是可以提高检测效率、降低成本,避免鉴定人员主观因素的影响,使鉴定过程更加客观、科学。模式识别的方法主要分为两种,即基于特征选取和分类器设计、基于深度学习。前者方法已在变造文件鉴定中有相关应用。Megahed等^[4]提取了字迹图像的RGB三个颜色通道强度平均值、标准差及偏度值9个特征,通过计算方均根误差(RMSE)对变造字迹进行检验。Dansena等^[5]通过提取银行手写支票笔迹图像的各个颜色通道的统计特征,利用多层感知器对支票中添改字迹笔墨进行检测分类,准确率为93.5%。而在深度学习方法中,近几年卷积神经网络在图像识别领域大放异彩,最早由日本学者Fukushima^[6]基于视觉皮层细胞的感受野提出,实现了第一个卷积神经网络;自2012年以来,深度学习的研究与应用开始爆发,从Krizhevsky团队^[7]提出的AlexNet到2015年Szegedy等^[8]提出的GoogLeNet,及2016年Silver等^[9]研发的AlphaGo都证明了卷积神经网络在图像识别领域中的巨大优势与潜力。同样,卷积神经网络在法庭科学领域也已有重大突破,主要包括现场勘查与追踪^[10-12]、声纹识别^[13-14]、微量物证鉴定^[15-16]等,但在添改变造文件检验中的应用较少。

本文着重研究基于卷积神经网络的手写变造数字检测,建立6类手写变造数字样本数据库,基于Caffe深度学习框架,在AlexNet的基础上改进设计网络结构,由于原始网络参数数量的庞大会给网络宽带的数据传输带来很大负担,因此引入Fire Module结构,以简化网络复杂度,同时探究第一层

全连接输出参数对网络性能的影响。通过调试网络参数提出基于变造数字检测的卷积神经网络(FNNet),以期推动卷积神经网络在文档变造检测中的应用,为文书鉴定提供智能化解决方案。

2 手写变造数字检测方法

2.1 卷积神经网络基本原理

卷积神经网络可以图像数据形式直接作为网络的输入,避免了传统特征算法选取与数据重建的过程。在图像特征提取过程中,整个网络持续地进行迭代并反馈出计算的误差,同时对网络的权值进行更新,得到最优的网络模型,其结构主要包括卷积层、池化层、全连接层及输出层。

通过卷积核,卷积层在原始图像上以一定的步长平移滑动来提取特征,每一个特征就是一个特征映射,是卷积神经网络中的核心组成部分。每个卷积层包含许多特征平面,组成同一特征平面的神经元的共享权值就是卷积核。卷积核不仅可以减少网络各层之间的连接,同时又降低了过拟合的风险。设定输入图像 a_c 大小为 $k_c \times l_c$,设定任意输入卷积层卷积核 a_s 尺寸大小为 $k_s \times l_s$,卷积核个数为 n ,得到多个卷积特征图 $a_i (i=1, 2, 3, \dots, n)$,其卷积计算过程为

$$a_i = \sigma[w^{(1)} * a_s + b^{(1)}], \quad (1)$$

式中: $w^{(1)}$ 为权重; $b^{(1)}$ 为偏执; σ 为非线性映射激活函数。本文采用ReLU函数,在每一层卷积操作后使用,相比传统的Sigmoid函数,ReLU函数可以防止梯度弥散,使网络收敛更快。Sigmoid函数的导数只有在0附近时有比较好的激活性,在正负饱和区的梯度都接近于0,所以这会造成梯度弥散;而ReLU函数在输入值大于0时,梯度为常数,所以不会发生饱和现象,而且只需要一个阈值就可以得到激活值,可提高网络训练速度。

则输出特征图像 a_i 的尺寸大小为

$$S(a_i) = \{[(k_c + 2 \times C_{\text{padding}} - k_s) / l_{\text{stride}}] + 1\} \times \{[(l_c + 2 \times C_{\text{padding}} - l_s) / l_{\text{stride}}] + 1\} \times n, \quad (2)$$

式中: C_{padding} 为边缘扩展参数,默认值为0; l_{stride} 为卷积核滑动步长,默认值为1。

池化层的主要目的是对特征图进行空间分辨率的降维,通过汇聚特征后稀疏参数来减少计算参数量,降低网络的复杂度,进而防止网络过拟合。常见的池化层包括最大值池化和平均值池化。这里使用最大值池化的方法来输出特征区域的最大值。全连接层一般在网络结构的末端,其作用是将所有卷积

层和池化层的输出特征整合输出为一维特征向量^[17],再经 Softmax 分类器进行特征分类。

2.2 损失函数

在机器学习中,对训练样本中不一致的预测值与真实值进行惩罚,得到损失。预测值和真实值映射到某一空间存在误差,而损失函数就是这种误差的描述形式,通过采用一定的优化算法对损失进行最小优化,得到合理的网络权值,进而衡量模型和参数与训练样本的符合度,损失函数越小说明模型与训练样本符合度越高。对手写变造数字进行识别是一个分类问题,即识别样本是否被变造,所以选择交叉熵损失函数,而本文本质上为二分类问题,又称为逻辑回归损失函数。在 Caffe 深度学习框架中使用 SoftmaxWithLoss 层实现了损失函数的计算,因为在卷积神经网络中的输出常为向量,而交叉熵描述两个概率分布的差异,所以将经 Softmax 函数的输入向量归一化为概率分布,再经交叉熵损失函数计算其损失。

2.3 网络权值初始化

为了让卷积神经网络在训练过程中学习到有用的信息,在各层激活值不会出现饱和现象和各层激活值不为 0 的条件下,应在每层网络训练之前对各层的参数进行初始化。在 AlexNet 中使用 Gaussian 方法进行初始化,然而在本实验中使用 Gaussian 方法会出现损失值溢出的问题,网络输出数据分布的方差会随着输入神经元个数而改变,导致网络不收敛,最终网络训练失败。而为了解决这个问题,本实验采用 Xavier 初始化方法,该方法对方差大小进行规范化,使各层激活值的方差与传播过程中的方差保持一致^[18],解决了 Gaussian 方法初始化后激活值方差发生改变,造成数据不稳定的问题。Xavier 初始化的实现满足一定的均匀分布,表达式为

$$W \sim U \left[-\sqrt{\frac{6}{n_k + n_{k+1}}}, \sqrt{\frac{6}{n_k + n_{k+1}}} \right], \quad (3)$$

式中: n_k 为参数所在第 k 层的输入维度; n_{k+1} 为参数所在第 $k+1$ 层的输出维度; W 为网络权值。激活值方差和传播梯度方差与层数无关,由于 ReLU 无法控制数据幅度,所以本实验将其与 Xavier 初始化搭配使用,得到了稳定收敛的网络模型。

2.4 Fire Module 结构

卷积神经网络从 2012 年开始一直在追求模型的识别准确率,从 AlexNet 到 16-VGGNet、19-VGGNet,网络层数线性堆叠不断加深,Top-5 错误

率大幅度下降,从 15.3% 到 7.3%。但网络中隐含层数增加,网络参数总量也在不断增加,使得计算量变得异常庞大,19-VGGNet 参数量达 144×10^6 ,加大了计算机数据传输压力,极易出现过拟合现象。为了解决以上问题,2016 年 UC Berkeley 和 Stanford 团队提出了 SqueezeNet 模型^[19],旨在简化网络复杂度,其中核心部件就是 Fire Module 结构。SqueezeNet 使用大量的 1×1 卷积核代替 AlexNet 中常用的 3×3 卷积核,这使得卷积核输出特征数量减少了 $1/9$,但为了不影响网络的识别性,保留了部分 3×3 的卷积核,从而引入 Fire Module 结构,其结构如图 1 所示。由图 1 可知,输入到 Fire Module 结构的图像大小为 $H \times W \times M$,其中 H 、 W 分别代表图像的高、宽, M 代表通道数。该结构将卷积层拆解成一个 squeeze 层和一个 expand 层,其中 squeeze 层中有 $s_{1 \times 1}$ 个 1×1 的卷积核,经过 ReLU 激活函数后,其特征图通道数为 $s_{1 \times 1}$;再将其分别输入到 expand 层中的 1×1 的卷积层和 3×3 的卷积层中,其中 1×1 卷积核个数为 $e_{1 \times 1}$, 3×3 的卷积核个数为 $e_{3 \times 3}$,输出特征图通道数分别为 $e_{1 \times 1}$ 、 $e_{3 \times 3}$;最后对 expand 层中的经 concat 层的两个卷积层的输出特征进行特征融合,最后输出通道数为 $e_{1 \times 1} + e_{3 \times 3}$ 的特征图,将其再作为下一层的输入特征。由此可见,经 Fire Module 结构输出的图像的分辨率保持不变,变的是图像维数,为了减少输入特征的通道数,减少参数量,需使 $s_{1 \times 1} < e_{1 \times 1} + e_{3 \times 3}$ 。

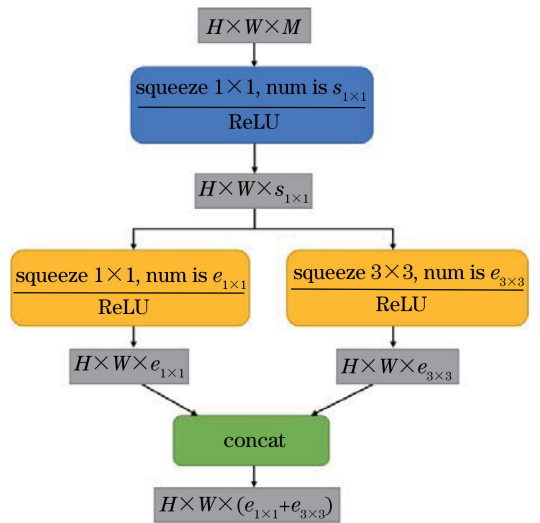


图 1 Fire Module 结构

Fig. 1 Fire Module structure

2.5 FNNNet 结构与模型参数

经 Fire Module 结构的输入特征维度将大大减少,在 SqueezeNe 模型中,使用了 9 层 Fire Module

结构。使用 Fire Module 结构时由于分类任务较简单,过深的网络层数并不利于图像特征提取,所以提出 FNNet,其使用 5 层 Fire Module 模块代替 AlexNet 中具有固定卷积核大小的卷积层,其具体网络结构如图 2(a)所示。同时为了直观地对 FNNet 与 AlexNet 结构进行对比,图 2(b)给出了 AlexNet 结构,可知 AlexNet 使用 5 个卷积层,其中第一个卷积层的卷积核大小为 11×11 ,第二个为 5×5 ,其余为 3×3 。为了控制图像的输出维度,在 FNNet 整个结构的首尾分别使用一个卷积层,中间穿插了 3 个卷积核为 3×3 、步长为 2 的 Max pooling 层。SqueezeNet 模型在最后一层使用

average pooling 层代替全连接层(FC),但在本文中若使用平均池化层直接代替全连接层,由于计算机对非均匀稀疏数据的计算效率差,模型分类准确率将会下降。而由图 2(b)可知,AlexNet 中使用了两层输出维度为 4096 的全连接层,这使网络参数过于庞大,所以为了更好地优化网络,增强其泛化能力但又不产生数据维度灾难,在网络末端使用了一全连接层,在后续经实验分析确定其输出数目为 1000。同时引入 Dropout 技术,输入到 Dropout 的神经元之间共享权重,降低了复杂的适应关系,避免过拟合,最后再将结果输入到 Softmax 分类器,该网络结构在简化了网络复杂度的同时保证了特性信息的充足性。

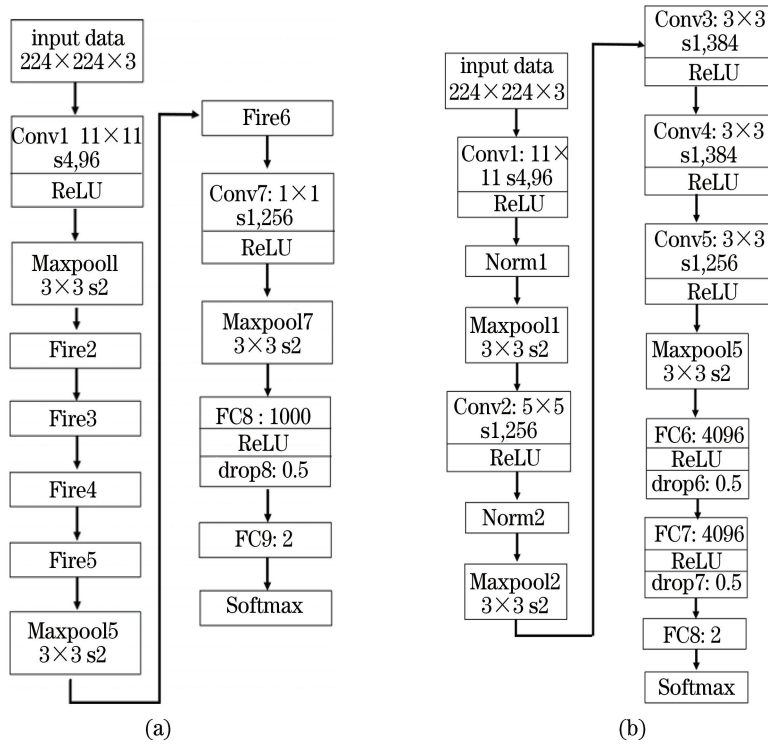


图 2 网络结构。(a) FNNet 结构;(b) AlexNet 结构

Fig. 2 Network structure. (a) FNNet structure; (b) AlexNet structure

FNNet 模型的网络维度参数如表 1 所示。设置笔迹图像的输入尺寸为彩色三通道 224×224 大小,在每个 Fire Module 结构中涉及到具有 3 个维度的超参数,即 $s_{1 \times 1}$ 、 $e_{1 \times 1}$ 、 $e_{3 \times 3}$,使用 5 个 Fire Modules,即一共 15 个超参数。第 j 个 Fire Module 结构的超参数表达式为

$$e_j = N_{\text{base},e} + N_{\text{incr},e} \times \frac{j}{N_{\text{freq}}}, \quad (4)$$

$$s_{j,1 \times 1} = R_s \times e_j, \quad (5)$$

式中: e_j 为第 j 个 Fire Module 中 expand 层卷积核的数量; $N_{\text{base},e}$ 为第一个 Fire Module 中 expand 层

的卷积核个数; N_{freq} 为 Fire Module 的个数; $N_{\text{incr},e}$ 为在 N_{freq} 个 Fire Module 后增加的 expand 层的卷积核个数; R_s 为压缩比,即 squeeze 层中卷积核个数占 Fire Module 中卷积核总数的比例。表 1 列出每个 Fire Module 结构中的 squeeze 层和 expand 层的具体参数。为了保证 expand 层的 $\text{expand}1 \times 1$ 和 $\text{expand}3 \times 3$ 的图像输出尺寸一致,使其能够在相同维度上进行特征融合,在 $\text{expand}3 \times 3$ 的卷积核边缘补充 pad 为 1。表 1 最后一列为 FNNet 中每层网络的具体参数维度。

表 1 FNNet 的维度参数
Table 1 Dimension parameters of FNNet

Layer name	Output size	Filter size/stride	$s_{1 \times 1}$	$e_{1 \times 1}$	$e_{3 \times 3}$	Dimension
Input image	$224 \times 224 \times 3$					
Conv1	$54 \times 54 \times 96$	$11 \times 11 / 4$				34,944
Maxpool1	$27 \times 27 \times 96$	$3 \times 3 / 2$				
Fire2	$27 \times 27 \times 256$		32	128	128	44,320
Fire3	$27 \times 27 \times 384$		48	192	192	104,880
Fire4	$27 \times 27 \times 384$		48	192	192	111,024
Fire5	$27 \times 27 \times 512$		64	256	256	188,992
Maxpool5	$13 \times 13 \times 512$	$3 \times 3 / 2$				
Fire6	$13 \times 13 \times 512$		64	256	256	197,184
Conv7	$15 \times 15 \times 256$	$1 \times 1 / 1$				131,328
Maxpool7	$7 \times 7 \times 256$	$3 \times 3 / 2$				
FC8	$7 \times 7 \times 1000$					12,545,000
FC9	$1 \times 1 \times 2$					2002

3 实验与分析

3.1 实验数据制备与采集

在现代日常书写活动中,中性笔书写流畅、手感好,成为书写的主要工具,同时黑色和蓝色墨水是书写活动中常使用的颜色,所以实验收集市面上常见的 50 种品牌型号的中性笔,其中黑色和蓝色各 25 支,分别对其进行 P1~P50 编号,同时召集 50 名志愿者分别与 50 支笔匹配,使用相对应编号的笔进行样本书写。实验样本是依据实际文件检验鉴定中遇到的几类典型数字变造问题而制备的。

50 名志愿者分别使用其对应的中性笔以正常用力书写 6 类正常数字样本“3”、“4”、“6”、“7”、“8”、“9”,每人每种品牌中性笔书写每类数字样本 15 个,得到正常数字样本 $6 \times 50 \times 15$ 共计 4500 个。再在上述相同书写条件下书写待添改字迹样本,依次选取黑笔 P1~P25、蓝笔 P26~P50 中的两种笔进行随机组合分别得到 300 种组合。为了保证实验样本的代表性和样本采集的可行性,50 名志愿者匹配相对应的中性笔编号,使用每种组合中对应编号其中一支笔书写待添改数字样本“0”、“1”、“2”、“5”。在日常添改行为中,由于待添改字迹与添改字迹之间存在时间差,因此在本实验中,将待添改字迹书写完毕后,笔迹样本在自然环境下经过 12 h 后再由每种组合中性笔编号对应的另一志愿者将待添改数字“0”改为“6”和“9”、“1”改为“4”和“7”、“2”改为“3”、“5”改为“8”,每种组合书写每类字迹样本 2 个,得到变造添改字迹样本 $6 \times (300 + 300) \times 2$ 共计 7200 个。

数据采集时使用配置 85 mm 定焦微距镜头的尼康相机 D7200,在暗环境下将相机固定在翻拍架

上同一高度位置,笔迹样本置于相机镜头正下方,光源为白炽灯,45°均匀配光。将相机调至光圈优先模式并调整白平衡,相机参数设置为光圈值 $f/11$ 、感光度 ISO-3200、曝光补偿+2,在该模式下依次对数字笔迹样本进行拍照采集。

3.2 实验样本数据预处理

从采集到的样本图像中挑选字迹清晰、拍摄质量好、添改痕迹不明显的字迹样本,分别保存每类添改与未添改字迹样本 600 张左右,6 类数字样本共 7200 余张。

使用 Matlab 软件对所有样本图像依次进行笔迹截图,裁剪掉图像中多余的留白区域,保证每张图像的长宽比相同,设置图像输出为 bmp 格式,从而保证样本图像数据信息的原始性和充足性。图像的分辨率过高会影响特征采集的速度,因而在保证字迹图像信息不丢失的前提下,选择合适的图像分辨率,设置其大小为 224×224 。对每类数字样本生成正常和变造的数据标签,并随机抽每类实验图像的 80% 作为训练集进行网络学习,其他的 20% 作为测试集进行测试,同时将图片格式转换为 LMDB 格式文件,该数据集的具体情况如表 2 所示,预处理后部分样本如图 3 所示。

3.3 实验环境配置及网络参数设置

实验环境为 64 位 Win10 操作系统,处理器为英特尔 Core i5-8250U @ 1.60GHz,四核,内存 16GB,Nvidia GeForce GTX1050 显卡。使用深度学习 Caffe 框架搭建网络,基础学习率为 0.02,学习率下降策略为“poly”,梯度更新权重为 0.9,权重衰减为 0.0005,网络各层权值初始化方法为 Xavier。训练批量数目 batch_size 为 64,迭代 epoch 为 65。

表 2 数据集组成

Table 2 Composition of dataset

Dataset	3	4	6	7	8	9
Normal handwriting	540	593	599	599	604	597
Forgery handwriting	621	612	616	614	600	614

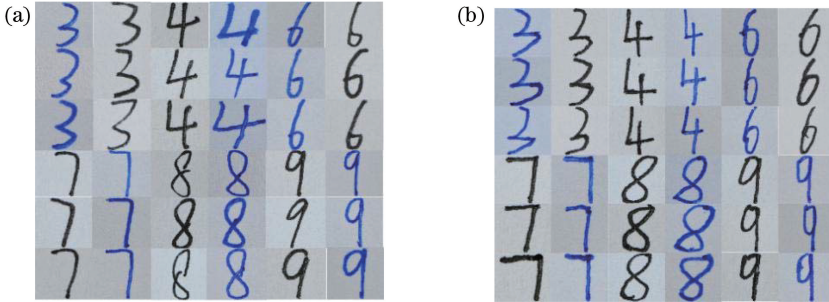


图 3 部分数据集样本。(a)正常样本;(b)变造样本

Fig. 3 Part dataset examples. (a) Normal samples; (b) forgery samples

3.4 结果分析

对全连接层中的每个神经元与其前一层的所有神经元进行全连接,全连接层可以对前序卷积池化层中提取到的各类图像区分性的有效信息进行整合,为分类器提供图像的特征向量。但由于其连接节点数量直接影响输出特征维数,本实验为了探究第一层全连接层输出神经元个数对网络性能的影响,以手写变造数字“4”的网络模型为例,分别设置其输出数量为 500,1000,2000,实验结果如表 3 所示。由表 3 可知:随着输出个数数量的增加,模型测试准确率从 500 的 95.28% 增加到 1000 的 96.32%,随后当输出数目为 2000 时,准确率下降;三类输出中,1000 的损失值最小;同时三类输出模型的训练时间相当。所以本实验设置第一层全连接输出个数为 1000,这在保证网络测试性能的同时,相比 AlexNet 两层输出数目为 4096 的全连接层,参数量大大减少。

表 3 全连接层中神经元数量的分类性能与损失性能

Table 3 Classification and loss performances of number of neurons in fully connected layer

Parameter	500	1000	2000
Test accuracy /%	95.28	96.32	93.98
Loss value	0.14	0.11	0.25
Train time /min	13	13	14

6 类变造数字的网络性能和检测精度如表 4 所示。在网络训练过程中,当网络每遍历整个训练集一次,即对其测试一次,训练完成后由此得到网络模型的训练准确率,为了验证模型的性能,最后通过测试集对其进行测试。由表 4 可知:每类变造字迹的

测试精度均在 96% 以上,最高为添改数字“8”,其测试准确率达 99.60%;每类变造字迹网络模型的损失值均趋近于 0,说明训练网络的数据拟合程度好,模型结构稳定。由损失值和测试准确率可以看出,损失值越小,网络越稳定,其检测精度越高。6 类变造数字的平均测试准确率为 98.36%。另外各类变造字迹的准确率之间略有差异,其主要原因是字体字形的结构、添改笔画特征、添改部位笔画特征不同,卷积神经网络在提取字迹特征信息时产生差异。

表 4 所提网络的训练和测试结果

Table 4 Results of training and testing of proposed network

Forgery numeral	Train accuracy /%	Loss value	Test accuracy /%
3	99.60	0.03	99.20
4	96.40	0.11	96.32
6	98.40	0.07	98.32
7	98.40	0.04	98.72
8	99.20	0.01	99.60
9	97.80	0.06	98.00

本实验对每类添改数字笔迹的检测本质上为二分类问题,为了评判网络模型的分类型能力,绘制了 6 类变造数字网络模型在测试集上的 ROC(receiver operating characteristic)曲线,如图 4 所示。横坐标代表样本的假正类率(FPR),纵坐标为样本的真正类率(TPR),ROC 曲线下方的面积代表模型分类性能的高低,面积越大表示该模型分类性能越好。由图 4 可知,添改数字“8”的 ROC 曲线与横坐标轴围成的面积最大,而添改数字“4”的 ROC 曲线与横坐标轴围成的面积最小,说明在这 6 类模型中,添改数

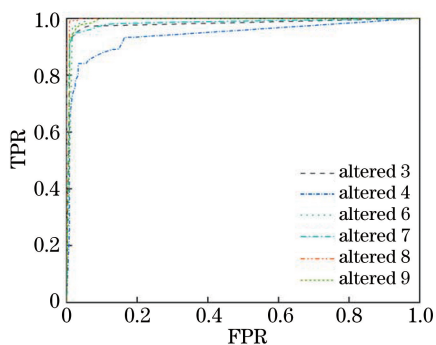


图4 六类变造数字的 ROC 曲线

Fig. 4 ROC curves of six forged figures

字“8”的网络模型分类性能最高,这与模型的测试精度结果一致。

由于所提 FNet 模型是在 AlexNet 的基础上进行改进的,在 6 类手写变造数字笔迹数据库上对这两种网络模型进行训练测试,对比结果如图 5 所示。由图 5 可知,相比 AlexNet 测试结果,FNet 模型在每类添改变造笔迹上的测试准确率都显著高于 AlexNet。

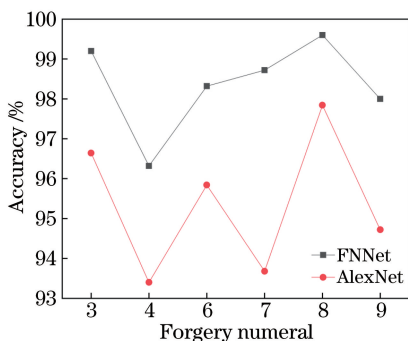


图5 FNet 与 AlexNet 的测试准确率比较

Fig. 5 Comparison of test accuracy between FNet and AlexNet

同时对 FNet 模型与 AlexNet 在 6 类变造数字笔迹上的平均测试准确率与网络参数总量进行比较,如表 5 所示。由表 5 可知,FNet 对 6 类变造数字的平均测试准确率比 AlexNet 高 3.01 个百分点,同时模型的参数量约为 AlexNet 的 1/5。

表5 FNet 与 AlexNet 的参数比较

Table 5 Parameter comparison of FNet and AlexNet

Network	Mean test accuracy /%	Parameter quantity
AlexNet	95.35	58,289,538
FNet	98.36	13,228,346

最后,对基于卷积神经网络的方法与传统的基于特征选取加分类器设计的方法进行比较。选择图像处理中常用的两种特征提取方法,即局部二值模式(LBP)和全局特征描述子(GIST),LBP 主要提取图像

的局部纹理特征,而 GIST 主要提取图像的全局信息。分别提取字迹图像的全局特征与局部特征后使用支持向量机(SVM)作为分类器,对图像特征进行分类,本实验中通过交叉验证设置 SVM 分类器的惩罚系数 $c=2$,核函数的幅宽 $g=0.4$ 。三种方法的测试准确率如图 6 所示。可知,所提 FNet 在任何一类添改数字中的准确率都比其他两种方法高,而 SVM 寻找字迹图像特征的最优超平面时存在困难,从而导致其分类精度较低。实验结果说明,卷积神经网络在检测手写变造数字时体现出明显的优势。

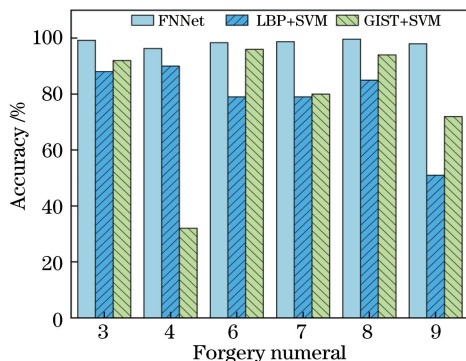


图6 FNet 与其他算法的检测准确率

Fig. 6 Test accuracy of FNet and other algorithms

4 结 论

创新性地将模式识别方法应用于变造笔迹鉴定,首次使用卷积神经网络对手写变造数字进行检测,解决了在文件司法鉴定中依赖鉴定人员的经验知识、通过笔迹笔画间的异常特征来对手写变造数字进行判断的问题。在 AlexNet 的基础上通过对网络参数结构进行改进,引入 Fire Module 结构,提出 FNet 模型,同时通过收集不同志愿者的笔迹样本,建立具有代表性的手写变造数字样本数据库。实验结果表明,FNet 模型的 6 类手写变造数字的平均检测准确率达 98.36%,明显高于 AlexNet,同时也充分证明卷积神经网络对手写变造数字笔迹识别的可行性与可靠性。但本文只涉及单一背景下中性笔添改笔迹图像的研究,在实际案例中存在书写背景复杂的情况,在今后的研究中,会着重研究在复杂背景下添改变造数字的情况,同时增加中性笔种类以进一步扩大样本数据库,探究更先进的优化网络模型,为识别变造笔迹开辟新方法。

参 考 文 献

- [1] Throckmorton G J. Disappearing ink: its use, abuse, and detection[J]. Journal of Forensic Sciences, 1990,

- 35(1): 12820J.
- [2] Harris J. Developments in the analysis of writing inks on questioned documents[J]. *Journal of Forensic Sciences*, 1992, 37(2): 13269J.
- [3] Reed G, Savage K, Edwards D, et al. Hyperspectral imaging of gel pen inks: an emerging tool in document analysis[J]. *Science & Justice*, 2014, 54(1): 71-80.
- [4] Megahed A, Fadel S M, Han Q, et al. Handwriting forgery detection based on ink colour features[C]//2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), November 24-26, 2017, Beijing, China. New York: IEEE, 2017: 141-144.
- [5] Dansena P, Bag S, Pal R. Differentiating pen inks in handwritten bank cheques using multi-layer perceptron[M]//Shankar B U, Ghosh K, Mandal D P, et al. *Pattern recognition and machine intelligence. Lecture notes in computer science*. Cham: Springer, 2017, 10597: 655-663.
- [6] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. *Biological Cybernetics*, 1980, 36(4): 193-202.
- [7] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // *Proceedings of the 25th International Conference on Neural Information Processing Systems*, December 5-8, 2012, Lake Tahoe, Nevada, USA. New York: Curran Associates, 2012(1): 1097-1105.
- [8] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions [C] // *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 1-9.
- [9] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484-489.
- [10] Yan S D, Gao S H, Tang Y Q, et al. Plier and scissor mark recognition based on convolutional neural network [J]. *Science Technology and Engineering*, 2019, 19(31): 227-232.
严圣东, 高树辉, 唐云祁, 等. 基于卷积神经网络的钳剪工具痕迹识别[J]. *科学技术与工程*, 2019, 19(31): 227-232.
- [11] Yang M J, Tang Y Q, Jiang X J. Research on shoe pattern recognition method based on convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(17): 171501.
杨孟京, 唐云祁, 姜晓佳. 基于卷积神经网络的鞋型识别方法研究[J]. *激光与光电子学进展*, 2019, 56(17): 171501.
- [12] Li Z R, Wang K X, He X L, et al. Heel-strike event detection algorithm based on convolutional neural networks [J]. *Laser & Optoelectronics Progress*, 2019, 56(21): 211503.
李卓容, 王凯旋, 何欣龙, 等. 基于卷积神经网络的足跟着地事件检测算法[J]. *激光与光电子学进展*, 2019, 56(21): 211503.
- [13] Abdel-Hamid O, Mohamed A R, Jiang H, et al. Convolutional neural networks for speech recognition [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2014, 22(10): 1533-1545.
- [14] Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks [C]//*2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1725-1732.
- [15] Zhang W, Zhang H X. Classification of main components of textile fabrics based on convolutional neural network [J]. *Journal of Zhejiang Sci-Tech University*, 2019(1): 1-8.
张玮, 张华熊. 基于卷积神经网络的纺织面料主成分分类[J]. *浙江理工大学学报*, 2019(1): 1-8.
- [16] Jiang X J, Gao S H. Automatic classification of microscopic hair images based on improved convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231002.
姜晓佳, 高树辉. 基于改进卷积神经网络的毛发显微图像自动分类[J]. *激光与光电子学进展*, 2019, 56(23): 231002.
- [17] Sainath T N, Mohamed A R, Kingsbury B, et al. Deep convolutional neural networks for LVCSR[C]//*2013 IEEE International Conference on Acoustics, May 26-31, 2013, Vancouver, BC, Canada*. New York: IEEE, 2013: 8614-8618.
- [18] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[J]. *Journal of Machine Learning Research*, 2010: 249-256.
- [19] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size [EB/OL]. (2016-11-04) [2020-02-26]. <https://arxiv.org/abs/1602.07360>.