

# 基于改进卷积神经网络的手写数字识别

杜洋, 王展青

(武汉理工大学 理学院, 武汉 430070)

**摘要:** 针对试卷卷头手写分数识别问题, 提出一种改进卷积神经网络 (CNN) 的手写数字识别算法; 该算法, 在网络提取图像特征的过程中改用多层特征融合和混合采样的方式, 实现图像特征信息的充分提取; 在网络训练阶段采用学习速率自适应的调整策略和增加动量项, 提高网络训练速度; 最后, 采用 SVM 分类器进行分类识别; 实验结果表明, 该算法有效的提高了手写数字的识别率, 减少了训练时间, 识别性能优于普通的 CNN 模型和其它对比方法; 在 MNIST 数据集上的仿真实验充分验证了算法的有效性。

**关键词:** 手写数字识别; 卷积神经网络; SVM 分类器

## Handwritten Numeral Recognition Based on Improved Convolutional Neural Network

Du Yang, Wang Zhanqing

(School of Science, Wuhan University of Technology, Wuhan 430070, China)

**Abstract:** A handwritten digit recognition algorithm based on improved convolutional neural network (CNN) is proposed to solve the problem of handwritten score recognition on the examination paper. This algorithm uses multi-layer feature fusion and mixed sampling in the process of network extracting image features, to achieve the full extraction of image feature information. Further the adjusting strategies of learning rate self-adapting and the methods of increasing the momentum items are used to improve the network training speed in the network training stage. Finally, the SVM classifier has been used for classification identification. The experimental results show that the algorithm effectively improves the recognition rate of handwritten digits and shortens the learning time, which is much better than the regular CNN model and the other compared methods. Simulation experiments on the MNIST data sets show the effectiveness of the proposed algorithm.

**Keywords:** handwritten digital recognition; convolutional neural network; SVM classifier

### 0 引言

手写数字识别在教育、邮政、银行以及商业等诸多领域有着重要的应用。目前研究者们提出了许多手写数字识别的方法, 如多尺度特征和神经网络相融合的方法<sup>[1]</sup>、基于原型生成技术的方法<sup>[2]</sup>、基于 AP 和 BP 神经网络的方法<sup>[3]</sup>、基于概率测度支持向量机 (support vector machine, SVM) 的方法<sup>[4]</sup>等, 但是上述方法对特征的表达力不足, 容易受到外界环境的影响, 不能达到高识别率的要求。

近年来, 卷积神经网络 (convolutional neural network, CNN) 在手写数字识别方面取得了很好的表现, 它对图像识别具有自动提取特征的能力, 避免了传统识别方法中复杂的特征提取和数据重建的过程; 它的局部感受野、权值

共享和下采样特性保证了图像对位移、缩放和扭曲的鲁棒性。从功能上 CNN 可以看作是特征提取器与分类器的结合体, 深层结构和大量的学习参数决定了 CNN 是很好的特征提取器。CNN 在算法和结构上与 BP 神经网络都比较相似, 因此具有与 BP 神经网络相同的缺点, 如训练时间长、训练可能收敛到局部最优、容易出现过拟合现象等。为了使 CNN 变得更加完善、识别效果更好, 近来不少学者对 CNN 提出了结构改进和算法优化的策略。文献 [5] 提出一种基于孪生网络的深度神经网络模型的手写字符识别方法, 识别率达到 98%, 但是该模型没有很好地获取样本的差异化特征。文献 [6] 提出一种二进制 CNN 深度学习模型用于行人再识别, 并取得了很好地识别效果, 但该模型不能很好地学习样本的高级特征。文献 [7] 提出一种多通道输入的 CNN 改进方法。在进行网络训练之前, 先对输入图像进行多尺度超像素分割, 得到 3 个通道的输入, 再将 3 个通道的信息输入网络进行训练。实验证明, CNN 的多通道改进方法在图像识别、图像匹配等方面效果明显。文献 [8] 在利用 CNN 进行图像识别时, 提出将卷积核设定为加权 PCA 矩阵的形式, 完成隐层神经元之间的映射, 充分利用每层的映射结果, 通过生成码本的方式产生最终的特征向量。

收稿日期: 2017-12-13; 修回日期: 2018-01-08。

基金项目: 中央高校基本科研业务费专项资金 (2017IB014); 湖北省重点基金资助项目 (Grant NO. 2015CFA059)。

作者简介: 杜洋 (1990-), 女, 湖北省黄冈人, 硕士研究生, 主要从事图像处理, 模式识别方向的研究。

王展青 (1965-), 男, 湖北省仙桃人, 博士, 教授, 硕士生导师, 主要从事图像处理, 模式识别方向的研究。

传统 CNN 在完成特征提取后,多采用 RBF 或 softmax 分类器进行分类识别,但随着 SVM、稀疏表示和流形学习等浅层分类器的不断发展,它们的分类性能也有了很大的提高,因此研究者将 CNN 模型与性能更优的分类器进行结合。文献 [9] 提出将 CNN 与 SVM 相结合的方法进行手写数字识别,通过在 MNIST 数据集上的实验证明了 CNN 分类层的分类能力不如一些传统的分类器。文献 [10] 提出一种 CNN 层间特征融合的方法并与流形分类器相结合解决年龄识别问题。

鉴于此,本文以试卷卷头手写分数为例,提出一种基于改进 CNN 的算法 (CNN-1) 来研究手写数字识别问题。先通过多层特征融合和混合采样的方式,充分提取图像特征信息;再对训练算法进行改进,提高网络的训练速度,减少识别所需时间;最后通过 SVM 分类器进行分类识别,提高识别率。实验结果表明,该算法有效的减少了训练时间,提高了手写数字的识别率。

## 1 预处理

试卷卷头手写分数识别的首要环节是通过预处理获取单个手写分数字符。它的具体步骤包括灰度化、图像倾斜矫正、手写分数区域提取、二值化、去框线、去噪声、字符分割、归一化等,主要过程是手写分数区域的提取和字符分割。

### 1.1 主要预处理过程

主要预处理过程方法如下。

(1) 手写分数区域的提取。鉴于手写分数区域是表格框线结构,所以采用 Hough 变换<sup>[11]</sup>检测手写分数区域。提取手写分数区域时,考虑到书写随意性导致的字符出分数线下框线的情况,需提取手写分数出下框线部分的区域,保留字符信息。

(2) 字符分割。字符分割分为每题分数的分割和两位分数的分割,对于每题分数的分割选择与垂直框线相同的位置进行分割。由于教师书写的随意性,两位手写分数存在粘连和未粘连的情形。对于未粘连分数,采用连通域分割法<sup>[12]</sup>;对于粘连的分数,采用改进的滴水算法<sup>[13]</sup>,该算法首先是对字符的笔划粘连部分利用距离变换提取中心线,然后计算字符笔划的倾斜角度来指导水滴在中心线上端和下端的渗漏方向,最终形成一条分割路径,实现粘连字符的分割。

### 1.2 其它预处理过程

其它预处理过程方法如下。

(1) 灰度化、二值化、归一化。对原图像 RGB 三个分量加权平均得到灰度图像。采用全局阈值的 Ostu 算法实现最佳阈值 T 的自动选取,进行二值化处理。为了增加特征提取的准确性,将分割完成后的字符图像统一大小归一化到  $32 \times 32$  像素。

(2) 图像倾斜矫正、去框线、去噪声。由于人为操作等因素,通过摄像头获取试卷卷头图像时可能存在一定的倾斜,采用 Hough 变换检测直线并调整倾斜。同时检测手

写分数区域的水平框线和垂直框线,通过滤波去除图像框线。框线去除后,与框线存在粘连的字符会在粘连处发生断裂,根据框线与字符在粘连处的行值和列值相等的特点对断裂字符进行补充。对于手写分数区域离散噪声点的消除,求待测像素  $f(x,y)$  周围 8 个像素的平均值  $a$ ,若  $|f(x,y) - a|$  的值大于或等于 127.5,则该点是噪声点,将其值设置为 255,若小于 127.5,则保持原值不变。

## 2 基于改进 CNN 的手写数字识别

### 2.1 CNN 基本原理

如图 1 所示,典型的 CNN 结构主要由输入层、卷积层 (C1、C3、C5)、池化层或下采样层 (S2、S4)、全连接层 (F 层) 和输出层组成。

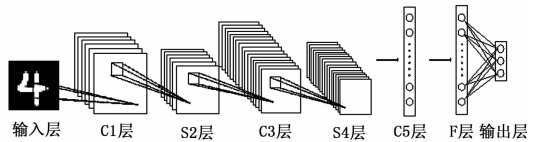


图 1 典型的卷积神经网络结构图

CNN 通过“卷积+下采样”的结构进行特征提取。输入层为单个手写数字图像  $T$ ,用  $X_i$  表示网络第  $i$  层的特征图 ( $X_0 = T$ )。若  $X_i$  为卷积层,则  $X_i$  表示为:

$$X_i = f(X_{i-1} * K_i + b_i) \quad (1)$$

其中  $K_i$  为可训练的卷积核,  $b_i$  为偏置,“\*”表示卷积运算,  $f(x)$  为激活函数。

常用的激活函数有 sigmoid 函数、tanh 函数和 ReLu 函数。本文选用 ReLu 函数<sup>[14]</sup>作为卷积层的激活函数,其表达式为:

$$f(x) = \max(0, x) \quad (2)$$

ReLu 函数较其它激活函数的优点是:(1) 它使网络一部分神经元的输出为 0,从而使网络具有一定的稀疏性、减少计算复杂度;(2) 使网络具有更强的分类能力,减小了参数间的相互依赖性,缓解了过拟合问题的发生;(3) 它的导数只为 0 或 1,在误差反向传播时较少出现梯度消失。

下采样层是对卷积层得到的每个特征图中  $n \times n$  区域内的像素进行采样操作,使特征图的尺寸减小。常见的采样操作有均值采样和最大值采样。若  $X_i$  为下采样层,则均值采样和最大值采样表达式分别为:

$$X_i = f(W_i P_{mean}(X_{i-1}) + b_i) \quad (3)$$

$$X_i = f(W_i P_{max}(X_{i-1}) + b_i) \quad (4)$$

其中:  $P_{mean}(x)$ 、 $P_{max}(x)$  分别表示均值采样操作和最大值采样操作,  $W_i$  为权值,  $b_i$  为偏置。

下采样的主要作用是:(1) 在保留原始特征信息的前提下,对特征图进行降维,避免陷入维数灾难,并且对平移、旋转等变形具有很好的鲁棒性;(2) 对特征图进行压缩,降低计算复杂度。

### 2.2 CNN 特征提取

本文对 CNN 中“卷积+下采样”的图像特征提取方式进行以下两个方面的改进:

(1) 传统 CNN 是图像经过一层层的映射, 映射到最后只将输出层作为特征提取结果, 这样在中间层就会丢失很多重要的图像特征, 从而影响识别率。本文受 Wang 等<sup>[10]</sup>层间特征融合思想的启发, 在输出图像特征提取结果时, 不仅仅只采用最后一层的输出结果, 而是对输入图像在网络中每一层的映射结果都进行输出, 再利用主成分分析 (Principal Component Analysis, PCA) 法<sup>[15]</sup>对提取的全部输出特征进行降维处理, 将其融合为多层深度特征。

(2) 在下采样过程中通过引入均值采样和最大值采样相结合的方式 (混合采样) 来增加采样层的多样性。这种方式可以兼顾到均值采样和最大值采样获得的特征值, 从而提取更丰富的图像特征, 稳定性较高, 混合采样表达式为:

$$T_i = \frac{1}{2}(P_{mean}(X_{i-1}) + P_{max}(X_{i-1})) \quad (5)$$

$$X_i = f(W_i T_i + b_i) \quad (6)$$

通过对特征提取方式的改进来弥补原来单一的输出层和采样层特征提取的不足, 实现图像特征的充分提取与融合, 改进的 CNN 特征提取结构, 如图 2 所示。

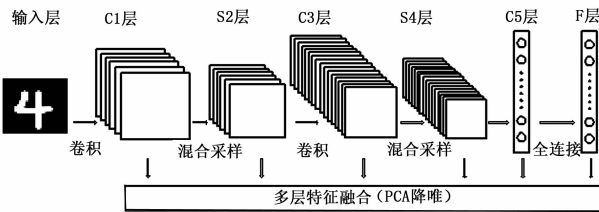


图 2 改进的 CNN 特征提取结构图

### 2.3 训练算法的优化

卷积神经网络的训练过程采用反向传播 (BP) 算法, CNN 通过正向计算和反向传播, 不断调整权值和阈值使网络的均方误差函数不断减小, 直至达到网络可以接受的精度要求。但是 BP 算法存在一些缺陷, 如训练效率不高、训练时间长、容易出现过拟合现象; 训练过程振荡、收敛速度慢、易陷入局部最优等。因此本文采用以下方法克服 BP 算法在网络训练过程中存在的缺陷:

(1) 在全连接层添加 dropout<sup>[16]</sup>随机隐退。在网络训练过程中, 随机的将一部分神经元的输出值以概率  $P$  设置为 0, 减少模型参数, 简化模型, 可以有效的防止过拟合现象, 提高网络泛化能力。

(2) 采用增加动量项的梯度下降法避免网络陷入局部最优。由于 BP 算法在权值调整过程中, 只按照  $t$  时刻的误差梯度方向调整, 没有考虑到  $t-1$  时刻的误差梯度方向, 从而在训练过程中出现动荡、收敛速度慢等现象。为此, 在权值调整公式中通过增加动量项来反映权值调整经验的积累, 含有动量项的权值调整<sup>[17]</sup>公式为:

$$\Delta W_{ij}(t) = \eta \delta_i p_j + \alpha \Delta W_{ij}(t-1) \quad (7)$$

其中  $\alpha$  为动量因子 ( $0 < \alpha < 1$ ),  $\eta$  为学习速率 ( $0 < \eta < 1$ ),  $\delta_i$  为误差项,  $p_j$  为节点的输出,  $\Delta W_{ij}(t)$  为第  $t$  次训练

后权值的调整量。从 (7) 式中可以看出, 增加动量项的作用是将前一次的权值调整量取一部分迭加到本次的权值调整量中, 避免出现局部最优。

(3) 学习速率自适应调整策略加快网络收敛速度。BP 算法收敛速度慢的一个重要原因是学习速率选择不当, 选择合适的学习速率, 可以加快网络的训练速度。本文通过学习速率自适应调整策略, 动态改变学习速率, 使学习速率能够根据误差的变化趋势自适应的调整大小。将 (7) 式中的学习速率  $\eta$  改进为以  $t$  为变量的时间序列函数, 则自适应学习速率  $\eta$  的调整公式为:

$$\eta(t+1) = \begin{cases} \beta_1 \cdot \eta(t), E(t) < E(t-1) \\ \beta_2 \cdot \eta(t), E(t) > \xi \cdot E(t-1) \\ \eta(t), E(t-1) \leq E(t) \leq \xi \cdot E(t-1) \end{cases} \quad (8)$$

其中  $\beta_1$ 、 $\beta_2$  分别为学习速率增加和减小的比例因子 ( $\beta_1 > 1$ 、 $0 < \beta_2 < 1$ ),  $\xi$  为允许反弹误差系数 ( $\xi > 1$ ),  $E(t)$  为网络误差。经过一次权值调整之后, 如果网络误差减小, 则在下次调整时可以增加学习速率, 从而加快网络收敛速度; 如果网络误差增大且大于前一次误差的  $\xi$  倍时则需要减小学习速率, 其它情况则保持学习速率不变。

### 2.4 本文算法

本文算法分为网络训练和图像识别两部分, 网络训练流程如图 3 所示。

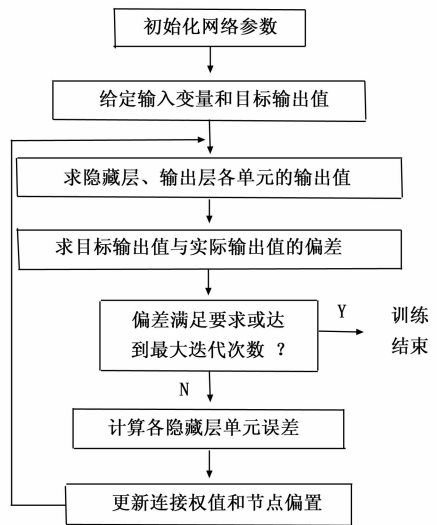


图 3 网络训练流程图

主要算法步骤如下。

(1) 将训练样本图片进行预处理后输入改进的卷积神经网络;

(2) 用符合高斯分布的小随机数对网络中的连接权值、阈值进行初始化, 并设定精度控制参数;

(3) 逐层计算网络各层的输出值:

$$a^{(l)} = f(z^{(l)}) \quad (9)$$

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \quad (10)$$

其中  $a^{(l)}$  表示第  $l$  层的输出值,  $z^{(l)}$  表示第  $l$  层的激活值,  $f(x)$  为激活函数,  $W^{(l)}$  为第  $l$  层的权值,  $b^{(l)}$  为第  $l$  层的偏置;

(4) 计算各网络层的残差  $\delta^{(l)}$ :

$$\delta^{(l)} = \begin{cases} -(y - a^{(l)}) \cdot f'(z^{(l)}), & l = n \\ ((W^{(l)})^T \delta^{(l+1)}) \cdot f'(z^{(l)}), & l = 2, 3, \dots, n-1 \end{cases} \quad (11)$$

其中  $y$  为目标输出值,  $f'(x)$  为激活函数的导数;

(5) 利用增加动量项的梯度下降法对网络参数权值  $W$  和偏置  $b$  进行微调, 公式如下:

$$\begin{cases} \Delta W_{ij}(t+1) = \eta \delta_i p_j + \alpha \Delta W_{ij}(t) \\ \Delta b_i(t+1) = \eta \delta_i + \alpha \Delta b_i(t) \end{cases} \quad (12)$$

其中  $\Delta W_{ij}$  为权值变化量,  $\Delta b_i$  为偏置变化量;

(6) 更新训练样本, 并判断是否达到要求的网络精度或迭代次数。若达到, 则网络训练完毕; 若没有达到, 则继续之前的步骤, 反复迭代训练网络参数, 直到满足要求;

(7) 将测试样本图片进行预处理后输入已经训练好的网络中, 并由 SVM 分类器得出分类结果。

### 2.5 SVM 分类识别

手写体数字识别是一个多分类问题, 包含 0 到 9 这 10 个数字, 常用的 SVM 多分类方法有“一对一”、“一对多”和有向无环图 (DAG) 法, 本文选用 SVM 一对一多分类法。

在训练  $i, j$  两类样本间的分类器时, 需要解决如下最优化问题:

$$\begin{aligned} & \min_{w, b, \xi} \frac{1}{2} \|w^{ij}\|^2 + C \sum_{i=1}^n \xi_i^j \\ & s. t. \begin{cases} w^{ij} \cdot \varphi(x_i) + b^{ij} \geq 1 - \xi_i^j, \text{ 若 } y_i = i \\ w^{ij} \cdot \varphi(x_i) + b^{ij} \leq -1 + \xi_i^j, \text{ 若 } y_i \neq i \\ \xi_i^j \geq 0, i, j = 1, 2, \dots, n \end{cases} \end{aligned} \quad (13)$$

其中  $w^{ij}$  为权重向量,  $C$  为惩罚因子,  $\xi_i^j$  为松弛变量,  $\varphi$  是将  $x_i$  映射到高维特征空间的映射函数,  $b^{ij}$  是偏置。

选取适当的核函数  $K(x, x) = (\varphi(x), \varphi(x))$ , 可以使 SVM 在高维核空间中通过分类函数将样本进行线性分类。常用的核函数有: 线性核函数、多项式核函数、sigmoid 核函数、RBF 核函数等, 本文选用 RBF 核作为 SVM 分类器的核函数:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (14)$$

采用“投票法”进行类别的判断: 对待测样本  $x$ , 依次用 45 个分类函数进行判别。  $i, j$  两类间的最优分类函数为  $f(x) = \text{sgn}((w^{ij})^T \varphi(x) + b^{ij})$ , 若  $f(x) \geq 0$ , 则  $i$  类得一票; 反之, 则  $j$  类得一票。当遍历完所有分类函数后, 根据得票数来确定待测样本  $x$  所属的类别。

### 2.6 算法复杂度分析

本文算法计算量主要来源于 CNN 的图像卷积过程, 图像尺寸、卷积核的数量与大小、卷积层的数量等因素都会影响算法的计算速度。

单个卷积核对一幅图像进行卷积操作计算量为:

$$S = \sum_{a=1}^{W-m+1} \sum_{b=1}^{K-n+1} \left( \sum_{i=1}^m \sum_{j=1}^n k_{ij} T_{a+i-1, b+j-1} \right) \quad (15)$$

其中  $k_{ij}$  表示大小为  $m \times n$  的卷积核,  $T$  表示大小为  $W \times$

$K$  的输入图像。

CNN 通常采用多个卷积层以及各层采用不同数量的卷积核来提取输入图像的不同特征, 当卷积层数量和卷积核数量增加时, 计算量也会增加。因此, CNN 对输入图像进行卷积操作计算量为:

$$C = \sum_{p=1}^P \prod_{q=1}^Q S_{pq} \quad (16)$$

其中:  $P$  表示卷积层数量,  $Q$  表示卷积核数量。

## 3 仿真实验及性能分析

为了验证算法的性能, 实验分为两部分: 在自制试卷卷头手写分数数据集上进行研究性实验和在 MNIST 数据集上进行验证性实验。实验环境为 Windows 10 操作系统, 仿真平台为 MATLAB R2014a,

各种算法的实验结果均为运行 10 次取平均值。

### 3.1 参数设置及评价指标

为了提高识别率和保证较好的识别效果, 结合文献 [18] 中的经验参数和多次实验结果, 实验参数设为: 动量因子  $\alpha = 0.5$ , 初始自适应学习速率  $\eta = 0.01$ , 学习速率增加的比例因子  $\beta_1 = 1.05$ , 学习速率减小的比例因子  $\beta_2 = 0.7$ , 允许反弹误差系数  $\xi = 1.04$ , 惩罚因子  $C = 128$ , 核参数  $\delta = 0.2$ 。

实验是在网络的不同迭代次数下, 通过识别率和运行时间 (训练时间和测试时间) 来评价算法的性能。识别率计算公式如下:

$$\text{rate} = \frac{n_c}{N} \times 100\% \quad (17)$$

其中:  $n_c$  为测试时正确识别的样本数,  $N$  为测试时识别的样本总数。

### 3.2 算法比较及结果评价

实验 1 为了研究算法的性能, 进行自制样本集试卷卷头的手写分数识别。首先通过摄像头获取试卷卷头样本图像并对其进行预处理得到单个手写分数样本集。预处理后的样本大小为  $32 \times 32$ , 部分预处理及样本集图像如图 4 (a) - (h) 所示。

本实验采用 2 000 张图片作为训练样本, 1 000 张图片作为测试样本。为了探究不同的采样操作和 dropout 层对实验的影响, 在本文算法 (CNN-1) 其它件都相同的情况下, 分别在采样层采用均值采样 (CNN-2) 和最大值采样 (CNN-3), 在采样层采用混合采样但在全连接层不添加 dropout (CNN-4), 将以上方法进行对比实验, 实验结果如表 1 和图 5 所示。

表 1 几种不同方法下的实验结果比较

实验方法	识别率/%	运行时间/s
CNN-1	98.92	72.3
CNN-2	93.48	71.8
CNN-3	95.55	69.1
CNN-4	98.80	88.5

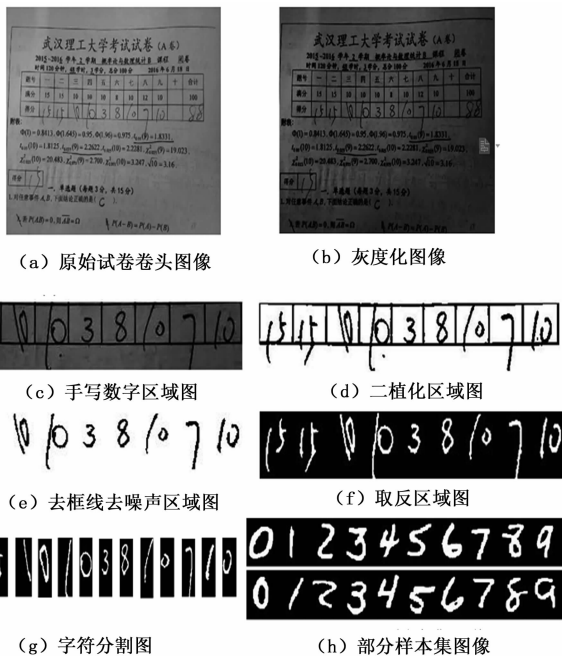


图 4 部分预处理和样本集图像

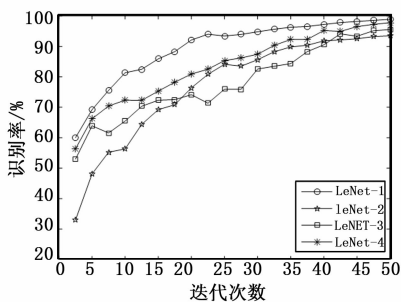


图 5 不同算法的识别结果对比图

由表 1 和图 5 的结果可知：CNN-1、CNN-2、CNN-3 的运行时间相差不多，但 CNN-1 的识别率却明显高于 CNN-2 和 CNN-3，说明混合采样优于均值采样和最大值采样，可以更加充分的获取图像特征信息，提高识别率。实验时设置 CNN-1 中 dropout 层的 dropout\_ratio 值为 0.5，即经过 dropout 层后每个输出节点的值都以 50% 的概率置为 0，CNN-4 较 CNN-1 没有在全连接层添加 dropout 随机隐退，识别率比 CNN-1 低 0.12%，运行时间比 CNN-1 多 16.2s，说明 dropout 层对本实验的识别效果有一定的贡献。

由于不同的核函数和分类方法具有不同的适应性和识别效果，因此选用不同的多分类方法和核函数可以构造不同的 SVM 分类器。为了研究 SVM 多分类方法和核函数对识别率的影响，采用几种常用的多分类方法和核函数进行对比实验，结果如表 2 所示。

从表 2 中可以看出：不同多分类方法及核函数的选择对识别率和运行时间都有一定的影响。对于分类方法来说，采用一对一多分类方法比采用一对多和 DAG 时的识别率要

表 2 基于不同多分类方法与核函数的实验结果对比

核函数	多分类方法					
	一对一		一对多		DAG	
	识别率/%	运行时间/s	识别率/%	运行时间/s	识别率/%	运行时间/s
线性核	86.51	93.0	84.57	152.1	83.92	103.4
多项式核	86.46	104.1	84.25	153.5	84.64	123.5
sigmoid 核	93.13	112.7	88.23	172.1	90.42	152.3
RBF 核	98.02	87.9	92.63	132.3	93.51	115.4

高，而采用一对多和 DAG 分类方法的识别率相差不大，在运行时间上一对一多分类方法也具有一定的优势。对于同种分类方法不同的核函数来说，采用 RBF 核函数时的识别率最高，并且明显优于其它三种核函数。所以 SVM 采用一对一多分类法和 RBF 核函数相结合的策略识别效果最好。

实验 2 为了证明算法的普适性，在 MNIST 数据集上进行验证性实验。MNIST 数据集由手写数字 0 到 9 的图像组成，包含 6 万张训练样本和 1 万张测试样本，而且每张图像都经过尺度归一化，大小均为  $28 \times 28$  像素。下面将 CNN-1 算法与 LeNet-5、文献 [9] 的算法进行对比实验，实验结果如表 3 和图 6 所示。

表 3 CNN-1 与 LeNet-5、文献[9]算法

实验方法	识别率/%	运行时间/min
CNN-1	99.05	28.66
LeNet-5	88.59	49.43
文献[9]	94.18	37.14

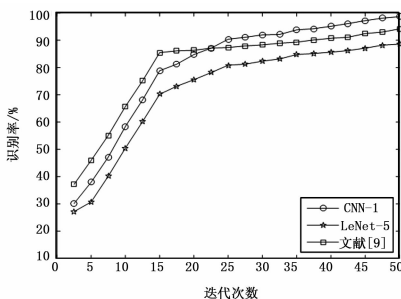


图 6 不同算法在 MNIST 上的识别结果对比图

LeNet-5 与文献 [9] 算法的区别在于分类层分别使用 softmax 和 SVM 分类器，由表 3 可知它们的运行时间分别为 49.43min 和 37.14min，使用 SVM 分类器的时间明显比 softmax 的短。这是因为对于 softmax 分类器来说，当某类的概率大于 0.9 时就意味着分类正确，但是损失函数会继续计算，直到概率接近于 1，这样会增加运行时间；SVM 分类器则更专注于分类错误的样本，对已经分类正确的样本不再处理，从而能大幅度的提高训练速度和识别率。由表 3 和图 6 可知，CNN-1 在文献 [9] 算法的基础上对网络结构和训练算法做的针对性改进，提高了识别率也减少了运行时间，更有利于解决手写数字的识别问题。实验证明：CNN-1 算法在 MNIST 数据集上相比其它算法同样具

有优越性, 从而验证了 CNN-1 算法的普适性和有效性。

#### 4 结语

本文针对试卷卷头的手写分数识别问题提出了改进卷积神经网络的手写数字识别算法。分别从网络特征提取方法和训练算法上对传统卷积神经网络进行了改进, 通过多层融合深度特征和混合采样层充分提取图像的特征和信息, 增加动量项和自适应学习速率的方式来减小训练时间、提高训练速度, 采用 SVM 分类器进行分类识别, 提高识别率。在 MNIST 数据集上进行验证性实验, 实验结果表明该算法有效的提高了网络的训练速度和识别率, 减少了运行时间。下一步工作将考虑如何通过多尺度池化的思想进一步改进卷积神经网络, 使网络能够接受任意尺寸的输入图像, 同时保持较好的识别效果和时效性。

#### 参考文献:

[1] 赵元庆, 吴 华. 多尺度特征和神经网络相融合的手写体数字识别 [J]. 计算机科学, 2013, 40 (8): 316 - 318.

[2] 任美丽, 孟 亮. 基于原型生成技术的手写体数字识别 [J]. 计算机工程与设计, 2015 (8): 2211 - 2216.

[3] 朱婷婷, 魏海坤, 张侃健. 基于 AP 和 BP 神经网络算法的手写数字识别 [J]. 中国科技论文, 2014 (4): 479 - 482.

[4] 方 向, 陈思佳, 贾 颖. 基于概率测度支持向量机的静态手写数字识别方法 [J]. 微电子学与计算机, 2015 (4): 107 - 110.

[5] Hosseiniasl E, GuhaA. Similarity - based text recognition by deeply supervised Siamese network [A]. Proceedings of Future Technologies Conference [C]. USA: IEEE Press, 2015: 1 - 7.

[6] Ahmed E, Jones M, Marks T K. An improved deep learning architecture for person re - identification [A]. Computer Vision and Pattern Recognition [C]. USA: IEEE Press, 2015: 3908 - 3916.

[7] S, Lau R W H, Liu W, et al. SuperCNN: A super pixel wise convolutional neural network for salient object detection [J]. International Journal of Computer Vision, 2015, 92 (3): 255 - 270.

[4] 李春虎, 罗光春, 李春豹. 基于斜帐篷混沌映射和 Arnold 变换的图像加密方案 [J]. 计算机应用研究, 2018, 35 (11): 1101 - 1109.

[5] 吕 群, 薛 伟. 结合混沌系统和动态 S - 盒的图像加密算法 [J]. 微电子学与计算机, 2017, 39 (3): 607 - 613.

[6] Ye G D, Zhao H Q, Chai H J. Chaotic image encryption algorithm using wave - line permutation and block diffusion [J]. Nonlinear Dynamics, 2016, 83 (4): 2067 - 2077.

[7] 杨建新, 王中叶. 双光束叠加与差异模的光学图像加密算法 [J]. 包装工程, 2017, 38 (23): 223 - 228.

[8] Wang Y, Quan C. Asymmetric optical image encryption based on an improved amplitude - phase retrieval algorithm [J]. Optics and Lasers in Engineering, 2016, 78 (10): 8 - 16.

115 (3): 330 - 344.

[8] Khalil - Hani M, Sung L S. A convolutional neural network approach for face verification [A]. 2014 International Conference on High Performance Computing & Simulation [C]. Bologna, Italy, 2014: 707 - 714.

[9] Niu X X, Suen C Y. A novel hybrid CNN - SVM classifier for recognizing handwritten digits [J]. Pattern Recognition, 2012, 45 (4): 1318 - 1325.

[10] Wang X, Guo R, Kambhamettu C. Deeply - learned feature for age estimation [A]. 2015 IEEE Winter Conference on Applications of Computer Vision [C]. USA, 2015: 534 - 541.

[11] 李慧鹏, 朱伟伟, 谭朦曦, 等. 基于改进 Hough 变换的直线检测算法 [J]. 半导体光电, 2017, 38 (4): 603 - 608.

[12] Shi X, Huang Y, Liu Y. Text on oracle rubbing segmentation method based on connected domain [A]. Advanced Information Management, Communicates, Electronic and Automation Control Conference [C]. IEEE, 2017: 414 - 418.

[13] 李兴国, 高 炜. 基于滴水算法的验证码中粘连字符分割方法 [J]. 计算机工程与应用, 2014, 50 (1): 163 - 166.

[14] Shopon M, Mohammed N, Abedin M A. Image augmentation by blocky artifact in deep convolutional neural network for handwritten digit recognition [A]. IEEE International Conference on Imaging, Vision & Pattern Recognition [C]. IEEE, 2017: 1 - 6.

[15] Shlens J. A tutorial on principal component analysis [J]. Statistics, 2014, 51 (3): 219 - 226.

[16] 李彦冬, 郝宗波, 雷 航. 卷积神经网络研究综述 [J]. 计算机应用, 2016, 36 (9): 2508 - 2515.

[17] 刘长征, 相文波. 基于改进卷积神经网络的肺炎影像判别 [J]. 计算机测量与控制, 2017, 25 (4): 185 - 188.

[18] Chowdhury A M S, Rahman M S. Towards optimal convolutional neural network parameters for bengali handwritten numerals recognition [A]. International Conference on Computer and Information Technology [C]. IEEE, 2017: 431 - 436.

[9] 刘效勇, 曹益平, 卢 佩. 基于压缩感知的光学图像加密技术研究 [J]. 光学技术, 2014, 34 (3): 91 - 99.

[10] Liu X Y, Cao Y P, Lu P. Optical image encryption technique based on compressed sensing and Arnold transformation [J]. Optik - International Journal for Light and Electron Optics, 2013, 124 (24): 6590 - 6593.

[11] Ronal Perez, Juan M. Vilarity, Carlos J. Jimenez. Non-linear image encryption system using the Gyrator transform and truncation operations [J]. Journal of Physics: Conference Series, 2017, 792 (1): 1012 - 1019.

[12] Juan M. Vilarity, Carlos J. Jimenez, Ronal Perez. Image encryption using the Gyrator transform and random phase masks generated by using chaos [J]. Journal of Physics: Conference Series, 2017, 850 (1): 12 - 22.