

文章编号 1004-924X(2021)07-1709-11

## 面向手写数字图像的压缩感知快速分类

肖术明<sup>1,2,3</sup>, 王绍举<sup>1,2,3\*</sup>, 常琳<sup>1,2,3</sup>, 冯汝鹏<sup>1,2,3</sup>

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100039;

3. 中国科学院 天基动态快速光学成像技术重点实验室, 吉林 长春 130033)

**摘要:**为了减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,同时保证手写数字图像的分类准确率,引入了压缩感知技术,提出了基于压缩感知和单隐层前馈网络(Compressive Sensing and Single Hidden Layer Feed-forward Network, CS-SHLNet)的手写数字图像快速分类算法。首先,利用高斯随机矩阵对具有稀疏性的手写数字图像进行线性观测,将高维图像信号投影到低维空间得到观测值;其次,通过误差反向传播(Error BackPropagation, BP)算法不断调整单隐层前馈网络权值建立适应于观测值的神经网络模型,将观测值嵌入神经网络中对图像进行特征提取;最后,采用单隐层前馈网络模型对手写数字进行图像分类,以训练计算耗时、最佳训练次数和分类准确率等指标对模型进行定量评估。实验结果表明:相比较单隐层神经网络和深度学习对 MNIST 手写数字数据集的高维图像信号图像分类,先通过 CS 技术利用观测数  $M=235$  的高斯随机矩阵线性观测得到图像的观测值,再利用单隐层前馈网络对观测值进行图像分类,网络模型的训练计算耗时缩短为 13.05 s,最佳训练次数缩短为 3 次,分类准确率保持 97.5%。该算法中的压缩感知线性观测可以有效减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,而且可以保证分类准确率。

**关键词:** 图像分类; 手写数字; 压缩感知; 神经网络

**中图分类号:** TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20212907.1709

## Compressive sensing fast classification for handwritten digital images

XIAO Shu-ming<sup>1,2,3</sup>, WANG Shao-ju<sup>1,2,3\*</sup>, CHANG Lin<sup>1,2,3</sup>, FENG Ru-peng<sup>1,2,3</sup>

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. University of Chinese Academy of Sciences, Beijing 100039, China;

3. Key Laboratory of Space-Based Dynamic & Rapid Optical Imaging Technology, Chinese Academy of Sciences, Changchun 130033, China)

\* Corresponding author, Email: wangshaoju@163.com

**Abstract:** To reduce the training calculation time and optimal training times of a neural network model and ensure high classification accuracy of handwritten digital images, compressive sensing technology was introduced, and a fast classification algorithm of handwritten digital images based on compressive sensing

收稿日期: 2020-12-02; 修订日期: 2021-01-06.

基金项目: 国家自然科学基金项目(No. 61805244); 吉林省重点科技研发计划项目(No. 20190303094SF); 钱学森空间技术实验室创新工作站开放基金(No. GZZKFJJ2020003)

and a single-hidden layer feedforward network (Compressive sensing and single-hidden layer feedforward network, CS-SHLNet) was proposed. First, a Gaussian random matrix is used to obtain a linear measurement of the handwritten digital image with sparseness, and the high-dimensional image signal is projected to the low-dimensional space to obtain the measurement value. Second, using the error backpropagation (BP) algorithm, the weights of the neural network are continuously adjusted to establish a single-hidden layer feedforward network model suitable for the measurement values, which are embedded into the neural network for image feature extraction. Finally, a single-hidden layer feedforward network model is used to classify handwritten digits, and the model is quantitatively evaluated by the time-consuming training calculations, the optimal training times, and the classification accuracy. Experimental results show that—in contrast to using a single-hidden layer neural network and deep learning for high-dimensional image signal classification of MNIST handwritten numeral datasets—through the CS technology, the Gaussian random matrix linear measurement number, i. e.,  $M=235$ , is first used to obtain the image measurement value; then, the single-hidden layer feedforward network is used for image classification. The training calculation time of the network model is reduced to 13.05 s, the best training times are reduced by a factor of three, and the classification accuracy is 97.5%. The compressive sensing linear measurement in the algorithm can effectively reduce the computation time of the training and the optimal training times of the neural network model for handwritten digital datasets and the classification accuracy can be ensured.

**Key words:** image classification; handwritten digits; compressive sensing; neural network

## 1 引 言

随着计算机技术的不断发展,图像分类算法取得了长足进步,并对人们的生活和社会的发展产生重要的影响。人机交互<sup>[1]</sup>、天文数据处理<sup>[2]</sup>、农产品成像<sup>[3]</sup>以及航空图像处理<sup>[4]</sup>等领域都离不开图像分类,因此图像分类算法有着重要的研究意义。手写数字图像分类算法指利用计算机图像处理技术将手写数字图像判定为0~9共10个类别的过程,该算法在财会、教育和金融等众多领域具有广泛的应用背景和使用价值。

基于神经网络模型的图像分类算法获得了很好的效果,使图像分类算法有了突破性的进展。对于手写数字图像分类问题也涌现了大量的相关研究工作:Hinton通过构造深层网络模型来学习有效特征的方法,完成了对手写数字集的认识,识别效果优于支持向量机(Support Vector Machine, SVM)。Lee提出了卷积深度置信网络,此网络能高效的学习和提取自然数字图像的特征。德国的Friedhelm Schweriker采用SV-RBF40对手写体数字进行识别,测试样本10000个,识别正确率为95.56%<sup>[5]</sup>。上海交通大学的团队采用压缩子结构特征的方式,用BP神经网络

分类,最终识别率为96.1%,误识率为1.04%。中科院自动化研究所的团队使用有限状态自动机方式对手写数字分类,识别率为95.2%。清华大学的团队利用SVM技术对邮政纸单上的手写数字做了识别工作,识别率为92%<sup>[6]</sup>。但是,上述算法都是利用网络模型对手写数字原始的高维数字图像进行图像分类,这极大地增加了网络模型训练计算耗时和最佳训练次数,同时还会影响图像分类的准确率,现有的研究<sup>[7-11]</sup>也都是对原始的高维数字图像进行处理。

近年来,针对信息量庞大的信号(图像)预处理问题,提出了一种新的信息采集方式,即压缩感知(Compressive Sensing, CS)理论。该理论表明,当信号(图像)具有稀疏性或可压缩性时,通过采集少量的信号(图像)观测值就可以实现信号的准确或近似重构<sup>[12]</sup>。文献[13]在光学相关断层扫描(Optical Coherence Tomography, OCT)成像技术中引入CS技术,在压缩感知重构算法中结合分块思想、引入正则项 $l_p$ 范数以及嵌入各项异性平滑算子,该改进的重构算法可以在一定程度上实现频域OCT图像重构效率和重构质量的平衡。文献[14]中提出了一种基于投影收缩的压缩感知锥束CT短扫描重建算法,该算

法将校正过程引入到压缩感知的图像重建中,有效抑制了条状伪影,清晰重建出边缘细节,极大提高了少量投影数据重建图像的质量。但是,上述的研究都是针对压缩采样信号的重构问题,而对于图像分类问题,可以使用压缩感知技术利用高斯随机矩阵对具有稀疏性的数字图像进行线性观测,少量的信号(图像)观测值保留了全部信息(特征),从而减少后续图像分类的数据处理量。

针对手写数字数据集信息量大,使用神经网络模型对其进行图像分类时训练计算耗时长和最佳训练次数多的问题,引入压缩感知技术,提出了一种基于压缩感知和单隐层前馈网络(Compressive Sensing and Single Hidden Layer Feedforward Network, CS-SHLNet)的手写数字图像快速分类算法。利用高斯随机矩阵对具有稀疏性的手写数字图像进行线性观测,将高维图像信号投影到低维空间得到观测值。将拥有图像全部信息(特征)的观测值嵌入神经网络中对图像进行特征提取。采用神经网络模型对手写

数字进行图像分类。实验结果表明,该算法可以有效减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,同时保证分类准确率。

## 2 手写数字图像快速分类算法

MNIST是一个手写数字识别领域的标准评测数据集,数据库中有70 000张灰度图像,每张图像大小为 $28 \times 28$ ,即784维<sup>[15]</sup>。因此,整个数据集的信息量庞大,传统的手写数字图像分类往往直接利用神经网络对其进行特征提取,训练计算耗时长最佳训练次数多。针对这个问题,利用CS-SHLNet对手写数字图像快速分类:首先利用压缩感知理论对数据集进行线性观测,得到手写数字的低维空间观测值;然后,将手写数字的观测值作为神经网络模型的输入,利用网络模型进行特征提取;再利用神经网络的学习能力对图像标签之间的相关性进行挖掘,输出最终的图像类别。总体的流程图如图1所示。

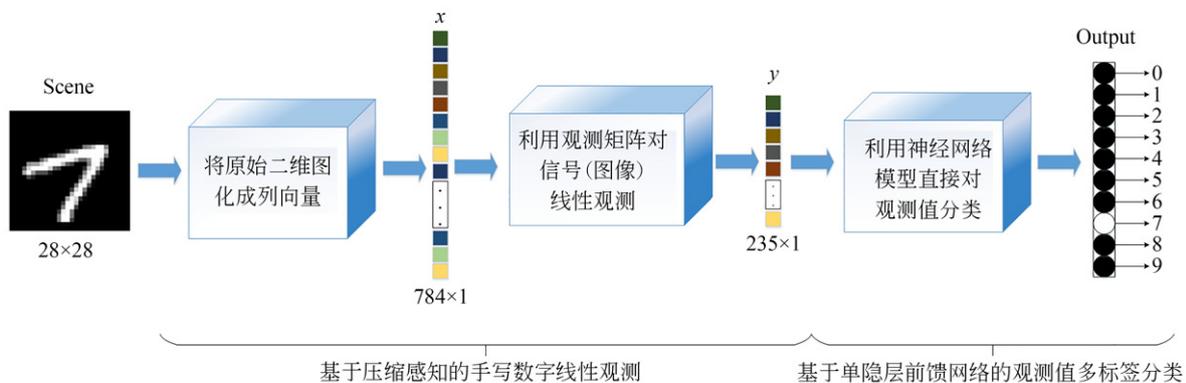


图1 手写数字图像分类流程图

Fig. 1 Flowchart of handwritten digital image classification

### 2.1 基于压缩感知的手写数字线性观测

手写数字数据集数据量庞大,为了减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,并且不影响图像分类准确率,利用压缩感知理论对数据集进行线性观测。首先将原始的二维图像列向量化,然后利用观测矩阵对信号线性观测得到观测值。

压缩感知理论假设 $N$ 维信号 $x_{N \times 1}$ 在空间正交基 $\{\psi_i\}_{i=1}^N$ 下的展开是 $K(K \ll N)$ 稀疏的<sup>[16]</sup>,即:

$$x_{N \times 1} = \psi_{N \times N} \theta_{N \times 1}, \quad (1)$$

其中 $\psi_{N \times N} = [\psi_1, \dots, \psi_N]$ 为稀疏矩阵; $\theta_{N \times 1}$ 为信号 $x_{N \times 1}$ 在稀疏域 $\psi_{N \times N}$ 下的展开系数。

如图2所示,MNIST手写数字数据集数字图片中非零数据数远小于总像素点数图像具有稀疏性,适合利用压缩感知进行图像分析。稀疏度指的是矩阵中非零元素的个数。首先对手写数字数据集图像进行二值化,然后进行稀疏度计算,为了结果的准确性,取所有数据集图像的稀

疏度平均值,计算得到稀疏度  $K$ 。

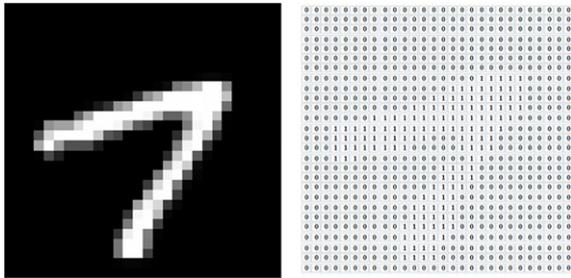


图2 MNIST手写数据集数字7数据特征图

Fig.2 MNIST handwritten data set figure 7 data feature diagram

采用一个与稀疏矩阵  $\psi_{N \times N}$  不相关的测量矩阵  $\Phi_{M \times N}$  对信号  $x_{N \times 1}$  进行线性观测,得到  $M$  ( $M \ll N$ ) 维观测值  $y_{M \times 1}$ ,即:

$$y_{M \times 1} = \Phi_{M \times N} x_{N \times 1} = \Phi_{M \times N} \psi_{N \times N} \theta_{N \times 1} = \Theta_{M \times N} \theta_{N \times 1}, \quad (2)$$

$$\begin{bmatrix} 0.0094 & -0.0685 & 0.0026 & \dots & -0.0553 & 0.0427 \\ -0.0840 & -0.0053 & 0.0376 & \dots & \dots & 0.0014 & -0.0788 \\ 0.0683 & 0.0426 & 0.0416 & \dots & \dots & -0.0690 & -0.0630 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0.0067 & 0.0227 & -0.0076 & \dots & 0.0061 & 0.0478 \\ -0.0472 & 0.0827 & 0.0513 & \dots & \dots & -0.0435 & -0.0767 \end{bmatrix}_{235 \times 784}$$

MNIST 手写数字数据集中的样本数字 5 的线性观测过程如图 3 所示。从线性观测的过程图可以看成,采用  $M=235$  的随机高斯矩阵线性观

其中,  $\Theta_{M \times N}$  为感知矩阵,  $y_{M \times 1}$  为稀疏信号  $\theta_{N \times 1}$  在感知矩阵  $\Theta_{M \times N}$  下的线性观测值。

随机高斯矩阵与任何稀疏矩阵都不相关,由约束等距性 (Restricted Isometry Property, RIP) 条件可知,它可以作为稀疏信号的观测矩阵,并且观测值会保留原始信号 (图像) 的特征。文献 [17] 中给出  $M \times N$  阶测量矩阵  $\Phi_{M \times N}$  的测量边界公式为:

$$M \geq C_0 K \log\left(\frac{N}{K}\right), \quad (3)$$

其中:  $C_0$  是一个常数  $C_0 < 1$ ,  $K$  为稀疏度。

由式 (3) 可知,当  $N=784$  时,可以计算得到  $M$  的取值范围为  $M \geq 87.48$ , 本文考虑到后续神经网络数据输入尽量少并且不能丢失图像分类所需特征等问题,选取  $M=235$ , 可合理实现使用较少的输入数据量,同时保留原始图像特征、保证图像分类准确率。随机高斯矩阵的结构式如下:

测,数据从 784 维减少到 235 维,数据量明显减少。

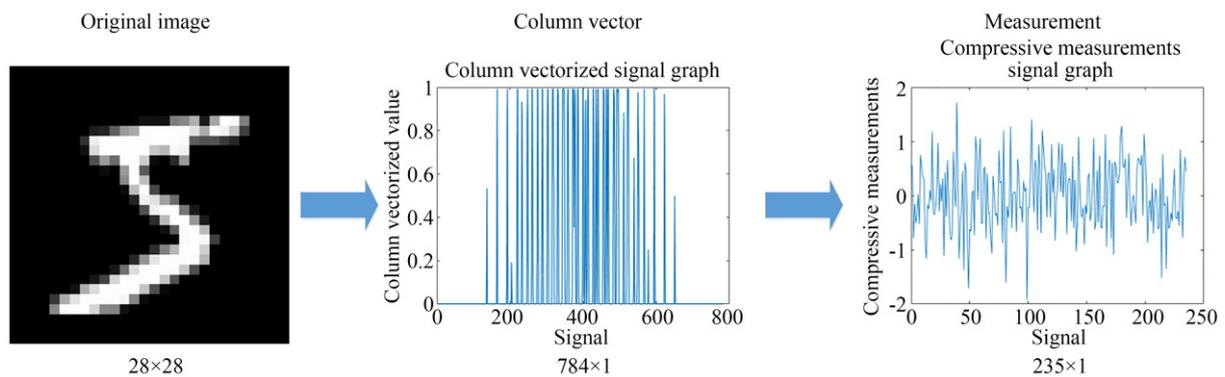


图3 MNIST 手写数字 5 的线性观测

Fig.3 Measurements of MNIST handwritten number 5

2.2 基于单隐层前馈网络的观测值多标签分类

对手写数字图像线性观测得到观测值后,需要将观测值嵌入神经网络中对图像进行特征提取,因此手写数字图像分类的第二步是建立适应于观测值的神经网络模型。

2.2.1 神经网络模型的构建

基于神经网络模型的图像分类算法中神经网络架构层出不穷,出现了很多复杂的网络结构<sup>[18-21]</sup>。考虑到网络架构的复杂度和训练耗时的影响,对于手写数字数据集的观测值多标签分类,采用单隐层前馈网络架构。设置神经网络模型为两个全连接层:第一层是 1 024 个神经元密集连接的带有 ELU 激活的全连接层,第二层是 10 个神经元密集连接的带有 Softmax 激活的全连接层。神经网络模型结构图如图 5 所示。神经网络结构参数如表 1 所示。

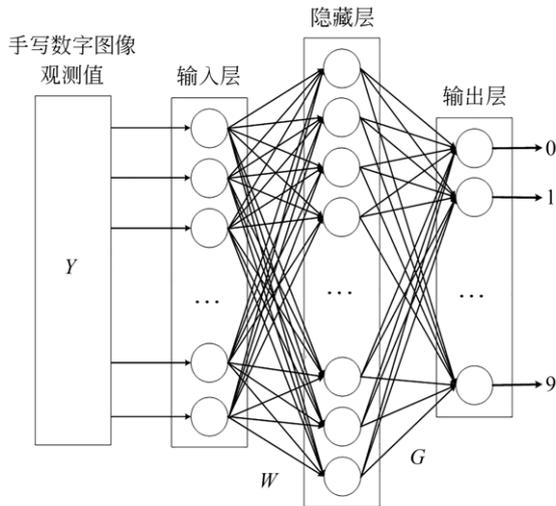


图 4 单隐层前馈神经网络模型结构图

Fig.4 Single hidden layer feedforward network model structure diagram

表 1 单隐层前馈神经网络结构参数表

Tab.1 Neural network structure parameter table

层名	输出尺寸	参数数量
dense_1 (Dense)	(None,1 024)	241 664
dense_2 (Dense)	(None,10)	10 250

神经网络的输入为  $Y=(y_1, y_2, \dots, y_{235})^T$ ,  $y_1, y_2, \dots, y_{235}$  为手写数字的观测值,  $n$  为输入层神经元个数。输入层输出向量为  $A=$

$(a_1, a_2, \dots, a_{235})^T$ ,  $a_1, a_2, \dots, a_{235}$  为每个输入层神经元的输出值; 隐含层输出向量为  $B=(b_1, b_2, \dots, b_m)^T$ ,  $b_1, b_2, \dots, b_m$  为每个隐含层神经元的输出值,  $m$  为隐含层神经元的个数; 输出层输出向量为  $C=(c_1, c_2, \dots, c_{10})^T$ ,  $c_1, c_2, \dots, c_{10}$  为每个输出层神经元的输出值。输入层到隐含层的权值矩阵为  $W_{m \times 235}$ , 其中  $w_{i,j}$  为权值矩阵的第  $i$  行、第  $j$  列元素; 隐含层到输出层的权值矩阵为  $G_{10 \times m}$ , 其中  $g_{p,q}$  为权值矩阵的第  $p$  行、第  $q$  列元素。

给定输入层神经元的权值为 1, 激励函数为  $f_a(y)=y$ , 其中  $y$  为输入, 则输入层的输出向量与训练样本相等, 即  $A=Y$ ; 隐藏层采用 ELU 激励函数  $f_b(a)=\begin{cases} a & a \geq 0 \\ \alpha(e^a - 1) & a < 0 \end{cases}$ ,  $\alpha$  为超参数; 输出

层采用 Softmax 激励函数  $S_v = \frac{e^{b_v}}{\sum_{i=1}^w e^{b_i}}$ ,  $w$  为输出层神经元个数。

此时, 神经网络得到的图像分类标签值就是输出层概率值最高的神经元所对应的分类结果。

2.2.2 神经网络模型的训练

为了调整神经网络的权值, 需要根据手写数字训练集的观测值对网络进行训练。如果只将数据标签以 0~9 的形式输入模型, 会出现不同类别元素之间距离不一的情况, 影响分类的准确率, 为了保证每个类别的样本彼此独立等距需要对标签进行 one-hot 编码。样本标签与 one-hot 编码的对应关系, 如表 2 所示。

表 2 样本标签与 one-hot 编码的关系表

Tab.2 Table of sample label and one-hot encoding

样本标签	one-hot 编码	样本标签	one-hot 编码
0	0 000 000 000	5	0 000 100 000
1	1 000 000 000	6	0 000 010 000
2	0 100 000 000	7	0 000 001 000
3	0 010 000 000	8	0 000 000 100
4	0 001 000 000	9	0 000 000 010

对于训练样例  $(y, lab)$ ,  $lab$  为训练样例的标签, 神经网络的输出为  $C=(c_1, c_2, \dots, c_{10})^T$ , 那么对于手写数字观测值, 神经网络模型得到的均方误差为:

$$E^{(l)} = \frac{1}{2} \sum_{h=1}^U (c_h^{(l)} - lab_h)^2, \quad (5)$$

其中:  $c_h^{(l)}$  为神经网络权值在  $l$  次调整后得到的手写数字图像分类标签值。

根据 BP 算法不断调整神经网络权值, 使得神经网络模型得到的均方误差  $E$  最小, 权值调整的流程如图 6 所示, 权值调整量公式为:

$$(\Delta g_{p,q}^{(l)}, \Delta \omega_{i,j}^{(l)}) = \left( -\eta \frac{\partial E^{(l)}}{\partial g_{p,q}^{(l)}}, -\eta \frac{\partial E^{(l)}}{\partial \omega_{i,j}^{(l)}} \right), \quad (6)$$

其中:  $\Delta g_{p,q}^{(l)}$  为权值矩阵  $G$  中第  $p$  行、第  $q$  列元素的变化量;  $\Delta \omega_{i,j}^{(l)}$  为权值矩阵  $W$  中第  $i$  行、第  $j$  列元素的变化量;  $l$  为权值调整次数;  $\eta$  为学习率,  $\eta \in [0, 1]$ 。

根据链式法则以及神经网络各层间的激励函数, 可将式(6)中的权值调整量整理为:

$$\begin{aligned} \Delta g_{p,q}^{(l)} &= -\eta (c_h^{(l)} - lab_h) n c_h^{(l)} (1 - c_h^{(l)}) b_i^{(l)} \\ \Delta \omega_{i,j}^{(l)} &= -\eta (c_h^{(l)} - lab_h) n c_h^{(l)} (1 - c_h^{(l)}) g_i b_i^{(l)}, \quad (7) \\ &\quad (1 - b_i^{(l)}) a_j^{(l)} \end{aligned}$$

那么调整之后的权值为:

$$(g_{p,q}^{(l+1)}, \omega_{i,j}^{(l+1)}) = (g_{p,q}^{(l)} + \Delta g_{p,q}^{(l)}, \omega_{i,j}^{(l)} + \Delta \omega_{i,j}^{(l)}), \quad (8)$$

当训练误差小于给定误差时, 网络模型停止训练, 根据式(4)可以得到手写数字图像分类标签值。

### 3 实验与结果分析

#### 3.1 神经网络模型的构建与训练

通过实验验证神经网络模型对 MNIST 手写数字图像分类的快速性和有效性, 分析了线性观测值对神经网络模型训练计算耗时、最佳训练次数和分类准确率的影响。

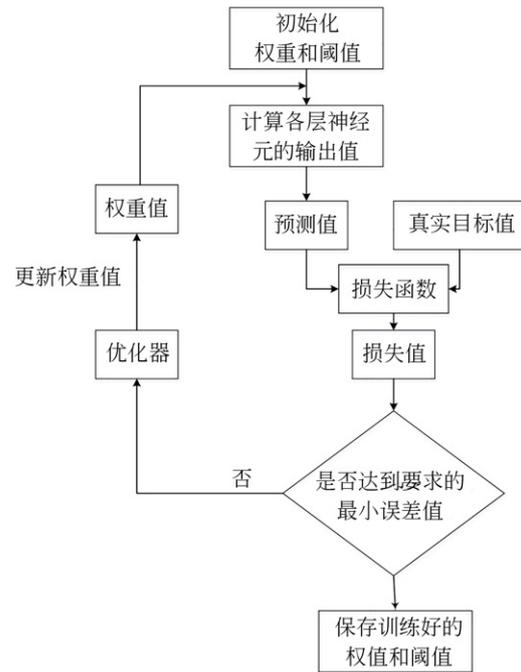


图 5 单隐层前馈网络模型权值调整流程图

Fig. 5 Flow chart of weight adjustment of single hidden layer feedforward network model

实验环境为 Windows 7 64 位操作系统、Intel (R) Core(TM) i7-7700 CPU @ 3.60 GHz 处理器、32.0G 内存、显卡 NVIDIA GeForce GTX 750Ti 的普通 PC 机, 编程环境 PyCharm Community Edition 2019.1.3 x64, 编程语言 Python 3.5.2 (64 位), 使用 TensorFlow\_GPU-1.4.0 的 Keras 高级 API 构建的一个神经网络模型, 来解析图像分类的实现过程。手写数字图像的分类是多分类问题, 神经网络模型的损失函数选择分类交叉熵 (Categorical Crossentropy), 优化器选择 rmsprop, 指标选择 accuracy。

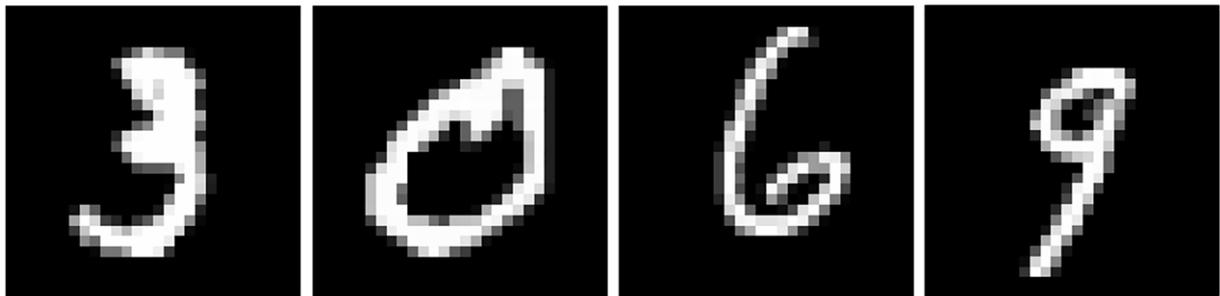


图 6 MNIST 手写数据集部分样本

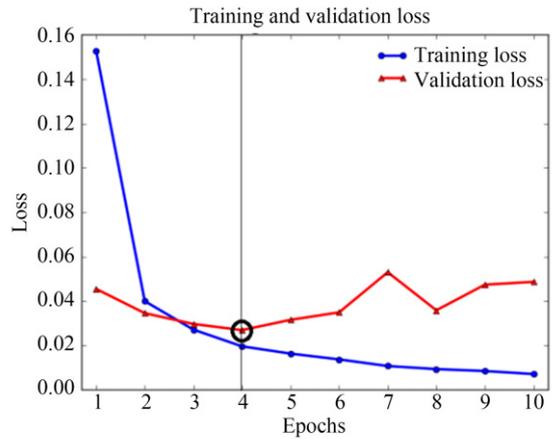
Fig. 6 Partial sample of MNIST handwritten data set

MNIST 数据库中有 70 000 张图像,一共包含 10 类阿拉伯数字,选择 50 000 张图像进行训练,10 000 张进行验证,10 000 张进行测试,MNIST 的每张图像大小为  $28 \times 28$ 。图 6 列出了本文中使用的部分数据集图像。

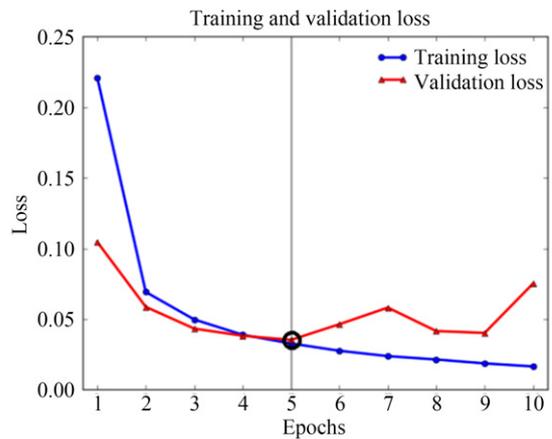
### 3.2 实验结果分析

为了评估本文算法的性能,将 CS-SHLNet 与单隐层神经网络、卷积神经网络(Convolutional Neural Networks, CNN)以及深度学习网络做了对比。其中单隐层神经网络的网络结构与本文 CS-SHLNet 中的神经网络部分结构相同;CNN 的网络结构主要由 3 个卷积层、2 个最大池化层和 2 个全连接层组成,卷积核大小为  $3 \times 3$ ,网络中加入 ReLU 激活函数;深度学习网络包括用于字符识别高效的 LeNet-5<sup>[22]</sup>和在 ImageNet 图像上训练的 VGG16<sup>[23]</sup>预训练模型。实验评价标准为最佳训练次数、训练用时和分类准确率。对手写数字数据集分别通过它们进行训练、验证和测试。

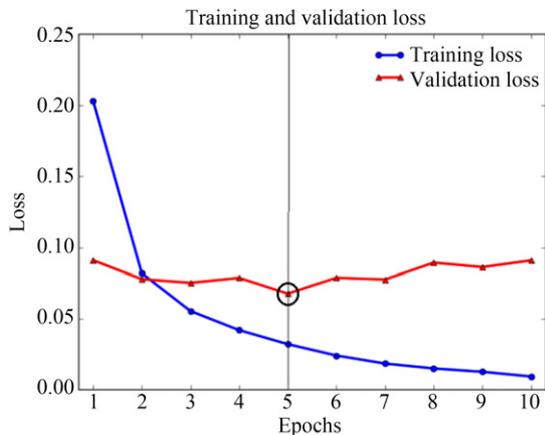
训练集和验证集的性能值(分类交叉熵)如图 7 所示,训练集分类交叉熵值每轮都在降低,这是梯度下降优化的预期结果,但是验证集分类交叉熵值在达到全局最小值后由于过拟合的原因会出现上升的情况。分类交叉熵值达到全局最小值时所对应的轮次为数据集在网络中的最佳训练次数。从图 7 中的验证集分类交叉熵值曲



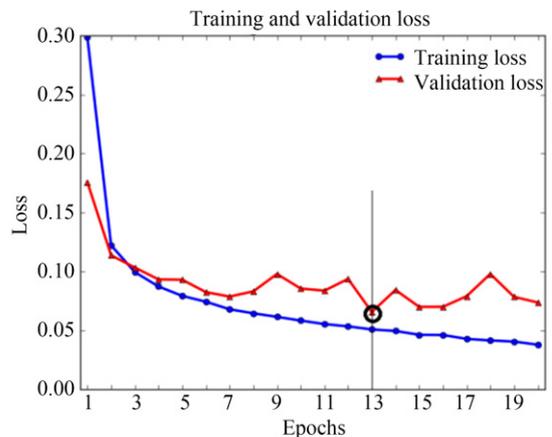
(b) CNN 分类交叉熵曲线图  
(b) Classification cross entropy of CNN



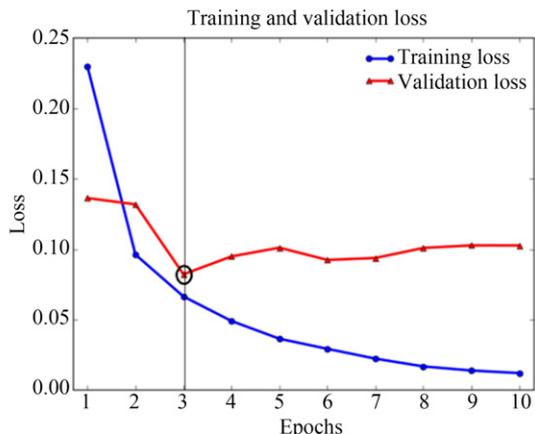
(c) LeNet-5 分类交叉熵曲线图  
(c) Classification cross entropy of LeNet-5



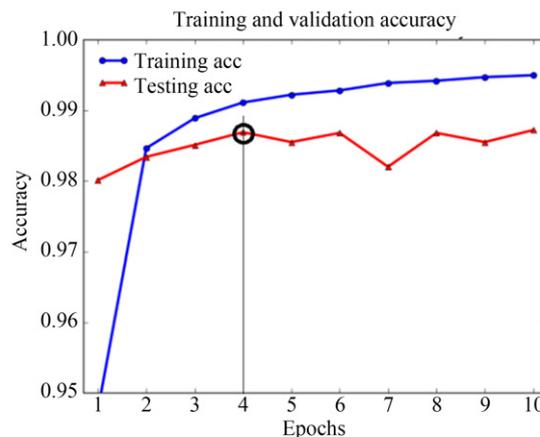
(a) 单隐层神经网络分类交叉熵曲线图  
(a) Classification cross entropy of the single hidden layer neural network



(d) VGG16 分类交叉熵曲线图  
(d) Classification cross entropy curve of VGG16



(e) CS-SHLNet分类交叉熵曲线图  
(e) Classification cross entropy curve of CS-SHLNet



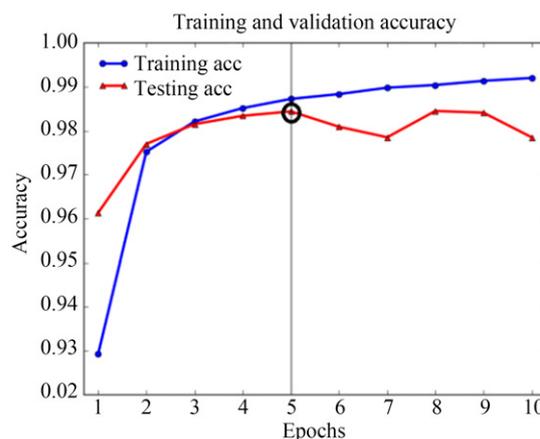
(b) CNN分类准确率曲线图  
(b) Classification accuracy curve of CNN

图 7 训练集和验证集分类交叉熵曲线图

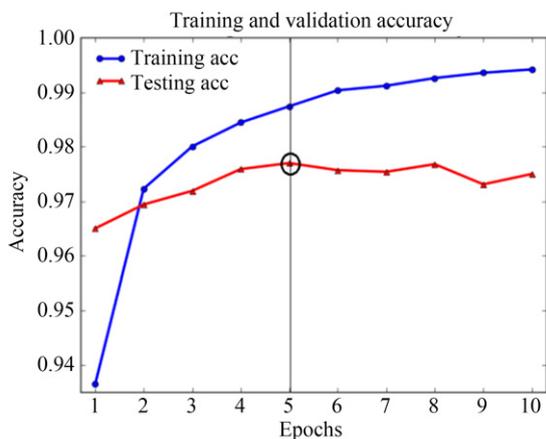
Fig.7 Classification cross-entropy curve diagram of training set and validation set

线可以得到,手写数字数据集的单隐层神经网络最佳训练次数为 5 次,CNN 最佳训练次数为 4 次,LeNet-5 最佳训练次数为 5 次,VGG16 最佳训练次数为 13 次;手写数字的 CS-SHLNet 最佳训练次数为 3 次。

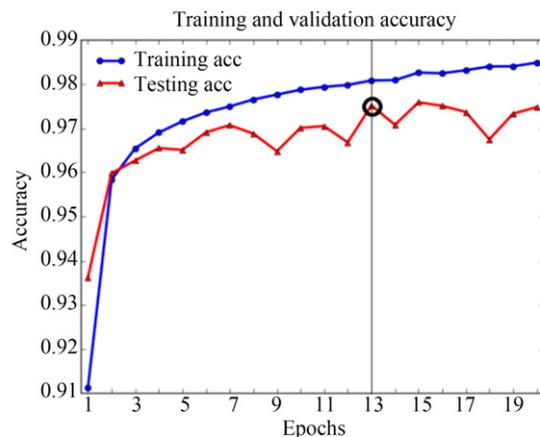
训练集和测试集的测试精确率如图 8 所示,测试集的精确率开始时保持上升趋势,在最佳训练次数时达到最大值。从图 8 中的测试集测试精确率曲线可以得到,手写数字数据集的单隐层神经网络最高测试精确率为 97.54%,CNN 最高测



(c) LeNet-5 分类准确率曲线图  
(c) Classification accuracy curve of LeNet-5

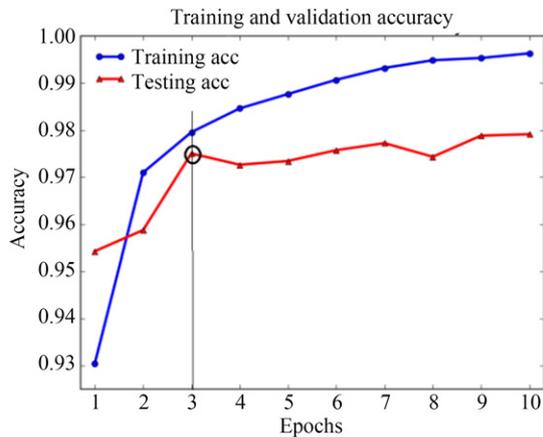


(a) 单隐层神经网络分类准确率曲线图  
(a) Classification accuracy curve of the single hidden layer neural network



(d) VGG16 分类准确率曲线图  
(d) Classification accuracy curve of VGG16

试精确率为 98.83%,LeNet-5最高测试精确率为 98.66%,VGG16最高测试精确率为 97.71%;手写数字的 CS-SHLNet 最高测试性能值为



(e) CS-SHLNet分类准确率曲线图  
(e) Classification accuracy curve of CS-SHLNet

图 8 训练集和测试集分类准确率曲线图

Fig. 8 Classification accuracy curve of training set and test set

97.53%,表明本文提出的算法在减少最佳训练次数、训练时间的情况下,分类准确率基本无影响。

表 3 为本文 CS-SHLNet 方法与其他方法的实验结果对比。从表 3 可以看出,网络模型的训练计算耗时缩短为 13.05s,最佳训练次数缩短为 3 次,分类准确率保持在 97.5%。表明该算法可以有效减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,并且不影响分类准确率。

### 4 结 论

本文针对手写数字数据集信息量庞大导致

表 3 实验结果对比

Tab. 3 Comparison of experimental results

实验数据集	算法	最佳训练次数	训练用时	分类准确率
MNIST 手写数字数据集	单隐层神经网络	5	25.92	97.54
	CNN	4	52.46	98.83
	LeNet-5	5	35.95	98.66
	VGG16	13	1219.99	97.71
	CS-SHLNet	3	13.05	97.53

神经网络模型对其训练计算耗时长、最佳训练次数多的问题,提出了基于 CS-SHLNet 的手写数字图像快速分类算法。通过压缩感知理论,利用高斯随机矩阵对具有稀疏性的手写数字图像进行线性观测,将高维图像信号投影到低维空间得到观测值。通过 BP 算法不断调整单隐层前馈网络权值建立适应于观测值的神经网络模型,将观测值嵌入神经网络中对图像进行特征提取。最后,采用单隐层前馈网络模型对手写数字进行图像分类。实验结果表明:相比较单隐层神经网络和深度学习对高维图像(信号)分类,利用 CS-SHLNet 对 MNIST 手写数字数据集进行图像分类,网络模型的训练计算耗时缩短为 13.05 s,最佳训练次数缩短为 3 次,分类准确率保持 97.5%。表明了该算法可以有效减少神经网络模型对手写数字数据集的训练计算耗时和最佳训练次数,同时保证分类准确率。但是该方法是在空域中基于压缩感知理论对手写数字图像进行线性观测,如何进一步减少图像的稀疏度,并且在减少数据量的同时保证特征数据不减少、分类精度进一步提高是下一步的研究重点。

### 参考文献:

[1] 雷安琪,冯玉田,唐子成. 基于 LBFGS-EKF 算法的三维空间目标跟踪研究[J]. 电子测量技术,2017, 40(10):99-103.  
LEI A Q, FENG Y T, TANG Z C. Research on tar-

get tracking based on limited-memory BFGS and extended kalman filter in 3D-space[J]. *Electronic Measurement Technology*, 2017, 40 (10) : 99-103. (in Chinese)  
[2] 刘丹枫. 深度学习用于天文图像空间碎片和恒星的分类[D]. 太原理工大学,2018.

- LIU D F. *Classification of Space Debris and Stars in Astronomic Images Based on Deep Learning*[D]. Taiyuan University of Technology., 2018. (in Chinese)
- [3] 彭俊. 基于卷积神经网络的交易环境下蔬果图像识别研究[D]. 浙江农林大学, 2017.  
PENG J. *Research of Fruit and Vegetable Image Classification in the Trading Environment Based on Convolutional Neural Network*[D]. Zhejiang Agriculture and Forestry University, 2017. (in Chinese)
- [4] 陈科峻, 张叶. 循环神经网络多标签航空图像分类[J]. 光学精密工程, 2020, 28(6):1404-1413.  
CHEN K J, ZHANU Y. Recurrent neural network multilabel aerial images classification[J]. *Opt. Precision Eng.*, 2020, 28(6):1404-1413. (in Chinese)
- [5] 白天毅. 基于神经网络的手写体数字识别关键技术研究[D]. 西北工业大学, 2014.  
BAI T Y. *Key Figures Handwriting Recognition Technology Based on Neural Networks*[D]. Northwestern Polytechnical University, 2014. (in Chinese)
- [6] 刘辰雨. 基于卷积神经网络的手写数字识别研究与设计[D]. 成都理工大学, 2018.  
LIU C Y. *Research and Design of Handwritten Digit Recognition Based on Convolutional Neural Network*[D]. Chengdu University of Technology, 2018. (in Chinese)
- [7] Support Vector Machines; Investigators at Semnan University Report Findings in Support Vector Machines (Persian Handwritten Digit Recognition Using Combination of Convolutional Neural Network and Support Vector Machine Methods)[J]. *Journal of Engineering*, 2020.
- [8] Yekta Said Can, Erdem KabadayıM. Automatic CNN-Based Arabic Numeral Spotting and Handwritten Digit Recognition by Using Deep Transfer Learning in Ottoman Population Registers[J]. *Applied Sciences*, 2020, 10(16).
- [9] Machine Learning; New Machine Learning Findings from Qassim University Described (Handwritten Digit Recognition: Hyperparameters-based Analysis)[J]. *Computers Networks & Communications*, 2020.
- [10] Hui-huang Zhao, Han Liu. Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition [J]. *Granular Computing*, 2020, 5(3).
- [11] 王爱丽, 薛冬, 吴海滨, 等. 基于条件生成对抗网络的手写数字识别[J]. 液晶与显示, 2020, 35(12): 1284-1290.  
WANG A L, XUE D, WU H B, *et al.* Handwritten digit recognition based on conditional generative adversarial network[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(12): 1284-1290. (in Chinese)
- [12] 叶浩. 基于压缩感知的深空图像压缩采样与恢复算法研究[D]. 重庆邮电大学, 2016.  
Ye H. *Research on Image Compressed Sampling and Recovery Algorithm Based on Compressive Sensing in Deep Space*[D]. Chongqing University of Posts and Telecommunications, 2016. (in Chinese)
- [13] 陈明惠, 王帆, 张晨曦, 等. 基于压缩感知的频域 OCT 图像稀疏重构[J]. 光学精密工程, 2020, 28(1):189-199.  
CHEN M H, WANG F, ZHANU C X, *et al.* Sparse reconstruction of frequency image based on compressed domain OCT sensing[J]. *Opt. Precision Eng.*, 2020, 28(1):189-199. (in Chinese)
- [14] 杨宏成, 高欣, 张涛. 应用投影收缩的压缩感知锥束 CT 短扫描重建[J]. 光学精密工程, 2014, 22(3):770-778.  
YANG H C, GAO X, ZHANG T. Compressing-sensing cone-beam CT short scan reconstruction based on projection-rcontraction [J]. *Opt. Precision Eng.*, 2014, 22(3):770-778. (in Chinese)
- [15] 黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述[J]. 计算机学报, 2014, 37(6):1225-1240.  
HUANU K Q, REN W Q, TAN T N. A review on image object classification and detection[J]. *Chinese Journal of Computers*, 2014, 37(6): 1225-1240. (in Chinese)
- [16] 刘洲洲, 王福豹. 基于离散萤火虫压缩感知重构的无线传感器网络多目标定位[J]. 光学精密工程, 2014, 22(7):1904-1911.  
LIU Z Z, WANG F B. Multiple target localization in WSNs via CS reconstrucyion based on discrete GSO algorithm[J]. *Opt. Precision Eng.*, 2014, 22(7):1904-1911. (in Chinese)
- [17] DAVENPORT M A. *Random observations on random observations: Sparse signal acquisition and processing*[D], 2010.
- [18] 马俊, 张荣福, 郭天茹, 等. 基于迁移学习的 VGG-16 网络芯片图像分类[J]. 光学仪器, 2020, 42(3): 21-27. (in Chinese)

- MA J, ZHANG R F, GUO T R, *et al.* Image classification of migration learning chip based on VGG-16 network[J]. *Optical Instruments*, 2020, 42(3): 21-27. (in Chinese)
- [19] 芦玥, 马瑜, 王慧, 等. 基于U-Net的多图谱标签融合算法[J]. *液晶与显示*, 2019, 34(11): 1091-1103.
- LU Y, MA Y, WANG H, *et al.* Multi-atlaslabel fusion based on U-Net[J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(11): 1091-1103. (in Chinese)
- [20] 陈彦彤, 陈伟楠, 张献中, 等. 基于深度卷积神经网络的蝇类面部识别[J]. *光学精密工程*, 2020, 28(7): 1558-1567.
- CHEN Y T, CHEN W N, ZHANG X Z, *et al.* Fly facial recognition based on deep convolutional neural network[J]. *Opt. Precision Eng.*, 2020, 28(7): 1558-1567. (in Chinese)
- [21] 余永维, 韩鑫, 杜柳青. 基于Inception-SSD算法的零件识别[J]. *光学精密工程*, 2020, 28(08): 1799-1809.
- YU Y W, HAN X, DU I Q. Target part recognition based Inception-SSD algorithm[J]. *Opt. Precision Eng.*, 2020, 28(8): 1799-1809. (in Chinese)
- [22] LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [23] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. *Arxiv Preprint Arxiv*: 1409. 1556, 2014.

## 作者简介:



肖术明(1997—),男,山东潍坊人,硕士研究生,2019年于东北农业大学获得学士学位,主要从事计算机视觉,压缩感知,机器学习方面的研究。E-mail:15636041235@163.com

## 通讯作者:



王绍举(1983—),男,山东聊城人,研究员,博士,2005年于东北师范大学获得学士学位,2010年于长春光机所获得博士学位,主要从事微纳卫星总体设计,空间光学相机控制和图像处理技术研究。E-mail:wangshaoju@163.com