第 18 卷　第 34 期　2018 年 12 月
1671—1815(2018)034-0078-07

科　学　技　术　与　工　程
Science Technology and Engineering

Vol. 18　No. 34　Dec. 2018
ⓒ 2018　Sci. Tech. Engrg.

引用格式:汪贵平, 盛广峰, 黄　鹤, 等. 基于改进 LeNet-5 网络的交通标志识别方法[J]. 科学技术与工程, 2018, 18(34): 78—84
Wang Guiping, Sheng Guangfeng, Huang He, et al. Traffic sign recognition method based on improved LeNet-5 network[J]. Science Technology and Engineering, 2018, 18(34): 78—84

基于改进 LeNet-5 网络的交通标志识别方法

汪贵平[1,2]　盛广峰[1]　黄　鹤[1,2*]　王会峰[1,2]　王　萍[1]
(长安大学电子与控制工程学院[1],陕西省道路交通智能检测与装备工程技术研究中心[2],西安 710064)

摘　要　针对传统 LeNet-5 卷积神经网络用于交通标志等多种类识别任务中,存在识别正确率低、网络容易过拟合以及梯度消失等问题进行改进。引入 Inception 卷积模块组来提取目标丰富的特征,同时增加网络的深度。引入 BN (batch normalization)层对输入批量样本进行规范化处理;同时改用性能更好的 Relu 激活函数,并使用全局池化层代替全连接层,合理改变卷积核的大小和数目。研究结果表明,改进 LeNet-5 网络能够有效解决过拟合和梯度消失等问题,具有较好的鲁棒性;网络识别率达到 98.5% 以上,相比 CNN (convolutional neural network) + SVM (support vector machine)提高了约 5%,比传统的 LeNet-5 网络提高了 3%。可见,改进后的 LeNet-5 网络图像识别的准确率得到显著提高。

关键词　交通标志　LeNet-5 网络　卷积神经网络　准确率
中图法分类号　TP391.4;　　**文献标志码**　A

随着城市现代化建设的发展,汽车保有量急剧增加,交通拥堵,交通事故频繁发生,因此提高车辆行驶的安全性已成为现代城市智能交通系统的首要问题[1]。近年来,交通标志识别逐渐得到国内外研究人员的广泛关注。目前,交通标志分类识别的主要算法包括统计分类法[2]、模板匹配法[3]、稀疏编码法[4]、神经网络法[5,6]和遗传算法[7]等。其中,基于卷积神经网络的交通标志识别算法因其自学习能力以及权值共享等优点得到国内外研究人员的高度关注。

LeNet-5 是一种经典的卷积神经网络结构[8,9],极大带动了 CNN(convolutional neural network)的发展。然而 LeNet-5 网络结构对交通标志图像等多类别目标进行分类识别时,准确率并不高[10]。针对这些问题,本文在 LeNet-5 及其他经典卷积神经网络分析的基础上,提出了一种新的改进 LeNet-5 网络模型。改进的 LeNet-5 网络引入了卷积模块组、并适当更改卷积核数目和卷积核大小,同时引入 BN(batch normalization)层对输入数据进行规范化处理,使用 Relu 激活函数替换 Sigmoid 函数并使用全局池化层替换全连接层,经过实验验证,达到预期效果。

2018 年 8 月 3 日收到　　国家自然科学基金(61402052)、陕西省科技计划[重点产业创新链(群)](2018ZDCXL-GY-05-04)和长安大学中央高校基本科研业务费专项资金(300102328204, 300102328101, 300102328501)资助

第一作者简介:汪贵平(1963—),男,汉族,湖北麻城人,博士,教授。E-mail:gpwang@chd.edu.cn。
*通信作者简介:黄　鹤(1970—),男,汉族,河南南阳人,博士,副教授。E-mail:8793141@qq.com。

1　卷积神经网络

1.1　卷积层

在卷积神经网络中,卷积层用于特征提取。上一层的 feature map 输入后,每个卷积核都与它作卷积,卷积核以一定步长在特征图上进行滑动,每滑动一次作一次卷积运算,最终就得到此层的一个 feature map,这样每个特征图就与上层的几个特征图建立了某种关系。每一个卷积核可提取一种特征,有 n 个卷积核就能提取 n 种特征,得到 n 个 feature map。一般卷积层的计算公式如下:

$$x_j^l = f\left(\sum_{i\in K_j} x_i^{l-1} * w_{ij}^l + b_j^l\right) \tag{1}$$

式(1)中:l 代表第几层;w_{ij} 代表卷积核;k_j 代表输入层的感受野。每个输出图有个偏置值 b。具体卷积过程如图 1 所示。

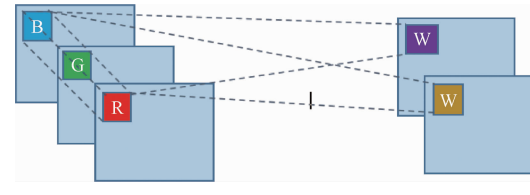


图 1　卷积示意图
Fig. 1　Convolution diagram

1.2　池化层

池化层也称为下采样层，其结构与卷积层有点类似，但不会生成更多的特征，有一个池化尺寸和池化步长。CNN 的池化方法很多，有均值池化（mean pooling）、最大值池化（max pooling）、重叠池化（overlapping）等方法，特征图经过池化后，特征维数降低，计算量减少。其中，均值池化和最大值池化是最常用的两种方法，计算公式如下：

$$AvePoolOut_i = f\left(\sum_{m=1}^{c}\sum_{n=1}^{c} ConvOut_i^{mn} + b_2\right) \tag{2}$$

$$MaxPoolOut_i = f\left(\sum_{m=1,n=1}^{c} ConvOut_i^{mn} + b_2\right) \tag{3}$$

式中，$ConvOut_i^{mn}$ 表示第 i 张输入特征图中 (m,n) 位置的元素，c 表示降采样池化域，b_2 为偏置。最大值池化和平均池化方式具体过程如图 2 所示。

1.3　BN 层

在卷积神经网络训练时，每一层的输入分布都会发生变化，这会使网络训练变得困难，所以必须使用较小的学习率来解决该问题。BN 层会对网络的每一个输入的批数量数据进行内部标准化处理，使输出规范化到 $N(0,1)$ 的正态分布。使用 BN 层可以提高学习率，在达到同等效果时，大大减小迭代次数，提升收敛速度。此外，添加 BN 层可以起到一定的正则化作用，防止过拟合和梯度消失等问题。

1.4　梯度下降算法

梯度下降算法是常用的优化算法，包括批量梯度下降法（BGD）、随机梯度下降法（SGD）、小批量梯度下降法（MBGD）。其目的是根据反向传播求出梯度，寻找梯度最小的地方，实现损失函数最小化。采用 MBGD 对样本进行迭代更新，如图 3 所示，使用固定数目的样本进行迭代更新，避免 BGD 样本更新速度较慢的问题以及 SGD 迭代更新过程中“噪音”较多的问题。

图 3 的黑色曲线即为 MBGD 下降时的轨迹，假设从点 x 开始，向下移动 Δh，将位置更新为 $x-\Delta h$，并且继续保持一致，直到达到底部，实现损失函数的最小化。

2　传统的 LeNet-5 网络

2.1　LeNet-5 网络结构

传统 LeNet-5 网络的结构如图 4 所示。

LeNet-5 网络结构一共 7 层，其中有 2 个卷积层、2 个池化层和 3 个全连接层。第 1 层是卷积层 Conv1，卷积核大小为 5×5，卷积核数目为 6，可以输出 6 个 28×28 的 feature map；第 2 层是池化层 Pool2，可以输出 6 个 14×14 的 feature map；第 3 层为卷积层 Conv3，卷积核大小为 5×5，卷积核数目为 16，输出 16 个 10×10 的 feature map；第 4 层为池化层 Pool4，输出 16 个 5×5 的 feature map；第 5、6、7 层都为全连接层，使用 Sigmoid 激活函数，其中 Fc1、Fc2、Fc3 的节点数目分别为 120、84、10，Fc3 的节点数目也是手写数字集 minist 的类别数。

2.2　在识别交通标志中存在的问题

LeNet-5 适用于手写数字图片的识别，但在识别交通标志图像时，损失值很高，识别率低，主要存在以下问题。

（1）交通标志图片尺寸为 64×64，比较大，使用原有的卷积核进行特征提取效果并不好。

（2）交通标志图片的背景、识别对象都较手写数字图片更为复杂，所以，使用传统网络识别率较低。

（3）手写数字种类只有 10 种，而常见交通标志种类有 62 种，种类更多，难度更大。

除了考虑以上问题，还需对网络其他部分进行调整、优化。为使整个网络性能得到更大程度的提

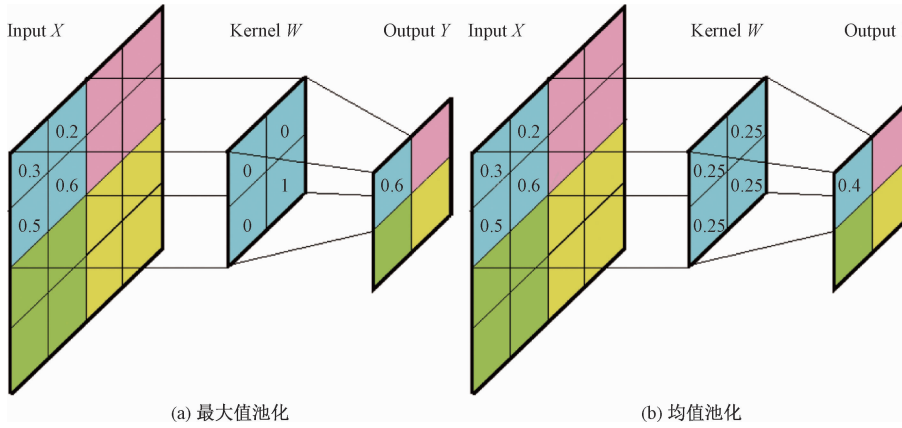


图 2　池化示意图

Fig. 2　Pooling diagram

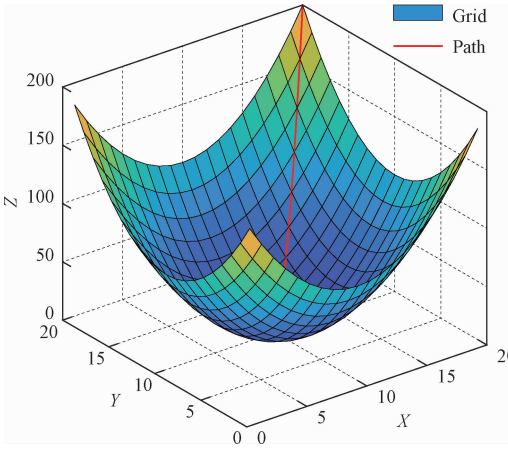


图 3 梯度下降示意图

Fig. 3 Gradient descent diagram

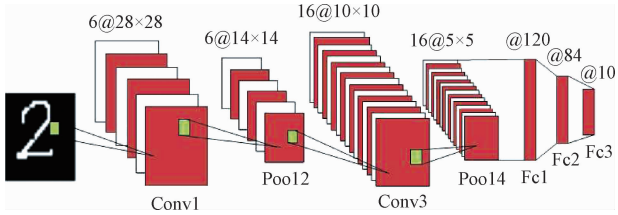


图 4 LeNet-5 网络结构图

Fig. 4 Lenet-5 network structure diagram

升,达到较高的准确率,必须对传统的 LeNet-5 网络进行改进。

3 基于改进 LeNet-5 网络的交通标志识别

3.1 LeNet-5 网络模型的改进

针对传统 LeNet-5 网络在识别交通标志出现的问题,做了如下改进。

(1)使用卷积模块组 Inception_conv3 和 Inception_conv5 分别取代 LeNet-5 网络第 3 层卷积层 Conv3 和第 5 层 Conv5,在提取到目标更多丰富特征的同时,增加网络深度。两个卷积模块组中分别包含 2 ~ 3 个卷积层。

(2)卷积核的选取关系着能否提取到有效特征,原有 5 × 5 的卷积核用于交通标志特征提取效果并不好,改变卷积核大小,在 Inception_conv3 和 Inception_conv5 模块组里用两个 3 × 3 卷积核取代原有卷积核。

(3)传统 LeNet-5 网络,每层卷积核数量较少,对于 62 分类的交通标志数据来说,不能充分提取到目标的丰富特征,因此,对卷积核数量进行合理改变。

(4)使用 BN 层,将输入的批样本进行规范化处理,改善神经网络的输入,在一定程度上可以提升网络训练速度,防止过拟合和梯度消失。

(5)传统 LeNet-5 采用的 Sigmod 函数不具有通用性,Relu 函数具有阻止负信号通过的特性,效率更高,可以提升收敛速度,所以选用效果更好的 Relu 激活函数。

(6)全连接层的参数占网络 80% 左右,这大大增加了网络训练时间。实验证明,全连接层对分类并没有效果提升,所以采用全局池化层 Max_pool8 来取代全连接层。

3.2 改进 LeNet-5 网络结构描述

改进 LeNet-5 网络结构图如图 5 所示。

输入层就是交通标志图片,尺寸为 64 × 64 的 RGB 图像。第 1 层是卷积层 Conv1,卷积核大小为 5 × 5,卷积核数目为 32,经过 Conv1 可以得到 32 个 60 × 60 的 feature map。第 2 层是最大值池化层 Max_pool2,池化尺寸为 2 × 2,步长为 2,经过 Max_pool2 可以得到 32 个 30 × 30 的 feature map。第 3 层是卷积模块组 Inception_conv3,有四个分支,分别为

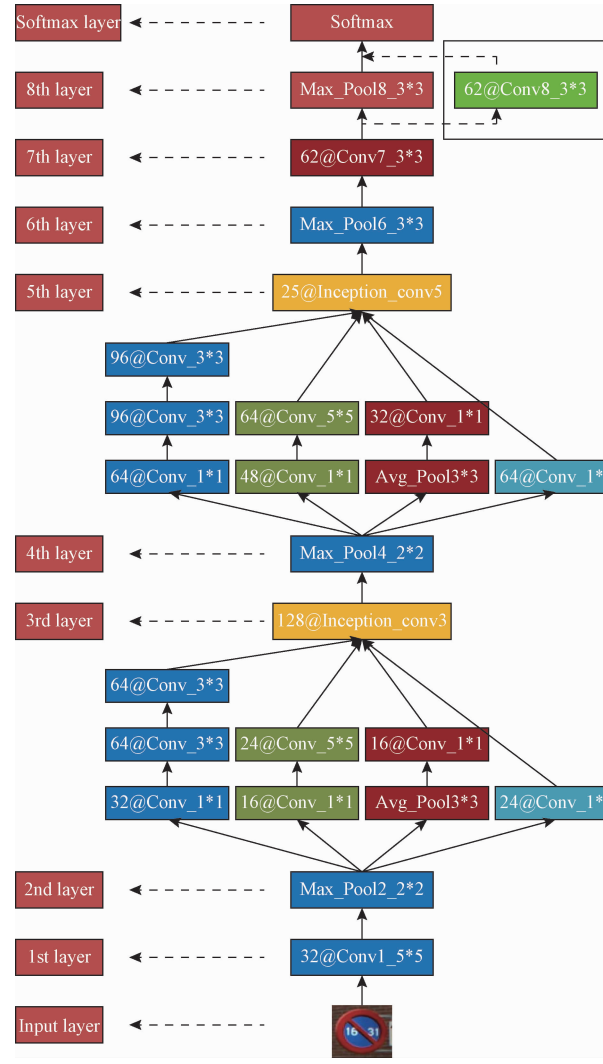


图 5 改进 LeNet-5 网络结构图

Fig. 5 Improve LeNet-5 network structure diagram

Branch_0、Branch_1、Branch_2、Branch_3,其中 Branch_0 使用 24 个 1×1 的卷积核;Branch_1 使用两个级联的卷积核,分别为 16 个 1×1 的卷积核和 24 个 5×5 的卷积核;Branch_2 将 Branch_1 中的5×5 的卷积核用两个 3×3 的卷积核替代,目的是减少参数量的同时提取到了目标更丰富的特征,即 Branch_2 由 32 个 1×1、64 个 3×3、64 个 3×3 的三级卷积核级联组成;Branch_3 由一个 3×3 的均值池化和 16 个 1×1 的卷积级联组成,经过卷积模块组 Inception_conv3 可以得到 128 个 30×30 的 feature map。第 4 层是一个最大值池化层 Max_pool4,最大值池化尺寸为 2×2,步长为 2,经过 Max_pool4 可以得到 128 个 15×15 的 feature map。第 5 层是卷积模块组 Inception_conv5,它与卷积模块组 Inception_conv3 的结构完全相同,同样有四个分支,分别为 Branch_0、Branch_1、Branch_2、Branch_3,不同的是卷积核数目发生了改变,分别为 64、64、96、32,经过卷积模块组 Inception_conv5,可以得到 256 个 15×15 的 feature map。第 6 层是最大值池化层 Max_pool6,池化尺寸为 3×3,步长为 3,经过 Max_pool6 可以得到 256 个 5×5 的 feature map。第 7 层是卷积层 Conv7,卷积核大小为 3×3,卷积核数目为 62,经过 Conv7 可以得到 62 个 3×3 的 feature map。第 8 层是全局池化层,池化尺寸为 3×3,将 62 个 3×3 的特征图池化为 1×1,组成特征向量输入到 Softmax 层。

图 5 中虚线框内的是卷积层 Conv8,卷积核大小为 3×3,卷积核数目为 62。尝试使用虚线框内的卷积层 Conv8 替换传统 LeNet-5 网络的全连接层进行网络改进,但是经过实验验证,引入 Conv8 的效果不如引入全局池化层 Max_pool8 的效果好。

改进 LeNet-5 网络参数如表 1 所示。改进后网络层数表面为 8 层,实际为 12 层(不包括输入层和 Softmax 层),因为卷积模块组 Inception_conv3 和 Inception_conv5 都为三层,增加了网络深度。

4　实验结果及分析

4.1　BelgiumTSC 数据集

为验证改进 LeNet-5 网络模型结构在交通标志识别方面的有效性,选用 BelgiumTSC 交通标志数据集作为实验数据进行对比实验。BelgiumTSC 是比利时一个公开的交通标志数据集,共有 62 种不同类型的交通标志,包含 3 773 个训练样本和 2 032个测试样本,图片大小为 64×64。数据集中的 62 类图片均为行驶车辆从真实道路场景拍摄到所得,涵盖了不同天气、背景、污损程度、光照条件的交通标志。

4.2　实验环境以及超参数的设置

实验基于 Linux 系统下的 Tensorflow 深度学习平台,处理器为 Intel core i7—7700k,主频 4.2 GHz,内存为 16 GB,显卡为 NVdia GTX1080TI,GPU 显存为 11 GB。

为了充分发挥 GPU 的性能,Mini-batch 大小设置为 2^n,Mini-batch 设为 256,迭代次数 Train_num 设为 1 000 次,初始学习率 Learn_rate 设为 0.05,使用指数衰减法对学习率进行衰减,衰减因子 Deacy 设为 0.96,BN 层的权重衰减项设为 0.005。

4.3　实验结果

实验利用 CNN + SVM(support vector machine)、传统 LeNet-5 网络、以及改进 LeNet-5 网络分别对 BelgiumTSC 交通标志数据集进行验证。

4.3.1　改进 LeNet-5 网络的可视化卷积特征图

为直观显示网络调优效果,利用可视化工具 Tensorboard,对卷积层 Conv2、Conv8 以及卷积模块组 Inception_conv3、Inception_conv5 处理后的特征图进行可视化显示,如图 6 所示。

表 1　改进 LeNet-5 网络参数

Table 1　Improve LeNet-5 network parameters

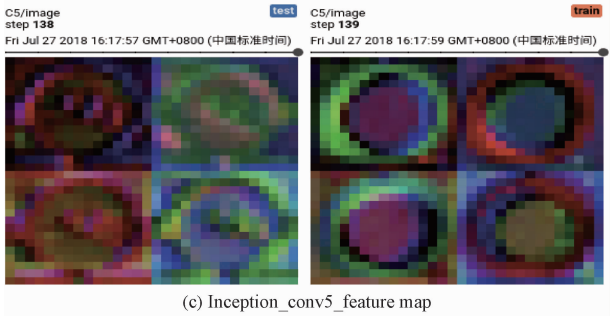
Layer 层	特征图数目	输出特征图尺寸	卷积核/池化尺寸	步长
Input	3	64×64	–	–
Conv1	32	60×60	5×5	1
Max_Pool2	32	30×30	2×2	2
Inception_conv3	24 + 24 + 64 + 16 = 128	30×30	–	1
Max_Pool4	128	15×15	2×2	2
Inception_conv5	64 + 64 + 96 + 32 = 256	15×15	–	1
Max_Pool6	256	5×5	3×3	3
Conv7	62	3×3	3×3	1
Max_Pool8	62	1×1	3×3	3



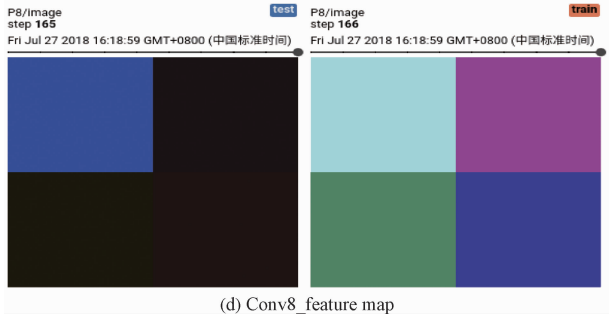
(a) Conv2_feature map



(b) Inception_conv3_feature map



(c) Inception_conv5_feature map



(d) Conv8_feature map

图 6 改进 LeNet-5 网络的可视化特征图

Fig. 6 Visual feature map of improve Lenet-5 network

图 6 中可以明显看到每一幅图像训练过程中的特征提取情况,越靠后的卷积层提取目标的特征更抽象,这是实验中使用 Inception 模块组替换 Conv1 和 Conv3 的效果不如替换 Conv3 和 Conv5 的效果好的原因。

4.3.2 不同迭代次数的训练结果

对 LeNet-5 网络加载数据训练时,将数据集随机分成 80% 的训练集和 20% 的验证集,同时对网络权重进行随机初始化,这会造成训练结果产生小幅度波动。所以训练网络时,进行了不同迭代次数的训练,每次迭代又采取三组实验,并求平均值,不同迭代次数的训练结果如表 2 所示。

表 2 不同迭代次数的训练结果

Table 2 Training results of different iterations

迭代次数	训练结果/%			
	第一次	第二次	第三次	平均值
135	98.622	98.334	98.538	98.498
150	98.338	98.784	98.417	98.513
200	98.546	98.536	98.423	98.501
215	98.156	98.714	98.664	98.511
155	98.742	98.546	98.223	98.503

由表 2 可知,5 组实验的平均结果均达到 98.5% 左右。训练结果虽有不同,但差别很细微,说明样本训练的收敛情况已趋于稳定,不会有大的波动。

4.3.3 算法实验结果比较

从表 3 结果中可以看出,改进后的 LeNet-5 网络的识别正确率高于传统的 LeNet-5 以及 CNN + SVM 网络的识别正确率。相比 CNN + SVM 及传统的 LeNet-5 网络分别提高了约 5% 和 3%,显著提高了图像识别的准确率。

表 3 不同网络的性能比较

Table 3 Performance comparison of different networks

网络类型	识别正确率/%
CNN + SVM	93.74
传统 LeNet-5	95.48
改进 LeNet-5	98.51

4.3.4 改进 LeNet-5 网络 loss 曲线和学习率衰减曲线

改进 LeNet-5 网络的 loss 曲线如图 7 所示。从图 7 中可以看出,在前 100 次迭代中 loss 曲线快速下降,随后趋于稳定,并逐渐接近于 0,说明改进 LeNet-5 网络已得到充分训练。网络的收敛速度非常快,这是因为 BN 层的使用加速了网络的训练过程,同时避免了过拟合和梯度消失的问题。

学习率衰减曲线的变化过程如图 8 所示。

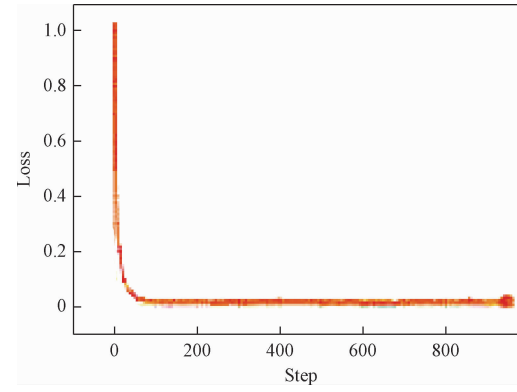


图 7 Loss 曲线

Fig. 7 Loss curve

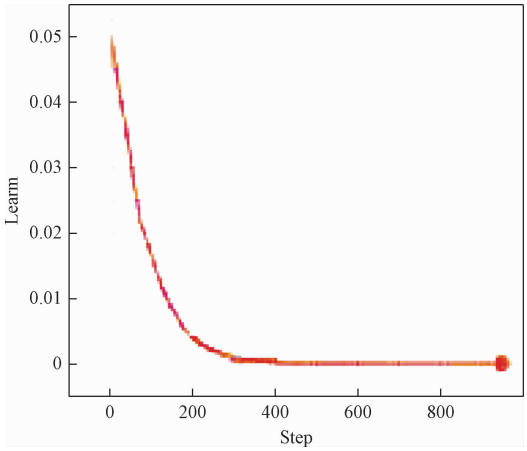


图 8　学习率衰减曲线

Fig. 8　Learning rate decay curve

从图 8 中可以看到学习率随着迭代次数的增加逐步衰减，最后趋近于 0。这样做的好处是避免了学习率过大可能导致的无法找到全局最优解，以及学习率过小导致训练时间较长的问题。

4.3.5　改进 LeNet-5 网络训练 Accuracy 曲线以及终端日志信息输出

图 9 中在前 100 次迭代中 Accuracy 快速上升，之后小幅度震荡，随着迭代次数的增加，趋于稳定，并逐渐接近于 100%，说明网络得到充分训练，网络调优效果较好。

从图 10 Linux 终端输出的日志信息中可以看到：迭代到 150 次的时候，网络已经得到充分训练，训练集正确率已经达到 100%，同时测试集正确率也达到 98.51%。

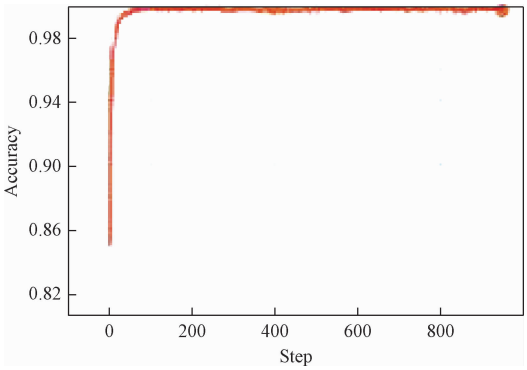


图 9　训练集正确率曲线

Fig. 9　Training set correct rate curve

```
After 145 training step(s), train loss 0.002761.
After 145 training step(s), train acc 1.000000.
After 150 training step(s), train loss 0.002746.
After 150 training step(s), train acc 1.000000.
After 150 testing step(s), test loss 0.144238.
After 150 testing step(s), test acc 0.985197.
```

图 10　Linux 终端输出的日志信息

Fig. 10　Log information output by the Linux terminal

5　结论

在对交通标志识别问题的研究上引入了卷积神经网络方法，并针对传统 LeNet-5 网络识别正确率低、全连接层参数繁多、容易过拟合和梯度消失等问题，在传统 LeNet-5 的基础上进行了改进。经过实验验证可以得到以下结论。

(1)改进网络在交通标志数据集 BelgiumTSC 上的识别率达到 98% 以上，与 CNN + SVM 网络、传统 LeNet-5 网络相比，改进网络正确率得到显著提升。

(2)通过 BN 层的使用，对输入的批量样本进行规范化处理以改善神经网络的输入，从而在加速网络训练过程的同时，有效避免了网络过拟合和梯度消失等问题。

(3)使用全局池化层代替参数繁多的全连接层，减少了网络的参数量。

(4)引入卷积模块组 Inception，增加网络深度的同时，增强了网络特征提取的能力。

参　考　文　献

1　党　倩，马　苗，陈昱莅. 基于二级改进的 LeNet-5 交通标志识别算法. 陕西师范大学学报，2017;45(2)：24—28
Dang Qian, Ma Miao, Cheng Yuli. Traffic sign recognition algorithm based on two-stage improved LeNet-5. Journal of Shaanxi Normal University, 2017;45(2): 24—28

2　郝　岩，白艳萍，张校非，等. 卷积神经网络在 SAR 目标识别中的应用. 重庆理工大学学报，2018;32(5)：204—209
Hao Yan, Bai Yanping, Zhang Xiaofei, *et al.* Application of convolution neural network in SAR target recognition. Journal of Chongqing University of Technology, 2018;32(5): 204—209

3　王　洋. 一种基于模板匹配的交通标志识别方法. 长春：吉林大学，2013
Wang Yang. A traffic sign recognition method based on template matching. Changchun: Jilin University, 2013

4　姜嘉琳. 基于结构稀疏的目标识别方法. 西安：电子科技大学，2013
Jiang Jialin. Target recognition method based on sparse structure. Xi'an: University of Electronic Science and Technology, 2013

5　李新叶，黄　腾. 基于多尺度跃层卷积神经网络的精细车型识别. 科学技术与工程，2017;17(11)：246—249
Li Xinye, Huang Teng. Fine vehicle identification based on multi-scale convolution neural network. Science Technology and Engineering, 2017;17(11): 246—249

6　胡　杰，李少波，于丽娅，等. 基于卷积神经网络与随机森林算法的专利文本分类模型. 科学技术与工程，2018;18(6)：268—272
Hu Jie, Li Shaobo, Yu Liya, *et al.* Patent text classification model based on convolutional neural network and random forest algorithm. Science Technology and Engineering, 2018;18(6): 268—272

7　李思琴，林　磊，孙承杰. 基于卷积神经网络的搜索广告点击率预测. 智能计算机与应用，2015;5(5)：22—28

Li Siqin, Lin Lei, Sun Chengjie. Click rate prediction of search ads based on convolution neural network. Intelligent Computers and Applications, 2015;5(5): 22—28

8 吴阳阳, 彭广德, 吴相飞. 基于 LeNet-5 改进的卷积神经网络图像识别方法. 信息与电脑, 2018;(7): 127—130
Wu Yangyang, Peng Guangdei, Wu Xiangfei. Image recognition method based on improved LeNet-5 convolution neural network. Information and Computer, 2018;(7): 127—130

9 王秀席, 王茂宁, 张建伟, 等. 基于改进的卷积神经网络 LeNet-5 的车型识别方法. 计算机应用研究, 2018;35(7): 2215—2218
Wang Xiuxi, Wang Maoning, Zhang Jianwei, *et al*. Vehicle identification method based on improved convolutional neural network LeNet-5. Computer Application Research, 2018;35(7): 2215—2218

10 李 丹, 沈夏炯, 张海香. 基于 Lenet-5 的卷积神经网络改进算法. 计算机时代,2016;30(8): 4—6,12
Li Dan, Shen Xiajiong, Zhang Haixiang. Improved convolution neural network algorithm based on lenet-5. Computer Generation, 2016;30(8): 4—6,12

Traffic Sign Recognition Method Based on Improved LeNet-5 Network

WANG Gui-ping[1,2], SHENG Guang-feng[1], HUANG He[1,2*], WANG Hui-feng[1,2], WANG Ping[1]

(School of Electronic and Control Engineering, Chang'an University[1], Shaanxi Road Traffic Intelligent Detection and Equipment Engineering Technology Research Center[2], Xi'an 710064, China)

[**Abstract**] For the traditional LeNet-5 convolution neural network used in traffic signs and other kinds of recognition tasks, the problems such as low recognition accuracy, easy network over-fitting and gradient disappearance are improved. Inception convolution module group was cited to extract rich features of the target while increasing the depth of the network, the BN (batch normalization) layer was introduced to normalize the input batch samples to improve the input of the neural network. At the same time, the better Relu activation function was used, and the global pooling layer was used instead of the full connection layer, and the size and number of convolution kernels were reasonably changed. The research results show that the improved LeNet-5 network can effectively solve the problems of over-fitting and gradient disappearance, and has better robustness. At the same time, compared with CNN (convolutional neural network) + SVM (support vector machine) and traditional LeNet-5 network, the accuracy of the improved network classification can be up to 98.5%, which is 5% higher than CNN + SVM and 3% higher than traditional LeNet-5 network. The accuracy of image recognition is improved significantly.

[**Key words**] traffic signs LeNet-5 net work convolutional neural network accuracy