

基于优化卷积神经网络结构的交通标志识别

王晓斌¹, 黄金杰¹, 刘文举^{2*}

(1. 哈尔滨理工大学 自动化学院, 哈尔滨 150080; 2. 中国科学院 自动化研究所, 北京 100190)

(* 通信作者电子邮箱 lwj@nlpr.ia.ac.cn)

摘要: 现有算法对交通标志进行识别时, 存在训练时间短但识别率低, 或识别率高但训练时间长的问題。为此, 综合批量归一化(BN)方法、逐层贪婪预训练(GLP)方法, 以及把分类器换成支持向量机(SVM)这三种方法对卷积神经网络(CNN)结构进行优化, 提出基于优化CNN结构的交通标志识别算法。其中: BN方法可以用来改变中间层的数据分布情况, 把卷积层输出数据归一化为均值为0、方差为1, 从而提高训练收敛速度, 减少训练时间; GLP方法则是先训练第一层卷积网络, 训练完把参数保留, 继续训练第二层, 保留参数, 直到把所有卷积层训练完毕, 这样可以有效提高卷积网络识别率; SVM分类器只专注于那些分类错误的样本, 对已经分类正确的样本不再处理, 从而提高了训练速度。使用德国交通标志识别数据库进行训练和识别, 新算法的训练时间相对于传统CNN训练时间减少了20.67%, 其识别率达到了98.24%。所提算法通过对传统CNN结构进行优化, 极大地缩短了训练时间, 并具有较高的识别率。

关键词: 卷积神经网络; 批量归一化; 贪婪预训练; 支持向量机

中图分类号: TP391.41 **文献标志码:** A

Traffic sign recognition based on optimized convolutional neural network architecture

WANG Xiaobin¹, HUANG Jinjie¹, LIU Wenju^{2*}

(1. School of Automation, Harbin University of Science and Technology, Harbin Heilongjiang 150080, China;

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: In the existing algorithms for traffic sign recognition, sometimes the training time is short but the recognition rate is low, and other times the recognition rate is high but the training time is long. To resolve these problems, the Convolutional Neural Network (CNN) architecture was optimized by using Batch Normalization (BN) method, Greedy Layer-Wise Pretraining (GLP) method and replacing classifier with Support Vector Machine (SVM), and a new traffic sign recognition algorithm based on optimized CNN architecture was proposed. BN method was used to change the data distribution of the middle layer, and the output data of convolutional layer was normalized to the mean value of 0 and the variance value of 1, thus accelerating the training convergence and reducing the training time. By using the GLP method, the first layer of convolutional network was trained with its parameters preserved when the training was over, then the second layer was also trained with the parameters preserved until all the convolution layers were trained completely. The GLP method can effectively improve the recognition rate of the convolutional network. The SVM classifier only focused on the samples with error classification and no longer processed the correct samples, thus speeding up the training. The experiments were conducted on Germany traffic sign recognition benchmark, the results showed that compared with the traditional CNN, the training time of the new algorithm was reduced by 20.67%, and the recognition rate of the new algorithm reached 98.24%. The experimental results prove that the new algorithm greatly shortens the training time and reached a high recognition rate by optimizing the structure of the traditional CNN.

Key words: Convolutional Neural Network (CNN); batch normalization; Greedy Layer-wise Pretraining (GLP); Support Vector Machine (SVM)

0 引言

随着我国经济的发展, 交通越来越发达, 同时随着私家车的增多, 交通事故频发, 安全驾驶问题越来越突出, 于是智能交通系统应运而生。交通标志识别是智能交通系统的重要组成部分, 在交通安全运行方面具有重大的现实意义, 它能够减少驾驶员的驾驶疲劳, 更好地保证出行安全。在汽车运行过

程中, 交通标志识别往往是在复杂的室外交通环境下进行的, 比一般静止的事物识别更难, 主要体现在: 复杂多变的光照条件; 背景环境干扰; 交通标志遮挡; 交通标志位置倾斜^[1]。受干扰图例如图1所示。为了解决这些问题, 大量的复杂算法被提出, 造成识别率高但训练速度过慢, 或者训练速度快但识别率低的问题, 因此交通标志快速训练与准确识别是迫切需要解决的问题。

收稿日期: 2016-07-28; 修回日期: 2016-09-21。 基金项目: 国家自然科学基金资助项目(61573357, 61503382, 61403370, 61273267)。

作者简介: 王晓斌(1990—), 男, 山东潍坊人, 硕士研究生, 主要研究方向: 图像识别、目标检测; 黄金杰(1967—), 男, 山东莱阳人, 教授, 博士, 主要研究方向: 智能建模、优化控制、模式识别; 刘文举(1960—), 男, 北京人, 教授, 博士, 主要研究方向: 机器学习、语音增强、语音识别、声源定位、声音事件检测、图像识别。

目前已经有很多传统分类器,包括支持向量机(Support Vector Machine, SVM)^[2]、贝叶斯分类器^[3]、随机森林分类器等被用于交通标志识别。这些方法均先利用典型的特征描述方法如尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)^[4]、加速鲁棒特征(Speeded Up Robust Feature, SURF)^[5]、方向梯度直方图(Histogram of Oriented Gradient, HOG)^[6]等对训练样本进行低层次特征提取,包含角点、边缘、纹理等,然后对这些特征进行训练分类操作。这些特征描述方法主要是由人工设计的,设计它们需要大量的时间,尽管这些算法训练时间很短,但是识别率很低,不能实现准确应用。

深度学习是目前在目标识别方面应用最多的方法。深度学习的实质是通过构建具有很多隐层的机器学习模型和训练数据,自发地学习更高层次的特征,从而最终提升分类或预测的准确性。和传统的学习方法相比,深度学习更能够刻画数据丰富的内在信息。本文主要使用卷积神经网络(Convolutional Neural Network, CNN)^[7]方法。

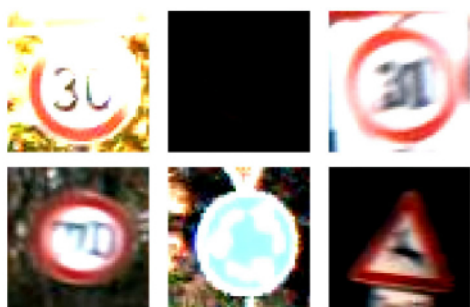


图1 受干扰图例

Fig. 1 Examples of interfered signs

CNN 是受生物的视觉系统启发而形成的,是一个多层的神经网络结构,每个卷积层由一个滤波器层、一个非线性层、一个空间采样层组成。CNN 的主要优点就是可以将数据直接输入 CNN 模型而不进行任何处理,它可以自动学习图像特征,并且具有很强的鲁棒性。Ciresan 等^[8]提出了多列卷积神经网络算法,一共有 25 个卷积网络结构,每个卷积网络用不同的方法进行训练,结果取 25 个网络的平均值,之后用这个模型对测试集进行测试;Sermanet 等^[9]提出了多规模卷积网络,把卷积网络的第二层特征和第三层特征联合送入全连接层进行分类;Krizhevsky 等^[10]提出用深度卷积神经网络对图片进行分类。在图像数据库比赛中,这些算法都获得了比较高的识别率;但是这些算法的共同缺点是训练时间过长,需要大量的计算资源,其中 Ciresan 等^[8]提出的方法训练时间达到 37 h,浪费了大量资源。

针对上述方法的缺点,本文对 CNN 结构进行了优化,包括加入批量归一化(Batch-Normalization, BN)^[11]层,使用逐层贪婪预训练(Greedy Layer-wise Pretraining, GLP)^[12]方法进行训练,同时最后的分类层使用支持向量机(SVM)进行分类。其中 GLP 能够提高识别率,BN 和 SVM 能够加速深度学习训练过程。

1 卷积神经网络

1.1 卷积神经网络结构

CNN 的两大优点就是局部感受野和权值共享。在传统的神经网络结构中,一般都是全连接,导致参数非常巨大,训练困难,速度很慢。比如一幅 $n \times n$ 像素图像输入全连接网络,如果第一层神经元个数是 m 个,因为是全连接,每个像素

与每个神经元全连接,就会有 $n \times n \times m$ 个参数,导致参数灾难,训练会非常困难,因此必须想办法减少参数。对于 CNN 来说,一幅 1000×1000 像素的图像,如果局部感受野范围是 50×50 ,那么仅仅会有 50×50 个参数,这是一种特征,如果要识别一种图像,可以设计 100 个滤波器,参数也仅仅是 $50 \times 50 \times 100$,所以参数大幅减少,训练速度也会大大加快。

CNN 结构由输入层、卷积层和采样层组成。图 2 所示是交通标志识别 CNN 结构,在图中,输入的是交通标志的图像,先经过第一层卷积 C1,滤波器个数是 8 个大小为 5×5 ,形成 8 个特征映射图,每个特征映射图大小为 44;然后再经过采样层 S2,使用的采样核大小为 2×2 ,得到 8 个特征图,特征图大小是 22;第二卷积层 C3 产生 14 个特征图,卷积核大小 3×3 ,产生 14 个特征图大小是 20;再经过采样,采样核大小是 2×2 ,产生 14 个特征图,每个特征图大小是 10,把 14 个特征图压缩成一维连接起来与 C5 层 120 个神经元全连接;最后一层是分类器,采用全连接方式进行分类输出。这就是一个完整的卷积神经网络结构。

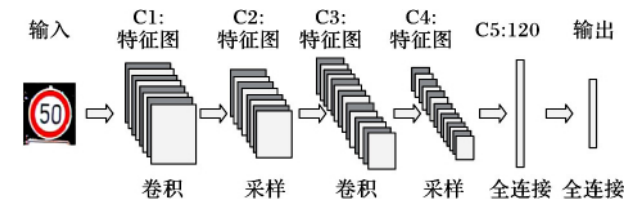


图2 卷积神经网络结构

Fig. 2 Architecture of CNN

1.2 卷积神经网络的训练

CNN 的训练过程主要采用的是反向传播算法,首先输入数据,进行前向计算,然后再反向计算误差,并对误差求各个权值和偏置的偏导数,依靠这个偏导数调整各个权值和偏差。在开始训练前,所有的权值都应该用不同的随机数进行初始化。选取的随机数应较小,以保证网络不会因权值过大而进入饱和状态,从而导致训练失败。“不同”则用来保证网络可以正常地学习。下面是用交通标志训练 CNN 的过程:

第一阶段,向前传播阶段:

- 1) 从交通标志样本中选取一部分标志输入网络;
- 2) 通过前向算法计算相应的实际输出标志类别。

第二阶段,向后传播阶段:

- 1) 把前面计算的类别与真实交通标志类别进行比较,求偏差;

- 2) 按极小化误差的方法反向传播调整权矩阵。

下面是具体的公式推导:

1) 卷积层。

a) 卷积计算。假设第 l 层为卷积层,第 $l+1$ 层为下采样层,则第 l 层第 j 个特征图 x_j^l 的计算公式如下:

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

其中: x_j^l 表示在 l 层(卷积层)中的第 j 个特征图; M_j 表示被选择的输入图集合; $*$ 号实质是让卷积核 k 在第 $l-1$ 层所有关联的特征图上作卷积运算,然后求和,再加上一个偏置参数 b ,用激活函数 $f(\cdot)$ 最终得到激励值的过程。

b) 残差计算。卷积层的下一层是下采样层,采用的是一对一非重叠采样,故残差计算更为简单。第 l 层第 j 个特征图的残差计算公式如下:

$$\sigma_j^l = \beta_j^{l+1} (f'(u_j^l) \circ \text{up}(\sigma_j^{l+1}))$$

其中:第 l 层为卷积层,第 $l+1$ 层为下采样层,下采样层与卷

积层是一一对应的; $\text{up}(x)$ 是将第 $l+1$ 层的大小扩展为和第 l 层大小一样; 下采样层中的权重全都等于 β (一个常量); $f'(\cdot)$ 表示激活函数 $f(\cdot)$ 的偏导数 “ \circ ” 表示每个元素相乘。

c) 梯度计算。偏置参数 b 的导数为:

$$\frac{\partial E}{\partial b_j} = \sum_{u \neq j} (\sigma_j^l)_{uw}$$

参数 k 的导数为:

$$\frac{\partial E}{\partial k_{ij}^l} = \sum_{u \neq j} (\sigma_j^l)_{uw} (p_i^{l-1})_{uw}$$

其中: k_{ij}^l 表示第 l 层第 j 个特征图与第 $l-1$ 层第 i 个特征图相连所对应的核, 是一个矩阵; $(p_i^{l-1})_{uw}$ 表示 x_i^{l-1} 中的区域块。

2) 采样层。

a) 卷积计算。设第 l 层为下采样层, 第 $l-1$ 层为卷积层, 由于是一对一采样, 假设采样大小为 2×2 , 故计算公式为:

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l)$$

其中: $\text{down}(x)$ 是将 x 中 2×2 的大小中像素值进行求和。计算过程为对上一层卷积层 2×2 大小进行求和然后乘以权重 w , 再加上一个偏置, 再取激活函数。

b) 残差计算。在计算下采样层梯度时, 需要找到残差图中给定像素对应于上一层的残差图中哪个区域块, 这样才可以将残差反向传播回来。另外, 需要乘以输入区域块和输出像素之间连接的权值, 这个权值实际上就是卷积核的权值。上述过程可以使用下面的 Matlab 函数公式实现:

$$\sigma_j^l = f'(u_j^l) \circ \text{conv2}(\sigma_j^{l+1} \text{rot180}(k_{ij}^{l+1}), \text{full})$$

其中: 第 l 层为下采样层, 第 $l+1$ 层为卷积层。

c) 梯度计算。偏置 b 的导数, 其公式的推导过程与卷积层一样。权重 β 的导数计算公式如下:

$$d_j^l = \text{down}(x_j^{l-1})$$

$$\frac{\partial E}{\partial \beta_j} = \sum_{u \neq j} (\sigma_j^l \circ d_j^l)_{uw}$$

2 加速训练模块和识别率模块

2.1 加速训练模块——批量归一化方法

交通标志识别主要使用 CNN 方法, CNN 学习过程本质就是为了学习数据分布, 深度卷积网络的训练是一个复杂的过程, 只要网络的前面几层发生微小的改变, 后面几层就会被累积放大下去。一旦网络某一层的输入数据的分布发生改变, 那么这一层网络就需要学习这个新的数据分布。训练过程中, 训练数据的每一层分布一直在发生变化, 并且每一层所需要的学习率不一样, 通常需要使用最小的那个学习率才能保证损失函数有效下降, 因此将会影响网络的训练速度。而 BN 算法就是要解决在训练过程中, 中间层数据分布发生变化的情况, BN 算法使每一层的数据都归一化为均值为 0、标准差为 1, 使数据稳定, 因此可以使用较大的学习率进行训练, 使网络加快收敛, 提高训练速度。BN 算法主要使用如下公式进行归一化:

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{\text{Var}[x^{(k)}]}} \quad (1)$$

其中 $E[x^{(k)}]$ 指的是每一批训练数据 $x^{(k)}$ 的平均值; 分母就是每一批数据 $x^{(k)}$ 的一个标准差。后面的处理也将利用式 (1) 对某一个层网络的输入数据作归一化处理。需要注意的是, 训练过程中采用批量随机梯度下降方法。

如果是仅仅使用式 (1) 对网络某一层输出数据作归一化, 然后再送入下一层, 会影响到本层网络学习到的特征。因

此 BN 算法的关键之处就是引入参数 γ 和 β , 公式为: $y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)}$ 。其中: $\hat{x}^{(k)}$ 表示归一化为均值为 0、标准差为 1 的数据; γ 和 β 的作用是用来保持模型的表达能力, 例如, 在 sigmoid 激活函数的中间部分, 函数近似于一个线性函数, 使用 BN 后会使归一化后的数据仅分布在线性部分, 使模型泛化能力变差, 引入 γ 和 β 使数据不仅分布在线性部分, 也可以分布在非线性部分, 增加模型的泛化能力。当 $\gamma^{(k)} = \sqrt{\text{Var}[x^{(k)}]}$, $\beta^{(k)} = E[x^{(k)}]$ 时, 完全可以恢复出原始的某一层数据。 γ 和 β 需要利用反向传播算法进行训练求出。下面公式是 BN 层的数据归一化过程:

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$y_i \leftarrow \gamma \hat{x}_i + \beta = \text{BN}_{\gamma\beta}(x_i)$$

其中: 输入值是 x_1, x_2, \dots, x_m , m 是批量化的数目; μ_B 是均值, σ_B^2 是方差; $y_i = \text{BN}_{\gamma\beta}(x_i)$ 是输出。

2.2 识别率模块——逐层贪婪训练

逐层贪婪预训练是一个层网络预训练结构, 逐层贪婪算法的主要思路是每次只训练网络中的一层, 即首先训练一个只含一个卷积层的网络, 仅当这层网络训练结束之后才开始训练一个有两个卷积层的网络, 以此类推。在每一步中, 把已经训练好的前 $k-1$ 层固定, 然后增加第 k 层 (也就是将已经训练好的前 $k-1$ 的输出作为输入)。每一层的训练是有监督的 (例如, 将每一步的分类误差作为目标函数), 这些各层单独训练所得到的权重被用来初始化最终的深度网络的权重, 然后对整个网络进行“微调” (即把所有层放在一起优化有标签训练集上的训练误差), 这就是一个完整的逐层贪婪预训练 (GLP) 过程。

3 实验与分析

本次实验所用的计算机配置是 CPU Xeon 2.60 GHz, centos 6.5 操作系统, 同时还使用了 K40c 显卡内存 12 GB 来加速训练。

3.1 数据准备

本次实验使用德国交通标志数据集 (German Traffic Sign Recognition Benchmark, GTSRB) [13], 该数据集包含 43 类从德国真实交通环境中采集的交通标志, 整个数据集共 51 839 幅交通标志图像, 其中训练集 39 209 幅, 测试集 12 630 幅, 每幅图像只包含一个交通标志; 且数据集中包括大量低分辨率、不同光照强度、局部遮挡、视角倾斜、运动模糊等各种不利条件下的图像, 能够较全面地反映现实情况, 以及算法的应用能力。GTSRB 数据集包括六大类: 限速标志、其他禁令标志、解除禁令标志、指示标志、警告标志、其他标志。

由于数据集中的原始图像每幅图像只包含一个交通标志, 而且有 10% 的冗余边, 每个图像大小尺寸并不一样, 为了提高识别率, 需对其进行预处理。即先参考交通标志说明文件把图片中 10% 的冗余边剪除, 使交通标志充满整个图像; 然后将图像进行旋转、平移和缩放, 旋转角度为 -5° 和 5° , 左、右各平移 4 个像素, 缩放因子为 0.9 和 1.1, 扩展成 155 517 张图片。这样更有利于训练过程, 同时对测试集进行更好的预测。接着将 RGB 图像转化成灰度图像。笔者进行

了相关实验,发现 RGB 图像识别率和灰度图像识别率大体一致,因此为了提高计算速度,使用灰度图像进行实验。由于本文使用的是卷积神经网络方法,所有的图像都必须大小一致,所以按照以往的经验把所有的图像都转化成 48×48 大小。卷积神经网络方法具有很强的鲁棒性,因此并没有使用具体的去噪方法进行去噪。

3.2 网络结构选择

在进行实验之前需要了解多少层卷积结构才能使识别率最高,因为对于深度学习而言,随着层数增加,结构能力越强,将会导致数据过拟合;但如果层数太少,结构能力太弱,又会导致数据欠拟合。这两者都会导致训练错误率和测试错误率升高。为了选择卷积层数,使用逐层增加的方法,直到随着层数增加,识别率不再明显变化,就可以确定最优层数。这里使用的数据还是德国交通标志数据库,分为训练集和测试集。在选择之前,每一层滤波器都是使用 108 个,滤波器随机初始化,实验结果如图 3 所示。从图 3 中可以看出,随着层数增加,识别率逐渐增加,随后不变或者下降。因此关于卷积神经网络结构选择三层最为合适,识别率也最高。

而关于滤波器个数选择,同样也是先参考文献[9-10],选择默认卷积层数为 3 层,然后在这个基础上再进行实验,结果如表 1 所示。从表 1 可以看出,尽管滤波器越多正确率越高,但是如果选择 90-108-108 结构,会因为滤波器过多而使得计算复杂,训练时间过长。综合考虑,选择 90-108-108,其中 90 代表第一层滤波器个数,以此类推。

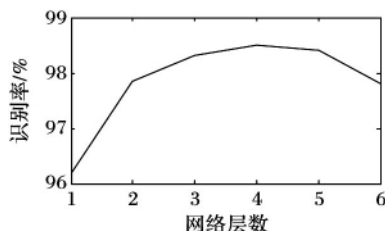


图3 识别率随层数变化曲线

Fig. 3 Curve of recognition rate varied with the number of layers

表1 滤波器个数与正确率的关系

Tab. 1 Relationship between the number of filters and correct rate

尺寸	正确率/%	尺寸	正确率/%
50-90-130	95.83	108-130-150	96.41
90-108-108	96.22	130-160-190	96.57

在选择好层数和滤波器个数后,在本次实验中使用隐层数为 1;参考文献[14]本次实验使用隐层神经元个数为 500。

3.3 训练速度对比

本次实验设计的网络结构是三层卷积神经网络、一个隐层和一个分类层,分类层分别使用 softmax 和 SVM 分类器,网络训练时间分别是 830.38 min 和 664.91 min。可以看出,使用 SVM 训练时间明显比 softmax 训练时间短,因为对于 softmax 来说,当某类的概率已经大于 0.9 时,这就意味着分类正确,然而损失函数会继续计算,直到概率接近 1,这样会造成计算资源的浪费;而 SVM 分类器只专注于那些分类错误的样本,对已经分类正确的样本不再处理,从而能大幅提高训练速度,同时还能提高模型的泛化能力,提高识别率。因此后续实验中的分类层采用 SVM 分类器。

接下来实验使用的网络结构包括三个卷积层、两个隐层和一个分类层;而对比结构也是三层卷积网络结构,只是每一层卷积后都使用 BN 算法,然后是两个隐层和一个分类层,其

他参数两个网络结构完全相同。结果显示:使用 BN 算法前的训练时间为 716.74 min,而使用 BN 算法后的训练时间为 568.6 min。可以看出,在网络结构中加入 BN 层后,训练时间变短,原因在于:1) BN 算法对卷积后的输出数据进行了归一化操作,使它们的均值为 0、方差为 1,因此不必像以前那样迁就维度规模,可以使用大的学习率对网络结构进行训练,收敛加快,达到最优的训练时间也就变短;2) 归一化后使得更多的权重分界面落在了数据中,降低了过度拟合的可能性,因此一些防止过拟合但会降低速度的方法,例如 dropout 和权重衰减就可以不使用或者降低其权重,从而提高训练速度。

3.4 识别率对比

为了验证加入逐层贪婪预训练 (GLP) 方法后比优化前的三层卷积网络更优,设计两个结构完全一致的网络,均包括三层卷积、一个隐层和一个分类层,只是在训练方法上不同。结果显示:优化前的识别率为 97.36%,优化后的识别率为 98.03%。可以看出,使用 GLP 方法的识别率比优化前的三层卷积识别率要高。在训练深层网络时,经常会遇到局部极值问题或者梯度弥散问题,当使用反向传播方法计算导数时,随着网络深度的增加,反向传播的梯度(从输出层到网络的最初几层)的幅度值会急剧地减小,结果就造成了整体的损失函数相对于最初几层的权重的导数非常小。这样,当使用梯度下降法时,最初几层的权重变化非常缓慢,以至于它们不能够从样本中进行有效的学习。因此,当使用贪婪预训练方法时,相比随机初始化而言,各层初始权重会位于参数空间中较好的位置上,然后可以从这些位置出发进一步微调权重。从经验上来说,以这些位置为起点开始梯度下降更有可能收敛到比较好的局部极值点,从而使用逐层贪婪训练方法能够更好地训练网络,使网络达到最优。

尽管本文算法还没有达到官网公布的最优识别率,但是该算法对原始卷积神经网络进行了优化,用逐层贪婪方法预训练网络,使网络跳出局部极值;而且在卷积网络中加入 BN 层,把原来的 softmax 分类器换成了 SVM 分类器,极大地缩减了训练时间,能快速收敛到最优值。

图 4 是本次实验训练测试曲线,可以看出,网络在测试之前参数已经训练到稳定状态,证明结论是可信的。虽然 Cirean 等^[8]提出的 Committee of CNNs 算法有很高的识别率,然而该网络需要用 4 块 GPU 训练 37 h,训练前需进行大量的预处理,操作复杂;而 Random Forests 算法、ANN 算法、SVM 算法尽管训练时间短,但识别率并不是很高(各算法的识别率对比如表 2 所示)。因此本文算法解决了识别率低、训练时间长的问題。交通标志识别错误部分,主要是由自然条件比较差、分辨率太低造成的,今后应该加强数据预处理,使数据更加清晰,易于识别。

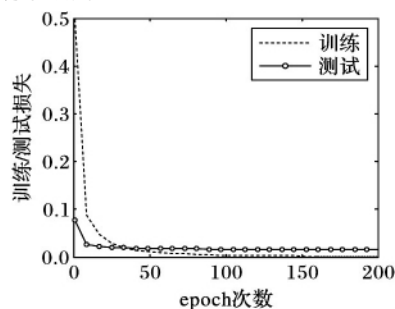


图4 训练测试曲线

Fig. 4 Curve of training and testing

表 2 各算法识别率对比

Tab. 2 Comparison of recognition rates of different algorithms

算法	识别率	算法	识别率
SVM	81.47	Conv + GLP + BN + SVM	98.24
ANN	89.63	Committee of CNNs	99.46
Random Forests	96.14		

4 结语

本文提出了基于优化卷积神经网络结构的交通标志识别算法,并在德国交通标志识别数据库上进行实验,实验结果表明该算法可以提高交通标志的训练速度,同时提高识别率。本文算法主要对卷积神经网络结构进行了三部分的优化:第一,在每一层卷积后面加入了批量归一化层;第二使用逐层贪婪预训练方法对网络结构进行训练;第三使用支持向量机进行分类。下一步工作的目标是在更复杂的背景中,首先准确地检测出交通标志,然后再识别出交通标志,真正实现交通标志的实时检测与识别。

参考文献 (References)

- [1] 杨心. 基于卷积神经网络的交通标志识别研究与应用 [D]. 大连: 大连理工大学, 2014: 3-4. (YANG X. Research and application of traffic sign recognition based on convolutional neural network [D]. Dalian: Dalian University of Technology, 2014: 3-4.)
- [2] GREENHALGH J, MIRMEHDI M. Real-time detection and recognition of road traffic signs [J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1498-1506.
- [3] MEUTER M, NUNNY C, GORMER S M, et al. A decision fusion and reasoning module for a traffic sign recognition system [J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1126-1134.
- [4] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] BAY H, TUYTELAARS T, VAN GOOL L. SURF: Speeded Up Robust Features [C]// ECCV 2006: Proceedings of the 9th European Conference on Computer Vision, LNCS 3951. Berlin: Springer-Verlag, 2006: 404-417.
- [6] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]// CVPR 05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2005, 1: 886-893.
- [7] MATSUGU M, MORI K, MITARI Y, et al. Subject independent facial expression recognition with robust face detection using a convolutional neural network [J]. Neural Networks — 2003 Special issue: Advances in Neural Networks Research — IJCNN 03, 2003, 16(5/6): 555-559.
- [8] CIREŞAN D, MEIER U, MASCI J, et al. Multi-column deep neural network for traffic sign classification [J]. Neural Networks, 2012, 32: 333-338.
- [9] SERMANET P, LeCUN Y. Traffic sign recognition with multi-scale convolutional networks [C]// IJCNN 2011: Proceedings of the 2011 International Joint Conference on Neural Networks. Piscataway, NJ: IEEE, 2011: 2809-2813.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1097-1105.
- [11] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// ICML 2015: Proceedings of The 32nd International Conference on Machine Learning. New York: ACM, 2015: 448-456.
- [12] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks [C]// NIPS 06: Proceedings of the 19th International Conference on Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2007: 153-160.
- [13] HOUBEN S, STALLKAMP J, SALMEN J, et al. Detection of traffic signs in real-world images: the German traffic sign detection benchmark [C]// IJCNN 2013: Proceedings of the 2013 International Joint Conference on Neural Networks. Washington, DC: IEEE Computer Society, 2013: 715-722.
- [14] QIAN R, ZHANG B, YUE Y, et al. Robust chinese traffic sign detection and recognition with deep convolutional neural network [C]// ICNC 2015: Proceedings of the 2015 International Conference on Natural Computation. Washington, DC: IEEE Computer Society, 2015: 791-796.

This work is partially supported by the National Natural Science Foundation of China (61573357, 61503382, 61403370, 61273267).

WANG Xiaobin, born in 1990, M. S. candidate. His research interests include image recognition, object detection.

HUANG Jinjie, born in 1967, Ph. D., professor. His research interests include intelligent modeling, optimization control, pattern recognition.

LIU Wenju, born in 1960, Ph. D., professor. His research interests include machine learning, speech enhancement, speech recognition, sound source localization, sound event detection, image recognition.

(上接第 529 页)

- [18] FANG H-L, ROSS P, CORNE D. A promising genetic algorithm approach to job-shop scheduling, re-scheduling and open-shop scheduling problems [C]// Proceedings of the 5th International Conference on Genetic Algorithms. San Mateo, CA: Morgan Kaufmann Publishers, 1993: 375-382.
- [19] MORAD N, ZALZALA A M S. Genetic algorithms in integrated process planning and scheduling [J]. Journal of Intelligent Manufacturing, 1999, 10(2): 169-179.
- [20] FUJI N, INOUE R, UEDA K. Integration of process planning and scheduling using multi-Agent learning [C]// Proceedings of the 41st CIRP Conference on Manufacturing Systems. Berlin: Springer-Verlag, 2008: 297-300.

This work is partially supported by the Scientific Research Fund by Shaanxi Provincial Education Department (15JK1311), Research Start-up Fund for Doctor of Xi'an Polytechnic University (BS1301), Innovation Fund for Graduate Students of Xi'an Polytechnic University (CX201628).

SONG Shuanjun, born in 1974, Ph. D., associate professor. His research interests include optimization of production system, supply chain management.

YANG Peili, born in 1992, M. S. candidate. Her research interests include optimization of production system.

SHI Wenli, born in 1991, M. S. candidate. Her research interests include supply chain management.