

# 基于 CNN 和 SVM 融合的交通标志识别

王新美,丁爱玲,雷梦宁,康 盟  
(长安大学 信息工程学院,陕西 西安 710000)

**摘 要:**针对传统的卷积神经网络对小样本分类易产生过拟合等问题,在卷积神经网络(CNN)和支持向量机(SVM)融合模型的基础上,提出对 CNN 网络结构提取的特征进行归一化处理,提高泛化能力,并将其应用到交通标志识别。该方法构建了一种 CNN-SVM 模型,将卷积神经网络和支持向量机结合起来,使用从 ImageNet 数据集初始化的网络进行特定域的微调,截取网络内层来提取交通标志图像特征,并对特征进行归一化处理,最后采用 SVM 进行识别,从而有效解决交通标志分类过拟合问题。仿真结果表明,通过 CNN 内层建立的特征映射模型,所传递的特征经过归一化处理后,在交通标志分类任务中具有良好的特征表示能力,较好地提升了 SVM 分类性能,表现出更好的分类精度以及泛化性能。

**关键词:**CNN;SVM;迁移学习;归一化;交通标志识别

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2020)06-0007-06

doi:10.3969/j.issn.1673-629X.2020.06.002

## Traffic Sign Recognition Based on Combination of CNN and SVM

WANG Xin-mei, DING Ai-ling, LEI Meng-ning, KANG Meng  
(School of Information Engineering, Chang'an University, Xi'an 710000, China)

**Abstract:** Aiming at the problem that the traditional convolutional neural network tends to over-fit the classification of small samples, we propose a new method to normalize the features extracted from CNN network structure based on the model of CNN and SVM combination, so as to improve its generalization ability and apply it to traffic sign recognition. The method builds a CNN-SVM model which combines convolutional neural network and support vector machine, and fine-tunes the network initialized from ImageNet dataset on specific domain and intercepts its inner layer to extract the image features from traffic signs. The features are normalized and finally identified by SVM, which effectively solves the over-fitting problem toward traffic signs' classification. The simulation results show that the feature mapping model established through the inner layer of CNN, after the transmitted features being normalized, have a superb feature presentation ability in traffic sign classification tasks, improve the SVM classification performance, and show better classification accuracy and generalization performance.

**Key words:** CNN; SVM; transfer learning; normalization; traffic sign recognition

## 0 引言

在自动驾驶、驾驶员安全与协助、交通场景分析等应用中,交通标志识别是必不可少的一部分,能够辅助驾驶员准确高效地识别道路交通标志,在一定程度上减轻驾驶疲劳,从而保证出行安全,有效缓解道路交通安全问题<sup>[1]</sup>。因此,交通标志识别研究具有重大的意义。交通标志识别往往是汽车行驶途中在复杂的室外环境下进行的,与静止的事物相比,识别难度更大,会出现运动模糊、标志褪色、人为损坏、光照以及天气变化等现象,且具有交通标志类别多、各国标准不一等特征,这些因素都对交通标志的识别造成影响。

针对以上这些问题,在交通标志识别的研究中,大量复杂的算法被提出。传统的算法主要依赖于图像形态学,采用“人工提取特征+机器学习”进行识别,通过分割和提取并整合交通标志图像的 LBP、Gabor 和 HOG 三类特征,利用 SVM、AdaBoost 完成交通标志的识别,但真实情况中的识别情况复杂,传统算法应用难度大。与之相比,深度学习是目前在交通标志识别中效果较好、应用较多的方法。其中的卷积神经网络(CNN)更是在提取交通标志特征中表现出了良好的性能<sup>[2]</sup>。在多个图像分类的比赛中,以卷积神经网络为基础改进的深度学习网络一次次刷新图像分类的记

收稿日期:2019-08-29

修回日期:2019-12-30

基金项目:国家青年科学基金项目(61806023)

作者简介:王新美(1995-),女,硕士研究生,CCF 会员(D3915G),研究方向为数字图像处理、深度学习;丁爱玲,博士,教授,研究方向为数字图像处理、智能信号与信息处理。

录,奠定了深度学习在计算机视觉中的地位。

CNN 之所以具有强大的分类能力,是因为其能够从大量的样本中学习得到特征表示,整个网络表达了原始图像像素与其类特征之间的映射关系。然而,在交通标志分类的实际应用中,硬件条件相同的情况下,这种技术需要一个巨大的数据集来训练和测试交通标志分类技术,以避免出现过拟合情况。此外,任何对交通标志设计的更改,只添加或删除一类交通标志,都需要重复整个训练过程,而这样的训练过程需要大量的处理时间,这都使得在本地样本少的实际应用中具有局限性。

针对上述问题,文献[3-4]将已在大量其他类别上训练过的分类器迁移到目标域,以克服特定训练样本数量不足的问题,在目标域得到了良好的使用效果。迁移学习并不局限于一种特殊的算法或模型,而是在与目标域具有相关性的源域上对知识或者模型进行迁移。利用 CNN 进行迁移学习已被广泛应用于图像识别<sup>[5-6]</sup>和自然语言处理<sup>[7-8]</sup>。此外,已有研究<sup>[9-11]</sup>探讨了另一种用支持向量机(SVM)替代 softmax 函数在网络模型中的分类作用。研究表明,在人工神经网络(ANN)结构中使用 SVM 比使用传统的 softmax 函数产生了更好的结果。

受上述研究启发,文中利用 CNN 迁移学习方法,对目前分类效果优越的卷积神经网络——VGG16 网络进行结构改进,设计了一种计算交通标志特征表示的特征映射模型,结合 SVM 构建 CNN-SVM 模型,完成交通标志识别分类。此外,在 CNN 迁移学习建立特征映射模型的基础上,提出特征归一化操作。以此能够更准确地从实际监测数据中提取交通标志特征信息,实现更高精度的识别分类,同时增加模型的泛化性能,减小过拟合。

## 1 相关算法

### 1.1 卷积神经网络

卷积神经网络广泛应用于计算机视觉中的图像分类问题,属于深度前馈人工神经网络。卷积神经网络负责接收检测到的图像,通过训练集以及验证集每一轮的训练结果反向传递以调整网络结构参数,由卷积层、池化层、全连接层和 softmax 层组成。卷积层通过卷积核对图像卷积滤波依次提取图像特征,卷积的结果经过激活函数,形成本层的特征图(feature map)输出,除了特征值本身外,还包含相对的位置信息。卷积层公式为:

$$Y_j^l = f(\sum_{i \in M_j} Y_j^{l-1} * W_{ij}^l + b_j^l) \quad (1)$$

其中,  $Y_j^l$  是第  $l$  层中第  $j$  个卷积输出的特征子图,“\*”

表示二维卷积运算,  $b_j^l$  是第  $l$  层中第  $j$  个卷积偏置项,  $W_{ij}^l$  为第  $l$  层中第  $j$  个卷积的权值,  $f(\cdot)$  是非线性激活函数。

$f(\cdot)$  加入非线性因素,将线性输入转化为非线性输出,解决线性模型所不能解决非线性的问题,以此增强神经网络的表达能力。卷积神经网络常用的激活函数为 ReLU,公式为:

$$f = \text{Max}(0, x) \quad (2)$$

池化层对卷积层得到的特征图进行压缩,以此简化网络计算的复杂度,对特征的压缩进一步提取了主要特征,去除冗余特征。但同时,特征图的缩小可能会影响到网络的准确度,因此需要通过增加网络深度来弥补。池化操作一般分为两种:平均池化、最大池化。最大池化对邻域内取特征值最大的点,能更多地保留图像纹理信息,文中采用的网络结构应用最大池化进行特征图压缩。

全连接层将卷积层学到的分布式特征表示映射到样本标记空间,本质是由一个特征空间变换到另一个特征空间。在卷积神经网络中,全连接层通常出现在最后几层,整合卷积层和池化层具有类别区分性的局部信息,进行特征加权。同时,全连接层在模型迁移学习的过程中,起到了特征表示“防火墙”的作用<sup>[12]</sup>,在迁移学习任务中,全连接层可以保证卷积神经网络特征表示能力的迁移。softmax 层是特殊的全连接层,区别在于具有激活功能<sup>[13]</sup>,用于实现最终分类。其功能是将多个神经元的输出,映射到 (0,1) 区间内,可以以概率来理解,softmax 层每一个神经元输出值对应交通标志图像各类别的概率大小,从而实现多分类。卷积神经网络强大的特征表示学习能力,对于平移、旋转、尺度缩放等形式具有不变性,使得其具有良好的性能。使用卷积神经网络的另一优点在于,卷积神经网络能有效减弱数据的个体差异和明显的噪声,不需要对图像进行预处理,可直接输入检测到的交通标志图进行特征提取。

本研究中使用的 CNN 以 VGG16 网络模型为基本结构,VGG16 是一个 16 层的卷积神经网络,由 5 组  $3 \times 3$  大小的卷积滤波器组成。其网络结构的主要思想是增加网络深度,减小卷积核尺寸,以更加细致地提取图像的局部特征,减少参数个数,提高图像识别准确度。这一独特的框架在 2014 年“ImageNet 大尺度视觉识别挑战”(ILSVRC)定位分类任务中分别获得第一名和第二名,是深度学习中分类性能优秀的网络模型。文中 VGG16 模型超参数通过迁移学习获取,以源域为 ImageNet 数据集上预训练过的 VGG16 网络进行微调,微调了输出层的结构,使其输出图像的特征表示,将其学习到丰富的图像特征表示方法迁移到目标

域,并针对目标域进行微调。

## 1.2 支持向量机

1995 年 Vapnik<sup>[14]</sup>首次提出支持向量机并用于二值分类,该方法是一种采用结构风险最小化准则的学习方法,相较于传统的采用经验风险最小化准则的学习方法,支持向量机具有更强大的泛化能力,可以广泛地应用于统计分类以及回归分析。对于训练集  $S$ , 设  $n$  为样本数量,  $x$  是训练样本的特征向量,  $y$  代表训练样本的标签,其中  $y_i = 1$  和  $y_i = -1$  分别代表训练样本的正类点和负类点。文中数据通过 CNN 特征映射模型趋于线性可分数据,同时特征归一化处理可以使线性 SVM 具有非线性分类器的能力, SVM 接收的数据近似线性可分,故分类器选择线性 SVM。

SVM 的目的在于找到一个划分超平面,其不仅能正确区分两类不同的训练样本,同时两个异类样本距离超平面最近的训练样本点所组成的支持向量,到此划分超平面具有最大间隔。在  $P$  维空间中若存在线性判别函数:  $f(w, x) = w \cdot x + b$ , 可以将两类数据分开。则分离超平面为:

$$w \cdot x + b = 0 \quad (3)$$

其中,  $w$  垂直于超平面,  $x$  表示输入向量,  $b$  表示偏移量。此线性可分向量机的优化约束问题如下:

$$\begin{cases} \min_{w,b} J(w) = \frac{1}{2} w_2^2 \\ \text{s. t. } y_i(w \cdot x + b) - 1 \geq 0, i = 1, 2, \dots, N \end{cases} \quad (4)$$

其对偶问题为:

$$L(w, b, a) = \frac{1}{2} w_2^2 - \sum_{i=1}^N a_i y_i (w \cdot x + b) + \sum_{i=1}^N a_i \quad (5)$$

其中,  $a_i > 0$  为拉格朗日系数。上述原始问题的求解转化为对偶问题求解:

$$\begin{aligned} \max_a L(w, b, a) &= \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j (x_i \cdot x_j) \\ \text{s. t. } \sum_{i=1}^N a_i y_i &= 0, a_i \geq 0, i = 1, 2, \dots, N \end{aligned} \quad (6)$$

求解上述问题得到的最大间隔分离平面为:

$$\sum_{i=1}^N a_i^* y_i (x \cdot x_i) + b^* = 0 \quad (7)$$

分类决策函数为:

$$f(x) = \text{sgn} \left( \sum_{i=1}^N a_i^* y_i (x \cdot x_i) + b^* \right) \quad (8)$$

其中,  $\text{sgn}(\cdot)$  为符号函数,在求得最优解参数  $w^*$  和  $b^*$  后,相应的最大间隔分离平面即可用于划分样本。式(7)是在数据完全线性可分情况下的最优分类面,成为“硬间隔”分类超平面。考虑到文中数据近似线

性可分,基于最大间隔原则引入松弛变量  $\xi$  和惩罚参数  $C$ , 允许拟合误差,构造“软间隔”的分类超平面来达到最优分类的效果。优化目标可写为:

$$\begin{cases} \min_w \Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s. t. } y_i(w \cdot x + b) - 1 + \xi_i \geq 0, i = 1, 2, \dots, N \end{cases} \quad (9)$$

其中,  $C$  是对分类错误的惩罚参数,用于调整置信区间和经验误差之间的均衡,惩罚参数越大,说明错误惩罚越严重。优化式(9)问题的对偶问题除去拉格朗日系数条件变为:  $0 \leq a_i \leq C$ , 其余与线性可分情况下的问题几乎完全相同。

## 1.3 特征归一化

为进一步提高 CNN-SVM 模型的泛化能力,即提取到的特征表示不受本地描述符数量的影响,使其能更好地应用于小样本分类。经过 CNN 内层  $\varphi(\cdot)$  获取的特征  $x$  需要进一步归一化,归一化的方法分为以下几种:

$\ell_1$  归一化,将特征  $x$  除以其  $\ell_1$  范数:  $x / \|x\|_1$ 。

$\ell_2$  归一化,将特征  $x$  除以其  $\ell_2$  范数:  $x / \|x\|_2$ 。

幂归一化,对提取到的特征  $x$  每一个维度应用:  $f(z) = \text{sign}(z) |z|^\alpha$ , 其中归一化参数  $\alpha$  的范围为(0,1)。

可以将幂归一化与  $\ell_1$  归一化或  $\ell_2$  归一化结合,提高模型泛化能力,不易出现过拟合。此外,文献[15]指出,对于不同的编码方法和局部特征,最终观察到的结果  $\ell_2$  归一化优于  $\ell_1$  归一化。它提出,归一化方法的选择实际上与最终分类器使用的内核相关。在本研究使用的线性向量机中,核函数是  $k(x, y) = x^T y$ , 选择  $\ell_2$  归一化,可以保证  $k(x, x) = \text{const}$  以及  $k(x, x) \geq k(x, y)$ , 即确保了一致性标准:将  $k(x, y)$  视作相似度值,则  $x$  为与自己相似度值最高的点。然而选择  $\ell_1$  归一化无法保证该点与自身最相似,同时可能会导致 SVM 训练过程不稳定。综上研究可得出,在使用线性 SVM 的应用中,  $\ell_2$  归一化优于  $\ell_1$  归一化。本研究将幂归一化与  $\ell_2$  归一化结合,使线性 SVM 具有非线性分类器的能力,是一种有效提高最终识别率的方法。

通过文献[15-16]启发,对经过 CNN 内层  $\varphi(\cdot)$  获取的特征  $x$ , 先通过幂归一化,对 CNN 内层提取的特征进行“非稀疏化”<sup>[16]</sup>, 即平滑直方图,提高特征表示性能。避免特征向量过于稀疏,丢失特征信息。此后进行  $\ell_2$  归一化处理,消除提取的特征对背景的依赖,提高泛化能力。特征处理过程如下:

$$x \leftarrow \varphi(I) \quad (10)$$

$$y \leftarrow \text{sign}(x) |x|^\alpha \quad (11)$$

$$z \leftarrow y / \|y\|_2 \quad (12)$$



## 2 基于 CNN 与 SVM 交通检测与识别模型

文中模型的思想是将 CNN 在一个具有大量带注释的通用数据集中训练,保留其各层参数,截取 CNN 内部层(卷积层、池化层)作为一个特征映射模型。有研究表明<sup>[17-18]</sup>,当特定领域的图像数据不足时,通过预训练,从额外的数据中可以获取特征表达增益,以此保证 CNN 内层具有强大的特征表示能力,以获取丰富的特征表示。特征映射模型将交通标志图像作为输入,图像特征表示作为输出。通过特征归一化处理以及结合 SVM 分类器,利用到本地目标任务中。

CNN-SVM 模型的优点在于:SVM 通过支持向量来构建最优超平面,其分类性能主要受支持向量的影响,CNN 卷积的训练过程就是一个对线性不可分的数据增加其线性可分程度的过程,当 CNN 的特征提取生效后,原本线性不可分的分类样本趋于线性可分,结合 SVM 只利用部分支持向量样本的分类原则优势,提升分类精度以及泛化性能。此外,文中提出特征归一化处理,对 CNN 内层输出的特征表示进行归一化,进一步提升模型泛化能力以及最终识别率。

文中特征映射模型是基于 VGG-16 网络模型进行结构改造,网络超参数由通用图像数据集 ImageNet

的分类训练得到。在截取 CNN 模型的内部层基础上,增加式(11)、(12)归一化处理,添加一层全连接层,能够在提升模型泛化能力的同时保证特征表示能力,其中全连接层的参数随机初始化。最后,这些特征表示将由 SVM 分类。CNN-SVM 模型具体构建步骤如下:

步骤一:将已预训练的 VGG-16 网络作为 CNN 模型,移除训练后的 CNN 模型三个全连接层及 softmax 层,对其网络模型的内部层保持结构参数固定,输入训练数据得到特征表示  $x$ ;

步骤二:对步骤一获取的特征  $x$ ,进行幂归一化(式(11))、 $\ell_2$  归一化(式(12))处理;

步骤三:添加一层全连接层,随机选取初始参数,在目标域采用以上步骤结果训练此全连接层,以保证模型表示能力的迁移;

步骤四:通过以上步骤获得特征映射模型,提取特征向量,对 SVM 进行训练和测试。

文中提出的交通标志分类方法的原理如图 1 所示,主要包括两个过程:利用 CNN 迁移学习建立特征映射模型和基于 SVM 的特征分类。

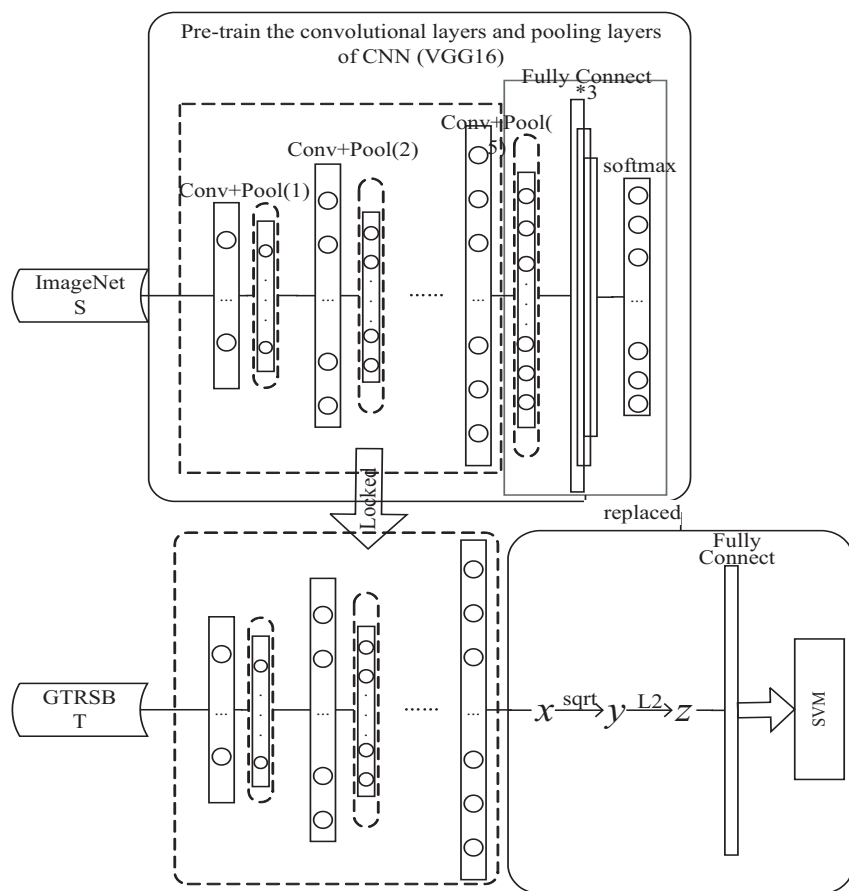


图 1 CNN-SVM 交通标志分类原理

与传统卷积神经网络分类相比,删除了 VGG-16 网络的三个全连接层以及 softmax 层,添加特征归一

化处理,添加一个全连接层,通过在目标域微调 CNN 模型参数,将其映射的特征作为新的输出,建立了特征

映射模型。其次,与源任务另一不同之处在于,目标任务中的样本不是直接分类,而是通过建立的映射模型映射到一个新的特征空间中,将提取的图像特征训练 SVM,实现交通标志分类。

3 仿真实验

在本节中,通过对交通标志样本识别检测,在 CNN 迁移学习以及特征归一化处理的基础上,通过交通标志数据集训练学习全连接层以及 SVM 的超参数,验证基于 CNN 迁移学习的特征映射模型的合理性以及所提出的特征归一化的泛化性能。根据经验选

择<sup>[16]</sup>,幂归一化(式(10))参数 $\alpha$ 取值为0.5时,能有效保留图像信息,在实验中表现出良好的效果。实验分为两部分:

(1)与传统卷积神经网络对比,验证 CNN-SVM 模型用于交通标志分类的有效性;

(2)验证特征归一化泛化能力。实验考虑训练精度和测试精度。

使用 GTSRB (德国交通标志数据集)对 CNN-SVM 和传统 CNN-softmax 模型进行比较。CNN-softmax 模型以及 CNN-SVM 模型在 GTSRB 中的训练精度以及训练损失如图 2(a)、图 2(b)所示。

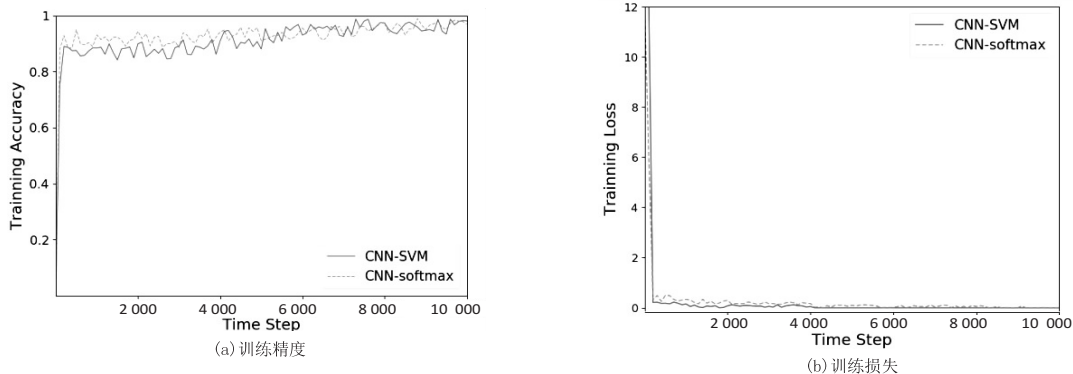


图2 模型对比

其中 CNN-softmax 模型的平均训练精度为 94.29%,CNN-SVM 的平均训练精度为 91.68%。在训练 10 000 步后,保留参数模型。表 1 为使用 GTSRB

数据集采用传统 CNN 模型以及文中添加特征归一化的 CNN-SVM 模型的最终训练精度以及测试精度。

表1 训练精度和测试精度对比(GTSRB 数据集) %

指标	CNN-softmax	CNN-SVM
训练准确率	97.69	98.48
测试准确率	97.53	98.31

为验证特征归一化泛化作用,建立了本地数据集,包括 6 种交通标志,一共 2 461 张,图 3 为本地数据集

图片。



图3 本地交通标志数据集示例

分别使用经过特征归一化与未经过特征归一化的 CNN-SVM 模型,在本地数据集预训练,保存参数模

型。对本地交通标志数据集进行测试。实验模型训练精度以及测试精度如表 2 所示。

表2 训练精度和测试精度对比(本地数据集) %

指标	CNN-SVM	CNN-SVM
	(without normalization)	(with power normalization & $\ell_2$ -normalization)
训练准确率	99.83	98.91
测试准确率	97.24	98.62

以上实验结果证明,对特征进行归一化(幂归一化以及  $\ell_2$  归一化),能有效避免过拟合,提高模型泛化能力。

#### 4 结束语

采用 CNN-SVM 模型,将其应用于交通标志图像识别任务中,此模型结合了 CNN 自适应提取特征和 SVM 泛化能力强的特点,使用深度卷积神经网络提取特征可以减弱数据的个体差异和明显的噪声,增强不同类别间的差异,提高 SVM 分类方法的速度和精度。此外,提出对 CNN 内层提取的图像特征归一化以及保留全连接层处理,以此保证泛化能力以及特征提取能力,识别形状和颜色各异的不同类型交通标志。实验结果验证了基于 CNN 迁移学习的特征映射模型用于交通标志图像分类的合理性,以及特征归一化操作能有效提高 CNN-SVM 模型的泛化性能,使其在本地小样本识别任务中具有优秀的识别精度。

#### 参考文献:

- [1] 葛笑寒. 基于机器视觉的交通标志识别研究[J]. 安徽职业技术学院学报, 2019, 18(2): 12-15.
- [2] 樊星, 赵祥模, 刘占文, 等. 基于多尺度卷积神经网络的交通标志识别方法[J]. 现代电子技术, 2019, 42(15): 134-138.
- [3] TOMMASI T, ORABONA F, CAPUTO B. Learning categories from few examples with multi model knowledge transfer[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(5): 928-941.
- [4] AYTAR Y, ZISSERMAN A. Tabula rasa: model transfer for object category detection[C]//International conference on computer vision. Barcelona, Spain; IEEE, 2011: 2252-2259.
- [5] 张雪芹, 陈嘉豪, 诸葛晶晶, 等. 基于深度学习的快速植物图像识别[J]. 华东理工大学学报: 自然科学版, 2018, 44(6): 887-895.
- [6] OQUAB M, BOTTOU L, LAPTEV I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C]//2014 IEEE conference on computer vision and pattern recognition (CVPR). Columbus; IEEE, 2014: 1717-1724.
- [7] CONLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch[J]. Journal of Machine Learning Research, 2011, 12: 2493-2537.
- [8] 魏晓聪. 面向评论文本的迁移学习研究及应用[D]. 大连: 大连理工大学, 2017.
- [9] 黄孝喜, 李晗雨, 王荣波, 等. 基于卷积神经网络与 SVM 分类器的隐喻识别[J]. 数据分析与知识发现, 2018, 2(10): 77-83.
- [10] ALASHEKMUBARAK A, SMITH L S. A novel approach combining recurrent neural network and support vector machines for time series classification[C]//International conference on innovations in information technology (II T). Abu Dhabi, AI-Ain; IEEE, 2013: 42-47.
- [11] 涂曼姝, 张艳, 颜永红. 基于 CNN-SVM 和转发树的微博事件情感分析[J]. 情报工程, 2017, 3(3): 77-85.
- [12] ZHANG C L, LUO J H, WEI X S, et al. In defense of fully connected layers in visual representation transfer[C]//Pacific rim conference on multimedia. Cham; Springer, 2017: 807-817.
- [13] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [14] CORTES C, VAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20: 273-297.
- [15] PENG X, WANG L, WANG X, et al. Bag of visual words and fusion methods for action recognition: comprehensive study and good practice[J]. Computer Vision and Image Understanding, 2016, 150: 109-125.
- [16] PERRONNIN F, SÁNCHEZ J, MENSINK T. Improving the fisher kernel for large-scale image classification[C]//Proceedings of the 11th European conference on computer vision. Berlin; Springer, 2010: 143-156.
- [17] CIMPOI M, MAJI S, KOKKINOS I, et al. Describing textures in the wild[C]//IEEE conference on computer vision and pattern recognition. Colombia; IEEE, 2014: 3606-3613.
- [18] DONAHUE J, JIA Y, VINYALS O, et al. DeCAF: a deep convolutional activation feature for generic visual recognition[C]//International conference on international conference on machine learning. Beijing: [s. n.], 2014: 647-655.