

# 复杂背景下车型识别分类器

张洁<sup>1,2</sup>, 赵红东<sup>1,2</sup>, 李宇海<sup>2\*</sup>, 闫苗<sup>1</sup>, 赵泽通<sup>1</sup>

<sup>1</sup>河北工业大学电子信息工程学院, 天津 300401;

<sup>2</sup>光电信息控制和安全技术重点实验室, 天津 300308

**摘要** 细粒度车型图像类间特征差异小, 在复杂图片背景中识别干扰因素多。为提高模型在复杂背景中对图像的特征提取能力和识别准确度, 提出了基于支持向量机(SVM)和深度卷积神经网络(DCNN)的分类器集成模型 Softmax-SVM。它将交叉熵代价函数与 hinge 损失函数相结合, 代替 Softmax 函数层, 减少了过拟合的发生。同时, 设计了一个 10 层的 DCNN 提取特征, 避免了手工提取特征的难题。实验数据集为复杂背景下的 27 类精细车型图像, 尤其还包含同一汽车厂商的相近车型。实验结果表明, 在不进行大量预处理的前提下, Softmax-SVM 分类器识别 269 张测试样本能够获得 97.78% 的准确率, 识别每张样本的时间为 0.759 s, 明显优于传统模式识别方法和未改进前的 DCNN 模型。因此, 基于 DCNN 的 Softmax-SVM 分类器能够适应环境的复杂变化, 兼顾识别精度与效率, 为复杂背景下的细粒度车型分类提供了实际参考价值。

**关键词** 机器视觉; Softmax-SVM; 深度卷积神经网络; 复杂背景; 细粒度车型

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP56.041501

## Classifier for Recognition of Fine-Grained Vehicle Models under Complex Background

Zhang Jie<sup>1,2</sup>, Zhao Hongdong<sup>1,2</sup>, Li Yuhai<sup>2\*</sup>, Yan Miao<sup>1</sup>, Zhao Zetong<sup>1</sup>

<sup>1</sup>School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China;

<sup>2</sup>Science and Technology Electro-Optical Information Security Control Laboratory, Tianjin 300308, China

**Abstract** The feature difference among the images of fine-grained vehicle models is small and there exist many factors disturbing recognition under complex image background. To improve the feature extraction ability and the recognition accuracy of images under complex background, a classifier named Softmax-SVM is proposed based on deep convolutional neural network (DCNN) and support vector machine(SVM), in which the cross-entropy cost function is combined with the hinge loss function to replace the Softmax function layer, so that the over-fitting is avoided. Meanwhile, a 10-layer DCNN is designed to extract features automatically and the problem of manual extraction of features is also avoided. The experimental dataset consists of the images of 27 types of fine-gained vehicle models under complex background, especially of the similar models from the same car manufacturer. The experimental results show that the Softmax-SVM classifier can be used to recognize the 269 sample images without much emphasis on the pre-processing stages, and in the identification process, the accuracy rate is 97.78% and the time to identify each image is 0.759 s. The above model performs more efficiently than the traditional recognition methods and the unimproved DCNN models. In consequence, the Softmax-SVM classifier based on DCNN can adapt to the complex changes of environment and give consideration to both the recognition accuracy and efficiency, which provides practical reference value in the classification field of fine-gained vehicle models under complex background.

**Key words** machine vision; Softmax-SVM; deep convolutional neural network; complex background; fine-grained vehicle models

**OCIS codes** 150.0155; 150.1135; 100.3008; 100.5010

收稿日期: 2018-08-16; 修回日期: 2018-08-21; 录用日期: 2018-09-04

基金项目: 光电信息控制和安全技术重点实验室基金(614210701041705)

\* E-mail: 695263295@qq.com

## 1 引言

随着我国经济快速发展,汽车保有量逐年上升,智能交通系统逐渐成为现代社会的研究重点。车型识别系统作为智能交通系统的重要组成部分,在交通监管、事故判罚等领域起到了关键的作用。相比于粗粒度车型识别<sup>[1]</sup>,细粒度分类是一个较新的研究领域<sup>[2-4]</sup>,尤其是当细粒度车型图像中包含大量复杂环境背景时,识别易受环境、光照、噪声等因素干扰,在不进行大量图片预处理工作的前提下,研究面临很多难点。

细粒度车型识别主要包括特征提取和分类两部分,传统识别方法主要依靠尺度不变特征转换(SIFT)特征提取算法<sup>[5]</sup>、梯度直方图(HOG)特征提取算法<sup>[6]</sup>、稳定特征加速算法(SURF)与积分通道特征提取算法<sup>[7]</sup>等,通过人工设定抽取特征,并结合支持向量机(SVM)进行分类。由于细粒度车型图像的外观形式多样,类间差异小<sup>[8]</sup>,因此,难以通过上述方法提取有效特征,从而造成 SVM 分类器无法准确判别。深度卷积神经网络(DCCN)以原始图像作为输入,能够通过卷积操作自主训练,完成图像处理工作并提取图像中的语义和特征,避免了人工选择特征的难题<sup>[9]</sup>。此外,为减少计算量,深度卷积神经网络还引入了空间子采样、局部感受野和权值共享概念。深度卷积神经网络由自动特征提取器和训练分类器 Softmax 组成,是多层次化的网络,在图像分类方面具有特征表达能力<sup>[10]</sup>和识别性能<sup>[11]</sup>,因此被广泛应用于人脸识别<sup>[12]</sup>、图像分类<sup>[13]</sup>、目标检测<sup>[14]</sup>等机器视觉任务。

深度卷积神经网络虽然善于学习不变量特征,但是在训练过程中,Softmax 层的损失函数会不断更新已经正确分类的样本,在图像背景复杂的情况下,容易发生过拟合。SVM 由 Vapnik 和 Cortes<sup>[15]</sup>提出,是一种功能强大的分类器<sup>[16]</sup>。它定义的损失函数不仅避免了上述问题,并且善于从良好的特征向量生成决策面<sup>[17]</sup>。因此,为了提高在复杂背景中的细粒度车型识别准确率,本文有针对性地研究了深度卷积神经网络的 Softmax 函数层和 SVM,提出了基于深度卷积神经网络的 Softmax-SVM 混合模型。该模型将深度卷积神经网络作为原始图像的自动特征提取器,训练达到训练集识别准确率阈值后,提取全连接层(FC1 层)数据用于 SVM 分类器训练,最终在识别复杂背景下的细粒度车型方面,Softmax-SVM 分类器有效避免了过拟合,增强了模

型的泛化能力,提高了识别精度和识别速度。

## 2 深度卷积神经网络与支持向量机

### 2.1 深度卷积神经网络

#### 2.1.1 卷积层

卷积层的表达式为

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

式中: $M_j$  为上一层选择输入的 map 的组合;\* 代表卷积运算; $x_i^{(l-1)}$  为第  $l-1$  层输入的数据中的第  $i$  个特征图; $k_{ij}^{(l)}$  对应  $l$  层的第  $j$  个特征图与第  $l-1$  层的第  $i$  个特征图进行卷积的卷积核; $b_j^{(l)}$  对应第  $l$  层的第  $j$  个特征图,为第  $l-1$  层卷积后的加性偏置。

#### 2.1.2 池化层

卷积神经网络中常用的池化方法包括平均池化、最大池化、随机池化等,池化后特征图数目保持不变,特征维数降低,计算量减少。通用公式为

$$x_j^{(l)} = f \left[ \beta_j^{(l)} \text{down}(x_j^{(l-1)}) + b_j^{(l)} \right], \quad (2)$$

式中: $\text{down}(\cdot)$  为下采样函数; $\beta_j^{(l)}$  为第  $l-1$  层卷积后的乘性偏置。每个输出的 map 均对应一个属于自己的乘性偏置  $\beta$  和一个加性偏置  $b$ 。

#### 2.1.3 批量标准化层

批量标准化处理使每个卷积层提取的深层特征信息都在有效范围内并向下一层传递,抑制因网络层数增加而带来的传递扰动。若对某一层的输出数据只进行批量标准化后就直接输入下一层,误差反向传播过程中会影响本层网络学习到的特征。因此,需引入可学习参数,使网络可以恢复出该层数据的原始特征分布。前向传递公式可表示为

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \xi^{(k)}, \quad (3)$$

$$\mu_B \leftarrow \frac{\sum_{i=1}^n x_i}{n}, \quad (4)$$

$$\sigma_B \leftarrow \frac{\sum_{i=1}^n (x_i - \mu_B)^2}{n}, \quad (5)$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, \quad (6)$$

$$y_i \leftarrow \gamma \hat{x}_i + \xi, \quad (7)$$

式中: $n$  为样本总数; $\hat{x}^{(k)}$  为归一化后的值; $x^{(k)}$  为输入  $x$  的第  $k$  维数据; $y^{(k)}$  为对  $x^{(k)}$  批量标准化后的结果; $\mu_B$  为均值; $\sigma_B$  为标准差; $\gamma$  和  $\xi$  为可学习重构参数; $\epsilon$  为一个很小的常数,防止分母为 0。

### 2.1.4 全连接层

全连接层能够把高维特征变到低维,同时,将前面所有的有用信息进行整合提取,通过激活函数的非线性映射,多层全连接可以用于模拟任何非线性变换。

### 2.1.5 Softmax 函数层

深度卷积神经网络训练是一个端到端的过程,最后一层的 Softmax 分类器将分类结果映射到概率域,目标是 minimized 预测类别的概率分布和损失。常用的损失函数有均方差代价函数(Square-loss)、指数代价函数(Exp-loss)和交叉熵代价函数(Cross-entropy-loss)。均方误差代价函数和指数代价函数分别表示为

$$L_s(Y_i, y_i) = \frac{\sum_{i=1}^n (Y_i - y_i)^2}{n}, \quad (8)$$

$$L_e(Y_i, y_i) = \frac{\sum_{i=1}^n \exp(-Y_i y_i)}{n}, \quad (9)$$

式中:  $y_i$  为第  $i$  个输入的分类标签预测值;  $Y_i$  为标签真实值。

这里引入交叉熵代价函数

$$L_c(Y_i, y_i) = -\frac{1}{n} \sum_{i=1}^n [Y_i \ln y_i + (1 - Y_i) \ln(1 - y_i)]. \quad (10)$$

交叉熵代价函数是凸函数,满足在随机化初始权重之后,分类结果越差梯度越大的性质。Softmax 分类器算法为

$$y_i = \frac{\exp p_i}{\sum_n \exp p_n}. \quad (11)$$

式中:  $p_i$  为第  $i$  个输入 Softmax 层的分类标签预测值的概率。当 Softmax 分类器与交叉熵损失函数结合时,可求得输出层误差为

$$\frac{\partial y_i}{\partial p_i} = \begin{cases} -\frac{\exp p_i}{\sum_n \exp p_n} \cdot \exp p_i = -y_j \cdot y_i, i \neq j \\ \frac{\exp p_i \cdot (\sum_n \exp p_n) - (\exp p_i)^2}{(\sum_n \exp p_n)^2} = y_j \cdot (1 - y_i), i = j \end{cases}, \quad (12)$$

$$\frac{\partial L_c}{\partial p_i} = \sum_n \frac{\partial L_c}{\partial y_n} \cdot \frac{\partial y_n}{\partial p_i} =$$

$$-Y_i(1 - y_i) + \sum_{n \neq j} \frac{Y_n}{y_n} \cdot (y_j \cdot y_i)_i = y_i - Y_i. \quad (13)$$

## 2.2 SVM 分类器

SVM 分类器具有全局最优、简洁和灵活的特性,已经被广泛应用于多种分类预测<sup>[18]</sup>和模式识别任务,取得了良好的效果<sup>[19]</sup>。不同于传统神经网络最小化经验风险, SVM 的基础是最小化结构风险<sup>[20]</sup>。因此, SVM 分类器和 Softmax 函数层最大的区别在于损失函数不同, SVM 的损失函数可以表示为

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + \Delta), \quad (14)$$

式中:  $s_{y_i}$  为真实类别  $y_i$  的得分;  $s_j$  为其他类别的得分;  $\Delta$  为边界值;  $L_i$  为某输入图像的损失。由(14)式可以看出,当正确类别结果获得的得分与不正确类别获得的得分之差大于固定的边界大小时,损失为 0,不再修正该样本的分类结果。

## 3 改进的 Softmax-SVM 模型

### 3.1 Softmax-SVM 分类器

在 DCNN 训练时, Softmax 函数层对于分类正确的输入样本仍会耗费时间计算损失,且反向传播使得该样本的分类概率更接近于 1,故容易造成过拟合。由于 SVM 分类器在一定范围内忽略了与真实值间的误差,因此对已经正确分类的样本不再进行修正,只须继续训练分类错误的样本。这里提出 Softmax-SVM 分类器,将 DCNN 作为特征提取器,提取 FC1 层特征用于训练 SVM 分类器,在节省计算时间的同时避免了过拟合。Softmax-SVM 分类器流程如图 1 所示,算法步骤如下。

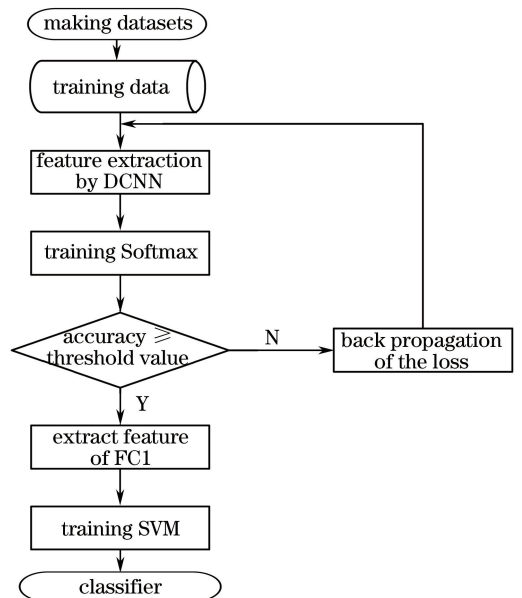


图 1 Softmax-SVM 分类器流程

Fig. 1 Flow chart of Softmax-SVM classifier

第一步: 输入的训练样本尺寸归一化为  $96 \text{ pixel} \times 96 \text{ pixel}$ , 采用 DCNN 模型提取特征, 由 Softmax 函数层结合交叉熵代价函数进行网络训练, 设置训练集识别准确率的阈值;

第二步: 当训练集的识别准确率达到阈值时 DCNN 停止训练;

第三步: 提取网络中的全连接层 FC1 的特征向量, 并接入到 SVM 分类器, 结合 hinge 损失函数继续训练, 生成最终的分类模型。

### 3.2 模型结构

基于 DCNN 的 Softmax-SVM 分类器模型结构如图 2 所示, 输入的训练样本为 RGB 彩色图像, 因此输入为 3 个通道层, 图像大小为  $96 \text{ pixel} \times 96 \text{ pixel}$ ; 卷积层 C1 含有 16 个大小为  $5 \times 5$  的卷积核, 滑动步长为 2; 卷积层 C2 含有 32 个  $3 \times 3$  的卷积核, 步长为 2; 最大池化层 S1 的核尺寸为  $2 \times 2$ , 滑动步长为 2; 卷积层 C3 含有 64 个  $3 \times 3$  卷积核, 步

长为 2; 卷积层 C4 含有 128 个  $3 \times 3$  卷积核, 步长为 2; 平均池化层 S2 的核大小为  $2 \times 2$ , 滑动步长为 2; 全连接层 F1 包含 729 个神经元; 全连接层 F2 包含 27 个分类单元, 数目与车型种类保持一致。此外, 还在每个卷积层后加入激活函数用于提高网络的表征能力; 加入批量标准化层, 抑制了因网络层数增加而带来的数据传递扰动。为降低底层特征丢失的可能性, 对每两层卷积层提取特征后进行一次池化操作, 以提高特征信息的完整性和关联性。

## 4 实验数据与结果分析

结合卷积神经网络经典参数模型和多次实验结果, 将深度卷积神经网络的初始学习率设置为 0.01, 动量因子设置为 0.9, 为避免过拟合, 加入正则规范化因子 L2, 设置为 0.0005。实验的笔记本内存 4 GB, CPU 运行在 Interli5 条件下, 操作系统为 Windows7, 软件平台为 Matlab2018a。

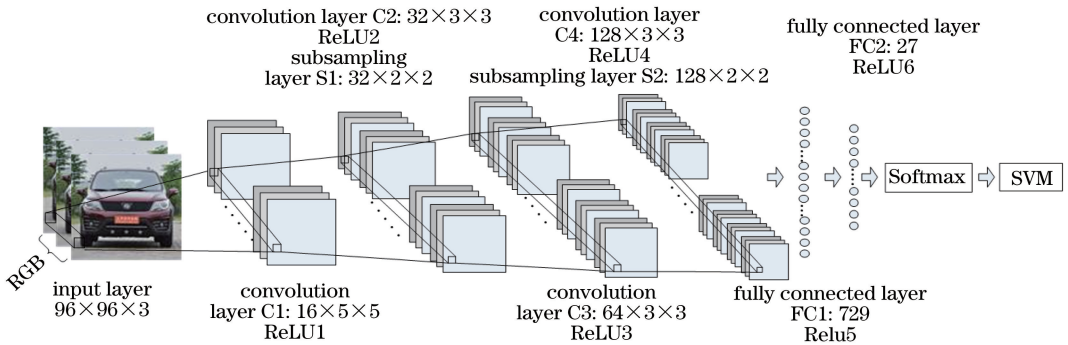


图 2 基于 DCNN 的 Softmax-SVM 分类器模型结构图

Fig. 2 Structural diagram of Softmax-SVM classifier based on DCNN

### 4.1 实验数据集

为验证基于深度卷积神经网络的 Softmax-SVM 分类器的分类性能, 采集了不同光照、不同噪声、不同背景等复杂环境下的共 27 种常见车型的正视图, 采用随机剪裁、反转等一系列随机变换的方法增大样本数量。为适应卷积神经网络的输入尺寸, 将图片的像素大小归一化为  $96 \times 96$ , 添加标签, 采用 ImageNet<sup>[21]</sup> 数据集构建方法, 建立了一个用于复杂背景下精细车型识别的标准数据库。该数据库包含奥迪、大众、别克等品牌, 共 2698 张图片, 尤其还包含了如启腾 V60、EX80, 景逸 X3、X5, 大众 CC、新宝来, 丰田凯美瑞、皇冠、RV4, BJ 20、战旗等同一汽车厂商的相近车型, 以增加实验的可行性。随机选取每种车型样本的 80% 作为训练集 (共 2160 张), 20% 作为测试集 (共 538 张), 27 种车型部分样本如

图 3 所示, 其中前 17 种车型为同一厂商相近车型展示。图像中汽车尺寸在整张图像中的占比涵盖了 60% 至 90% 之间的多种比例, 保证了汽车占比的多样性, 使得实验结果更具稳健性和可靠性。

### 4.2 结果与分析

训练传统 DCNN 模型达到阈值是训练基于 DCNN 的 Softmax-SVM 分类器的重要过程, 因此, 实验研究 Softmax 损失函数的选择。采用均方差代价函数、交叉熵代价函数、指数代价函数, 分析 DCNN 模型中训练集和测试集的识别精度和损失随训练次数的变化, 如图 4 所示。模型完成 350 次训练时, 不同代价函数的 DCNN 模型的识别准确率如表 1 所示, 不同代价函数的 DCNN 模型的损失如表 2 所示。

由图 4 可知, 采用交叉熵代价函数的 DCNN 模型的性能变化曲线收敛速度最快, 训练集和测



图 3 27 种车型部分样本

Fig. 3 Partial samples of 27 types of vehicle models

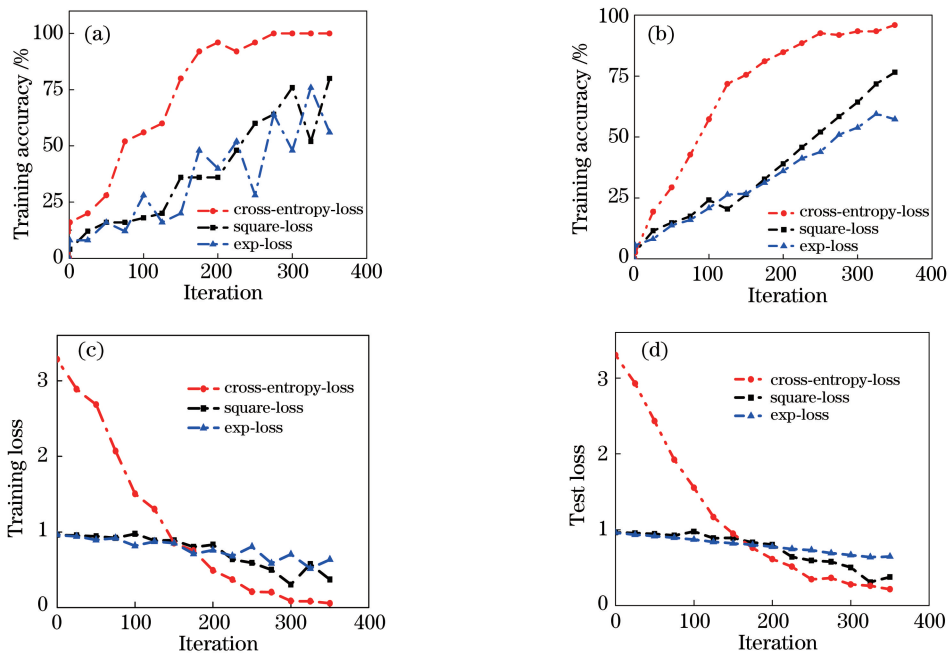


图 4 采用不同损失函数的 DCNN 性能随训练变化的曲线。(a) 训练准确率；(b) 测试准确率；(c) 训练损失；(d) 测试损失

Fig. 4 DCNN performance versus training under different loss functions. (a) Training accuracy;

(b) test accuracy; (c) training loss; (d) test loss

试集的损失曲线稳定下降, 识别精度曲线稳定上升, 在训练 200 次后, 识别准确率均已经高于 90%。而使用均方差代价函数和指数代价函数训练 DCNN 模型时, 尤其是训练集, 其识别准确率和损失的变化曲线均出现了明显的反复和波动。这是由于 Softmax 函数层为 Sigmoid 函数的多分类

形式, 易发生梯度消失现象。根据链式求导法则及损失函数导数形式可知, 在使用交叉熵代价函数为损失函数时, 满足梯度下降原则, 避免了梯度消失现象, 克服了权重更新过慢的问题, 提高了学习速度, 使得输出值能够更快更好地接近真实值, 模型训练更加稳定。

表1 不同代价函数的 DCNN 训练 350 次时的识别准确率  
Table 1 Recognition accuracies of DCNN training for 350 times under different loss functions

Loss function	Accuracy /%	
	Training set	Test set
Square-loss	80.0	76.58
Cross-entropy-loss	99.8	95.51
Exp-loss	56.0	57.28

表2 不同代价函数的 DCNN 训练 350 次时的损失  
Table 2 Losses of DCNN training for 350 times under different loss functions

Loss function	Loss/arb. units	
	Training set	Test set
Square-loss	0.3729	0.3886
Cross-entropy-loss	0.0810	0.2144
Exp-loss	0.6354	0.6448

此外,结合表1、表2在测试样本上的性能评估结果可知,经过350次训练后,使用交叉熵代价函数的DCNN模型的最高测试精度为95.51%,相比使用均方差代价函数和指数代价函数分别提高了18.93%和38.23%;使用交叉熵代价函数的DCNN模型的损失为0.2144,只有均方差代价损失的55%,指数代价损失的33%,证明交叉熵代价函数能够提高DCNN模型的拟合效果和稳定性。

为了验证在DCNN模型中引入Softmax-SVM分类器,能够在复杂环境背景的干扰下提高细粒度车型的识别效果,对分类器的识别精度和速度两个性能指标进行统计,实验同时对比了未改进的DCNN模型、HOG特征结合SVM、SURF特征结合词袋算法3种识别方法。表3为不同分类器模型测试的准确率,给出了采用4种不同模型识别269张测试样本的准确率。表4为不同分类器训练与识别全部测试样本的时间,统计了4种不同模型识别全部测试样本所耗费的时间。

表3 不同分类器模型测试准确率

Table 3 Test accuracies of different classifier models

Feature extraction	Classifier	Accuracy /%
HOG	SVM	87.59
SURF	Bag for word	46.00
DCNN	Softmax	95.51
DCNN	Softmax-SVM	97.78

由表3可知,在同样的图片预处理条件下,使用提出的基于DCNN的Softmax-SVM分类器模型的测试集识别精度达到了97.78%,而SURF特征结合词袋算法和HOG特征结合SVM分类器两种传统模式识别方法的测试准确率只有46%和87.59%,

表4 不同分类器训练与识别全部测试样本的时间  
Table 4 Time for training and recognition of all test samples by different classifiers

Feature extraction	Classifier	Time /s	
		Training set	Test set
HOG	SVM	195.212	2.127
SURF	Bag	279.534	8.869
DCNN	Softmax	312.386	2.441
DCNN	Softmax-SVM	239.690	0.759

未改进的DCNN模型的识别准确率为95.51%。实验结果表明,在特征提取方面,深度卷积神经网络的提取能力和去噪效果明显优于手工提取特征的方法;相比于使用Softmax函数层的DCNN,Softmax-SVM分类器的识别精度提高了2.27%,改善了Softmax层的过拟合现象,具有更好的泛化能力。

在速度方面,由表4可知,在训练过程中,基于DCNN的Softmax-SVM分类器训练所用的时间为239.690 s,只略高于结合HOG特征的SVM分类器的195.212 s,但比较分析表3中识别精度的数据可以看出,耗时的增加是有意义的。而在测试方面,识别全部269张测试样本,采用基于DCNN的Softmax-SVM分类器只需要0.759 s,明显低于为改进的DCNN的2.441 s和其他两种传统模型的2.127 s和8.869 s,在提高识别精度的基础上,实现了识别速度的大幅度提升,验证了Softmax-SVM分类器的可行性。

提取DCNN模型的不同层特征用于SVM训练,生成的Softmax-SVM分类器性能存在差异。表5为提取DCNN不同层特征训练SVM的Softmax-SVM性能,统计了识别测试样本的损失、精度、耗时的最优结果。由表5可知,当提取的特征为FC2、ReLU6、FC1、ReLU5层时,识别精度和速度均高于未经改进的DCNN模型。而且综合评估4个指标可见,提取FC1层特征输入SVM时Softmax-SVM分类器性能最好,损失最小值为1.0018,识别精度最高为97.78%,由于该全连接层的神经元是FC2层的27倍,因此识别时间略小于FC2层的结果。

表5 提取DCNN不同层特征训练SVM的Softmax-SVM性能

Table 5 Softmax-SVM performance in extracting features of different layers of DCNN to train SVM

Layer	Accuracy /%	Loss	Time /s
FC2	97.03	1.0024	0.4609
ReLU6	97.40	1.0020	0.7639
FC1	97.78	1.0018	0.7591
ReLU5	95.54	1.0027	4.5576

## 5 结 论

复杂背景下,细粒度车型识别的环境干扰因素较多,为避免卷积神经网络的分类层 Softmax 在训练过程中发生过拟合现象,从车型识别的精度和耗时角度出发,提出了基于深度神经网络的 Softmax-SVM 分类器,先将深度神经网络作为共性特征提取器,通过 Softmax 函数层,训练达到识别准确率阈值后,提取 FC1 层数据用于训练 SVM 分类器。针对复杂环境背景下精细车型进行识别实验。基于 DCNN 模型的 Softmax-SVM 分类器识别 269 张测试样本,识别精度达到了 97.78%,相比于传统模式识别方法和改进前的 DCNN 模型分别提高了 51.78%,10.19%,2.27%;耗时仅为 0.759 s,只有未改进的 DCNN 模型的 31%。因此,在复杂图像背景下,基于深度神经网络的 Softmax-SVM 模型特征提取能力强,识别精度高,运行速度快,优势明显,为类间差异小且环境、噪声、光照等因素干扰多的复杂背景下的细粒度车型识别提供了应用参考。

### 参 考 文 献

- [1] Zhang J, Zhang T, Yang Z L, *et al.* Vehicle model recognition method based on deep convolutional neural network [J]. *Transducer and Microsystem Technologies*, 2016, 35(11): 19-22.  
张军, 张婷, 杨正瓴, 等. 深度卷积神经网络的汽车车型识别方法[J]. *传感器与微系统*, 2016, 35(11): 19-22.
- [2] Zheng H L, Fu J L, Mei T, *et al.* Learning multi-attention convolutional neural network for fine-grained image recognition [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, Venice, Italy. New York: IEEE, 2017: 5219-5227.
- [3] Yan Y, Ni B, Yang X. Fine-grained recognition via attribute-guided attentive feature aggregation [C] // The 25th ACM International Conference on Multimedia, October 23-27, 2017, California, USA. New York: ACM, 2017: 1032-1040.
- [4] Huang K, Zhang B L. Fine-grained vehicle recognition by deep convolutional neural network [C] // 2016 9th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics, October 15-17, Datong, China. New York: IEEE, 2016: 465-470.
- [5] Khusnuliawati H, Faticah C, Soelaiman R. Multi-feature fusion using SIFT and LEBP for finger vein recognition [J]. *Telecommunication Computing Electronics and Control*, 2017, 15(1): 478-485.
- [6] Murtaza F, Yousaf M H, Velastin S A. Multi-view human action recognition using 2D motion templates based on MHIs and their HOG description [J]. *IET Computer Vision*, 2016, 10(7): 758-767.
- [7] Ling Y G, Hu W P. Vehicle type recognition based on SURF and integral channel features [J]. *Video Engineering*, 2016, 40(7): 139-143.  
凌永国, 胡维平. 基于 SURF 特征与积分通道特征的车型识别[J]. *电视技术*, 2016, 40(7): 139-143.
- [8] Fang J, Zhou Y, Yu Y, *et al.* Fine-grained vehicle model recognition using a coarse-to-fine convolutional neural network architecture [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(7): 1782-1792.
- [9] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1798-1828.
- [10] Zou Y B, Zhou W L, Chen X Z. Research of laser vision seam detection and tracking system based on depth hierarchical feature [J]. *Chinese Journal of Lasers*, 2017, 44(4): 0402009.  
邹焱飏, 周卫林, 陈向志. 基于深度分层特征的激光视觉焊缝检测与跟踪系统研究[J]. *中国激光*, 2017, 44(4): 0402009.
- [11] Zhou Y C, Xu T Y, Zheng W, *et al.* Classification and recognition approaches of tomato main organs based on DCNN [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2017, 33(15): 219-226.  
周云成, 许童羽, 郑伟, 等. 基于深度卷积神经网络的番茄主要器官分类识别方法[J]. *农业工程学报*, 2017, 33(15): 219-226.
- [12] Song J, Kim H I, Yong M. Fast and robust face detection based on CNN in wild environment [J]. *Journal of Korea Multimedia Society*, 2016, 19(8): 1310-1319.
- [13] Liu B, Yu X C, Zhang P Q, *et al.* A semi-supervised convolutional neural network for hyperspectral image classification [J]. *Remote Sensing Letters*, 2017, 8(9): 839-848.
- [14] Qu L, Wang K R, Chen L L, *et al.* Fast road detection based on RGBD images and convolutional neural network [J]. *Acta Optica Sinica*, 2017, 37(10): 1010003.

- 曲磊, 王康如, 陈利利, 等. 基于RGBD图像和卷积神经网络的快速道路检测[J]. 光学学报, 2017, 37(10): 1010003.
- [15] Cortes C, Vapnik V. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273-297.
- [16] Elleuch M, Maalej R, Kherallah M. A new design based-SVM of the CNN classifier architecture with dropout for offline Arabic handwritten recognition [J]. Procedia Computer Science, 2016, 80: 1712-1723.
- [17] Li X Q, Zhang Y, Liao D. Mining key skeleton poses with latent SVM for action recognition [J]. Applied Computational Intelligence and Soft Computing, 2017, 2017: 1-11.
- [18] Cheng L Y, Mi G Y, Li S, *et al.* Quality diagnosis of joints in laser brazing based on principal component analysis-support vector machine model [J]. Chinese Journal of Lasers, 2017, 44(3): 0302004.
- 程力勇, 米高阳, 黎硕, 等. 基于主成分分析-支持向量机模型的激光钎焊接头质量诊断[J]. 中国激光, 2017, 44(3): 0302004.
- [19] Huang F J, LeCun Y. Large-scale learning with SVM and convolutional for generic object categorization [C] // 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), June 17-22, New York, USA. New York: IEEE, 2006: 284-291.
- [20] Lapin M, Hein M, Schiele B. Loss functions for top-k error: analysis and insights [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, Las Vegas, NV, USA. New York: IEEE, 2016: 1468-1477.
- [21] Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3): 211-252.