

一种多任务的卷积神经网络目标分类算法

张苗辉^{1,2}, 张博^{1*}, 高诚诚¹

¹河南大学大数据分析处理河南省重点实验室, 河南 开封 475004;

²河南大学数据与知识工程研究所, 河南 开封 475004

摘要 提出一种基于细粒度图像和多属性融合的多任务卷积神经网络(MTCNN)。该网络主要包含几个关键环节,首先在网络中增加标签输入层,复制并分离输入的多个标签,通过全连接层与多个任务相匹配,增加与标签数量相应的 Softmax 损失函数,来对多个任务进行反向传播;然后,使用显著性检测与角点检测相结合的方法,提取出原始图像中的细粒度图像用于 MTCNN 的数据输入,使神经网络提取到的目标特征具有独特性和区分性;最后,使用非线性激活函数 PReLU,进一步提高网络的分类精度。通过在 Car Dataset 中进行多任务并行训练,测试精度较传统单个任务的分类精度提升 10%,实验结果表明,MTCNN 有较高的泛化能力,对于图像分类的精度有明显的提升。

关键词 机器视觉;多任务卷积神经网络;深度学习;细粒度图像

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.231502

Object Classification Based on Multitask Convolutional Neural Network

Zhang Miaohui^{1,2}, Zhang Bo^{1*}, Gao Chengcheng¹

¹Henan Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng, Henan 475004, China;

²Institute of Data and Knowledge Engineering, Henan University, Kaifeng, Henan 475004, China

Abstract This paper proposes a multitask convolutional neural network (MTCNN) based on fine-grained images and multi-attribute fusion. The network mainly includes the following key links. First, the label input layer is added to the network; the input multiple labels are copied and separated, and then matched to multiple tasks with a fully connected layer. The Softmax loss function corresponding to the number of labels is added to backpropagate multiple tasks. Then, a fine-grained image in the original image is extracted by the combination of saliency detection and corner detection, and used as the input of MTCNN. The target features extracted by the neural network are more unique and distinguishable. Finally, the MTCNN uses the nonlinear activation function PReLU to further improve the classification accuracy of the network. This paper uses the MTCNN to perform multi-task parallel training in the Car Dataset and achieves a 10% improvement in the classification accuracy over the traditional single task. The results show that the MTCNN has high generalization performance and the accuracy of image classification is obviously improved.

Key words machine vision; multi-task convolutional neural network; deep learning; fine-grained image

OCIS codes 150.1135; 100.4996; 330.7326

1 引言

基于卷积神经网络(CNN)的图像分类是近些年计算机视觉领域较为前沿的研究课题。通过使用

神经网络的算法结构,能够自动获取数据的多个抽象级别,发现数据集中隐含的复杂结构,大大推动目标检测与识别、语音识别、药物发现和基因组学等领域的发展^[1]。

收稿日期: 2019-04-12; 修回日期: 2019-05-08; 录用日期: 2019-05-27

基金项目: 国家自然科学基金(61802111)、中国博士后面上基金(2015M582182)、河南省青年骨干教师资助课题(2017GGJS019)、河南省博士后基金(001703007)、河南省教育厅科学技术研究重点项目(19A520002)

* E-mail: zhangbo208@163.com

传统的深度特征的抽取或是更深层的 CNN,大多仅对候选目标赋予单一的属性标签^[2-4],这忽略了物体本身具有多属性的基本特征,在神经网络训练的过程中并没有为这些属性分配独立的训练任务,或对各样本的属性间进行有效的融合,导致得到的样本信息较少,置信度不高。

多任务学习是机器学习中的一个重要的研究方向,其目的是提高神经网络的泛化能力,采用并行训练的方式可达到共享数据源中包含的不同信息的目的^[5]。多任务卷积神经网络(MTCNN)的关键在于其多个损失函数的反向传播,如 DeepID2 中提出一种联合训练方式,将人脸确认损失和人脸分类损失分别作用于神经网络的反向传播^[6],该方法将 SoftmaxLoss 与 ContrastiveLoss 相结合,分别用于人脸分类和人脸确认。与文献[6]相似,基于限定任务的深度卷积神经网络(TCDCN),通过 MTCNN 对面部关键点的定位和属性预测进行了有效的结合^[7-8]。Faster R-CNN 通过结合窗口回归和窗口分类,将 MTCNN 应用于目标检测,该算法在两个任务之间共享用于提取目标信息的卷积层^[9],取得了非常好的效果,是目前最有效的目标检测框架之一。

Part-based R-CNN 等一系列基于强监督学习

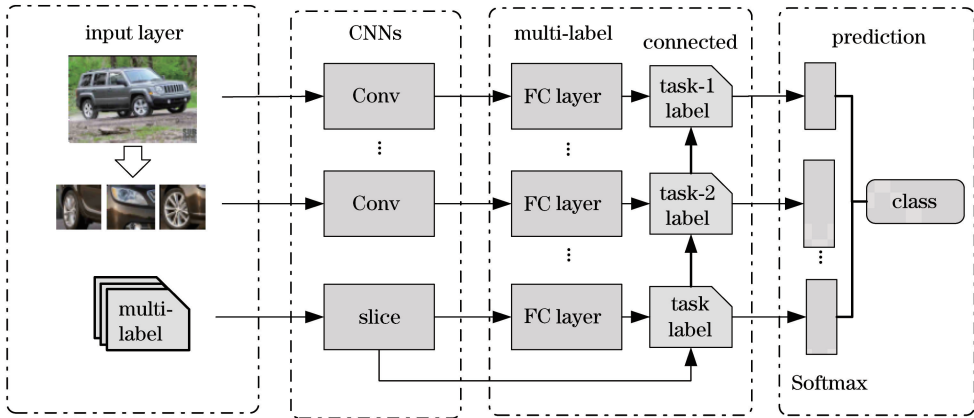


图1 MTCNN 的结构

Fig. 1 Structure of MTCNN

在图1中,输入图像由原始图像和细粒度图像组成,利用视觉注意机制对原始图像中的关键区域进行定位,自动筛选出样本的重要部分,引入细粒度图像中的卷积特征来使图像的原始信息得到增强。CNNs 的部分沿用了 AlexNet^[13] 的架构,由同样的卷积层和多个全连接层组成。不同的是,这里使用 PReLU 来替换原始卷积层之后的 ReLu 单元,增加相应任务数目的全连接层,达到与标签的结合,同时,多个 Softmax 损失函数用于反向传播,这些损

的细粒度神经网络^[10-12],在目标检测中取得了较好的效果,但基于强监督学习的细粒度图像分类需要花费更大的人工标注代价,因此本文提出一种基于显著性区域和角点检测相融合的方法,该方法能够无干预地提取用于神经网络训练和分类的图像细节区域,然后对得到的细粒度图像进行训练和分类。

针对传统 CNN 泛化能力弱,分类效率低的问题,本文提出一种 MTCNN,采用多任务并行训练的方法,将输入图像和标签作为两个不同的输入源,多个任务共享同一卷积网络,根据并行训练任务的个数添加相应数量的全连接层(FC)和损失层;同时,将显著性检测与角点检测相结合,提取出细粒度图像作为数据输入,使用图像中提取到的局部区域信息,加强对前景目标的识别;使用 PReLU 非线性激活函数进一步提高网络分类的精度,设计并实现了一种基于细粒度图像的 MTCNN。

2 MTCNN 的结构

MTCNN 总体架构包括 4 个部分,输入图像的生成和多任务标签的设定、CNNs 特征学习、全连接层与多标签的结合、损失函数及结果预测,如图 1 所示。

失函数的权重之和为 1。

2.1 MTCNN 并行训练的实现

CNNs 分类网络的输入由样本和其对应的标签组成,称为监督学习。通常用于图像分类的网络,只为每个样本赋予单一的标签,通过大量的迭代学习得到预测函数的权重,对新的待测样本赋予标签属性。CNNs 的框架为机器学习的发展打下了扎实的基础^[14-15]。但样本往往具有多属性的特征,神经网络的迭代训练需要结合更多信息,才能提升神经网络

络的分类精度和泛化能力^[16],因此,MTCNN 设定更多的标签让神经网络学习更多的内容。

对 Caffe 源码进行修改,使得 Caffe 深度学习框架实现对多标签数据格式的支持和多标签的分类脚本。网络结构中利用 Caffe 框架下的 Slice 层对输入的多标签进行复制,同时将多个标签分割成相互独立的层。最关键的地方在于每个标签对应各自损失函数的计算。Softmax 函数为

$$S_i = \frac{e^{S_{y_i}}}{\sum_{j=1}^c e^{S_j}}, \quad (1)$$

式中: S_{y_i} 为第 i 个样本的线性得分函数; S_j 为第 j 个类别的线性得分函数。在计算梯度时,该函数在数值计算上更为稳定,因此每一个标签 k 所对应的损失函数都由 Softmax 函数转化而来,可表示为

$$L_i^k = -\omega_k \ln \frac{e^{S_{y_i}^k}}{\sum_{j=1}^c e^{S_j^k}}, \quad (2)$$

式中, $S_{y_i}^k$ 为第 k 个标签的第 i 个样本的线性得分函数, j 为对应第 k 个任务的类别数, ω_k 为第 k 个任务损失函数的权重,初始化每个任务的权重为 $\omega_k = 1/k$,通过设定权重的大小能够调整多个任务的优先级。由于所有的输入样本共享同一 CNN,所以其损失函数的权重满足

$$\sum_{k=1}^K \omega_k = 1, \quad (3)$$

式中: K 为总任务数。

2.2 细粒度图像的产生

监督学习的方案分为强监督和弱监督,其中,将通过手工标注的 object bounding box 和 part annotation 等信息用于模型训练的方法称为基于强监督信息的细粒度图像分类,如 Part-based RCNN^[10]和 Mask-RCNN,这种标注信息的获取代价

十分昂贵。因此基于弱监督信息的细粒度图像分类成为人们关注的焦点,在不同程度上获得了较高的召回率,基于弱监督信息获取的关键是 part annotation 的有效获取^[17]。作为一种弱监督学习的模型优化方案,本文对图像细粒度与视觉注意力机制进行有效的融合,提出一种有效的 part annotation 标注机制。首先介绍视觉注意力机制。

对于人类来说,视觉注意机制就是面对一个场景时,人类首先会被该场景中最显著或者最引人注目的某一个局部所吸引,然后自动地对感兴趣区域进行处理而选择性地忽略不感兴趣区域,这些人们感兴趣区域被称为显著性区域^[18]。最后,根据获取到的显著性区域对相关目标进行分类识别。基于此,采用基于谱残差的显著性检测算法来获取样本图片中有较强区分性的图像区域。

2.2.1 基于谱残差的显著性检测

对于一幅样本图像,首先,通过视觉显著性检测提取样本中的感兴趣区域,该区域包含样本图像所属类别特有的属性。

该算法的核心是谱残差的概念,即

$$R(f) = L(f) - A(f), \quad (4)$$

式中, $L(f)$ 是图像的 log 振幅谱, $A(f)$ 是由均值滤波得到的平均 log 振幅谱, $R(f)$ 是图像的谱残差。

显著性图的计算式为

$$S(x) = g(x) * \mathcal{F}^{-1}[\exp(R(f) + P(f))], \quad (5)$$

式中, $g(x)$ 是高斯模糊函数, $\mathcal{F}^{-1}(\cdot)$ 是傅里叶逆变换, $P(f)$ 是图像相位谱。

首先对训练样本采用显著性检测方法,对包含该样本类别特有信息的区域进行预定位。通过谱残差法检测出的车辆样本的显著性区域包含车轮、车前脸及车的框架轮廓等,如图 2(b)所示,这和人类视觉所认为该图片是汽车的显著性判别区域基本一致。



图 2 图像细节区域的检测。(a)原图;(b)显著性检测;(c)角点检测

Fig. 2 Detection of image detail areas. (a) Original image; (b) saliency detection; (c) corner detection

2.2.2 细粒度图像的标定

基于弱监督信息的细粒度图像分类,设定相应

的规则,使计算机能够自主地对图像细节区域进行标注,减少了人工干预的程度。角点检测能够将像

素邻域点中灰度变化较大或亮度对比足够大的点定义为关键角点,具有计算简便、抗噪声能力强的特性。由图 2(c)可以看出,在给定合适的阈值条件下,角点检测能够很好地标注出数据集样本中的关

键位置,进一步筛选出图像的细节区域。对显著性区域检测中给定的视觉感兴趣区域进行角点检测,再将全部角点进行聚类,舍弃零散的角点,得到最终的细粒度图像,如图 3 所示。



图 3 细粒度图像

Fig. 3 Fine-grained images

ReLU 函数作为神经网络的激活函数首先在文献[13]中被提出,该函数增加了神经网络的非线性因素,使得神经网络可以应用于各种非线性模型。当 ReLU 函数的输入为负时,不能通过反向传播得到一个梯度,导致激活函数永久关闭(dying ReLU)的问题,为解决这一问题,He 等^[19]提出了一种参数修正的激活函数 PReLU,可表示为

$$f(y_i) = \begin{cases} y_i, & \text{if } y_i > 0 \\ a_i y_i, & \text{if } y_i < 0 \end{cases}, \quad (6)$$

式中, i 为通道数, a_i 是一个随动量 μ 和学习率 ϵ 不断更新的权重。更新后的权重为

$$\Delta a'_i = \mu \Delta a_i + \epsilon \frac{\partial \epsilon}{\partial a_i}. \quad (7)$$

在 PReLU 的初始化中,将 a_i 的初始值设置为

0.25。实验表明,随着迭代次数的不断增加, a_i 的值始终小于 1,稳定在 0.2~0.5 之间。在文献[17]中,使用 PReLU 作为网络模型的激活函数,在 ImageNet 2012 数据集上 top-1 的错误率仅为 24.27, top-5 的错误率为 7.38,较 VGG-16 与 GoogLeNet 模型在错误率上降低了 2%。为验证 PReLU 激活函数对 MTCNN 精度提升的有效性,将 PReLU 激活函数应用于 AlexNet,对车型和车辆颜色分别进行 2500 次的迭代训练,与传统的 AlexNet 进行对比,结果如图 4 所示。可以看出,传统的 AlexNet 对车型和车辆颜色的测试精度分别为 77% 和 74.05%,而在 PReLU 激活函数的作用下,对车型和车辆颜色的测试精度达到了 79.45% 和 77.7%,通过实验证明 PReLU 激活函数在实际

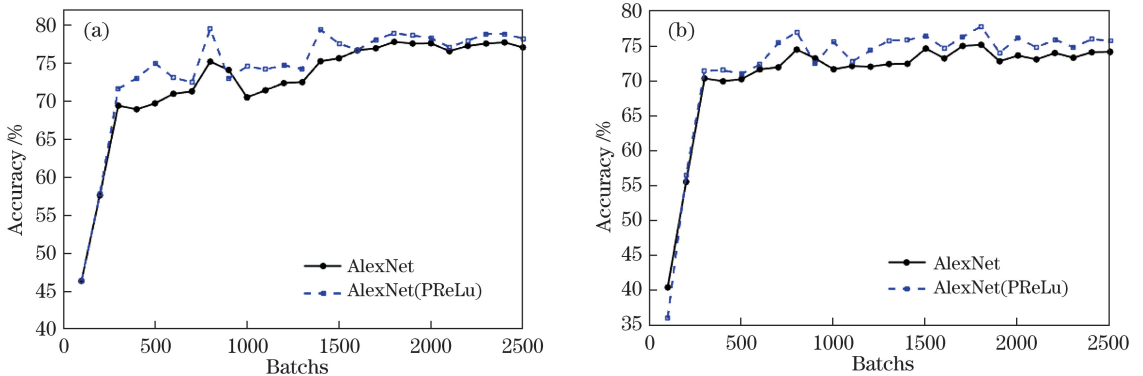


图 4 PReLU 作用下车型和车辆颜色数据集的精度变化。(a) 车辆类型精度变化;(b) 车辆颜色精度变化

Fig. 4 Accuracy changes of car-type and car-color datasets under PReLU. (a) Accuracy changes of car-type; (b) accuracy changes of car-color

应用过程中对于校验集精度的提升在 5% 左右。由此可知在 PReLU 作用下, 神经网络的收敛速度更快, 在达到最高精度后, 其振荡较小, 特别是对车型这种更高语义层次的输出有着很好的优化作用。

3 实验

3.1 数据集

在 Car Dataset 中, 首先根据车型将数据集划分为 SUV 和 Sedan, 然后设定第二个训练任务, 将车

辆颜色划分为黑、蓝、灰、红、白 5 种类型, 这样同一幅车辆图像被赋予 2 个标签, 用于 MTCNN 的训练, 如图 5 所示。该数据集有 8144 张训练样本(图片)和 8041 张测试样本, 这里使用全部的训练样本进行训练, 选取一半的测试样本作为校验集, 其余为测试集, 以下称该数据集为 Car-multi。同时, 将 Car Dataset 复制两份, 采用相同的比例划分数据集, 一份单独用于车型的训练, 称为 car-type; 另一份用于车辆颜色的训练, 称为 car-color。



图 5 Car Dataset

Fig. 5 Car Dataset

3.2 实验分析

在训练过程中, 环境为 GeForce GTX Titan X 12 GB GPU, 选用 Car Dataset 数据集进行测试, 训练设置相同的 batch size 为 256, 这个数值依赖于 12 GB 的 GPU 显存, 使得收敛速度更快。MTCNN 采用 Caffe 框架进行搭建, 设定每 100 次迭代就对验证集进行一次精度的估计, 并分析训练过程中精度的变化。

3.2.1 Car Dataset 在 MTCNN 上的测试

由于 Car Dataset 各类别训练样本的数量不一致, 所以在数据集的数量上做了一些调整。同时, 为验证细粒度图像对 MTCNN 精度的提升作用, 设定

MTCNN* 为加入细粒度图像的网络。将调整后的数据集分为 3 组进行测试, 如表 1 所示。其中, Train 和 Test 分别为该数据集下训练和测试的样本数, Accuracy 为该组实验获得的最高精度。

对实验结果进行分析, 在 MTCNN* 中, 对于车型和车辆颜色的测试分别获得了 90.23% 和 80.75% 的正确率, 而在 MTCNN 下, 对于车型和车辆颜色的测试分别获得了 81.03% 和 76.18% 的正确率。在同样的训练参数下, 使用经典的 AlexNet 对两个任务分别进行训练, 最高获得了 77.75% 和 75.06% 的正确率, 如图 6 和图 7 所示。

表 1 Car Dataset 分组实验表

Table 1 Group experiments of Car Dataset

Dataset	Model	Class	Train	Test	Accuracy / %
Car-multi(type)	MTCNN*	2	1599	697	90.23
Car-multi(color)	MTCNN*	5	1500	693	80.75
Car-multi(type)	MTCNN	2	1599	697	81.03
Car-multi(color)	MTCNN	5	1500	693	76.18
Car-type	AlexNet	2	1599	697	77.75
Car-color	AlexNet	5	1500	693	75.06

由图 6 可以看出, 在 MTCNN 的实验中, 对于车型和车辆颜色的分类任务测试精度分别为

81.03% 和 76.18%, 将车型和车辆颜色分别作为单个任务进行训练, 其测试精度最高为 77.75% 和

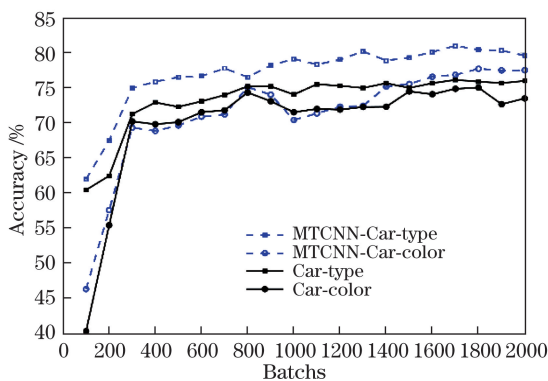


图6 MTCNN 训练 Car Dataset 的精度变化

Fig. 6 Accuracy changes of MTCNN training for Car Dataset

75.06%。在 MTCNN 的作用下每个任务的测试精度都有较大提升。这个优势除了有 PReLU 激活函数带来的些许提升以外,主要归功于多任务之间信息的共享,花费相同的训练时间,同时得到了物体的多个属性。

由图 7 可以看出,在采用细粒度图像作为输入后 MTCNN* 的精度有大幅度的提升,对于车型和车辆颜色测试的最高精度分别达到 90.23% 和 80.75%,在 MTCNN 的基础上又提升了 4%~9%,这比起传统的 AlexNet 有将近 10%~15% 的提升。

表 2 Car Dataset 在不同网络下的表现

Table 2 Performance of Car Dataset under different networks

Dataset	Model	Net-Deep	Batches	Model size /Mbit	Accuracy /%
Car-multi(type)	MTCNN*	8	2000	217	90.23
Car-multi(color)	MTCNN*	8	2000	217	80.75
Car-type	AlexNet	8	1500	202	77.75
Car-color	AlexNet	8	1500	202	75.06
Car-type	GoogleNet-v1	22	2500	48	90.15
Car-color	GoogleNet-v1	22	2500	48	78.78
Car-type	VGG-16	19	3000	527	88.56
Car-color	VGG-16	19	3000	527	77.17

由表 2 的实验结果来看,传统的 AlexNet 在迭代次数超过 1500 次之后就出现过拟合的现象,收敛速度较快但测试精度并不高;GoogleNet-v1 使用了不同尺度的卷积核,并且减少了卷积核的数量,导致其参数更少,模型更小。但该网络需要更多的迭代次数才能接近 MTCNN* 的测试精度;VGG-16 在精度上略低于 MTCNN*,该网络使用更多的卷积核来提升网络的深度,导致其参数量增多,模型大小成倍增加。从测试结果可以看出 MTCNN* 在车辆类型上获得了 90.23% 的精度,这一数值甚至超过了网络层数更深的 GoogleNet-v1。GoogleNet-v1 和 VGG-16 同 AlexNet 的对比来看,更深的网络层

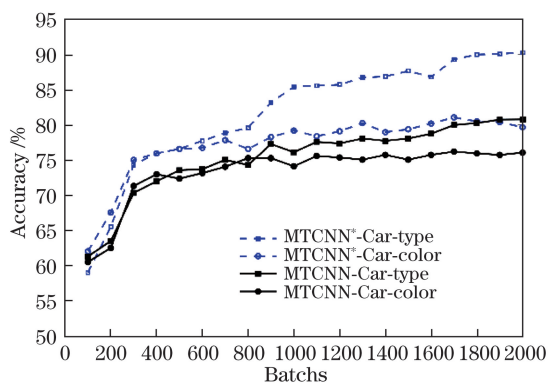


图7 MTCNN* 训练 Car Dataset 的精度变化

Fig. 7 Accuracy changes of MTCNN* training for Car Dataset

3.2.2 MTCNN 的扩展

对 AlexNet 的结构进行进一步改进,提出的 MTCNN 在图像分类中取得了明显的效果。将 MTCNN 同 GoogleNet-v1 和 VGG-16 模型在 Car Dataset 上进行对比测试,车型和车辆颜色的训练样本与测试样本与表 1 相同。在每个网络模型的训练测试实验中, batch size 均设定为 256,且迭代次数 Batches 等网络训练过程中的超参数,均为多次实验后调整的最优参数,测试结果如表 2 所示。

对于精度的提升有明显的效果,若使用 GoogleNet-v1 作为 MTCNN 的基础网络,将会获得更好的分类效果。

MTCNN 的架构不仅在识别精度方面有较大提升,而且可以实现多属性的预测输出。对于一张待检测的样本图片,仅通过一次神经网络的前项计算,就可以得到该样本对应的多个属性标签,如图 8 所示,要想实现测试样本的多属性预测,传统的算法框架需要多个模型的叠加才能实现。

4 结 论

设计了一种多任务输入、多属性输出的高精度



图 8 样本的多标签输出

Fig. 8 Multi-label outputs of samples

卷积神经网络算法框架。实验结果表明,多任务并行训练提升了网络的泛化能力,通过加入基于弱监督的图像细粒度和优化参数的 PReLU 激活函数,使得 MTCNN 在深层特征的挖掘中获得了很好的效果。接下来还会进行大量的移植实验,在卷积层数更深的网络中,采用 MTCNN 架构在标签数量和类别数量成倍增加的情况下,验证其精度的可靠性,使得 MTCNN 可以作为一种通用的多任务学习框架来对大多数现有网络进行扩展。

参 考 文 献

- [1] Al-Saffar A A M, Tao H, Talab M A. Review of deep convolution neural network in image classification[C] // 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), October 23-24, 2017, Jakarta, Indonesia. New York: IEEE, 2017: 26-31.
- [2] He Z C, Zhao L Z, Chen C. Convolution neural network with multi-resolution feature fusion for facial expression recognition[J]. Laser & Optoelectronics Progress, 2018, 55(7): 071503.
何志超, 赵龙章, 陈闯. 用于人脸表情识别的多分辨率特征融合卷积神经网络[J]. 激光与光电子学进展, 2018, 55(7): 071503.
- [3] Li J N, Zhang B H. Face recognition by feature matching fusion combined with improved convolutional neural network [J]. Laser & Optoelectronics Progress, 2018, 55(10): 101504.
李佳妮, 张宝华. 特征匹配融合结合改进卷积神经网络的人脸识别[J]. 激光与光电子学进展, 2018, 55

(10): 101504.

- [4] Abdullah, Hasan M S. An application of pre-trained CNN for image classification [C] // 2017 20th International Conference of Computer and Information Technology (ICIT), December 22-24, 2017, Dhaka, Bangladesh. New York: IEEE, 2017: 17575567.
- [5] Yang P, Zhao P L, Gao X. Robust online multi-task learning with correlative and personalized structures [J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(11): 2510-2521.
- [6] Sun Y, Chen Y H, Wang X G, *et al.* Deep learning face representation by joint identification-verification[C] // Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal, Canada. Canada: NIPS, 2014: 1988-1996.
- [7] Zhang Z P, Luo P, Loy C C, *et al.* Facial landmark detection by deep multi-task learning [M] // Fleet D, Pajdla T, Schiele B, *et al.* Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8694: 94-108.
- [8] Trottier L, Giguère P, Chaib-draa B. Multi-task learning by deep collaboration and application in facial landmark detection[J/OL]. (2018-05-15) [2019-04-11]. <https://arxiv.org/abs/1711.00111>.
- [9] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

- [10] Zhang N, Donahue J, Girshick R, *et al.* Part-based R-CNNs for fine-grained category detection [M] // Fleet D, Pajdla T, Schiele B, *et al.* Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8689: 834-849.
- [11] Peng Y X, He X T, Zhao J J. Object-part attention model for fine-grained image classification[J]. IEEE Transactions on Image Processing, 2018, 27(3): 1487-1500.
- [12] He K, Gkioxari G, Dollár P, *et al.* Mask R-CNN[C] // The IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2961-2969.
- [13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C] // Proceedings of the 25th International Conference on Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada Canada: NIPS, 2012: 1097-1105.
- [14] Szegedy C, Liu W, Jia Y Q, *et al.* Going deeper with convolutions [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 15523970.
- [15] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J/OL]. (2015-04-10)[2019-04-11]. <https://arxiv.org/abs/1409.1556>.
- [16] Zhang C, Bengio S, Hardt M, *et al.* Understanding deep learning requires rethinking generalization [J/OL]. (2017-02-26)[2019-04-11]. <https://arxiv.org/abs/1611.03530>.
- [17] Zhang J, Zhao H D, Li Y H, *et al.* Classifier for recognition of fine-grained vehicle models under complex background [J]. Laser & Optoelectronics Progress, 2019, 56(4): 041501.
张洁, 赵红东, 李宇海, 等. 复杂背景下车型识别分类器 [J]. 激光与光电子学进展, 2019, 56(4): 041501.
- [18] Zhang S L, Xie L B. Salient detection based on all convolutional feature combination [J]. Laser & Optoelectronics Progress, 2018, 55(10): 101502.
张松龙, 谢林柏. 基于全部卷积特征融合的显著性检测 [J]. 激光与光电子学进展, 2018, 55(10): 101502.
- [19] He K M, Zhang X Y, Ren S Q, *et al.* Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1026-1034.