



## 基于部件关注DenseNet的细粒度车型识别

陈立潮, 朝昕, 潘理虎, 曹建芳, 张睿

引用本文:

陈立潮,朝昕,潘理虎,曹建芳,张睿. 基于部件关注DenseNet的细粒度车型识别[J]. 智能系统学报, 2022, 17(2): 402–410.

CHEN Lichao, CHAO Xin, PAN Lihu, CAO Jianfang, ZHANG Rui. Fine-grained vehicle-type identification based on partially-focused DenseNet[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(2): 402–410.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202012012>

## 您可能感兴趣的其他文章

### 生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296–301 <https://dx.doi.org/10.11992/tis.201901004>

### 基于增强AlexNet的音乐流派识别研究

Music genre recognition research based on enhanced AlexNet

智能系统学报. 2020, 15(4): 750–757 <https://dx.doi.org/10.11992/tis.201909032>

### 基于时空约束密度聚类的停留点识别方法

Stay point recognition method based on spatio-temporal constraint density clustering

智能系统学报. 2020, 15(1): 59–66 <https://dx.doi.org/10.11992/tis.201910026>

### 认知视角下的舆论观点句情感计算

Research on computation of affect in public opinion sentences from the cognition viewpoint

智能系统学报. 2017, 12(4): 498–503 <https://dx.doi.org/10.11992/tis.201607023>

### 基于多粒度结构的网络表示学习

Network representation learning based on multi-granularity structure

智能系统学报. 2019, 14(6): 1233–1242 <https://dx.doi.org/10.11992/tis.201905045>

### 一种基于2D时空信息提取的行为识别算法

A behavioral recognition algorithm based on 2D spatiotemporal information extraction

智能系统学报. 2020, 15(5): 900–909 <https://dx.doi.org/10.11992/tis.201906054>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.202012012

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20210622.1126.006.html>

## 基于部件关注 DenseNet 的细粒度车型识别

陈立潮<sup>1</sup>, 朝昕<sup>1</sup>, 潘理虎<sup>1</sup>, 曹建芳<sup>1,2</sup>, 张睿<sup>1</sup>

(1. 太原科技大学 计算机科学与技术学院, 山西 太原 030024; 2. 忻州师范学院 计算机科学与技术系, 山西 忻州 034000)

**摘 要:** 针对细粒度车型识别率低, 车型区别主要集中在鉴别性部件上以及深度学习不能有效对部件进行关注的问题, 提出一种基于部件关注 DenseNet(part-focused DenseNet, PF-DenseNet) 的细粒度车型识别模型。该模型可以基于细粒度车型的车灯、车标等区分性部件进行有效分类, 通过处理层 (process layer) 对车型部件信息反复加强提取并进行最大池化下采样, 获取更多的车型部件信息, 然后通过密集卷积对特征通道进一步复用提取, 密集卷积前嵌入独立组件 (independent component, IC) 层, 获得相对独立的神经元, 增强网络独立性, 提高模型的收敛极限。实验结果表明, 该模型在 Stanford cars-196 数据集上的识别准确率、查全率和  $F_1$  分别达到 95.0%、94.9% 和 94.8%, 高于经典卷积神经网络, 并具有较小的参数量, 与其他方法相比实现了最高准确率, 验证了该车型识别模型的有效性。

**关键词:** 细粒度车型识别; 部件关注; 密集连接网络; 独立组件; 数据增强; 深度学习; 特征提取; 特征复用  
**中图分类号:** TP391    **文献标志码:** A    **文章编号:** 1673-4785(2022)02-0402-09

中文引用格式: 陈立潮, 朝昕, 潘理虎, 等. 基于部件关注 DenseNet 的细粒度车型识别 [J]. 智能系统学报, 2022, 17(2): 402-410.

英文引用格式: CHEN Lichao, CHAO Xin, PAN Lihu, et al. Fine-grained vehicle-type identification based on partially-focused DenseNet [J]. CAAI transactions on intelligent systems, 2022, 17(2): 402-410.

## Fine-grained vehicle-type identification based on partially-focused DenseNet

CHEN Lichao<sup>1</sup>, CHAO Xin<sup>1</sup>, PAN Lihu<sup>1</sup>, CAO Jianfang<sup>1,2</sup>, ZHANG Rui<sup>1</sup>

(1. School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China; 2. Department of Computer Science and Technology, Xinzhou Teachers University, Xinzhou 034000, China)

**Abstract:** Given that fine-grained model recognition rates are low and are mainly concentrated in the diagnostic model difference between parts and that deep learning cannot effectively concern parts, we put forward a fine-grained recognition model—the partially-focused DenseNet. The model can be classified effectively based on its discriminative parts, such as lights and marks of fine-grained vehicle models. First, through the Process Layer, the vehicle part information is repeatedly extracted, and the maximum pool sampling is carried out to obtain more vehicle part information. Then, feature channels are further extracted by multiplexing through dense convolution, and the independent component layer is embedded before dense convolution to obtain relatively independent neurons. This enhances network independence and improves the convergence limit of the model. Experiments show that the model's recognition accuracy, recall rate, and  $F_1$  on the Stanford cars-196 data set reach 95.0%, 94.9%, and 94.8%, respectively, which are higher than the classic convolutional neural network and have a smaller number of parameters. Compared with other methods, the highest accuracy rate is achieved, verifying the effectiveness of the vehicle recognition model.

**Keywords:** fine-grained vehicle type identification; part focus; dense connection network; independent component; data enhancement; deep learning; feature extraction; reuse of characteristics

细粒度车型识别是智能交通领域的重要组成部分, 旨在通过任意角度和光线环境下, 识别出

车辆型号、制造商、生产年份等精准有用的车辆信息。通常粗粒度车型识别提供的车辆信息不充足, 识别出的大类别数据, 无法做到对车辆的有效标识和追踪, 细粒度车型识别区别于一般车型识别, 能够对粗粒度的大类别车型进行精细的子

收稿日期: 2020-12-03. 网络出版日期: 2021-06-22.

基金项目: 山西省自然科学基金项目 (201901D111258); 山西省应用基础研究项目 (201801D221179).

通信作者: 潘理虎. E-mail: [panlh@tyust.edu.cn](mailto:panlh@tyust.edu.cn).

类别识别,精确标识车辆,在交通视频中有效地进行车辆监控,不仅节省了人力资源,提高交通执法效率,而且能够快速分析套牌车辆和外观改装车辆,在车牌信息无法获取时,细粒度的车型信息尤其重要,提高了有关车辆刑事案件的侦破效率,在智能交通、交通数据分析、安防等领域具有重要意义<sup>[1]</sup>。

针对细粒度车型识别问题,国内外学者已经做了大量研究,运用不同的方法提高细粒度车型识别准确率。主要有2种方法:基于车辆部件的方法和基于深度学习的方法。

目前已经有许多基于车型部件的细粒度车型识别方法,Liao等<sup>[2]</sup>分析了车辆部件在细粒度车型分类中的辨别能力,提出了一种新的将各部件的区分能力整合到分类框架中的模型。Krause等<sup>[3]</sup>在局部特征外观和位置的层面上,将两种最先进的2D目标建模表示提升到3D,实现对车型精细识别。Fang等<sup>[4]</sup>从整个车辆图像和检测区域中提取全局特征和局部部件特征,基于整体线索和局部部件特征的隶属层次变化,采用一对一的支持向量机分类器进行分类。车型部件方法的分类效果主要取决于对车型鉴别性部件的提取能力,上述方法的部件提取能力有限,成为了制约识别性能提升的主要瓶颈。

在细粒度车型识别领域,深度学习也得到广泛的研究与应用。Shi等<sup>[5]</sup>提出GLM(generalized large-margin)损失,并与Softmax损失级联训练明显地减少了类间相似性和类内方差,增强模型对车型的表征能力。Ke等<sup>[6]</sup>提出一种密集关注网络(dense attention network, DA-Net),通过其密集的连接和关注单元,使网络对车型的分类能力集中在重要区域和特征,提高了模型的识别能力。Fu等<sup>[7]</sup>基于区域检测和细粒度特征学习的关联性,提出一种递归注意力卷积神经网络(recurrent attention convolutional neural network, RA-CNN),在多尺度上以相互增强的方式递归学习区分区域注意力和基于区域的特征表示。马力等<sup>[8]</sup>提出稀疏化B-CNN,通过训练在特征通道中引入比例因子的B-CNN,并采用正则化按比例因子大小对其稀疏,将不重要特征通道按比例裁剪,消除网络过拟合提高关键特征的显著性。王阳等<sup>[9]</sup>提出双线性残差注意网络,用残差网络替换B-CNN的特征函数,并在残差单元间添加通道注意力和空间注意力模块,获取更为丰富的注意力特征。Valev等<sup>[10]</sup>系统比较和总结了现有深度学习方法在细粒度车型识别上的应用,在细粒度车型数据集Stanford Cars-196上进行训练测试,经典深度卷积网络(deep convolutional neural networks, DCNN)中

DenseNet-161获得较好的识别精度,DenseNet密集连接<sup>[11]</sup>是2017年Huang等提出的一种全新的卷积网络连接模式,以密集连接的方式实现特征复用,使网络学习到更具鉴别性的特征,并通过网络特征复用最大程度上降低参数的冗余,降低了网络随深度增加出现的网络退化等问题。相比基于部件的方法,基于深度学习的方法操作简单,能够直接对二维图像进行处理,从大量训练样本中有效地学习到相应特征,避免了复杂的特征提取过程<sup>[12]</sup>。但细粒度车型分类集中在微小但重要的车型部件特征上,上述方法无法对部件特征进行有效关注。

根据以上问题,结合部件和深度学习的方法,提出一种基于部件关注DenseNet的细粒度车型识别网络。通过提出的处理层(Process Layer)对车型部件信息进行重复加强提取关注,获取更多的部件信息,然后利用嵌入独立组件的密集连接网络进行特征复用,提取高维特征,独立组件降低模型过拟合风险,有效提高模型的收敛极限。提出的PF-DenseNet深度学习网络,能够有效获取细粒度车型的鉴别性部件信息,利用部件特征进行分类。

## 1 细粒度车型识别网络

细粒度车型识别网络模型如图1所示,主体框架采用DenseNet的密集卷积结构,并在框架前加入Process Layer,利用Process Layer直接对输入图像进行下采样处理,相对于一般DenseNet对输入图像的处理,Process Layer首先通过步长为1的卷积核对图像车辆细节特征重复提取加强,获取更加清晰的车型图像特征,保留足够多的车型部件特征细节信息,然后用最大池化进行下采样,降低后续计算量。Process Layer后主体框架采用密集连接网络中的密集卷积(Dense Block)和过渡层(Transition Layer),对图像特征进一步提取和复用,获得更加抽象高维的特征,Transition Layer引用在每一层Dense Block后面,对图像进一步下采样,并在Dense Block和Transition Layer中嵌入IC层,提高模型收敛极限。最后通过输出层(Output Layer)和全连接层(fully connected layers, FC),对图像进行分类,FC通过softmax激活函数对输出作类别概率预测,模型训练采用交叉熵损失函数进行参数更新优化。部件关注DenseNet(part-focused DenseNet, PF-DenseNet)的具体结构参数和输出大小如表1 PF-DenseNet的具体网络结构所示,层结构显示其对应名称的具体网络结构,如Process Layer的结构为 $7 \times 7 \times 96$ -BN-ReLu加上MaxPool( $2 \times 2/s=2$ ), $7 \times 7 \times 96$ 表示一个核大小为



7×7 通道数为 96 的卷积核, 然后进行一个批归一化<sup>[13]</sup>(batch normalization, BN) 和 ReLu<sup>[14]</sup> 激活处理, 最后经过一个核大小为 2×2 步长为 2 的最大池化, 输出 (96,111,111) 表示通道数为 96、宽高尺寸为 111×111 的输出特征图。表 1 中有 4 个

Dense Block<sub>x</sub>: Dense Block<sub>1</sub>、Dense Block<sub>2</sub>、Dense Block<sub>3</sub> 和 Dense Block<sub>4</sub>, 分别由 6、12、36 和 24 个基础层结构块构成, 如 Dense Block<sub>1</sub> 由 6 个连续的基础层结构: IC-ReLu-1×1×192 加上 IC-ReLu-3×3×48 构成。

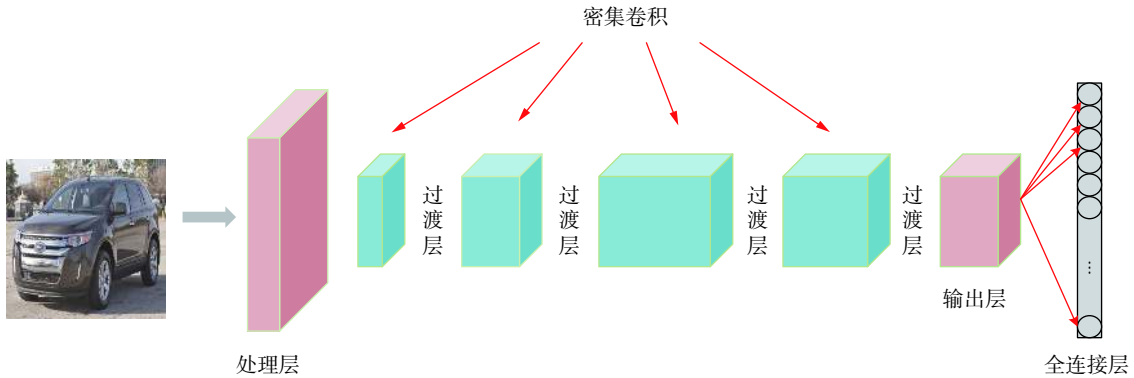


图 1 细粒度车型识别网络模型

Fig. 1 Fine-grained vehicle identification network model

表 1 PF-DenseNet 具体网络结构  
Table 1 PF-DenseNet specific network structure

名称	层结构	输出
Process Layer	7×7×96-BN-ReLu MaxPool(2×2/s=2)	(96,111,111)
Dense Block <sub>1</sub>	IC-ReLu-1×1×192 IC-ReLu-3×3×48 ×6	(384,111,111)
Transition Layer	IC-ReLu-1×1×192 AvgPool(2×2/s=2)	(192,55,55)
Dense Block <sub>2</sub>	IC-ReLu-1×1×192 IC-ReLu-3×3×48 ×12	(768,55,55)
Transition Layer	IC-ReLu-1×1×384 AvgPool(2×2/s=2)	(384,27,27)
Dense Block <sub>3</sub>	IC-ReLu-1×1×192 IC-ReLu-3×3×48 ×36	(2 112,27,27)
Transition Layer	IC-ReLu-1×1×1 056 AvgPool(2×2/s=2)	(1 056,13,13)
Dense Block <sub>4</sub>	IC-ReLu-1×1×192 IC-ReLu-3×3×48 ×24	(2 208,13,13)
Output Layer	BN-ReLu-AvgPool(13×13)	(2 208,1,1)
FC	Softmax	196

### 1.1 Process Layer 卷积特征

细粒度车型分类中, 类别间的差异极其微小, 相比粗粒度车型分类, 更关注图像中微小但十分重要的局部特征, PF-DenseNet 通过 Process Layer 对车型特征重复提取加强, 获取了更加有效清晰的部件信息, 使得网络能够基于车型部件像素信息进行类别预测。

Process Layer 的组成结构如表 1 所示, Pro-

cess Layer 主要由一个核大小为 7×7, 步长为 1 的卷积层以及尺度为 2×2, 步长为 2 的最大池化层组成, 卷积层后面进行 BN 和 ReLu 激活处理。图 2 所示为卷积层操作, 输入特征图通过卷积核, 每次对特征图上卷积核大小区域进行卷积操作输出一个像素值, 每次操作移动一步, 使特征图中心部位车辆的 9/10 像素被重复卷积提取 7 次, 加强了对车型部件的特征提取。然后经过 BN 和 ReLu 层处理, BN 对数据归一化加快网络收敛, ReLu 激活函数增加网络稀疏性。最后利用最大池化层进行下采样, 特征图大小变为原来的 1/2, 降低后续网络计算量。

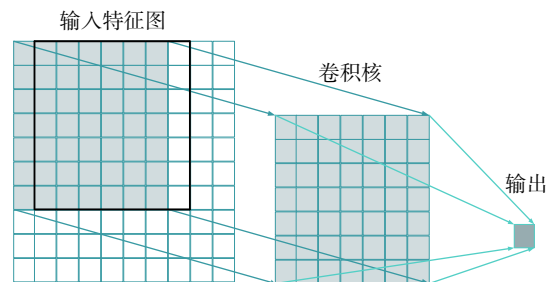


图 2 卷积操作示意

Fig. 2 Convolution operation diagram

如图 3 所示为不同 DenseNet 对同一车型图片输入的不同特征提取效果图, 图 3(a) 为 PF-DenseNet 中密集卷积前经过 Process Layer 提取的特征图, 图 3(b)、3(c) 分别为 DenseNet201 和 DenseNet161 密集卷积前对图像的特征提取图。图 3(a) 通过 Process Layer 对图像同一区域的重复卷积提取, 特征图更加清晰, 保存了更多的部件细节, 车标、车

灯等鉴别性部件信息被有效的提取出来, 使网络模型最后能够有效地根据车型部件像素进行分类。



(a) PF-DenseNet 特征图



(b) DenseNet201 特征图



(c) DenseNet161 特征图

图 3 特征提取图

Fig. 3 Feature extraction diagram

## 1.2 独立密集连接

密集连接基于跨层连接连通网络中前后层使得信号可以在输入层和输出层之间高速流通的思想, 将所有层都进行两两通道连接 (Cat), 使得网络中的每一层都接收前面所有层的特征作为输入, 尽可能地提取到车型的细节信息, 不仅大幅度减少了网络的参数量, 而且在一定程度上降低梯度消失的风险。神经科学家发现, 神经系统的表征能力与系统中独立神经元的数量呈正比, 独立神经元构成的独立神经网络具有更好的特征表达能力和更高的收敛极限, 在 CNN 中引入 IC 层是获得独立神经网络最快速有效的方式, 通过在密集连接中嵌入 IC 层构造独立密集连接, 获得更好的收敛极限。

独立密集连接的 Dense Block 基本结构如图 4

所示, 在密集卷积前嵌入独立组件 IC: {BN-Dropout}, 获得相对独立的网络模型, 提高模型收敛极限, 增加识别准确率<sup>[15]</sup>。

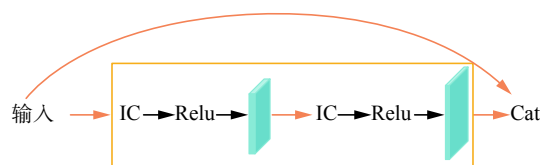


图 4 独立密集连接网络

Fig. 4 Independent dense connection network

IC 中, BN 通过降低内部协变量偏移, 减少内部神经元分布的改变, 加速网络的收敛, BN 公式如式 (1)~(4) 所示:

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i \quad (1)$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \quad (2)$$

首先通过式 (1) 求均值  $\mu$ , 式 (1) 中  $x_i$  为第  $i$  个神经元的输入,  $m$  表示神经元数目, 然后通过式 (2) 得到方差  $\sigma^2$ 。

$$\hat{x} = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} \quad (3)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (4)$$

然后式 (3) 将输入进行标准化处理, 使输入服从均值为 0, 方差为 1 的正态分布, 最后通过式 (4) 对标准化后的值进行平移缩放处理, 数据经过式 (2) 和式 (3) 的处理后, 上一层网络所学习到的特征分布被破坏, 式 (4) 引入可学习重构参数  $\gamma$  和  $\beta$ , 让网络在训练时恢复出原始网络所要学习的特征分布。

Dropout 以概率  $P$  线性减弱任何一对神经元之间的相关性, 降低网络每次训练时的参数量, 增强网络稀疏性, 构造相对独立的神经元激活<sup>[16]</sup>。

将 IC 嵌入到 Dense Block 中以及 Transition Layer 中, 采用 {IC-ReLu-Conv} 的形式, 能够有效提取车型的高层次信息, 提高模型收敛极限, 获得更好的识别准确率。

## 2 实验与分析

### 2.1 数据集

Stanford Cars-196<sup>[3]</sup> 数据集是目前最常用的细粒度车型识别数据集, 图像数据包含 196 类共 16185 张汽车图片, 其中训练集 8144 张, 测试集 8041 张。每张图片的标签有制造商、车辆型号和生产年份 3 个信息。

训练前对数据集进行了预处理, 将所有图像裁剪为  $224 \times 224$  大小, 训练时对训练集做了顺时

针 20° 的角度旋转、水平翻转的在线数据增强, 原图和增强后的图像样例分别如图 5 所示, 扩充后

的训练集为 20360 张, 通过数据增强提升模型的泛化能力, 增强网络的鲁棒性。



图 5 数据增强示例

Fig. 5 Sample data enhancement diagram

## 2.2 实验细节

硬件环境: Intel(R) Xeon(R) Silver 4110 CP 处理器, 内核数 32, GeForce GTX 1080Ti 10 GB 显卡, 单精度浮点数为 11.34TFLOPS, 双精度浮点数为 0.33TFLOPS。软件环境: Ubuntu 16 操作系统, Python3.6 程序设计语言, PyTorch 深度学习框架。

实验网络训练使用随机梯度下降 (stochastic gradient descent, SGD) 方法, 在 ImageNet 预训练权重上进行微调, 每训练一个周期测试一次准确率, 批次大小为 4, 初始学习率 (learning rate, Lr) 为 0.001, 10 个训练周期测试准确率不再上升时, Lr 下降为原来的 0.1, 采用早停策略, 50 个训练周期测试准确率没有上升, 训练自动结束。模型评价指标为准确率 (accuracy)、查全率 (recall) 和  $F_1$ 。准确率公式如式 (5) 所示:

$$\text{accuracy} = \frac{T}{R} \quad (5)$$

式中:  $T$  表示测试集中预测正确的样本数;  $R$  表示测试集样本总数。

$$\text{recall} = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{R_i} \quad (6)$$

式中:  $TP$  表示第  $i$  类预测正确的样本数;  $R_i$  表示第  $i$  类测试样本总数;  $n$  为样本类别数。查全率表示了样本类别的平均准确率。

$$\text{precision} = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FP_i} \quad (7)$$

$$F_1 = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (8)$$

式中: precision 表示查准率;  $FP_i$  表示预测为第  $i$  类的非  $i$  类样本数。式 (8) 中  $F_1$  是 accuracy 和 recall 的调和平均。

## 2.3 实验与结果分析

### 2.3.1 独立组件对实验的影响

提出的 PF-DenseNet 的密集卷积结构采用:

{IC-ReLu-Conv}, 融合了独立组件, 为了验证独立组件对提出的基于部件关注 DenseNet 在细粒度车型识别上的影响, 对 PF-DenseNet 控制单一变量 IC, 在同一训练环境下, 观察 PF-DenseNet 包含 IC 和不含 IC 时, 测试准确率随训练周期变化的情况。

如图 6 所示为独立组件对模型的影响, 初始阶段, 两种模型的测试准确率变化基本相同, 随着训练周期不断增加, 迭代 10 次后包含 IC 的模型具有更高的测试准确率, 60 次迭代后, 不含 IC 的模型测试准确率基本稳定不再上升, 而包含 IC 的模型依然缓慢上升, 分析可知独立组件嵌入文中模型后, 能够有效提高网络的收敛极限, 增加识别准确率。

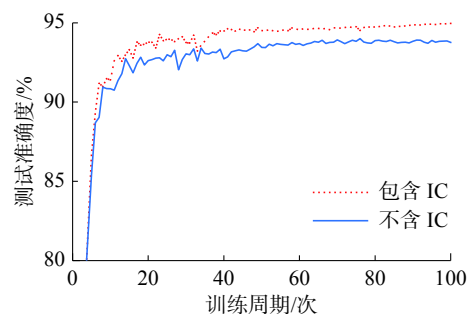


图 6 独立组件对 PF-DenseNet 的影响

Fig. 6 Impact of individual components on the model

### 2.3.2 与普通密集连接网络的对比

图 7 给出了车型 3 种不同姿态下 (侧面、正面和背面) 的 3 种 DenseNet 热力图, 观察不同密集连接网络 DenseNet201、DenseNet161 和 PF-DenseNet 的热力图, 图 7(b)、7(c) 分别为 DenseNet201、DenseNet161 的热力图, 通过训练关注车型的整体特征, 无法对鉴别性部件进行有效的关注, 提出的基于部件关注的密集卷积网络模型能够关注车辆的车灯、车标等鉴别性部位, 如图 7(d) 所示, 图 7(d) 侧面图中, 热力图深红色区域主要



集中在车灯和车标, 说明车灯、车标区域像素对最后分类结果的贡献最大, 图 7(d) 正面和背面热力图中, 车灯部件特征对分类起到了最大作用。

从图 7 比较可看出, 提出的模型能够基于这些关键性车型部件, 对车辆进行分类, 进而实现对细粒度车型的有效识别。

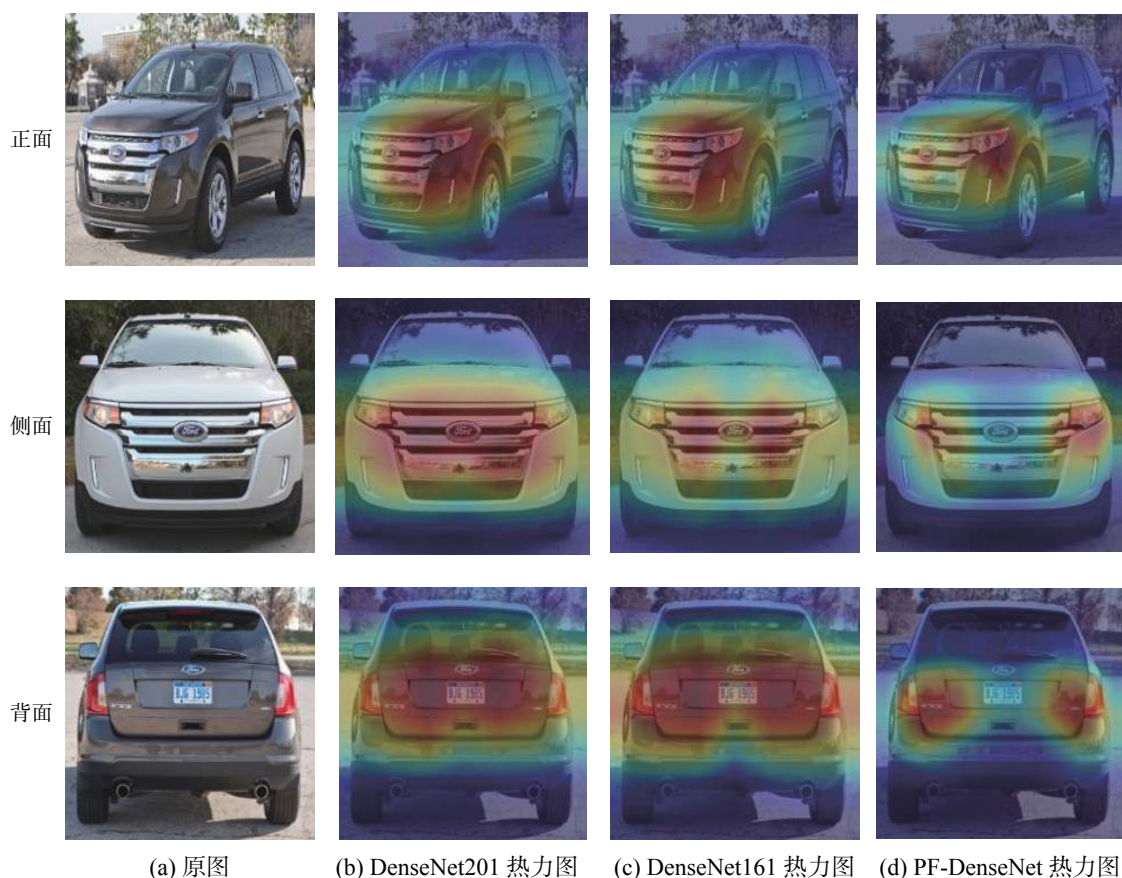


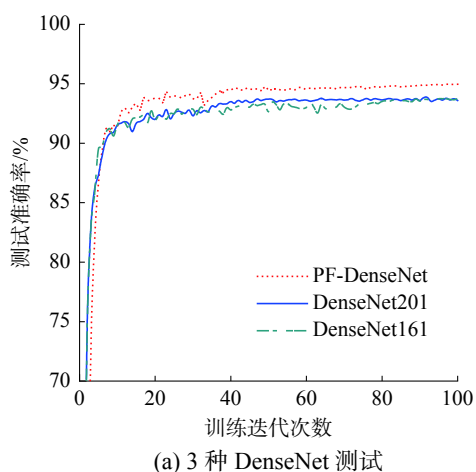
图 7 不同 DenseNet 热力图

Fig. 7 Different DenseNet heat maps

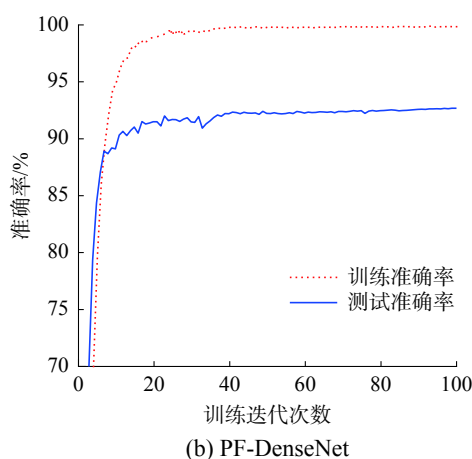
图 8 为 3 种 DenseNet 网络在 Cars-196 上的训练和测试准确率随训练周期变化的折线图。

图 8(a) 为 3 种 DenseNet 测试准确率变化图, 图中显示训练周期在 15 次以后 PF-DenseNet 的测试准确率提升速度明显高于 DenseNet201 和 DenseNet161, 在训练 40 次以后 3 种 DenseNet 的

测试准确率都趋于稳定, PF-DenseNet 的最终测试准确率分别比 DenseNet201 和 DenseNet161 提高了 1.2%、1%。在训练过程中相较于原始 DenseNet201 和 DenseNet161 网络, PF-DenseNet 的训练准确率和测试准确率的差距更小, 具有更好的拟合效果。



(a) 3 种 DenseNet 测试



(b) PF-DenseNet

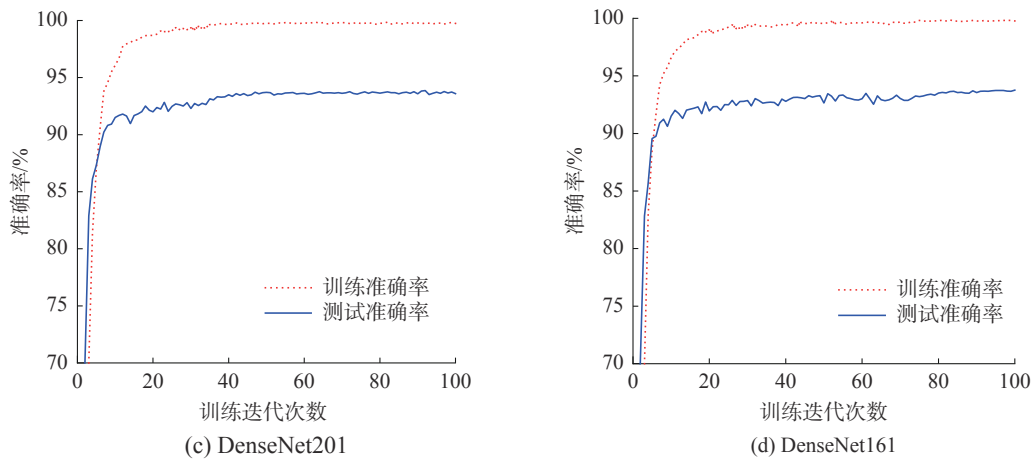


图 8 不同 DenseNet 准确率变化图

Fig. 8 Diagrams of DenseNet accuracy variation

### 2.3.3 不同方法对比

PF-DenseNet 和在本实验数据集 Cars-196 上的其他方法进行比较,如表 2 所示,Zhao 等<sup>[17]</sup>提出一个多样化的视觉注意力网络 (diversified visual attention networks, DVAN) 来解决细粒度车型分类的问题,在学习定位区分区域时大大减少了对强监督信息的依赖,追求注意力的多样性,最大限度地收集有区别的信息,但精度只有 87.1%,远低于本文方法。马力等提出 Sparse B-CNN<sup>[8]</sup>,通过在每个特征通道引入比例因子,按比例因子的大小判别特征通道的重要性并进行剪枝,减弱过拟合提高特征的显著性,准确率为 91.0%,由于网络训练中需要进行剪枝,操作较为复杂。Fu 等<sup>[7]</sup>采用 3 种尺度作为输入,以相互增强的方式递归学习车型区域注意力和区域的特征表示,准确率为 92.5%,比本文方法低 2.5%。Li 等<sup>[18]</sup>提出了一种迭代矩阵平方根归一化方法用于全局协方差共用网络的快速端到端训练,并运用在 ResNet101 上,获得了 93.3% 的精度,但该方法大大增加了网络参数,训练时间也过长。文献 [19] 提出互通道损失 (mutual channel loss, MC-Loss),由鉴别分量和多样性分量两个特定信道的分量组成,训练每个特征通道反映针对特定类的不同局部区分区域,以 B-CNN 为基础框架,获得了 94.4% 的准确率,准确率较高,但在分量中引入了需要额外寻优的超参数,模型训练较繁琐。PF-DenseNet 通过对车型区域特征反复提取,获得足够多的部件信息,使模型能够基于部件信息分类,再通过独立密集卷积得到抽象特征,提高收敛极限,获得了 95.0% 的准确率。表 2 还显示了 PF-DenseNet 与其他经典 DCNN 在 Cars-196 上查全率和  $F_1$  对比,PF-

DenseNet 较于 Inception-ResNet-v2、ResNet152、DenseNet201 和 DenseNet161 查全率分别提升了 3.7%、2.4%、1.1% 和 1%, $F_1$  分别提升了 3.7%、2.3%、1% 和 0.9%。

表 2 不同方法在 Cars-196 数据集上的准确率对比  
Table 2 Comparisons of accuracy of different methods in Cars-196 data set

方法	准确率	查全率	$F_1$
DVAN <sup>[17]</sup>	87.1	—	—
Sparse B-CNN <sup>[8]</sup>	91.0	—	—
RA-CNN <sup>[7]</sup>	92.5	—	—
iSQRT-COV with ResNet101 <sup>[18]</sup>	93.3	—	—
文献[19]	94.4	—	—
Inception-ResNet-v2	91.3	91.2	91.1
ResNet152	92.6	92.5	92.5
DenseNet201	93.8	93.8	93.8
DenseNet161	94.0	93.9	93.9
PF-DenseNet	95.0	94.9	94.8

图 9 所示为 PF-DenseNet 与其他 DCNN 准确率和参数量的对比,PF-DenseNet 的参数量远小于 ResNet152 和 Inception-ResNet-v2, DenseNet161 和 DenseNet201 的参数量略小于 PF-DenseNet,但精度远小于 PF-DenseNet, Inception-ResNet-v2、ResNet152、DenseNet201 和 DenseNet161 准确率分别为 91.3%、92.6%、93.8% 和 94.0%,分别比本文方法降低 3.7%、2.4%、1.2% 和 1%,综合准确率、查全率、 $F_1$  和参数量来看,文中方法获得了较好的效果。



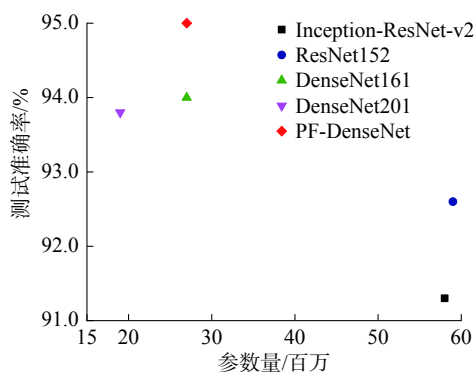


图9 PF-DenseNet与其他DCNN准确率和参数数量对比  
Fig. 9 Comparison of PF-PenseNet and other DCNN accuracy and number of parameters

### 3 结束语

为了实现细粒度车型识别,提高细粒度车型识别准确率,提出一种基于部件关注的DenseNet细粒度车型识别网络模型,通过Process Layer对车型部件特征反复提取加强,获取更多的部件细节信息,然后通过融合独立组件的密集卷积网络,进一步提取高维特征,嵌入独立组件提高模型收敛极限。该模型能够对细粒度车型的车灯、车标等鉴别性部件进行有效关注,根据车型部件像素信息进行分类,在Stanford Cars-196公开数据集上获得了目前最优的识别准确率、查全率和 $F_1$ ,验证了本文模型在细粒度车型识别上的有效性。但文中方法着重关注车型细节信息,忽视了全局特征对车型分类的作用,因此下一步工作是在关注车型鉴别性部件的基础上结合车型全局特征,进一步提高模型的识别准确率。

### 参考文献:

- [1] 杨娟,曹浩宇,汪荣贵,等.区域建议网络的细粒度车型识别[J].中国图象图形学报,2018,23(6):837-845.  
YANG Juan, CAO Haoyu, WANG Ronggui, et al. Fine-grained car recognition method based on region proposal networks[J]. Journal of image and graphics, 2018, 23(6): 837-845.
- [2] LIAO Liang, HU Ruimin, XIAO Jun, et al. Exploiting effects of parts in fine-grained categorization of vehicles[C]//Proceedings of 2015 IEEE International Conference on Image Processing. Quebec City, Canada, 2015: 745-749.
- [3] KRAUSE J, STARK M, DENG Jia, et al. 3D object representations for fine-grained categorization[C]//Proceedings of 2013 IEEE International Conference on Computer Vision Workshops. Sydney, Australia, 2014: 554-561.
- [4] FANG Jie, ZHOU Yu, YU Yao, 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.
- [5] SHI Weiwei, GONG Yihong, TAO Xiaoyu, et al. Fine-grained image classification using modified DCNNs trained by cascaded softmax and generalized large-margin losses[J]. IEEE transactions on neural networks and learning systems, 2019, 30(3): 683-694.
- [6] KE Xiao, ZHANG Yufeng. Fine-grained vehicle type detection and recognition based on dense attention network[J]. Neurocomputing, 2020, 399: 247-257.
- [7] FU Jianlong, ZHENG Heliang, MEI Tao. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 4438-4446.
- [8] 马力,王永雄.基于稀疏化双线性卷积神经网络的细粒度图像分类[J].模式识别与人工智能,2019,32(4):336-344.  
MA Li, WANG Yongxiong. Fine-grained visual classification based on sparse bilinear convolutional neural network[J]. Pattern recognition and artificial intelligence, 2019, 32(4): 336-344.
- [9] 王阳,刘立波.面向细粒度图像分类的双线性残差注意力网络[J].激光与光电子学进展,2020,57(12):121011.  
WANG Yang, LIU Libo. Bilinear residual attention networks for fine-grained image classification[J]. Laser & optoelectronics progress, 2020, 57(12): 121011.
- [10] VALEV K, SCHUMANN A, SOMMER L, et al. A systematic evaluation of recent deep learning architectures for fine-grained vehicle classification[C]//Proceedings of SPIE 10649, Pattern Recognition and Tracking XXIX. Orlando, USA, 2018: 1064902.
- [11] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 4700-4708.
- [12] 白琮,黄玲,陈佳楠,等.面向大规模图像分类的深度卷积神经网络优化[J].软件学报,2018,29(4):1029-1038.  
BAI Cong, HUANG Ling, CHEN Jianan, et al. Optimization of deep convolutional neural network for large scale image classification[J]. Journal of software, 2018, 29(4): 1029-1038.
- [13] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J/OL]. (2020-01-01)[2020-07-01]<https://arxiv.org/>

- [abs/1502.03167](#).
- [14] YAROTSKY D. Error bounds for approximations with deep ReLU networks[J]. *Neural networks*, 2017, 94: 103–114.
- [15] 陈立潮, 朝昕, 曹建芳, 等. 融合独立组件的 ResNet 在细粒度车型识别中的应用 [J]. *计算机工程与应用*, 2021, 57(11): 248–253.
- CHEN Lichao, CHAO Xin, CAO Jianfang, et al. Application of ResNet with independent components in fine-grained vehicle recognition[J]. *Computer engineering and applications*, 2021, 57(11): 248–253.
- [16] 周安众, 罗可. 一种卷积神经网络的稀疏性 Dropout 正则化方法 [J]. *小型微型计算机系统*, 2018, 39(8): 1674–1679.
- ZHOU Anzhong, LUO Ke. Sparse Dropout regularization method for convolutional neural networks[J]. *Journal of Chinese computer systems*, 2018, 39(8): 1674–1679.
- [17] ZHAO Bo, WU Xiao, FENG Jiashi, et al. Diversified visual attention networks for fine-grained object classification[J]. *IEEE transactions on multimedia*, 2017, 19(6): 1245–1256.
- [18] LI Peihua, XIE Jiangtao, WANG Qilong, et al. Towards faster training of global covariance pooling networks by iterative matrix square root normalization[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 47–955.
- [19] CHANG Dongliang, DING Yifeng, XIE Jiyang, et al. The devil is in the channels: mutual-channel loss for fine-grained image classification[J]. *IEEE transactions on image processing*, 2020, 29: 4683–4695.

### 作者简介:

陈立潮, 教授, 主要研究方向为人工智能、图像信息处理。主持山西省自然科学基金等项目 12 项, 获山西省科学技术奖二等奖 2 项。发表学术论文 180 余篇。

朝昕, 硕士研究生, 主要研究方向为智能图像信息处理。

潘理虎, 教授, 主要研究方向为智能软件工程理论与应用、人工智能、复杂系统仿真。主持省部级科研项目 10 余项。发表学术论文 60 余篇, 出版专著 1 部。