



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

双层残差语义分割网络及交通场景应用

谭睿俊, 赵志诚, 谢新林

引用本文:

谭睿俊,赵志诚,谢新林. 双层残差语义分割网络及交通场景应用[J]. 智能系统学报, 2022, 17(4): 780–787.

TAN Ruijun,ZHAO Zhicheng,XIE Xinlin. Double-residual semantic segmentation network and traffic scenic application[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(4): 780–787.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202106020>

您可能感兴趣的其他文章

改进MobileNet的图像分类方法研究

Research on the improved image classification method of MobileNet

智能系统学报. 2021, 16(1): 11–20 <https://dx.doi.org/10.11992/tis.202012034>

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

区域损失函数的孪生网络目标跟踪

Regional loss function based siamese network for object tracking

智能系统学报. 2020, 15(4): 722–731 <https://dx.doi.org/10.11992/tis.201910005>

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

基于跳跃连接金字塔模型的小目标检测

Skip feature pyramid network with a global receptive field for small object detection

智能系统学报. 2019, 14(6): 1144–1151 <https://dx.doi.org/10.11992/tis.201905041>



微信公众平台



期刊网址

DOI: 10.11992/tis.202106020

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20220608.1352.002.html>

双层残差语义分割网络及交通场景应用

谭睿俊, 赵志诚, 谢新林

(太原科技大学 电子信息工程学院, 山西 太原 030024)

摘要: 针对图像语义分割过程中特征提取网络的深度问题以及下采样池化层降低特征图分辨率等问题, 提出了一种基于双层残差网络特征提取的图像语义分割网络, 称为 DRResnet。首先提出一种双层残差网络, 对训练集各目标的细节进行特征提取, 提高网络对部分细节目标的感知能力; 其次在 Layer1 层开始跳跃特征融合, 并持续以 2 倍反卷积方法进行上采样操作, 融合底层特征与高层特征, 降低部分细节信息丢失对分割精度的影响; 最后使用网络分支训练法, 先训练图像上各目标的大致轮廓特征, 在此基础上再训练各目标的细节特征。结果表明: 该网络的平均交并比较全卷积网络相比, 在 CamVid 上由 49.72% 提升至 59.44%, 在 Cityscapes 上由 44.35% 提高到 47.77%, 该网络得到准确率更高、分割物体边缘更加完整的图像分割结果。

关键词: 双层残差网络; 细节特征提取; 跳跃特征融合; 上采样; 网络分支训练法; 图像语义分割; CamVid 数据集; Cityscapes 数据集

中图分类号: TP183 **文献标志码:** A **文章编号:** 1673-4785(2022)04-0780-08

中文引用格式: 谭睿俊, 赵志诚, 谢新林. 双层残差语义分割网络及交通场景应用 [J]. 智能系统学报, 2022, 17(4): 780-787.

英文引用格式: TAN Ruijun, ZHAO Zhicheng, XIE Xinlin. Double-residual semantic segmentation network and traffic scenic application[J]. CAAI transactions on intelligent systems, 2022, 17(4): 780-787.

Double-residual semantic segmentation network and traffic scenic application

TAN Ruijun, ZHAO Zhicheng, XIE Xinlin

(School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

Abstract: An image semantic segmentation network based on the double-residual network, named DRResnet, is proposed to address the depth problem of feature extraction network in semantic image segmentation and resolution reduction of feature maps by a down-sampling pooling layer. Firstly, a double-residual network is proposed to extract the detailed features of each target in the training set and improve the perception ability of the network for some detailed targets. Secondly, the jump feature fusion starts from Layer 1, and the up-sampling operation is continued by the 2×deconvolution method to fuse the low-level and the high-level features to reduce the impact of partial detail information loss on the segmentation accuracy. Finally, the network branch training method is adopted to train the outline features of each target on the image, followed by training the detailed features of each target. The results indicate that the network's P_{MIoU} improves from 49.72% to 59.44% on CamVid and from 44.35% to 47.77% on Cityscapes when compared to the full convolution network. The network can produce image segmentation results with higher accuracy and more complete edge segmentation.

Keywords: double-residual network; detail feature extraction; jump feature fusion; up sampling; network branch training; semantic image segmentation; CamVid data set; Cityscapes data set

近年来, 随着深度学习的不断发展, 计算机视

觉领域取得显著的成就。越来越多的交通场景需要精确且高效的分割技术, 实现交通场景的图像语义分割成为广大人员研究的问题之一。交通场景的图像语义分割, 是对交通场景图像底层的像

收稿日期: 2021-06-13. 网络出版日期: 2022-06-08.

基金项目: 山西省自然科学基金青年基金项目 (201901D211304).

通信作者: 赵志诚. E-mail: zhzhich@126.com.

素进行相应类别的分类,为自动驾驶中的车辆控制、路径规划等问题提供一个高效的解决方法。常见的图像分割算法有。传统方法的图像分割和深度学习的图像语义分割。传统的图像分割算法主要有阈值分割法、边缘检测法、交互式图像分割法等。阈值分割法主要是根据图像中像素的颜色或灰度值对不同的分割目标来设定不同的阈值,根据不同的阈值对每个像素点进行分类^[1]。在交通应用场景中,当图像像素的灰度值接近甚至重叠时,该方法很难得到准确的分割结果。

边缘检测法通过检测确定区域的边界,再根据边界把图像分割成不同的区域^[2]。该分割算法具有一定的局限性,适合分割边缘的灰度值变化较大且噪声比较小的图像。交互式图像分割法,是一种基于图划分的方法,该算法针对大量复杂交通场景的图像分割存在一定的难度^[3-4]。

将深度学习与神经网络引入交通场景的图像语义分割领域中成为发展的潮流,通用的深度神经网络主要有:文献[5]提出的 AlexNet 网络;文献[6]提出的卷积神经网络模型(vision geometry group,VGG)网络;文献[7]提出的卷积网络模型(GoogLeNet)网络;由微软实验室提出的 Resnet 网络,该网络由一系列的残差网络堆叠而成,可以在增加其网络深度的同时有效地解决梯度消失与梯度爆炸等问题^[8]。

文献[9]利用卷积神经网络(convolutional neural network, CNN)对使用类别定位过滤器对图像粗略地分割,对小目标的分割精度很低。文献[10]提出了全卷积神经网络(fully convolutional Network, FCN),该网络利用卷积层代替 CNN 的全连接层,进行上采样操作恢复输入图像的分辨率,并在 PASCAL VOC2012^[11]以及室内环境语义分割数据集 NYUDv2^[12]进行实验。实验结果发现 FCN 对于小目标边缘细节的分割模糊甚至识别错误。文献[13]提出的 U-net 网络,通过跳跃连接融合两路径的特征信息,实现图像语义分割,但由于该网络未充分考虑各像素与像素之间的关联,缺乏语义特征的空间一致性。

基于此,本文提出一种基于双层残差网络的图像语义分割网络(dresnet network, DResnet)。首先利用双层残差网络对图像进行更加全面的特征提取,获得更多高分辨率的细节信息,并将输出的不同尺度的特征图进行融合作为上采样网络的输入;其次通过深层跳跃连接将高层信息与底层特征融合,充分利用底层的细节特征;最后将 DResnet 网络应用到交通场景中,采用网络分支

训练方法对该网络进行训练以及测试。根据交通数据集 CamVid 和 Cityscapes 的实验结果表明,本文提出的 DResnet 网络对分割小目标物体具有一定的优势。

1 DResNet 语义分割网络

图像语义分割的任务主要包括准确地分类、精确地分割物体的边缘轮廓两方面,FCN 能够在一定程度上对图像各物体分类以及预测出物体的边缘轮廓,但对于边缘细节的分割,准确度很低,这是因为 FCN 中的池化操作,在增大感受野的同时使图像损失了部分边缘细节信息。

为此,本文提出了 DResnet,具体结构如图1所示。DResnet 由双层残差网络特征提取、跳跃特征融合网络组成。利用双层残差网络对训练集图像进行双层特征提取,并将每一层输出特征图进行相应位置融合,输出融合特征图 $X_1 \sim X_5$;在跳跃特征融合网络中,首先将最高层(第5层)的融合特征图进行2倍反卷积操作,从第4层到第1层,依次将融合特征图与其高一层的反卷积操作结果相加融合得到特征图 $W_1 \sim W_4$,然后对其进行2倍反卷积操作,从而添加更多的跳跃连接结构,将高层信息与底层信息进行充分融合,最终将不同分辨率的特征图上采样恢复至原输入图像的相同尺寸。

1.1 双层残差网络

从提高网络特征提取能力的角度出发,本文提出了双层残差网络代替原 FCN 网络的特征提取网络 VGG,双层网络在提取的特征图融合了更多的边缘细节特征。还可以解决由神经网络层数加深而引起的梯度消失或者梯度爆炸问题^[14-15]以及网络模型出现的退化问题^[16]。

具体来说,双层残差特征提取网络如图2所示。该网络包含两个相同的特征提取网络单元,每一单元均由5层组成。整体网络的输入为 RGB 三维数组,第1层由大小为 7×7 ,步长为2的卷积核进行卷积操作,输出的特征图尺寸为原输入的 $1/2$,卷积之后进行批量归一化处理、Relu 函数激活以及最大池化下采样操作,池化窗口大小为 3×3 ,步长为2;第2层到第4层由不同数量的 Block1 模块和 Block2 模块构成,对应的 Block 模块如图3所示,本文的主特征提取网络(第2层)由4个 Block1 构成;第3层由1个 Block2 和3个 Block1 构成;第4层由1个 Block2 和5个 Block1 构成;第5层由1个 Block2 和2个 Block1 模块构成。

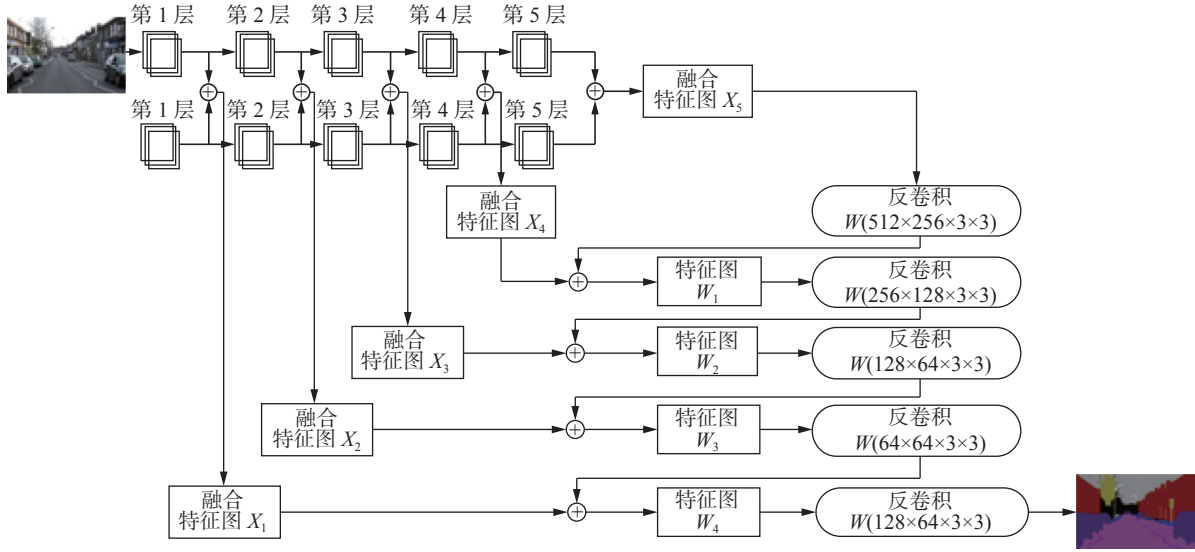


图 1 DRResnet 网络结构

Fig. 1 DRResnet network structure

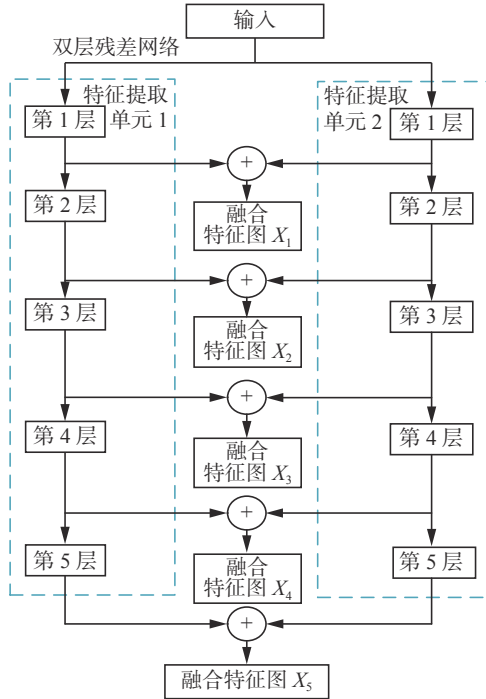


图 2 双层残差特征提取网络

Fig. 2 Double residual feature extraction network

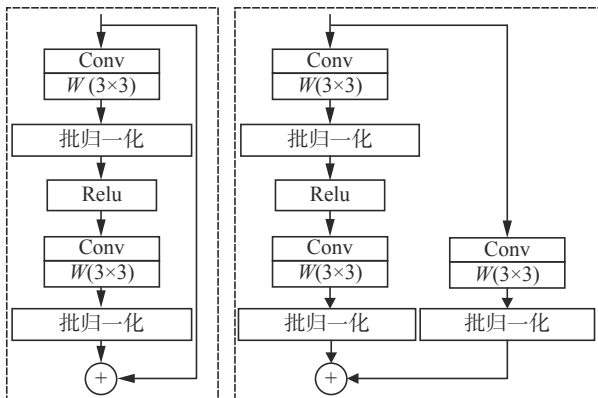


图 3 Block 模块结构

Fig. 3 Block module structure

1.2 跳跃特征融合网络

考虑原 FCN 网络模型存在池化操作, 在增大感受野的同时, 减小了特征图的分辨率, 从而导致这些高层语义特征和之前的底层细节特征分辨率不一致, 将底层特征与高层特征有效地融合, 对提高网络的分割精度至关重要。

本文在原 FCN 网络的基础上融合了更底层第 1 层与第 2 层的细节特征信息, 具体结构如图 1 所示。具体步骤: 将第 5 层输出的融合特征图 X_5 进行 2 倍反卷积操作, 将其大小恢复至原图的 $1/16$, 与第 4 层输出的融合特征图 X_4 进行叠加得到特征图 W_1 ; 再对特征图 W_1 进行 2 倍反卷积操作, 将其大小恢复至原图的 $1/8$, 与第 3 层输出的融合特征图 X_3 进行叠加得到特征图 W_2 ; 再对其进行 2 倍反卷积操作, 将其大小恢复至原图的 $1/4$, 与第 2 层输出的融合特征图 X_2 进行叠加得到特征图 W_3 ; 再进行 2 倍反卷积操作, 将图像恢复至原图的 $1/2$, 与第 1 层输出的特征图融合 X_1 进行叠加得到特征图 W_4 ; 最后进行 2 倍反卷积操作, 得到与输入相同大小尺寸的图像。

利用跳跃特征融合, 将高层的语义信息与底层的细节信息相结合, 降低了由上采样操作所带来的部分细节特征损失, 提高了交通场景下图像语义分割的精度。

1.3 网络分支训练

本文为了获得更高精度的语义分割结果, 在下采样过程加入了双层残差网络, 在上采样过程中加入了更多的卷积层, 整体增加了网络的深度, 为了使 DRResnet 网络得到充分的学习, 更好地对图像中的细节信息的训练, 本文提出了网络分

支训练法,先对整个训练集的目标的大体轮廓位置进行训练,其次,在此基础上对所训练的目标进行细节方面的准确训练,可以提高网络的训练准确度。

2 实验设计与结果分析

2.1 实验环境

本文所有实验是基于深度学习 Pytorch 开源框架进行的,实验环境的配置:操作系统为 64bit-Ubuntu16.04, GPU 为 GeForce RTX 2080 Ti, 内存为 64 GB, 网络镜像为 Pytorch 1.7.1, 数据处理为 Python 3.8。

2.2 实验数据集

本文采用数据集 CamVid 和 Cityscapes 验证本文 DResnet 网络的有效性。

CamVid 是常用的交通场景数据集之一,本文使用其中的 701 张的逐像素语义分割的图像数据集,其中训练集、测试集、验证集各 421、168、112 张,每张图像的分辨率为 720 dpi×960 dpi。

Cityscapes 是一个侧重于城市环境理解的数据集,主要包含 50 个城市不同场景、不同季节的 5000 张精细标注的图像,其中训练集、测试集、验证集各 2975、500、1525 张,每张图像的分辨率为 1024 dpi×2048 dpi。

数据处理:1)由于该数据集类别过多且部分类别与其它类别的界定相对比较模糊,所以参考文献[17-21],在实验时选择 CamVid 12 类和 Cityscapes 20 类, CamVid 12 类主要包括背景、天空、建筑物、电线杆、道路、人行道、树、信号标志、栅栏、汽车、行人、骑自行车的人; Cityscapes 20 类主要包括背景、道路、人行道、建筑物、墙、栅栏、杆、交通灯、交通标志、植被、地形、天空、人、骑手、汽车、卡车、公共汽车、火车、摩托车、自行车等。2)由于该数据集的分辨率过高以及 GPU 服务器显存有限,对 2 组数据集分别进行中心裁剪至 352 dpi×480 dpi、512 dpi×1024 dpi,再输入网络训练。

2.3 训练参数

为了得到交通场景分割较好结果的网络,进行了特征提取网络的替换、特征融合方法的相关实验,测试的对象为 CamVid 12 类和 Cityscapes 20 类。该实验在相同数据集、固定参数下进行实验,考虑到硬件原因,将两数据集中图片大小分别裁剪为 352 dpi×480 dpi 和 512 dpi×1024 dpi,采用反向传播和随机梯度下降法来进行训练,将训练批次大小设置为 4,设置基础学习率为 0.0001,

经过 10 次 epoch 迭代后,对应的学习率减半。冲量设置为 0.9,迭代次数为 200 次。在设置完超参数之后,分别对不同的卷积神经网络进行训练并保存训练最好的权重,作为网络模型最终的参数,然后对测试集进行测试,得到不同网络的最终分割结果。

2.4 评价指标

在图像语义分割技术中,为了合理评价改进网络的分割精度,本文采用平均交并比、准确率以及平均像素准确率 3 个评价指标。

1)平均交并比 P_{MIOU} (mean intersection over union, MIOU),作为图像语义分割最常用的评价指标,通过计算标签集合与预测结果集合的交集与并集之比,并计算所有类别的平均值,计算公式为

$$P_{\text{MIOU}} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}}$$

式中: k 为图像中的类别总数(不包含背景); P_{ii} 为真实像素类别为 i 被预测为像素为 i 的像素数量; P_{ij} 为真实像素类别为 i 被预测为像素为 j 的像素数量。

2)平均像素准确率 P_{MPA} (mean pixel accuracy, MPA),首先对每个类计算正确预测的像素比值,其次对所有的类取平均,计算公式为

$$P_{\text{MPA}} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \times 100\%$$

3)平均准确率 P_{ACC} (accuracy, ACC),所有分割正确的像素点与总像素点的比值,计算公式为

$$P_{\text{ACC}} = \frac{\sum_{i=0}^n p_{ii}}{\sum_{i=0}^n \sum_{j=0}^n p_{ij}}$$

3 实验结果分析

3.1 CamVid 实验结果

为了验证本文网络的可行性,本文在 CamVid 数据集上得到原 FCN 网络、不同深度 Resnet 网络、不同深度 DResnet 网络的测试分割结果。分别使用相同的数据划分对 FCN、Resnet-34、DResnet-34 不同的特征提取网络进行网络分支训练,得到表 1 所示的结果,由表 1 可以看出, DResnet 网络可以得到更高分割精度的结果。

表 1 不同网络在 CamVid 测试集结果对比
Table 1 Comparison of the results of different networks in the CamVid test set %

网络结构	P_{ACC}	P_{MIOU}	P_{MPA}
FCN	54.25	49.72	84.65
Resnet-34	57.63	53.91	85.99
DResnet-34	61.93	59.44	88.28

原 FCN 网络、不同深度 DResnet 网络的测试分割结果,具体如表 2 所示。可以看出,随着 DResnet 网络的层数不断加深,网络的 3 项测试评价指标的测试精度逐渐增加,但当网络达到 50 层时,测试的 P_{MIOU} 精度出现下降,这是因为在训练的过程中出现了过拟合,考虑到网络本身的参数问题、训练集的数目问题以及测试精度问题,本文通过实验分析,选择 DResnet-34 网络作为主网络的特征提取网络。

表 2 不同深度 DResnet 在 CamVid 测试集结果对比
Table 2 Comparison of the results of different depths of DResnet in the CamVid test set %

网络结构	P_{ACC}	P_{MIOU}	P_{MPA}
FCN	54.25	49.72	84.65
DResnet-18	61.48	57.68	87.74
DResnet-34	61.93	59.44	88.28
DResnet-50	62.33	59.09	88.09
DResnet-101	62.37	59.41	88.16

跳跃特征融合网络实验。为选取合适的上采样网络,在 DResnet-34 网络作为特征提取网络的

基础上,对融合不同尺度特征图网络 DResnet-2S 到 DResnet-32S 进行训练与测试,得到测试结果如表 3 所示。

表 3 不同跳跃特征融合在 CamVid 测试集结果对比
Table 3 Comparison of the results of different jump features fusion in the CamVid test set %

网络模型	P_{ACC}	P_{MIOU}	P_{MPA}
DResnet-32S	60.00	56.74	86.71
DResnet-16S	60.54	58.32	87.28
DResnet-8S	61.50	59.29	87.94
DResnet-4S	61.91	59.40	87.95
DResnet-2S	61.93	59.44	88.28

由表 3 可以看出,仅仅融合 Layer5 层的特征图, DResnet-32S 网络的平均交并比 P_{MIOU} 仅为 56.74%,测试的 P_{ACC} 为 60.00%,通过不断地融合不同 Layer 层的特征图,测试精度出现了上升,融合所有尺度的特征图, DResnet-2S 网络的测试准确度 P_{ACC} 高达 61.93%,平均交并比 P_{MIOU} 达到 59.44%。

从表 4 中可以看出,不同深度的 DResnet 网络对电线杆(Pole)、信号标志(SignSymbol)、栅栏(Fence)、行人(Pedestrian)、骑自行车的人(Bicyclist)等小目标的物体的分割精度提升较大,整体的 P_{MIOU} 由 49.72% 提升到 59.44%,通过增加 DResnet 网络的深度,可以改善部分小目标的分割精度。

表 4 不同网络在 CamVid 测试集的各类别精度
Table 4 Accuracy of each category of different networks in the CamVid test set %

类别	FCN	DResnet_18	DResnet_34	DResnet_50	DResnet_101
Background	58.35	65.20	69.04	62.26	64.88
Sky	91.34	93.24	93.25	93.12	93.22
Building	84.12	85.92	88.01	87.79	86.83
Pole	11.41	33.86	35.13	35.16	32.36
Road	94.38	95.13	95.66	94.00	94.93
Sidewalk	75.45	80.28	80.40	81.58	82.45
Tree	70.49	73.03	72.84	73.35	72.20
SignSymbol	10.47	13.21	13.08	14.86	15.68
Fence	16.61	21.63	23.11	22.15	22.24
Car	72.05	73.24	79.98	78.91	80.83
Pedestrian	22.38	41.48	43.56	42.40	40.47
Bicyclist	14.41	23.92	23.50	22.92	24.37
P_{MIOU}	49.72	57.68	59.44	59.09	59.41

3.2 Cityscapes 实验结果

为了进一步验证 DResnet 网络的可行性,在

另一交通数据集 Cityscapes 上进行实验,但考虑到该数据集图像较多以及显存硬件原因,本文选

取 DResnet34 网络作为主特征提取网络, 分别对 FCN、不同特征融合的 DResnet 网络进行网络分支训练以及测试, 测试结果见表 5。

表 5 不同跳跃特征融合在 Cityscapes 测试集结果对比
Table 5 Comparison of the results of different jump features fusion in the Cityscapes test set %

特征融合方法	P_{ACC}	P_{MIOU}	P_{MPA}
FCN	38.85	44.35	88.46
DResnet-32S	41.29	47.16	89.41
DResnet-16S	42.05	47.33	89.82
DResnet-8S	41.98	47.54	90.02
DResnet-4S	41.45	47.43	89.76
DResnet-2S	42.33	47.77	90.10

从表 5 中可以看出, DResnet-32S 网络仅融合特征图 X_5 , 平均交并比 P_{MIOU} 仅达到 47.16%, 测试准确度 P_{ACC} 达到 41.29%, DResnet-2S 网络融合所有尺度的特征图, 再对其进行上采样操作, 测试准确度 P_{ACC} 高达 42.33%, 平均交并比 P_{MIOU} 达到 47.77%。因此再次验证融合所有尺度特征图的融合方法可以提高整体的测试分割精度。

在 Cityscapes 测试集上部分预测结果如图 4 所示, 第 1 列为 Cityscapes 测试集的 RGB 图像, 第 2 列为对应的标签, 第 3 列为 FCN-8s 的测试结果, 第 4 列为 DResnet-18 的测试结果, 第 5 列为 DResnet-34 的测试结果。对比 FCN-8s 与 DResnet 网络的结果图, 可以明显地看出 DResnet-34 网络分割后的栅栏、杆、交通标志、人等小物体边缘更加精细, 分割边界更加规整。

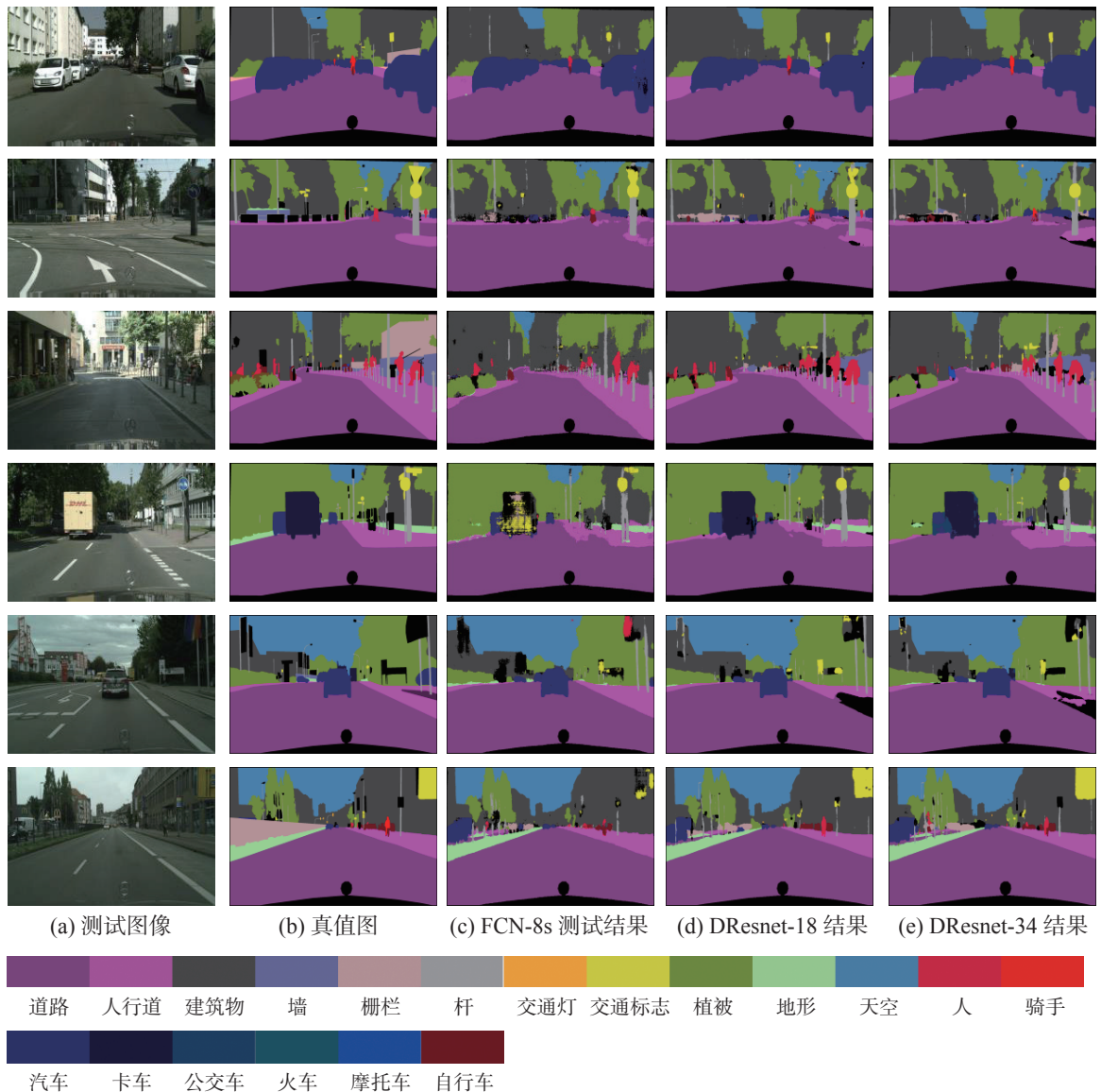


图 4 Cityscapes 测试集部分图像预测结果
Fig. 4 Cityscapes test set partial image prediction results

4 结束语

本文提出了一种基于双层残差网络特征提取、跳跃特征融合的图像语义分割网络(DResnet)。它能够解决由特征提取网络带来的分辨率减小以及空间细节信息损失问题。DResnet 网络构建双层残差特征提取网络与跳跃特征融合网络, 双层残差网络对标准训练集的图像同时进行两次特征提取操作, 以获得更全面的图像细节特征信息; 跳跃特征融合网络在第 1 层与第 2 层开始融合跳跃连接, 更好地融合图像的背景信息与语义信息, 使网络获得更高的特征提取能力与识别精度。通过交通场景数据集 CamVid 和 Cityscapes 的实验结果表明, 本文的 DResnet 网络模型在 CamVid 12 类上的评价指标 P_{MIU} 、 P_{MPA} 、 P_{ACC} 分别提高了 9.72%、3.63%、7.68%, 在 Cityscapes 20 类上的评价指标 P_{MIU} 、 P_{MPA} 、 P_{ACC} 分别提高了 3.42%、1.64%、3.48%。

参考文献:

- [1] 朱磊, 滕奇志, 龚剑. 基于改进模糊 C 均值聚类 and 区域合并的矿物颗粒分割方法 [J]. 科学技术与工程, 2020, 20(34): 14138–14145.
ZHU Lei, TENG Qizhi, GONG Jian. Mineral particle segmentation algorithm based on improved fuzzy C-means and region merging[J]. *Science technology and engineering*, 2020, 20(34): 14138–14145.
- [2] 许林, 孟娜, 袁静, 等. 基于边缘检测和图像分割的超声诊断机器人控制系统设计 [J]. 计算机测量与控制, 2020, 28(8): 125–129.
XU Lin, MENG Na, YUAN Jing, et al. Design of ultrasonic diagnosis robot control system based on edge detection and image segmentation[J]. *Computer measurement & control*, 2020, 28(8): 125–129.
- [3] 陈飞. 改进的交互式 Otsu 红外图像分割算法 [J]. 计算机测量与控制, 2020, 28(9): 248–251.
CHEN Fei. An improved interactive otsu infrared image segmentation algorithm[J]. *Computer measurement & control*, 2020, 28(9): 248–251.
- [4] 杨金鑫, 杨辉华, 李灵巧, 等. 结合卷积神经网络和超像素聚类的细胞图像分割方法 [J]. 计算机应用研究, 2018, 35(5): 1569–1572, 1577.
YANG Jinxin, YANG Huihua, LI Lingqiao, et al. Cell image segmentation method based on convolution neural network and super pixel clustering[J]. *Application research of computers*, 2018, 35(5): 1569–1572, 1577.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the acm*, 2017, 60(6): 84–90.
- [6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations, ICLR 2015-Conference Track Proceedings. San Diego: [s. n.], 2015.
- [7] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1–9.
- [8] TRONG V H, HYUN Y G, YOUNG K J, et al. Yielding multi-fold training strategy for image classification of imbalanced weeds[J]. *Applied sciences*, 2021, 11(8): 3331.
- [9] MÜLLER D, KRAMER F. MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning[EB/OL]. (2019–10–21) [2021–06–13]. <https://www.semanticscholar.org/paper/MIScnn%3A-A-Framework-for-Medical-Image-Segmentation-M%26C3%BCller-Kramer>.
- [10] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[C]//IEEE transactions on pattern analysis and machine intelligence. New York: IEEE, 2015: 640–651.
- [11] EVERINGHAM M, ESLAMI S M A, GOOL L, et al. The pascal visual object classes challenge: a retrospective[J]. *International journal of computer vision*, 2015, 111(1): 98–136.
- [12] SILBERMAN N, HOIEM D, KOHLI P, et al. Indoor segmentation and support inference from RGBD images [C]//European Conference on Computer Vision. Berlin: Springer, 2012: 746–760.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234–241.
- [14] XIAO Zhitao, LIU Bowen, GENG Lei, et al. Segmentation of lung nodules using improved 3D-UNet neural network[J]. *Symmetry*, 2020, 12(11): 1787.
- [15] GUAN Haixing, LI Hongliang, LI Rongqiang, et al. Face detection of innovation base based on faster RCNN[M]//2021 International Conference on Applications and Techniques in Cyber Intelligence. Cham: Springer International Publishing, 2021: 158–165.
- [16] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE In-

ternational Conference on Computer Vision. Santiago: IEEE, 2015: 1026–1034.

- [17] MANDAL R, AZAM B, VERMA B, et al. Deep learning model with GA-based visual feature selection and context integration[C]//2021 IEEE Congress on Evolutionary Computation. Kraków: IEEE, 2021: 288–295.
- [18] YANG Tao, WU Yan, ZHAO Junqiao, et al. Semantic segmentation via highly fused convolutional network with multiple soft cost functions[J]. *Cognitive systems research*, 2019, 53: 20–30.
- [19] JIA Fan, LIU Jun, TAI Xuecheng. A regularized convolutional neural network for semantic image segmentation[J]. *Analysis and applications*, 2021, 19(1): 147–165.
- [20] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 3213–3223.
- [21] KAZEROUNI I A, DOOLY G, TOAL D. Ghost-UNet: an asymmetric encoder-decoder architecture for semantic segmentation from scratch[J]. *IEEE access*, 2021, 9: 97457–97465.

作者简介:



谭睿俊, 硕士研究生, 主要研究方向为机器视觉与智能信息处理。曾获国家奖学金, “华为杯”第十七届中国研究生数学建模竞赛三等奖, “兆易创新杯”第十六届中国研究生电子设计竞赛三等奖, 全国大学生智能汽车竞赛—讯飞智慧餐厅组别线上国家三等奖, 第十六届全国大学生智能汽车竞赛分赛区讯飞智慧餐厅组比赛国家二等奖。



赵志诚, 教授, 主要研究方向为先进控制技术。承担并完成国家级、省部级以及企业委托项目 40 余项, 获省部级科技奖励 2 项, 发表学术论文 80 余篇, 出版专著 1 部, 授权发明专利 4 项。



谢新林, 讲师, 主要研究方向为机器视觉、深度学习。主持省级科研项目 2 项、横向课题 1 项, 发表学术论文 10 余篇、申请和授权专利 10 余项。

第七届认知系统与信息处理国际会议 (ICCSIP 2022) Seventh International Conference on Cognitive Systems and Information Processing (ICCSIP 2022)

由中国人工智能学会、中国自动化学会、IEEE 计算智能学会主办的第七届认知系统与信息处理国际会议 (ICCSIP 2022) 将于 2022 年 10 月 21 日至 23 日在中国福州召开。本次会议旨在汇聚人工智能领域的顶级人工智能专家和学者, 共同探讨人工智能系统、认知系统和智能信息处理领域的最新进展, 并展示最新研究成果和未来发展前景。

重要日期:

论文截止日期: 2022 年 9 月 1 日

录用通知时间: 2022 年 9 月 30 日

参会登记时间: 2022 年 10 月 10 日

会议日期: 2022 年 10 月 21-23 日

详情请关注:

第七届认知系统与信息处理国际会议官网: <http://iccsip.fzu.edu.cn>