

基于改进 DeepLabV3+ 的无人机高速公路护栏检测

王洋¹, 郭杜杜^{2*}, 王庆庆¹, 周飞¹, 秦音¹¹新疆大学智能制造现代产业学院, 新疆 乌鲁木齐 830017;²新疆大学交通运输工程学院, 新疆 乌鲁木齐 830017

摘要 针对现有语义分割方法检测高速公路护栏时存在预测速度慢、分割精度低的问题, 提出一种基于改进 DeepLabV3+ 的无人机高速公路护栏检测方法。首先, 采用 MobileNetv2 网络替换原模型的主干并输出中层特征, 减少参数量的同时恢复降采样过程中丢失的空间信息; 然后, 采用密集连接扩张卷积改进空洞空间金字塔池化, 以减少漏分割现象; 最后, 在编码器部分引入空间分组增强 (SGE) 注意力机制, 减少错分割现象。实验结果表明, 改进后模型平均交并比、平均像素准确率、画面每秒传输帧数达到了 79.20%、87.89%、52.59, 相比基础模型, 分别提高了 2.59%、2.93%、56.70%, 参数量降低了 78.85%, 能够在保障模型预测速度的同时提高对护栏的分割精度。

关键词 图像处理; 语义分割; 无人机; DeepLabV3+; 注意力机制

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP231270

UAV Highway Guardrail Inspection Based on Improved DeepLabV3+

Wang Yang¹, Guo Dudu^{2*}, Wang Qingqing¹, Zhou Fei¹, Qin Ying¹¹School of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830017, Xinjiang, China;²School of Traffic and Transportation Engineering, Xinjiang University, Urumqi 830017, Xinjiang, China

Abstract To address the problems of slow prediction speed and low segmentation accuracy of existing semantic segmentation methods for highway guardrail detection, an UAV highway guardrail detection method based on improved DeepLabV3+ is proposed. First, the MobileNetv2 network was used to replace the backbone of the original model and outputs the middle layer's features to reduce number of parameters while recovering the spatial information lost in the downsampling process; then an atrous spatial pyramid pooling was improved by the densely connected expansive convolution to reduce the phenomenon of missed segmentation; finally, the spatial group-wise enhance (SGE) attention mechanism was introduced in the encoder part to reduce the phenomenon of wrong segmentation. The experimental results show that the average intersection over union, average pixel accuracy, and frames per second transmission of the improved model can reach 79.20%, 87.89%, and 52.59, which are 2.59%, 2.93%, and 56.70% higher than the base model, respectively, and number of parameters is reduced by 78.85%. This method can thus improve the segmentation accuracy for the guardrail while guaranteeing the model's prediction speed.

Key words image processing; semantic segmentation; unmanned aerial vehicle; DeepLabV3+; attention mechanism

1 引言

高速公路护栏作为道路沿线设施的重要组成部分, 是高速公路巡检与养护的重要内容, 但目前高速公路护栏检测通常依赖人工巡检和手工测量, 存在周期长、效率低等问题。因此, 高精度、高效率的高速公路护栏检测是现在的研究热点, 具备较高的研究价值和应用前景。传统的高速公路护栏检测主要通过 3D 点

云建模、路侧设备信息传输或车载视频滤波处理等方法实现, 如 Vidal 等^[1]提出了一种使用移动激光扫描仪采集数据的护栏分割分类方法, 通过点云与护栏几何特征相结合实现对护栏的分割。Gomes 等^[2]提出了一种基于物联网的护栏碰撞检测系统, 通过物联网与路侧传感设备的相互连接实现对护栏的检测。Zhu 等^[3]将定向梯度直方图和局部二进制模式相结合进行特征融合, 然后通过滑动窗口及聚类方法实现对护栏的检

收稿日期: 2023-05-10; 修回日期: 2023-06-08; 录用日期: 2023-06-27; 网络首发日期: 2023-07-07

基金项目: 新疆维吾尔自治区重点研发计划(2022B01015-3)

通信作者: *guodd@xju.edu.cn

测。但 3D 点云数据分辨率不足,路侧设备及车载视频易受外界干扰影响且覆盖范围有限。

近年来,无人机技术发展迅速,因具备高机动性、低成本、随时悬停等特点,已广泛应用于智慧城市、空中安防等方面^[4-6],结合近年来爆发式发展的语义分割技术,可以为高速公路护栏检测提供新的解决思路。作为计算机视觉领域的基本问题之一,语义分割经历了从传统机器学习分类器到深度学习方法的演变^[7-8]。传统的语义分割方法每运行一次只能切割一次图片,对多目标图像的分割效率低且准确度不高。随着全卷积神经网络(FCN)^[9]的出现,深度学习正式进入图像分割领域,并以此为基础演化为专门池化网络(PSPNet)、编码器-解码器架构(Unet 和 DeepLab)、多尺度处理网络(DeepLab)等^[10]。其中 Google 提出的 DeepLabV1^[11]被认为是语义分割的经典结构之一,针对传统卷积神经网络(CNN)下采样降低分辨率的问题,提出了空洞卷积和条件随机场,提高了模型捕获细节的能力。DeepLabV2^[12]在此基础上,利用残差网络替换模型主干,并提出空洞空间金字塔池化(ASPP)扩大模型的感受野,有效地提高了模型的检测精度。DeepLabV3^[13]使用串行或并行的方式重新布局空洞卷积模块,采用多比例空洞卷积捕获多尺度背景,获得了更高的检测精度。DeepLabV3+^[14]引入语义分割常用的编码器-解码器结构,实现了特征层的跨尺度融合,并采用深度可分离卷积代替普通卷

积,获得了更加精细的物体边界,在高速公路护栏检测上取得了不错的结果,但该算法的预测速度较慢,且在复杂多变的高速公路场景下仍存在一定的错分割和漏分割问题。

为了解决 DeepLabV3+ 模型对高速公路护栏检测存在的上述问题,本文提出了一种基于改进 DeepLabV3+ 的无人机高速公路护栏检测算法,通过替换主干网络、增加中层特征融合、改进空洞空间金字塔池化模块和引入轻量级注意力机制,在有效改善模型错分割和漏分割问题的同时显著提升预测速度。

2 基本原理

2.1 DeepLabV3+原理

DeepLabV3+ 是由 Google 提出的语义分割经典模型,主要包含编码器和解码器,使用 Xception^[15] 主干网络。在编码器部分,利用主干网络进行图像特征提取并将其分为浅层语义特征和深层语义特征两部分。中间层输出的浅层语义特征被直接送入解码器,最后一层输出的深层语义特征将送入 ASPP 进行多尺度特征提取。在解码器部分,对 ASPP 输出的深层语义特征进行 4 倍线性插值上采样后将结果与主干网络中间层输出的浅层语义特征相融合,融合特征在经过 3×3 卷积和 4 倍线性插值上采样后,得到与原始图像相同分辨率的预测图像。DeepLabV3+ 网络架构^[14]如图 1 所示。

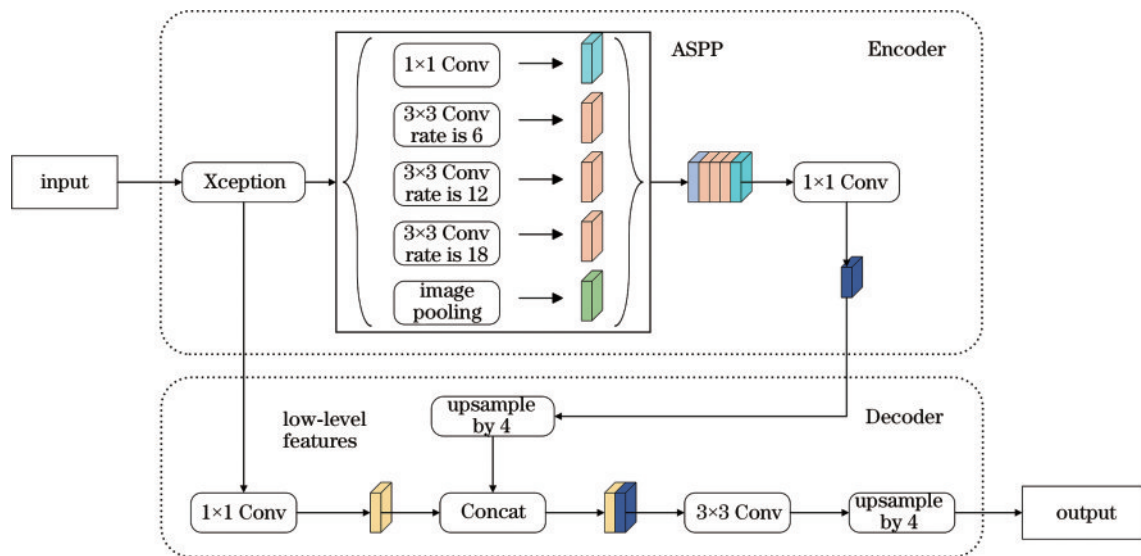


图 1 DeepLabV3+ 网络构架

Fig. 1 DeepLabV3+ network architecture

2.2 改进的 DeepLabV3+ 模型

为了提升高速公路护栏检测效率,改善错分割和漏分割的问题,通过 4 个方面改进 DeepLabV3+ 模型。首先,采用改进轻量级 MobileNetv2 网络^[16] 替换主干网络,有效减少模型参数数量的同时提高预测速度;其次,输出主干网络中的第 4 层、第 7 层特征并进行多尺

度特征融合,减少降采样过程中的特征丢失,提高模型的分割精度;然后,利用密集连接空洞卷积改进的 DenseASPP^[17] 替换原模型的 ASPP 模块,DenseASPP 中的空洞率(d)分别设置为 3、6、12、18、24,获得更大感受野的同时减少模型漏分割的现象;最后,在主干网络输出的多尺度特征和 DenseASPP 输出的深层特征

结果中加入空间分组增强(SGE)注意力机制^[18],将多尺度特征相加融合后对结果进行 3×3 卷积和 1×1 卷积操作,调整通道数并输出特征图张量,对深层特征进

行4倍线性插值上采样后与多尺度融合特征相加,并进行 3×3 卷积和第二次4倍线性插值上采样,输出预测图像。改进DeepLabV3+模型如图2所示。

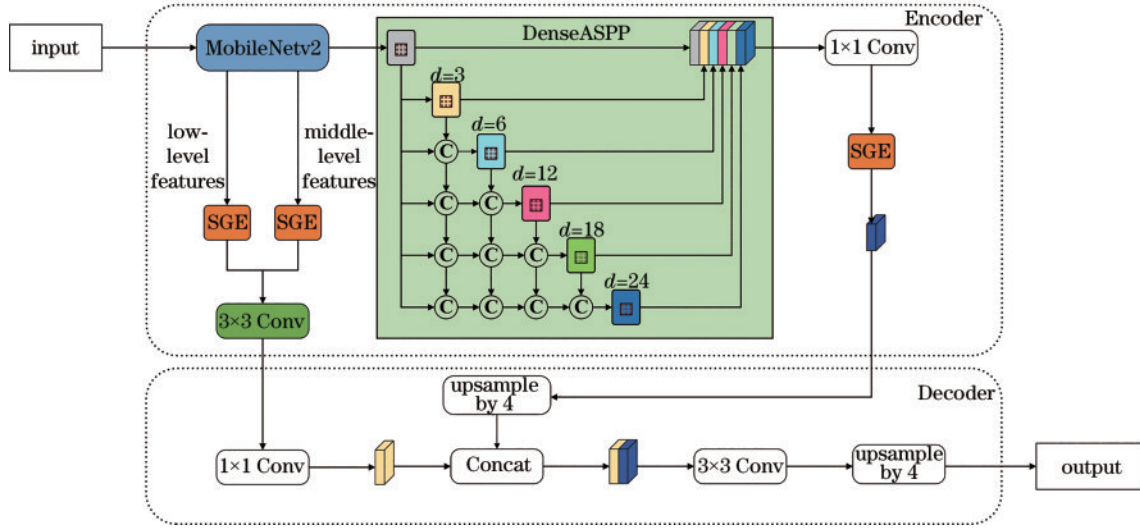


图2 改进DeepLabV3+网络构架
Fig. 2 Improved DeepLabV3+ network architecture

2.2.1 替换主干网络

DeepLabV3+的主干网络Xception模型参数量大、训练时间和预测时间较长,无法满足对护栏分割效率的需求。采用经过剪枝的轻量级主干网络MobileNet2替换原本的主干网络,在有效减少模型参数量和训练时间的同时提高了预测速度。

MobileNet2是Google提出的轻量级主干网络模型,在引入深度可分离卷积的MobileNet1的基础上,增加了线性瓶颈和反向残差结构来提高对特征的代表能力。MobileNet2经过5次下采样会使得特征图的尺寸进一步减小,导致图像丢失大量细节特征,这对精细边界敏感区域的分割是不利的,因此在DeepLabV3+等语义分割网络中通常不会进行5次下采样。为了保留更多的图像细节特征信息,本文对网络结构参数进行了修改,仅进行4次下采样^[19-20];并通过剪枝删除MobileNet2中用于分类的网络结构,进一步提高模型的运行效率。以 $640\times 640\times 3$ 的图像作为输入,改进后的MobileNet2网络结构如表1所示,其中 t 表示通道倍数, c 表示通道大小, n 表示重复次数, s 表示步长。

2.2.2 多尺度特征融合

DeepLabV3+将主干网络输出的经过空洞空间金字塔池化4倍上采样的深层特征直接与浅层特征融合,这种单一尺度的特征融合方式容易导致特征细节丢失。为了充分利用主干的不同层次特征,对主干网络输出的不同特征层进行可视化分析,结果如图3所示。

由图3可以看出:浅层特征与中层特征的分辨率更高,表现出不同的颜色、纹理、边缘等细节信息;深层

表1 改进MobileNet2网络结构

Table 1 Improved MobileNet2 network structure

Operator	t	c	n	s	Output size
Conv2d 3×3		32	1	2	$320^2\times 32$
Bottleneck	1	16	1	1	$320^2\times 16$
Bottleneck	6	24	2	2	$160^2\times 24$
Bottleneck	6	32	3	2	$80^2\times 32$
Bottleneck	6	64	4	2	$40^2\times 64$
Bottleneck	6	96	3	1	$40^2\times 96$
Bottleneck	6	160	3	1	$40^2\times 160$
Bottleneck	6	320	1	1	$40^2\times 320$

特征分辨率较低,但表现出更强的高层语义信息。为了减少深层特征降采样过程中的细节丢失,将主干网络输出的中层特征上采样后与浅层特征融合,获得更多的图像细节特征,并利用 3×3 卷积、ReLU激活函数和正则化操作调整通道数量,防止过多低层语义信息影响深层语义信息的表达,在仅增加少量参数的情况下,有效地提高了护栏分割的准确度。

2.2.3 改进的空洞空间金字塔池化

ASPP包含多个不同采样率的并行空洞卷积,通过不同空洞率构建不同感受野的卷积核,从而获得图像的多尺度信息。针对无人机等高分辨率图像,需要很大的空洞率才能获得足够的感受野,导致空洞卷积的衰退衰减无效。而DenseASPP采用密集连接的方式将空洞卷积的输出融合,在获得更大感受野的同时避免了空洞率过大导致卷积退化的问题。

空洞卷积的扩张能够在保留原有图像分辨率的情况下增加感受野,获得更详细的图像全局信息,从而减

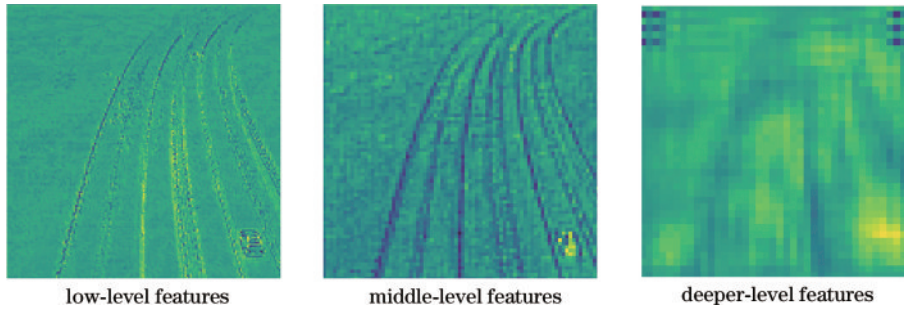


图3 不同尺度特征层可视化结果

Fig. 3 Visualization result of feature layers at different scales

小分割遗漏率。其中,感受野指卷积过程中某一层输出结果中某个元素对应的上一层区域的大小,计算方法为

$$R_n = dK_n - d + 1, \quad (1)$$

式中: R_n 为感受野大小; d 是空洞率; K_n 为卷积核大小。ASPP对不同空洞率的空洞卷积层进行并联,故最大感受野为并联空洞卷积层中最大空洞率所对应的感受野,以空洞率(6,12,18)为例,最大感受野为

$$R_{\max} = \max[k_{3,d=6}, k_{3,d=12}, k_{3,d=18}] = k_{3,d=18} = 37。 \quad (2)$$

将两个空洞卷积层堆叠连接将会获得更大的感受野,对于滤波层为 K_1 和 K_2 的两个空洞卷积层,其等效感受野叠加计算方法为

$$R_{\max} = R_n = \sum_{n=1}^N k_n - (n - 1)。 \quad (3)$$

故DenseASPP对应空洞率(6,12,18)的最大感受野为

$$R_{\max} = k_{3,d=6} + k_{3,d=12} + k_{3,d=18} - 2 = 72。 \quad (4)$$

空洞率相同的情况下,DenseASPP对应的最大感受野明显大于ASPP,可以获得更多尺度的图像语义信息并有效增强模型对不同尺度物体的敏感程度,从而减少漏分割的现象。

2.2.4 引入SGE注意力机制

注意力机制可以模拟人类视觉机制,帮助模型判断局部信息的影响力大小,是提升语义分割网络性能的重要方法之一。常见的注意力机制包括空间注意力、通道注意力、卷积注意力等,但注意力往往会增加模型参数量,降低网络的训练及检测速度。本文为了保障对无人机影像识别的速度,引入轻量级注意力网络SGE,其主要思想是将特征图分组,并认定每一组代表一个语义特征,之后通过语义组中各目标的空间位置生成注意力来调整每个子特征的重要性,结合局部特征与全局特征的相似性生成注意力掩码,增强语义特征的空间分布。SGE注意力网络结构^[18]如图4所示。

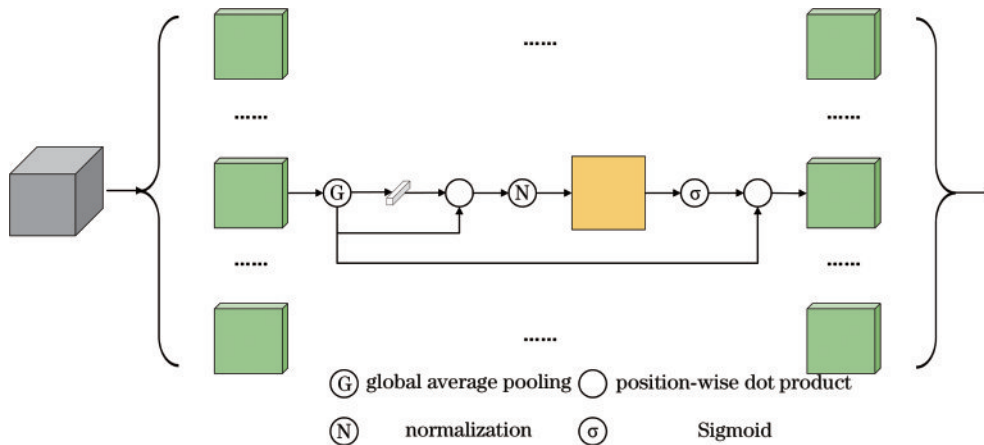


图4 SGE注意力网络结构

Fig. 4 SGE attention network structure

3 实验与结果

3.1 实验数据集

所用数据主要包含两部分:1)采集于新疆某高速公路,包含无人机飞行50 m正射视角、50 m前45°视角、70 m正射视角、70 m前45°视角等6段视频,每

30帧保留一张图像,图像分辨率为 3840×2160 和 1920×1080 ;2)收集于网络,包含正常环境和雾天环境等4段视频,每30帧保留一张图像,图像分辨率为 1920×1080 。为加快检测速度,提高数据集质量,采用随机窗口将图像裁剪为 640×640 分辨率,共计2023张。使用开源标注软件Labelme对图像进行标注,包括

2065 个 right-guardrail(右侧护栏)标签和 2014 个 left-guardrail(左侧护栏)标签,按照 8:2 划分为训练集和测

试集,部分数据集如图 5 所示。



图 5 数据集样例

Fig. 5 Sample dataset

3.2 实验环境和评价指标

实验环境如表 2 所示。

表 2 实验环境设置

Table 2 Experimental environment setting

Name	Parameter
Operating system	Windows 10
Processor	AMD Ryzen 9 5950X
Video card	NVIDIA GeForce RTX 3090Ti
RAM	32 G
Development language	Python 3.8
Development environment	PyCharm 2021
Web framework	PyTorch 1.7

为了减少模型的训练时间,利用迁移学习思想将训练过程分为冻结阶段和解冻阶段。冻结训练阶段,冻结模型的主干权重,将更多的资源放在训练模型后面部分的网络参数,冻结阶段的 Batchsize 设置为 24,训练轮数设置为 50;在解冻阶段,对模型的参数进行调整,Batchsize 设置为 12,训练轮数设置为 450。两阶段其余参数相同,最大学习率设置为 0.007,学习率下降方式为 cos;设置权值衰减为 0.0001,以防止模型过拟合;训练损失函数使用交叉熵损失函数 CE-Loss。值得注意的是,为了准确验证所提算法的有效性,各对比实验的参数设置相同。采用的评估指标为平均交并比(mIoU)、类别平均像素准确率(mpa),计算公式分别为

$$R_{\text{mIoU}} = \frac{1}{n+1} \sum_{j=0}^n \frac{m_{jj}}{\sum_{k=0}^n m_{jk} + \sum_{k=0}^n m_{kj} - m_{jj}}, \quad (5)$$

$$A_{\text{mpa}} = \frac{1}{n+1} \sum_{j=0}^n \frac{m_{jj}}{\sum_{k=0}^n m_{jk}}, \quad (6)$$

式中: $n+1$ 表示有 n 个目标类与 1 个背景; m_{jj} 表示分类正确的像素数量; m_{jk} 表示属于第 j 类却被分到第 k 类的像素数量; m_{kj} 表示属于第 k 类却被分到第 j 类的像素数量。

3.3 实验结果对比分析

为了验证所提算法的先进性,设计 3 组实验进行论证。首先在模型的相同位置增加不同的注意力模块,验证 SGE 注意力模块的优越性;之后根据所提改进方法设计消融实验,验证各模块的有效性;最后进行相同超参数的语义分割模型对比,验证所提改进 DeepLabV3+ 的高效性。

3.3.1 注意力机制对比实验

在采用 MobileNetv2 作为主干网络后,为了提高模型的检测精度,进行注意力对比实验。选择目前在语义分割领域表现较好的 CBAM^[21]、Coord^[22]、ECA^[23] 及本文使用的轻量级注意力机制 SGE 进行对比。输入图像为随机测试集图像,选取经空洞空间金字塔池化处理后的深层特征进行特征层可视化,可视化结果如图 6 所示。

由图 6 可以看出,SGE 注意力机制对护栏存在区域的关注度更明显。为了直观地对上述不同注意力机制的效果进行验证,对增加不同注意力机制的模型训练 500 轮次后的结果进行对比,比较分割精度与预测速度,实验结果如表 3 所示。

由表 3 可以看出,当主干网络为 MobileNetv2 时,引入不同注意力机制的网络的检测精度都有所提升,

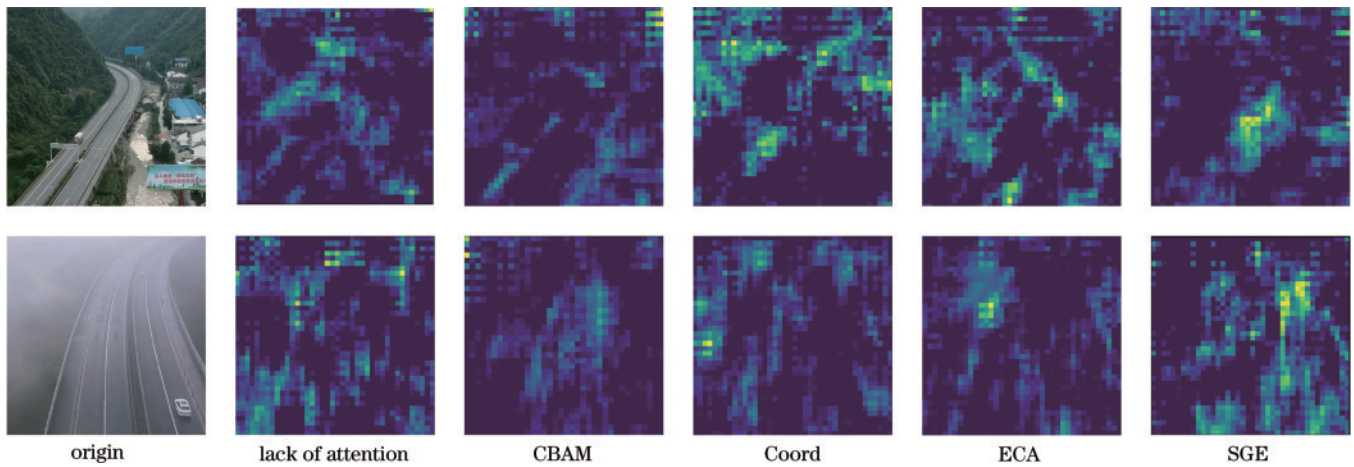


图 6 注意力机制对比

Fig. 6 Comparison of attentional mechanisms

表 3 注意力机制对比实验

Table 3 Comparison experiment of attentional mechanisms

Base model	Backbone network	Attention model	mpa / %	mIoU / %	Speed / (frame · s ⁻¹)
DeepLabV3+	MobileNetv2		87.35	77.82	64.33
DeepLabV3+	MobileNetv2	CBAM	87.33	78.20	54.10
DeepLabV3+	MobileNetv2	ECA	87.39	78.46	66.65
DeepLabV3+	MobileNetv2	Coord	87.06	78.39	62.33
DeepLabV3+	MobileNetv2	SGE	87.79	78.56	66.74
DeepLabV3+	Xception		84.96	76.61	33.56
DeepLabV3+	Xception	SGE	85.95	77.78	32.75

而SGE注意力机制具备最好的效果,mIoU、mpa、速度都取得了最优。为了进一步验证该注意力机制的有效性,将其加入原始网络的相同位置进行实验,同样取得了较好的效果。实验结果表明,SGE注意力机制可以在增加模型精度的同时保障模型的检测速度。

3.3.2 不同模块的消融实验

为了验证所提改进策略在护栏检测上具有优化作用,利用控制变量法设计7组消融实验。在DeepLabV3+网络的基础上分别进行替换MobileNetv2主干网络、融合中层特征、添加SGE注意力机制、引入DenseASPP的改进,然后对各模块进行依次累加,并与基础模型进行对比,以mIoU、mpa、速度、Params为评价指标,结果如表4所示,其中最优结

果用加粗形式表示。

对比①和②发现,改进MobileNetv2主干网络后,模型的mIoU、mpa与检测速度分别提高了1.21个百分点、2.39个百分点与91.69%,参数量降低了89.38%;对比②和③可以发现,进行中层特征多尺度融合后,模型的mIoU提高了0.44个百分点,检测速度仅降低8.32%,参数量仅增加0.34%;对比②和④可以发现,SGE模块的引入使得模型在参数量几乎不变的情况下,mIoU、mpa、检测速度分别提高了0.74个百分点、0.44个百分点、3.75%;对比②、⑤、⑥可以发现,虽然密集连接带来了参数的增加,但对mIoU的提升效果较好,且与其余模块融合时也能保持较好效果;从实验①和⑦可以看出,相较基础网络,所提算法

表 4 消融实验结果对比

Table 4 Comparison of ablation experiment results

Group	MobileNetv2	Feature fusion	SGE	DenseASPP	mpa / %	mIoU / %	Speed / (frame · s ⁻¹)	Params / 10 ⁶
①					84.96	76.61	33.56	54.71
②	✓				87.35	77.82	64.33	5.81
③	✓	✓			87.39	78.26	58.98	5.83
④	✓		✓		87.79	78.56	66.74	5.81
⑤	✓			✓	87.31	78.45	62.50	11.55
⑥	✓		✓	✓	87.34	78.81	59.72	11.56
⑦	✓	✓	✓	✓	87.89	79.20	52.59	11.57

的 mIoU 提高了 2.59 个百分点,mpa 提高了 2.93 个百分点,检测速度提高了 56.70%,参数量降低了 78.85%。综合分析表 4 可知,4 种改进模块均有提高算法精度的效果。

3.3.3 分割网络模型对比实验

对所提模型与近年来主流的语义分割模型 UNet^[24]、PSPNet^[25]、DeepLabV3+ 在所使用的数据集上进行对比,设置相同的超参数和训练轮次,它们的训练损失收敛情况如图 7 所示。

由图 7 可以看出,模型在经过 500 个 epoch 后都趋于收敛且保持稳定,所提算法收敛性更强,损失值更小。为了进一步验证所提算法的高效性,对各模型在测试集中的表现进行对比,结果如表 5 所示。

由表 5 可以看出,所提算法的 mpa、mIoU、训练速度及训练时间分别为 87.89%、79.20%、52.59、11 h 15 min,相较于 UNet 和 DeepLabV3+, mIoU 分别提高了 2.23 个百分点和 2.59 个百分点;mpa 分别提高了 3.10 个百分点和 2.93 个百分点;训练速度分别提高了 27.54 frame/s 和 19.03 frame/s;训练时间分别缩短了 6 h 39 min 和 9 h 2 min。所提算法虽然在训练速度和

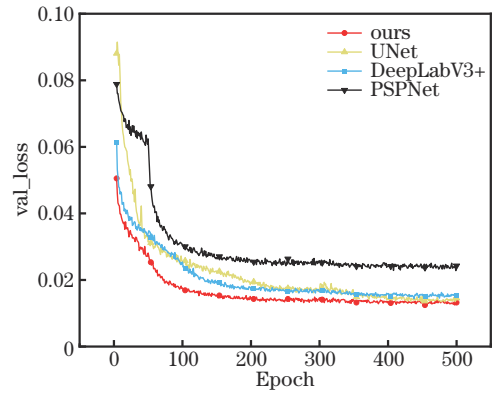


图 7 主流分割网络模型的损失变化

Fig. 7 Loss changes of mainstream segmentation network models

训练时间上的表现弱于 PSPNet,但 mIoU 和 mpa 高出 16.67 个百分点和 18.95 个百分点。为了进一步地展现算法分割效果的差异性,对各模型进行分割效果可视化验证,结果如图 8 所示。

由图 8 的②和③可以看出,PSPNet 对护栏的分割效果不佳,在随机选取的三张图像上都存在较大的漏分割现象;通过对比②、④和⑤可以看出,UNet 和

表 5 主流语义分割模型的实验结果对比

Table 5 Comparison of experimental results of mainstream semantic segmentation models

Network	Backbone network	mpa / %	mIoU / %	Speed / (frame · s ⁻¹)	Training time
UNet	VGG16	84.79	76.97	25.05	17 h 54 min
PSPNet	MobileNetv2	68.94	62.53	85.16	8 h 3 min
DeepLabV3+	Xception	84.96	76.61	33.56	20 h 17 min
Ours	MobileNetv2	87.89	79.20	52.59	11 h 15 min

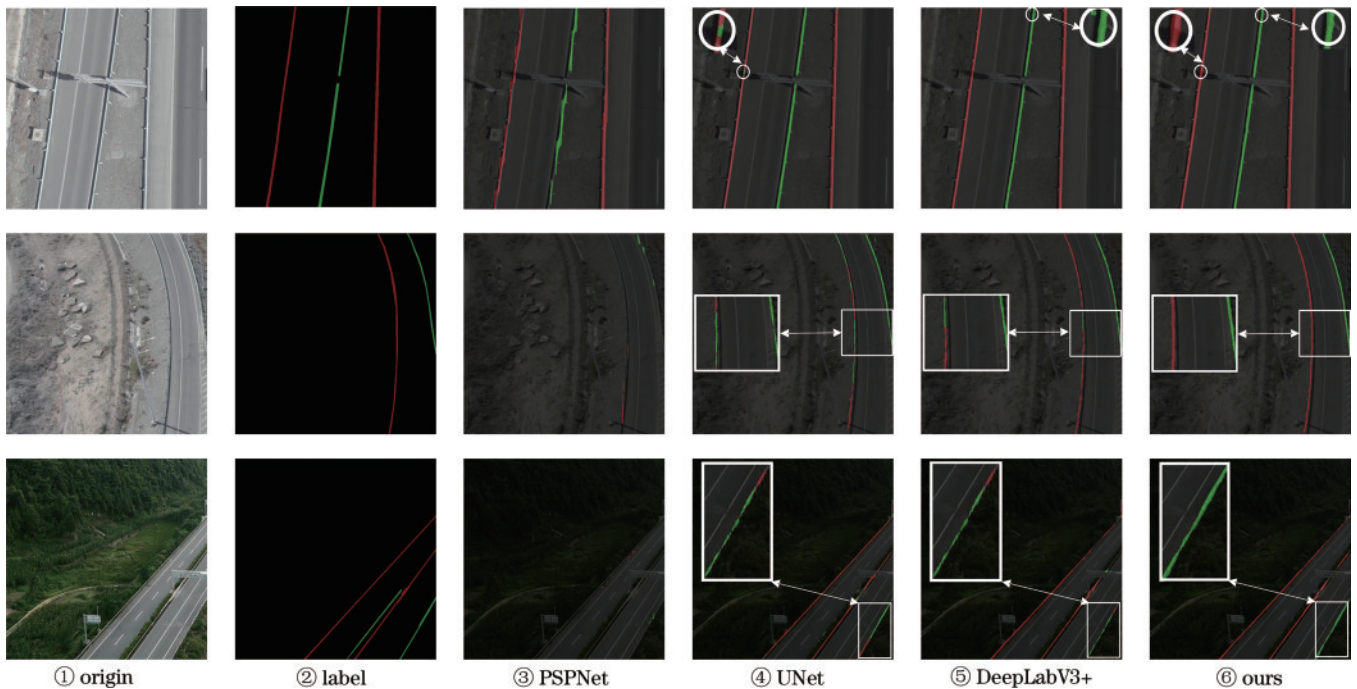


图 8 主流分割网络模型的分割效果

Fig. 8 Segmentation effect of mainstream segmentation network models

DeepLabV3+虽然漏分割现象较少,但UNet在分割left-guardrail(红色标签)、DeepLabV3+在分割right-guardrail(绿色标签)时存在明显的错分割现象;通过对比③、④、⑤、⑥可以看出,相较于其余三种模型,所提算法具备更优的高速公路护栏分割效果,且错分割和漏分割较少。

4 结 论

为了实现对高速公路护栏的高效检测,提出一种基于改进DeepLabV3+的无人机高速公路护栏检测方法。采用改进MobileNetv2替换原网络的主干,将输出的中层特征与原本的浅层特征融合;采用密集连接的DenseASPP替换原本的ASPP,使得模型生成更大的感受野,获得多尺度特征;最后通过对比不同的注意力模型,选择轻量级注意力机制SGE,获取更多的图像边缘特征。通过对比实验与消融实验发现,改进DeepLabV3+模型在显著降低参数数量的基础上,极大地提高了检测效果,明显改善了对高速公路护栏的错分割和漏分割等问题,且获得了更清晰的边缘特征。下一步工作将考虑高速公路中存在的柔性护栏、混凝土护栏等多种类护栏分割问题,标注更多的高速公路护栏数据,进一步增强改进DeepLabV3+算法的鲁棒性,并探究如何将网络模型部署至无人机,提高护栏巡检的效率与安全性。

参 考 文 献

- [1] Vidal M, Diaz-Vilarino L, Arias P, et al. Barrier and guardrail extraction and classification from point clouds [J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, 43: 157-162.
- [2] Gomes T, Fernandes D, Ekpanyapong M, et al. An IoT-based system for collision detection on guardrails [C]//2016 IEEE International Conference on Industrial Technology (ICIT), March 14-17, 2016, Taipei, China. New York: IEEE Press, 2016: 1926-1931.
- [3] Zhu H, Chen F R, Wang Z Y. A robust feature fusion method for camera-based highway guardrail detection [C]//2019 Chinese Automation Congress (CAC), November 22-24, 2019, Hangzhou, China. New York: IEEE Press, 2020: 1368-1373.
- [4] 苏昂, 陆伟康, 张仕林, 等. 基于目标运动模型的无人机对地视觉目标跟踪[J]. 激光与光电子学进展, 2022, 59(14): 1415022.
Su A, Lu W K, Zhang S L, et al. Visual ground target tracking of unmanned aerial vehicle based on target motion model[J]. Laser & Optoelectronics Progress, 2022, 59(14): 1415022.
- [5] 蒋镛圻, 叶泽聪, 彭月平, 等. 针对弱小无人机目标的轻量级目标检测算法[J]. 激光与光电子学进展, 2022, 59(8): 0810006.
Jiang R Q, Ye Z C, Peng Y P, et al. Lightweight target detection algorithm for small and weak drone targets[J]. Laser & Optoelectronics Progress, 2022, 59(8): 0810006.
- [6] 张宇, 张焱, 石志广, 等. 基于图像衍生的红外无人机图像仿真方法研究[J]. 光学学报, 2022, 42(2): 0210003.
Zhang Y, Zhang Y, Shi Z G, et al. Image simulation method of infrared UAV based on image derivation[J]. Acta Optica Sinica, 2022, 42(2): 0210003.
- [7] Chu X K, Zhang H, Tian Z Y, et al. Detection of laser-induced optical defects based on image segmentation[J]. High Power Laser Science and Engineering, 2019, 7(4): e66.
- [8] 陈兵, 贺晟, 刘坚, 等. 基于轻量化DeepLab v3+网络的焊缝结构光图像分割[J]. 中国激光, 2023, 50(8): 0802105.
Chen B, He S, Liu J, et al. Weld structured light image segmentation based on lightweight DeepLab v3+ network [J]. Chinese Journal of Lasers, 2023, 50(8): 0802105.
- [9] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [10] 王瑞绅, 宋公飞, 王明. 引入ECA注意力机制的U-Net语义分割[J]. 电光与控制, 2023, 30(1): 92-96, 102.
Wang R S, Song G F, Wang M. U-Net semantic segmentation by introducing ECA attention mechanism[J]. Electronics Optics & Control, 2023, 30(1): 92-96, 102.
- [11] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science, 2014(4): 357-361.
- [12] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [13] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-12-05) [2022-08-16]. <https://arxiv.org/abs/1706.05587>.
- [14] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 833-851.
- [15] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [16] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4510-4520.
- [17] Yang M K, Yu K, Zhang C, et al. DenseASPP for semantic segmentation in street scenes[C]//2018 IEEE/

- CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3684-3692.
- [18] Li X, Hu X L, Yang J. Spatial group-wise enhance: improving semantic feature learning in convolutional networks[EB/OL]. (2019-05-23) [2022-09-02]. <https://arxiv.org/abs/1905.09646>.
- [19] 黎颖, 吴清锋, 刘佳桐, 等. 引导性权重驱动的图表问答重定位关系网络[J]. 中国图象图形学报, 2023, 28(2): 510-521.
Li Y, Wu Q F, Liu J T, et al. Repositioning relational network of chart question-answering driven by guiding weight[J]. Journal of Image and Graphics, 2023, 28(2): 510-521.
- [20] 洛怡航, 赵振宇, 胡银记, 等. 基于孪生网络的轻量级高速跟踪算法[J]. 电光与控制, 2022, 29(1): 51-55, 79.
Luo Y H, Zhao Z Y, Hu Y J, et al. A lightweight high-speed tracking algorithm based on Siamese network[J]. Electronics Optics & Control, 2022, 29(1): 51-55, 79.
- [21] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [22] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [23] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [24] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [25] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.