

## 融合轻量化和注意力机制的语义分割算法

袁嫚嫚<sup>1</sup>, 陆 灏<sup>2</sup>

1. 安徽理工大学 计算机科学与工程学院, 安徽 淮南 232001

2. 安徽人防建筑设计研究院, 安徽 合肥 230022

**摘要:**目的 考虑现有的图像语义分割网络存在分割精度低、参数量大等问题,提出一种融合轻量化和注意力机制的语义分割算法。方法 该算法在 DeeplabV3+网络模型结构的基础上,使用 MobileNetV2 网络替换原始网络模型结构的 Xception 主干网络,构建轻量化语义分割网络结构,以此减少模型参数量和计算量,提高分割速度。同时,为了有效获取关注语义信息的正确特征,在编码阶段加入注意力模块机制,使网络模型在学习过程中只关注它所需要关注的点,提高图像分割精度,达到良好的分割效果。最后,在网络模型训练过程中引入 BCE loss(Binary Cross Entropy loss)和 Dice loss 损失函数相结合,加快网络的快速收敛,对模型更好的优化,以此提高模型的分割精度。结果 通过在数据集 PASCAL VOC2012 实验验证表明,该算法的分割精度提高了 2.82 个百分点,参数量降低了 14.46 M。同时,数据集 Cityscapes 的实验结果也验证了该算法的优越性。结论 优化后的 DeeplabV3+网络模型提高网络模型性能。

**关键词:**语义分割;DeeplabV3+;轻量化网络;通道注意力机制;损失函数

**中图分类号:**TP391 **文献标识码:**A **doi:**10.16055/j.issn.1672-058X.2025.0001.008

### A Semantic Segmentation Algorithm Integrating Lightweight and Attention Mechanisms

YUAN Manman<sup>1</sup>, LU Hao<sup>2</sup>

1. School of Computer Science and Engineering, Anhui University of Science and Technology, Anhui Huainan 232001, China

2. Anhui Civil Air Defense Architectural Design and Research Institute, Hefei 230022, China

**Abstract: Objective** Considering the problems of low segmentation accuracy and a large number of parameters in existing semantic image segmentation networks, a semantic segmentation algorithm integrating lightweight and attention mechanisms was proposed. **Methods** The algorithm replaced the Xception backbone network of the original network model structure with the MobileNetV2 network based on the DeeplabV3+ network model structure, and constructed a lightweight semantic segmentation network structure, so as to reduce the number of model parameters and computational volume and improve the segmentation speed. Additionally, an attention module mechanism was introduced during the encoding stage to effectively capture correct features of focused semantic information, enabling the network to focus only on relevant points during the learning process, thereby enhancing image segmentation accuracy and achieving satisfactory segmentation results. Finally, BCE loss (binary cross entropy loss) and Dice loss functions were combined in the network model training process to accelerate the rapid convergence of the network and better optimize the model, so as to improve the

**收稿日期:**2023-03-23 **修回日期:**2023-06-13 **文章编号:**1672-058X(2025)01-0057-07

**基金项目:**国家自然科学基金资助项目(51274011)。

**作者简介:**袁嫚嫚(1995—),女,河南商丘市人,硕士研究生,从事计算机视觉、智能优化研究。

**通讯作者:**陆灏(1991—),男,安徽怀远人,电气工程师,从事电气工程设计和计算机控制研究。Email: Hlukly001@qq.com.

**引用格式:**袁嫚嫚,陆灏.融合轻量化和注意力机制的语义分割算法[J].重庆工商大学学报(自然科学版),2025,42(1):57-63.

YUAN Manman, LU Hao. A semantic segmentation algorithm integrating lightweight and attention mechanisms[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(1): 57-63.

segmentation accuracy of the model. **Results** The experimental verification on the PASCAL VOC2012 dataset showed that the segmentation accuracy of the algorithm was increased by 2.82%, and the number of parameters was reduced by 14.46 M. Furthermore, experimental results on the Cityscapes dataset confirmed the superiority of the proposed algorithm. **Conclusion** The optimized DeeplabV3+ network model enhances the performance of the network model.

**Keywords:** semantic segmentation; DeeplabV3+; lightweight network; channel attention mechanism; loss function

## 1 引言

语义分割技术<sup>[1-2]</sup>是计算机视觉领域中的一个重要研究方向。其任务是对输入图像的像素进行分类,并且能从背景区域中准确地分割出目标对象。该技术被广泛应用于医学影像、无人驾驶、遥感图像分析、机器人导航等领域<sup>[3-6]</sup>,对人类生活方式的进一步便捷化具有重要的研究意义。

传统的语义分割算法通常根据图像的颜色、纹理等一些图像特征信息对其进行分割,然后对分割后的图像特征进行语义标注,从而完成图像分割。但这种分割效果往往不理想。近年来,一系列基于 FCN 的语义分割方法<sup>[7]</sup>取得了显著进展。FCN 为语义分割领域开辟了新的思路,但该网络没有考虑像素之间的关系,缺乏空间一致性,会导致分类误判,使得最终的分割结果不够精细。Zhao 等<sup>[8]</sup>提出的 PSPNet (Pyramid Scene Parsing Network, 金字塔场景解析网络)引入了模块金字塔池化,金字塔池化的加入可以使不同区域的上下文信息聚合在一起,来提高获取全局上下文信息的能力,以此来提高分割效果。Chen 等<sup>[9-11]</sup>提出的 DeepLab 系列算法具有较高的分割精度。DeepLabV1 引入了一个属性卷积操作,在不改变网络参数数量的情况下,增加了接受野的范围,得到了更好的分割结果。近年来,学者在 DeepLabV1 的基础上,陆续提出了 DeepLabV2、DeepLabV3、DeeplabV3+,通过优化网络结构,逐步提高了算法的分割性能。由于存储空间和计算能力资源的限制,如此庞大而复杂的神经网络模型难以应用于移动设备和嵌入式设备中,所以,Sandler Mark 等<sup>[12-13]</sup>提出了轻量化 MobileNet 系列的网络模型,以此来提升模型的计算效率。其中 MobileNetV1 网络模型采用了深度可分离卷积对图像特征进行提取,MobileNetV2 网络模型是在 MobileNetV1 网络模型的基础上添加了线性瓶颈层 (Linear Bottlenecks) 和倒残差结构 (Inverted Residual) 模块,对 MobileNetV1 网络模型进一步进行优化。虽然使用轻量化模型能够有效减少模型的参数数量和计算量,但也降低了模型的分割性能。Badrinarayanan 在 2015 年提出的 SegNet<sup>[14]</sup>采用轻量级网络和跳过连接

来实现对象的快速分割,但分割得到的结果不够精细,导致网络结构参数量变大。

针对以上分割算法中的不足,提出一种能降低模型参数量、提高分割精度的语义分割算法。在原始 DeeplabV3+网络结构的基础上对其进行优化,将原有的主干网络 Xception 网络替换为 MobilenetV2 轻量化网络,同时,在 ASPP 不同特征层进行合并之后引入注意力机制模块,并在模型训练的损失函数中引入 Binary cross entropy 和 Dice loss 损失函数相结合的方式对模型进行优化。实验结果也验证了该算法的优越性和准确性。

## 2 DeeplabV3+算法设计与优化改进

DeeplabV3+算法是目前最好的语义分割算法之一,结构如图 1 所示。该算法在 DeepLabV3 的基础上,通过增加简洁有效的解码器,形成编解码结构。在编码器中,首先通过主干网络 Xception 对图像进行特征提取;其次,利用 ASPP 模块的优势,对压缩 4 次得到的特征图像进行加强处理,并进行合并,以此来获得丰富的图像特征信息;最后,利用  $1 \times 1$  卷积对加强处理得到的特征图像进行通道压缩,获得编码器输出的高层语义信息。在解码器中,首先对主干网络提取的浅层特征利用  $1 \times 1$  卷积调整维度,同时,深层特征使用双线性插值进行 4 倍上采样,获得较低的输出步长。然后与具有相同分辨率的编码器的相应特征图进行融合。最后,进行两次卷积和一次 4 倍上采样操作得到最终的特征图。

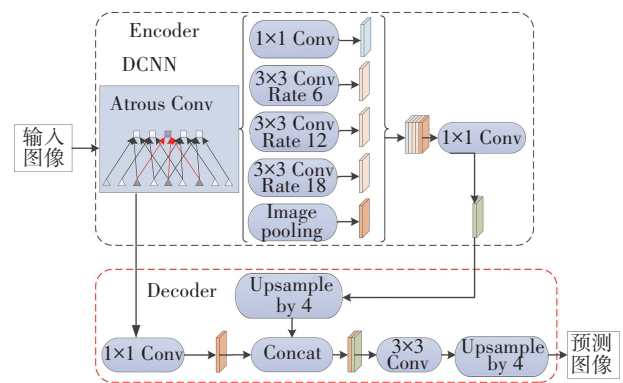


图 1 DeeplabV3+模型结构

Fig. 1 Structure of DeeplabV3+ model

本文在基础 DeeplabV3+网络结构上对其进行进一步

步优化,提出了一种融合轻量化和注意力机制的语义分割算法,该模型整体结构如图 2 所示。

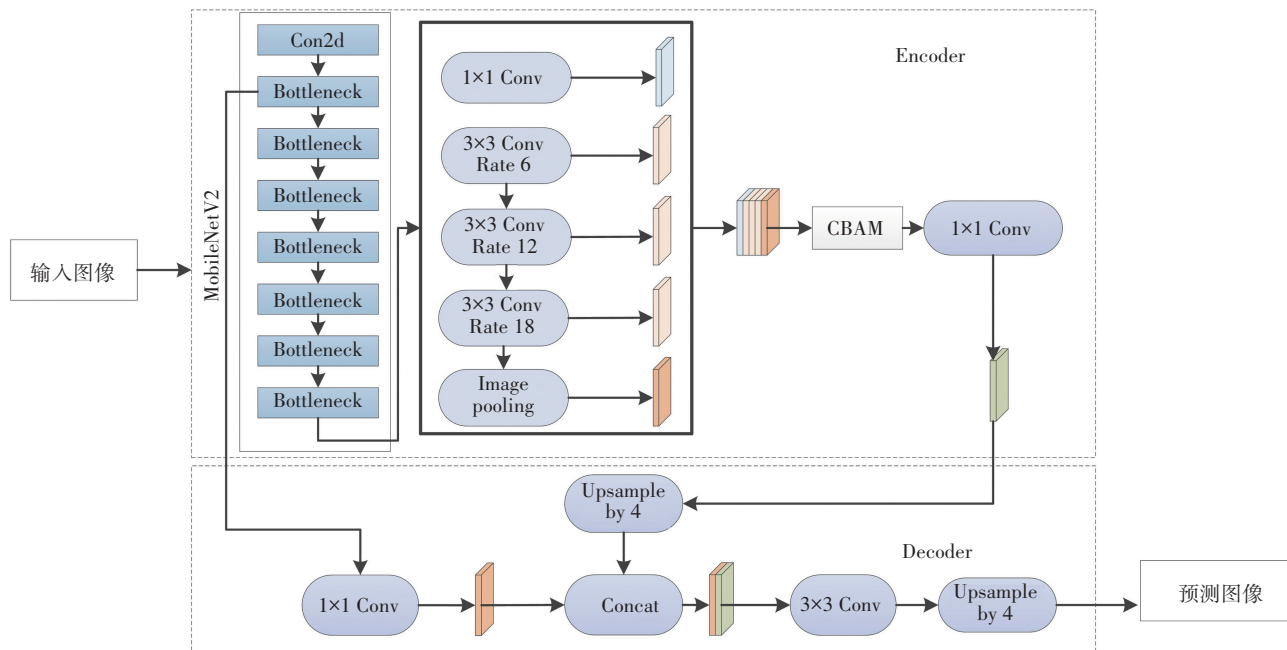


图 2 融合轻量化和注意力机制的语义分割模型结构

Fig. 2 Semantic segmentation model structure integrating lightweight and attention mechanisms

### 2.1 MobileNetV2 网络

MobileNetV2 网络是由 Google 团队在 2018 年提出的轻量化网络,MobileNetV2 是在 MobileNetV1 的基础上引入线性瓶颈层和倒残差结构,以此来提高网络的表征能力。线性瓶颈层的使用可以减少网络参数量和运算量;同时,倒残差结构的使用可以加深网络层数,避免梯度消失问题,减少参数量。相比 MobileNetV1, MobileNetV2 不仅使网络更加轻量化。而且对图像的分割精度也有较高地提升。因此,本文使用轻量化网络 MobileNetV2 作为 DeeplabV3+模型的主干网络。以此来减少网络模型的参数量和计算量,使网络模型获得较好的性能表现。同时,为了减少由通道数增加而造成的大量的计算资源浪费,本文只选取 MobileNetV2 网络结构的前 8 层,并对输入图像只进行 4 次下采样。

### 2.2 注意力机制模块

随着注意力机制(CBAM)的出现,注意力机制就被广泛用在深度学习中。注意力机制的核心是使网络关注到更为重要的信息。本文将注意力机制模块融入 DeeplabV3+网络的编码部分,用于大量融合的拼接特征之后,更加有效提取到关注图像的语义信息,保留关键特征,有利于图像的还原过程。CBAM 模型的结构如图 3 所示,主要由通道注意力模块(Channel Attention Module)和空间注意力模块(Spatial Attention Module)

进行结合,对输入进来的特征层分别进行通道注意力模块和空间注意力模块的处理。

注意力机制模块的流程如下:首先,输入特征图  $F$ ,  $F \in R^{C \times H \times W}$ ;其次,利用通道注意力模块生成一维通道注意力图  $M_c$ ,  $M_c \in R^{C \times 1 \times 1}$ ;最后,利用空间注意力模块生成二维空间注意力图  $M_s$ ,  $M_s \in R^{1 \times H \times W}$ 。公式如下:

$$F' = M_c(F) \otimes F \tag{1}$$

$$F'' = M_s(F') \otimes F' \tag{2}$$

其中,  $\otimes$  表示元素相乘,  $M_c$  表示通道注意力,  $M_s$  表示空间注意力。  $F'$  是由通道注意力模块与输入的特征图进行相乘得到,第二个计算过程则是通过  $F'$  在空间注意力模块上对特征图进行提取,再与第一计算过程的输出进行相乘,最终得到输出结果。

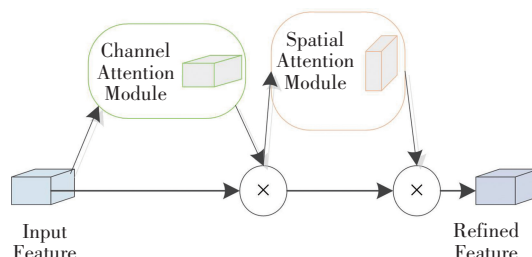


图 3 CBAM 注意力机制模块

Fig. 3 CBAM attention mechanism module

通道注意力模块首先对输入的特征图进行最大池化和平均池化,然后将得到的最大池化和平均池化的



结果送入到多层感知机 MLP 中学习,最后将 MLP 输出的两个特征图进行加操作,然后通过 Sigmoid 函数得到的权重系数对其进行处理,在将权重系数与输入的特征图进行相乘操作,得到最终的特征图通道注意力值。该模块中使用的两种池化,增加了最大池化保留分割目标判别特征的能力,使得图像能够更好地分割。其模块结构如图 4 所示,公式如式(3)所示:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (3)$$

其中,MaxPool 表示最大池化层,AvgPool 表示平均池化层,MLP 表示全连接层, $\sigma$  表示 Sigmoid 激活函数。

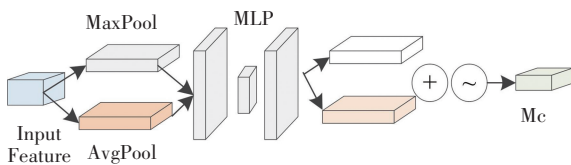


图 4 通道注意力模块

Fig. 4 Channel attention module

空间注意力模块主要关注的是输入图像的哪部分信息是更重要的,是对通道注意力模块的补充。空间注意力模块结构如图 5 所示。空间注意力模块的计算流程首先是把通道注意力模块的输出特征图作为空间注意力模块的输入,对其进行平均池化和全局池化,便于后面学习空间的特征;其次,将平均池化和全局池化得到的结果通过通道进行拼接,获得一个特征图维度;最后对拼接的结果进行卷积操作,接着使用 Sigmoid 激活函数,进而得到二维空间注意力图,即获得输入特征图中每个特征点的权重值,最后,将权重通过乘法逐通道加权到输入特征层上。

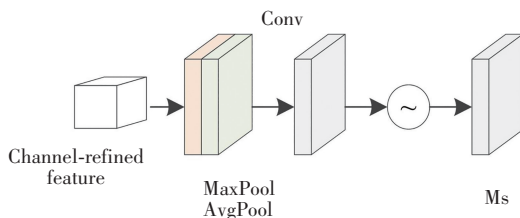


图 5 空间注意力模块

Fig. 5 Spatial attention module

### 2.3 改进的损失函数

为了提高分割结果的准确性,本文在基础 DeeplabV3+ 网络使用损失函数的基础上对其进行改进。由于数据集中图像信息的像素区域大小不一样,尤其是图片背景在整个图像中的比例较大,在计算损失的过程中,损失函数会陷入局部最小值,导致网络模型对数据进行预

测时,容易预测图像中像素较大的类别,从而会降低模型对小物体的分割精度。

经过以上分析,在原来损失函数的基础上引入二元交叉熵损失函数 BCE 和 Dice loss 进行结合,利用结合后的优势,帮助模型在训练过程中快速收敛,进一步对不同权重的目标进行分割,从而使图像分割地更加准确。

BCE Loss 能够有效地判断预测值和目标值的相近程度,即预测值与目标值越接近,Loss 越小;预测值与目标值差距越大,Loss 越大。数学表达式如式(4)所示:

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i)) \quad (4)$$

其中, $y$  是二元标签 0 或者 1, $p(y)$  是输出属于  $y$  标签的概率。作为损失函数,二元交叉熵函数被用来评判一个二分类模型预测结果的好坏程度,通俗讲,即对于标签  $y$  为 1 的情况,如果预测值  $p(y)$  趋近于 1,那么损失函数的值就趋近于 0。反之,如果此时预测值  $p(y)$  趋近于 0,那么损失函数的值就会变得非常大。

Dice loss 是由 Dice 系数而得名的,Dice 系数是一种用于评估两个样本相似性的度量函数,其值越大意味着这两个样本越相似,Dice 系数( $f_{\text{Dice}}$ )的数学表达式如式(5)所示:

$$f_{\text{Dice}} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (5)$$

其中, $|X \cap Y|$  表示  $X$  和  $Y$  之间交集元素的个数, $|X|$  和  $|Y|$  分别表示  $X$ 、 $Y$  中元素的个数。Dice Loss ( $L_{\text{Dice}}$ ) 表达式如式(6)所示:

$$L_{\text{Dice}} = 1 - f_{\text{Dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (6)$$

其中, $X$  表示真实分割图像的像素标签, $Y$  表示模型预测分割图像的像素类别, $|X \cap Y|$  近似为预测图像的像素与真实标签图像的像素之间的点乘,并将点乘的结果相加, $|X|$  和  $|Y|$  分别近似为它们各自对应图像中的像素相加。

## 3 仿真实验与结果分析

### 3.1 实验环境和数据集

本实验使用的是 Windows 10 操作系统,处理器为 Intel Core i7,8th Gen,基于 Python 3.7 的 Pytorch1.6 框架。该算法利用 Adam 优化训练器能够加快网络训练的收敛速度、帮助优化算法提高精度的优点对模型进

行训练。

为了验证本文算法的有效性和正确性,使用公开数据集 PASCAL VOC2012 和 Cityscapes 进行实验,并将其与原始 DeeplabV3+算法进行比较。

PASCAL VOC2012 数据集是 VOC2007 数据集的升级版,PASCAL VOC 2012 数据集有人、自然、交通工具、地面、动物和背景等 21 个分割类别,其扩充数据集由训练集和验证集组成,分别有 10 582 和 1 449 张图像,由于扩充后的 PASCAL VOC2012 数据集中没有划分测试集。因此,本文在训练过程中,选择数据集中的训练集用于模型训练,数据集中原本的验证集作为测试集。

Cityscapes 数据集包含大量真实、复杂的城市街景图像,该数据集共包含 25 000 张真实道路场景图像,其中有 5 000 张图像属于精细标注,20 000 张图像属于粗略标注,每张图像分辨率大小为 1 024×2 048。本文只使用精细标注的数据,精细标注的数据集由训练集、验证集和测试集组成,分别包括 2 975、500 和 1 525 张图像。

### 3.2 评估指标

为了验证改进算法的有效性,采用语义分割中常用的平均像素精度(mean Pixel Accuracy,  $f_{mPA}$ )和平均交叉比 mIoU (mean Intersection over Union,  $f_{mIoU}$ ) 作为分割精度的评价指标。

mPA 即计算每一类中正确分类的像素点所占比例的平均值( $f_{mPA}$ )。每类被正确标记的像素个数的比例,再求所有类的平均。mPA 值越大,说明对图像的分割性能越好。公式定义如式(7)所示:

$$f_{mPA} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (7)$$

mIoU 就是该数据集中的每一个类的交并比的平均, mIoU 直观地反映了分割效果的好坏。mIoU 值越大说明图像分割的精度越高。公式定义如式(8)所示:

$$f_{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (8)$$

### 3.3 网络模型改进实验

本文中算法在原始模型 DeeplabV3+的基础上进行优化,通过在 PASCAL VOC2012 和 Cityscapes 数据集上进行对比实验,来验证该算法的准确性和优越性。

#### 3.3.1 PASCAL VOC2012 数据集对比实验

为了验证本文算法的有效性,采用了对比方法,与

原始 DeeplabV3+算法进行分割对比实验,实验结果如表 1 所示。由表 1 的分割评估结果可知,改进后的 DeeplabV3+的 mIoU 达到了 79.32 个百分点,和原始的 DeeplabV3+算法相比提高了 2.82 个百分点。虽然,mPA 和原始的 DeeplabV3+算法相比降低了 2.83 个百分点,但是,改进后的 DeeplabV3+算法和改进前的参数量相比降低 14.46 M。实验结果表明,改进的 DeeplabV3+网络对 PASCAL VOC2012 数据集的图像分割是有效的。

表 1 改进前后算法的对比结果

Table 1 The comparison results of the algorithm before and after improvement

Module	Backbone	$f_{mPA}$ /%	$f_{mIoU}$ /%	Parameters /M
DeeplabV3+	Xception	86.26	76.5	71.18
本文算法	MobileNetV2	83.43	79.32	56.72

为了更直观地看出改进算法对 PASCAL VOC2012 数据集的图像分割效果,本文给出了改进算法和原算法分割结果对比图,如图 6 所示。图 6 是改进方法和原方法分割结果对比图,第 1 列为原图像,第 2 列为真实标签,第 3 列原始模型 DeeplabV3+的分割结果,第 4 列为改进后的 DeeplabV3+模型的分割结果。

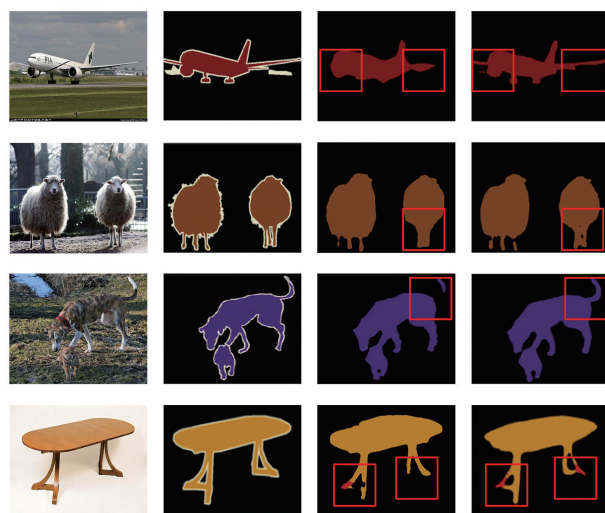


图 6 改进前后方法分割结果对比图

Fig. 6 Comparison of segmentation results before and after method improvement

从图 6 中第 1 行可以看出,本文算法在飞机的机头、机翼、机尾得到了有效地改进。通过图 6 中第 3 行的分割结果可以看出原始模型对图片进行分割时存在漏分割、分割不连续、细节不清晰等问题,而本文模型

在对图片进行分割时,大狗尾巴的分割结果明显更加连续,而且大狗后肢和小狗四肢都得到了更细化的分割效果。原始模型在对第 4 行小桌子进行分割时比较粗糙,通过比较第 4 行对小桌子的分割效果,更清楚地看出改进后的模型在对小桌子进行分割时,桌子腿之间的间隙分割得更加完整。通过以上分析,改进后的 DeeplabV3+模型相比原始模型 DeeplabV3+模型分割地更加准确,能够把图像特征中的信息完全分割出来。虽然,改进后的 DeeplabV3+算法不能对第 2 行的分割的结果完整的分割出来,但整体来说,改进后的 DeeplabV3+算法对图像的分割精度与原始 DeeplabV3+算法相比也有一定的提高。

为了更直接地看出改进后算法的训练效果,图 7 给出了本文算法在 PASCAL VOC 2012 数据集上的训练结果。其中,黑色的线条代表测试集的损失函数值,红色的线条代表训练集的损失函数值。由图 7 可知,在整个模型训练过程中,随着迭代次数的增加,训练集的损失函数值不断减小,在第 50 次时减小至 0.108,之后不断趋于稳定,最终在第 60 次训练时收敛于 0.103。同时,测试集的损失函数在训练过程中,在前期的迭代过程中减小至 0.17,随后又增加到 0.175,之后又不断减小,在第 50 次时减小至 0.1172,最终在第 60 次训练时收敛于 0.11。从以上分析可以得出,训练集和测试集的损失函数值在训练次数为 60 次时,两者均趋于稳定,由此说明该算法在训练次数为 60 次时,获得了较好的损失函数收敛性,也有利于该算法对图像中的特征信息进行更好地分割。

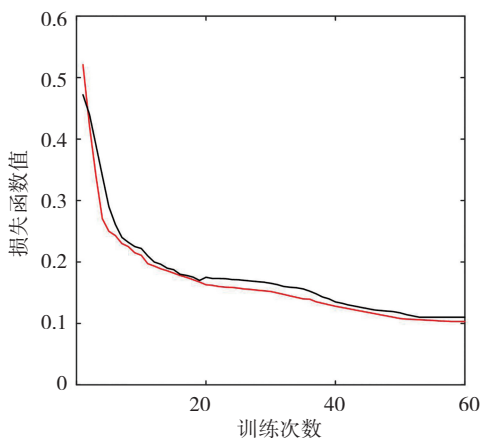


图 7 模型训练结果

Fig. 7 Model training results

### 3.3.2 Cityscapes 数据集对比实验

为了进一步验证本文算法的优越性,本文采用对

比的方法,与原始 DeeplabV3+算法进行分割对比实验。由表 2 实验结果可以看出改进后 DeeplabV3+算法的 mIoU 和 mPA 都有所提升。由图 8 给出的可视化分割结果对比图可以看出本文算法的优越性。图 8 中的第 1 列为原图像,第 2 列为真实标签,第 3 列原始模型 DeeplabV3+的分割结果,第 4 列为改进后的 DeeplabV3+模型的分割结果。

表 2 改进前后算法的对比结果

Table 2 The comparison results of the algorithm before and after improvement

Module	Backbone	$f_{mPA}/\%$	$f_{mIoU}/\%$
DeeplabV3+	Xception	84.29	74.37
本文算法	MobileNetv2	85.32	75.12

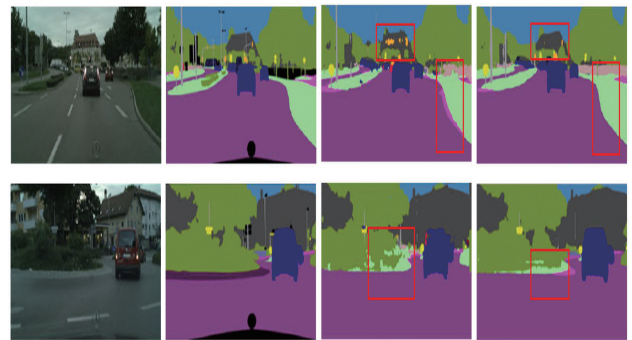


图 8 改进前后方法分割结果对比图

Fig. 8 Comparison of segmentation results before and after method improvement

## 4 结论

本文提出了一种融合轻量化和注意力机制的语义分割算法,该算法通过 MobileNetV2 构建的轻量化网络减少了模型的参数量,精简了上采样过程。同时,引入注意力机制之后,加强了网络模型对图像特征学习,通过网络对输入图像的不同信息特征分配给不同的权重,只关注网络所需要关注的部分,以此来提取更重要的信息。最后,改进的损失函数也对该算法进行了更好的优化。通过和改进前的算法性能对比,改进后的算法在 PASCAL VOC2012 数据集上得到的实验结果表明,该算法的分割精度和参数量都有所提高。同时,在 Cityscapes 数据集上得到的实验结果表明,该算法的分割精度和平均像素精度都有所提高。但改进后的算法对其他数据集和场景中图像的分割效果是否还有所提高,还需要进一步研究。



## 参考文献(References):

- [1] 田萱,王亮,丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440-468.  
TIAN Xuan, WANG Liang, DING Qi. Review of image semantic segmentation based on deep learning[J]. Journal of Software, 2019, 30(2): 440-468.
- [2] 胡云卿,潘文波,侯志超,等. 面向街景自动驾驶的 DSC-MB-PSPNet 语义分割技术研究[J]. 控制与信息技术, 2020(4): 1-9.  
HU Yun-qing, PAN Wen-bo, HOU Zhi-chao, et al. Research on DSC-MB-PSPNet semantic segmentation in street-scene autonomous driving[J]. Control and Information Technology, 2020(4): 1-9.
- [3] 张哲晗,方薇,杜丽丽,等. 基于编码-解码卷积神经网络的遥感图像语义分割[J]. 光学学报, 2020, 40(3): 46-55.  
ZHANG Zhe-han, FANG Wei, DU Li-li, et al. Semantic segmentation of remote sensing image based on encoder-decoder convolutional neural network[J]. Acta Optica Sinica, 2020, 40(3): 46-55.
- [4] 徐聪,王丽. 基于改进 DeepLabv3+网络的图像语义分割方法[J]. 激光与光电子学进展, 2021, 58(16): 225-232.  
XU Cong, WANG Li. Image semantic segmentation method based on improved DeepLabv3+network[J]. Laser & Optoelectronics Progress, 2021, 58(16): 225-232.
- [5] 杨鹤,柏正尧. CoT-TransUNet: 轻量化的上下文 Transformer 医学图像分割网络[J]. 计算机工程与应用, 2023, 59(3): 218-225.  
YANG He, BAI Zheng-yao. CoT-TransUNet: lightweight context transformer medical image segmentation network[J]. Computer Engineering and Applications, 2023, 59(3): 218-225.
- [6] 何淼楹,崔宇超. 面向自动驾驶的交通场景语义分割[J]. 计算机应用, 2021, 41(S1): 25-30.  
HE Miao-ying, CUI Yu-chao. Semantic segmentation of traffic scenes for autonomous driving[J]. Journal of Computer Applications, 2021, 41(S1): 25-30.
- [7] 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.
- [8] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//2017 IEEE conference on computer vision and pattern recognition. IEEE Computer society, 2017: 6230-6239.
- [9] CHEN L C, PAPANDEOU 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.
- [10] CHEN L C, ZHU Y, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 833-851.
- [11] CHEN L C, ZHU Y, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 833-851.
- [12] 孟球,徐磊,郭嘉阳. 一种基于改进的 MobileNetV2 网络语义分割算法[J]. 电子学报, 2020, 48(9): 1769-1776.  
MENG Lu, XU Lei, GUO Jia-yang. Semantic segmentation algorithm based on improved MobileNetV2[J]. Acta Electronica Sinica, 2020, 48(9): 1769-1776.
- [13] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4510-4520.
- [14] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.

责任编辑:陈芳