

BiSeNet轻量语义分割网络优化研究

张梦真

燕山大学理学院, 河北 秦皇岛

收稿日期: 2024年3月25日; 录用日期: 2024年4月22日; 发布日期: 2024年4月30日

摘要

语义分割是对图片上每一个像素的归类预测, 使得每个语义类别对应的预测区域得以分割显现, 是图像处理的重要方面。轻量级语义分割模型的研究点在于掌握性能与速度的天平, 使其能够投入移动设备的应用, 本文是对BiSeNet轻量语义分割模型的优化研究。首先, 本文介绍了BiSeNet模型的ResNet50主体上下文分支结构, 以及表层卷积辅助分支结构, 还有基于通道注意力机制的ARM特征加强模块和FFM融合模块作用和原理; 然后, 提出模型优化改进结构, 先在辅助分支表层卷积中以空洞卷积增强信息整体分析后, 然后以SAM空间注意力模块增强特征质量, 再利用ASPP金字塔加强辅助分支与主分支融合; 最后, 在VOC2012数据上, 得出改进前后BiSeNet模型对比结果, 在轻量性和正确性上, 验证优化结构合理性。

关键词

轻量级语义分割, BiSeNet模型, 空洞卷积, 注意力模块, ASPP金字塔

Research on BiSeNet Lightweight Semantic Segmentation Network Optimization

Mengzhen Zhang

College of Science, Yanshan University, Qinhuangdao Hebei

Received: Mar. 25th, 2024; accepted: Apr. 22nd, 2024; published: Apr. 30th, 2024

Abstract

Semantic segmentation is an important aspect of image processing, which is to classify and predict every pixel in an image, and to segment and display the prediction region corresponding to each semantic category. The research point of light-weight semantic segmentation model is to grasp the balance of performance and speed so that it can be applied to mobile devices. Firstly, this paper introduces the RESNET50 subject context branch structure of BiSeNet model and the surface con-

volution auxiliary branch structure, then, the function and principle of ARM feature enhancement module and FFM fusion module based on channel attention mechanism are put forward, in the surface convolution of auxiliary branches, the whole information is enhanced by void convolution, then the feature quality is enhanced by Sam spatial attention module, and then the fusion of auxiliary branches and main branches is enhanced by ASPP pyramid, based on the VOC2012 data, the comparison results of BiSeNet model before and after improvement were obtained, and the rationality of the optimized structure was verified in the light weight and correctness.

Keywords

Lightweight Semantic Segmentation, BiSeNet Model, Hole Convolution, Attention Module, ASPP Pyramid

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

现代轻量级语义分割是以深度学习[1]为理论基础,对图像逐像素预测其所属类别的一门技术,场景应用广泛。轻量级语义分割始于 FCN [2],其基于图像分类,提出 U 型编码器-解码器结构,在当时取得领先效果。随着移动设备的普及应用,我们对分割模型有更严格的轻量与精确性要求,以嵌入到载体中发挥实际作用。如何兼顾轻量性与准确率,发挥分割网络最大效用,是我们研究的课题。

主流轻量级语义分割网络常以 ResNet [3]、MobileNet [4]、DeepLab [5]等作为主体基础,以创新分支结构、增强特征提取、增强特征融合等手段提高模型准确率,同时注重采用轻量结构增强模型可内嵌性。在分支结构方面,有双路径互补结构的 BiSeNet [6],三路径级联汇合特征的 ICNet [7],以及多分支特征再利用结构的 DFFNet [8]等。在增强特征提取方面,常用注意力机制来筛选重要特征,加强信息利用,张铮等[9]将双空间注意力门分别安插在 Fast-SCNN 双路径末端,在像素颗粒上监督细小裂缝分割,增强鲁棒性;方家吉等[10]以瓶颈注意力模块(BAM)来同时加强通道与空间方向的注意,去除电力线分割噪点。增强特征融合方面,多尺度多层次是切入口,肖哲璇等[11]以双侧金字塔结构分别富集双分支对应的空间、语义信息;谢刚等[12]设计多尺度条形特征提取模块,加强对纤细条状物的分割识别。模型轻量化方面, 1×1 卷积、深度可分离卷积(DSC)等都是常见的轻量化手段。

虽然现代轻量级语义分割模型在平衡速度与精度的问题上取得较大进展,但没有任何一个模型同时达到最小内存与计算、最快速度与最大精度。轻量级语义分割仍有很大的优化提升空间,在轻量级与拟合性两个互斥性能中寻求平衡,实现优中取优,是本文的研究目的。本文基于双路径 BiSeNet 结构,以空洞卷积和轻量 SAM 模块增强特征提取,以轻量 ASPP 模块增强特征融合,实现对 BiSeNet 的优化。

本文内容安排为:第一章引言描述本文研究背景与目的,第二章介绍 BiSeNet 模型结构,第三章描述优化 BiSeNet 结构,第四章对比分析结果,最后得出结论。

2. BiSeNet 模型结构

2.1. ResNet50 主体

本文研究的 BiSeNet 模型骨干主体选用 ResNet50 网络,ResNet50 模型由不同 Bottleneck 瓶颈结构残差块有序搭建。Bottleneck 以恒等映射形式减小深层传播产生的信息偏离损失,以瓶颈结构轻量化模型。

如图 1 所示，以输入输出维度划分，Bottleneck 残差块分为两种结构，第一种 Bottleneck 结构先以 1×1 卷积减半输入通道数，然后以 3×3 等大卷积做特征提取，再以 1×1 卷积恢复输入通道数得到残差，最后将残差与输入直接进行 Add 操作完成恒等映射连接，输出高宽维度与通道深度均无变化；第二种 Bottleneck 结构先以 1×1 卷积保持输入通道，然后以步长为 2 的 3×3 卷积完成特征 2 倍数下采样，再以 1×1 卷积将通道数扩大 2 倍得到残差，最后将残差与 2 步长 1×1 卷积调整后输入恒等连接，输出 2 倍通道数的 2 倍下采样尺度特征图。

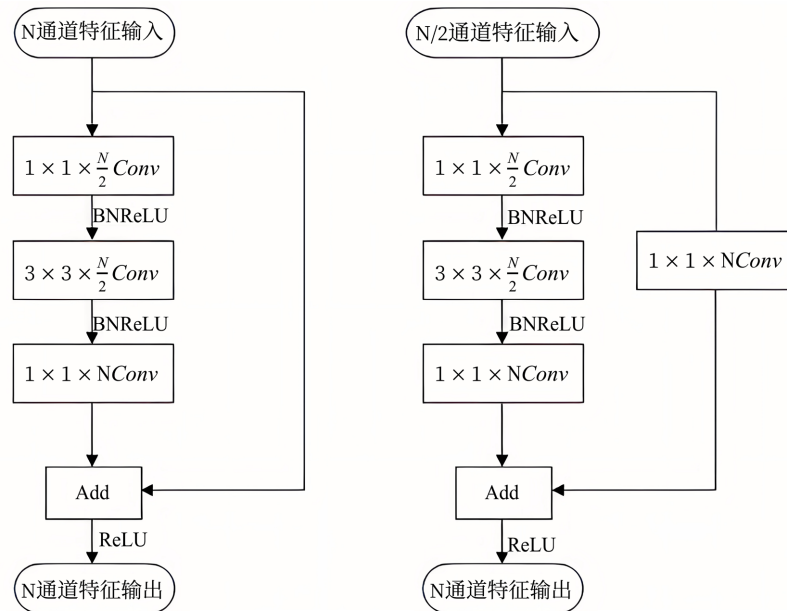


Figure 1. Bottleneck residual block structural diagram

图 1. Bottleneck 残差块结构图

ResNet50 有 50 个可训练层，本文采用的全连接层之前的 ResNet50 结构如表 1 所示，其中步骤 1 中的 7×7 卷积的填充数为 3 且步长为 2，步骤 2 中的 3×3 最大池化操作填充数为 1 且步长为 2。

Table 1. Structured data of ResNet50

表 1. ResNet50 结构数据

步骤	操作	恒等映射步长	输出尺度	输出通道数
1	7×7 卷积	—	1/2	64
2	3×3 卷积	—	1/4	64
3	Bottleneck $\times 3$	1	1/4	64
4	Bottleneck $\times 1$	2	1/8	128
5	Bottleneck $\times 3$	1	1/8	128
6	Bottleneck $\times 1$	2	1/16	256
7	Bottleneck $\times 5$	1	1/16	256
8	Bottleneck $\times 1$	2	1/32	512
9	Bottleneck $\times 2$	1	1/32	512
10	全局平均池化	—	1×1	512

2.2. 两分支结构

BiSeNet 网络在传统单路线编码器 - 解码器结构基础上, 提出了两分支结构优化其性能。BiSeNet 模型结构如图 2 所示。

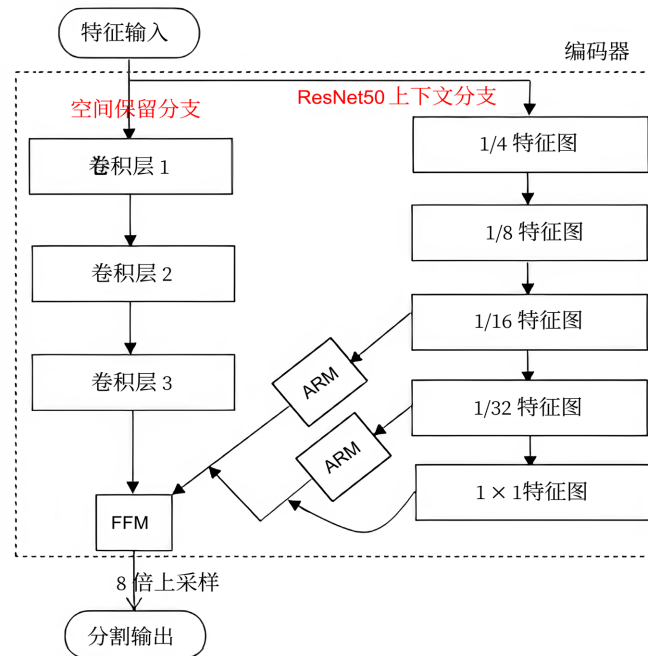


Figure 2. Two branch structure diagram

图 2. 两分支结构图

BiSeNet 将 ResNet50 作为主体分支, 图中上下文分支中 1/4 特征图、1/8 特征图、1/16 特征图、1/32 特征图、1 × 1 特征图分别对应 ResNet50 网络结构中步骤 3、步骤 5、步骤 7、步骤 9、步骤 10 的输出。考虑到 ResNet50 模型流程中特征尺度逐渐缩小至 1 × 1 维度, 若将其直接作为编码器连接解码器上采样, 必然无法还原编码过程中的过滤信息; 若将 1/8 特征图作为最终编码结果, 则上下文提取度不达标。基于以上内容, BiSeNet 模型引出第二条空间保留分支辅助编码, 第二条分支只有 3 层卷积, 以表层处理结构辅助编码保留原特征图 1/8 空间尺度, 同时减小解码上采样计算量, 空间保留分支结构如表 2 所示。

Table 2. Structured date of space reserved branch

表 2. 空间保留分支结构数据

层	操作	步长	扩充数	输出尺度
1	3 × 3 × 64 卷积、批标准化、ReLU 激活	2	1	1/2
2	3 × 3 × 64 卷积、批标准化、ReLU 激活	2	1	1/4
3	3 × 3 × 128 卷积、批标准化、ReLU 激活	2	1	1/8

2.3. ARM 模块与 FFM 模块

本节讲解 BiSeNet 模型结构中的相关模块, 包括 ARM 注意力模块和 FFM 分支融合模块。ARM 模块结构如图 3 所示, 为通道注意力模块, 其以 1 × 1 卷积核提取 GAP 特征图通道特征, 再将通道特征标准化处理, 经 Sigmoid 函数激活输出各通道概率权重, 最后以 Multiply 操作作用于输入图, 完成通道加

权，输出三个维度信息与输入无异。在上下文分支的 $1/16$ 特征图、 $1/32$ 特征图、 1×1 特征图级联融合编码输出 $1/8$ 特征图过程中，ARM 模块直接作用级联融合前的 $1/16$ 特征图和 $1/32$ 特征图，以通道加权形式加强其特征编码有效性。ARM 模块和 FFM 模块结构分别如图 3 和图 4 所示。

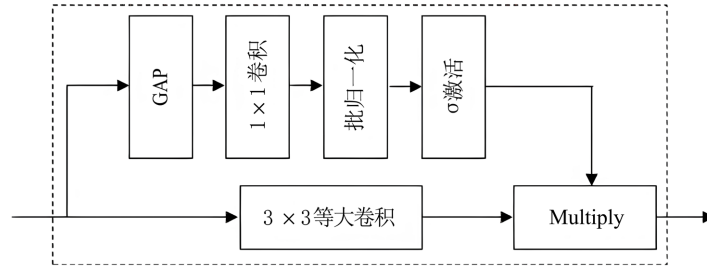


Figure 3. ARM module structure diagram
图 3. ARM 模块结构图

FFM 模块结构如图 4 所示，将上下文分支中级联输出 $1/8$ 特征图作为输入 1，将空间保留分支卷积层 3 输出的 $1/8$ 特征图作为输入 2，目的为融合两分支特征。融合过程中，首先以 1×1 卷积降噪二输入叠加数据，然后以连续两个 1×1 卷积核交互 GAP 特征图通道特征，激活后加权于降噪后叠加特征，最后将通道加权前后降噪叠加图相加，得到最终二分支融合输出，至此完成 BiSeNet 编码。

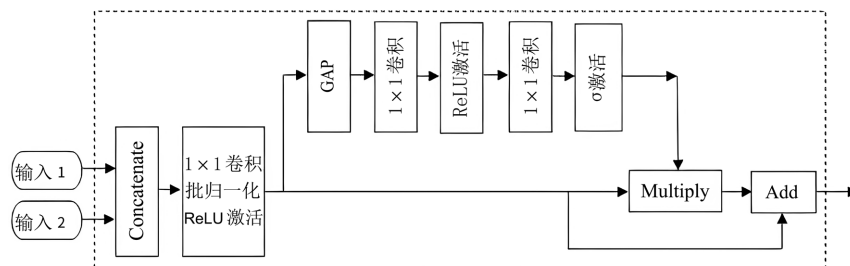


Figure 4. FFM module structure diagram
图 4. FFM 模块结构图

3. 改进 BiSeNet 模型结构

本章主要讲述基于 BiSeNet 模型的优化结构，包括倾向信息全局化提取的空洞卷积、仿照金字塔形状搭建提取尺度塔形成的 ASPP 二分支融合模块、以及增强空间特征质量的 SAM 注意力模块。

3.1. 空洞卷积

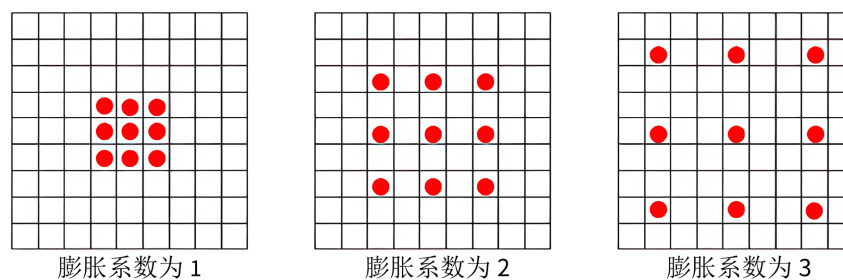


Figure 5. Atrous convolutional sketch map
图 5. 膨胀卷积示意图

空洞卷积将普通卷积稍做修改，便可得到相当可观的感受野，支撑全局化方向信息提取。空洞卷积的关键内容是其对卷积核的填充扩张操作，其余步骤与普通卷积无异。对卷积核的填充扩张操作均匀作用在其内部，以数字 0 填充卷积核扩张空位，以膨胀系数 r 来调控卷积核的扩张程度，膨胀系数为卷积核扩张后非零元素间的距离，普通卷积膨胀系数为 1。以 3×3 卷积为例，图 5 显示了当膨胀系数分别为 1、2、3 时的空洞卷积示意图，对应感受野分别为 3、5、7，膨胀系数越大，卷积核空洞越明显，信息提取区域越广。

假设特征高宽相同，不考虑通道深度，记输出维度、输入维度、卷积大小、步长、填充数、膨胀系数分别为 N 、 n 、 k 、 s 、 p 、 r ，则维度关系式见公式(1)：

$$N = \left[\frac{n + 2p - r(k-1) - 1}{s} + 1 \right] \quad (1)$$

本文将空洞卷积方法安排在 BiSeNet 模型的空间保留分支，以弥补该路径表层卷积带来的感受野局限。遵照膨胀锯齿形设置经验，本文以空洞卷积修改的空洞空间保留分支结构如表 3 所示。

Table 3. Structured date of atrous space reserved branch

表 3. 空洞空间保留分支结构数据

层	操作	膨胀系数	步长	扩充数	输出尺度
卷积层 1	$7 \times 7 \times 64$ 空洞卷积、批标准化、ReLU 激活	1	2	1	1/2
卷积层 2	$3 \times 3 \times 64$ 空洞卷积、批标准化、ReLU 激活	2	2	1	1/4
卷积层 3	$3 \times 3 \times 128$ 空洞卷积、批标准化、ReLU 激活	3	2	1	1/8

3.2. SAM 模块

BiSeNet 空间保留辅助分支输出 1/8 特征图具有较多空间像素特征，若将密集的空间信息无差别重视，一定程度会降低分割效率，易混淆不同语义像素。为增强模型的空间专注度，将特征按重要程度予以注意力，增强特征质量，提高分割效率，对照上下文分支中 ARM 模块，本文采用基于空间注意力机制的 SAM 模块来强化高分辨率特征，作用于空间保留分支的 1/8 特征图，输出空间加权的 1/8 特征图。SAM 模块结构如图 6 所示。

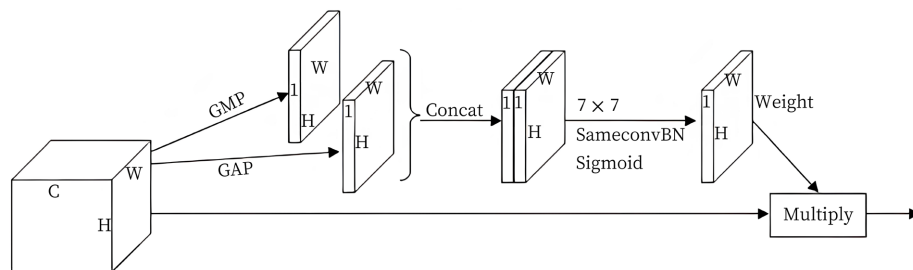


Figure 6. SAM module structural diagram

图 6. SAM 模块结构

SAM 模块结构如图所示，SAM 模块首先将并联的纵向 GAP 特征和纵向 GMP 特征图叠加，得到空间尺度不变、通道数为 2 的初步空间整合特征，然后以大的 7×7 卷积核平移运算得到空间相关性信息特征图，将其标准化后的 Sigmoid 激活值作为空间像素点权重，最后与原输入相乘完成空间加权。SAM 模块中 GMP、GAP 操作都具轻量性能，以大卷积核代替矩阵相关系数等密集运算操作，提取空间信息，同

样是轻量化的体现。

3.3. ASPP 模块

原 BiSeNet 模型 FFM 模块是一种通道注意力结构的残差融合模块，融合复杂度一般。本文以 ASPP 模块代替 FFM 模块，ASPP 模块在多尺度范围上挖掘信息，以增强对不同大小事物的分割能力。ASPP 模块结构如图 7 所示。

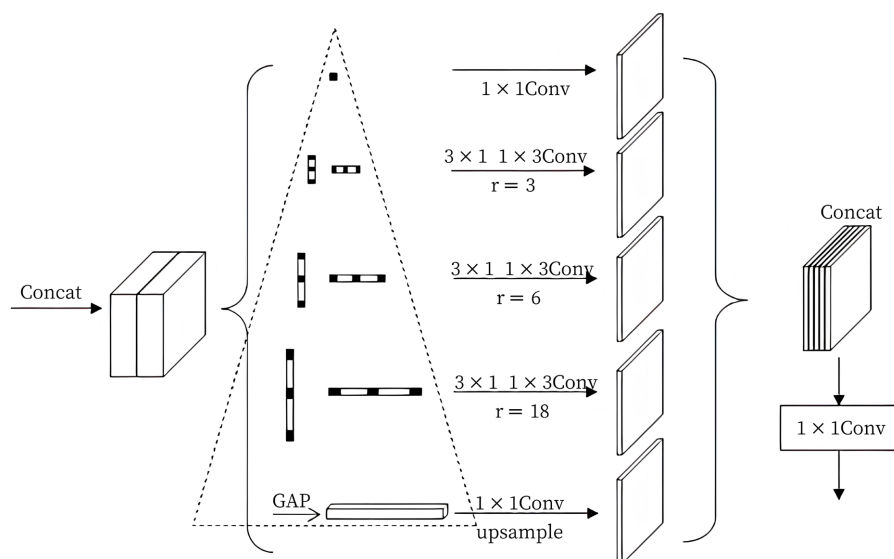


Figure 7. ASPP module structural diagram
图 7. ASPP 模块结构图

之所以说 ASPP 模块为金字塔形状，是因为其特征提取过程中，并行的卷积操作依次提取递增的尺度特征，递增尺度叠加为金字塔形状，图中三角虚线框住内容即为特征提取金字塔示意。特征提取金字塔共有 5 层，塔最顶层为提取单个像素尺度信息的 1×1 卷积，塔的第 2、3、4 层均为 3×3 卷积，以不同膨胀系数将对应层提取尺度边长分别控制为 13、25、37，塔最底层则以上采样 GAP 操作获取全局尺度信息。注意实际操作中，金字塔 2、3、4 层中的 3×3 卷积为两步骤分解形式，目的是发挥卷积分解的轻量化优势。

ASPP 模块首先将 BiSeNet 网络主分支和辅助分支输出叠加操作，然后将叠加图输入特征提取金字塔，再叠加金字塔每层并行输出的 $1/8$ 特征图，得到 5 通道 $1/8$ 特征图，最后逐点卷积调整通道数，至此完成分支融合与网络编码。

4. 实验结果对比分析

4.1. 数据集

本文选取开源 VOC2012 文件夹中 trainval.txt 文本文档所列出的 2913 张对应图片为训练测试集，设置 865 张数据用于测试，其余图片用于训练，同时采用在 Segmentation Class 子文件夹中顺序排列的对应 png 格式分割标签。该数据集有 21 个分割类，涉及室内、交通、动物、人类等丰富场景。

4.2. 性能评价指标

性能评价指标大多数是对多分类混淆矩阵的描述性统计，本文以像素准确率(PA)、交并比(IOU)、F1

值作为分割效果指标，此外以分割速度衡量模型是否具有轻量化特性。多分类混淆矩阵中，某元素值为真实行类别预测为列类别的像素个数。本文结果对应 21×21 混淆矩阵，以 n_{ij} 代表将 i 类别预测为 j 类别的像素个数。选定评价指标中，像素准确率代表总体类别上的简单正确率，如公式(2)所示：

$$PA = \frac{\sum_{c=1}^{21} n_{cc}}{\sum_{i=1}^{21} \sum_{j=1}^{21} n_{ij}} \quad (2)$$

交并比是某单一类别上，预测正确的像素数在真实值或预测值涉及到该类别的像素总数比值， $mIOU$ 则为各类交并比平均。第 c 类别的交并比如公式(3)所示：

$$IOU_c = \frac{n_{cc}}{\sum_{i=1}^{21} n_{ci} + \sum_{i=1}^{21} n_{ic} - n_{cc}} \quad (3)$$

F1 分数为某单一类别上的二次整合指标，均衡反方向的精确率 P 与召回率 R ， $mF1$ 则为 F1 的类别平均。第 c 类别的 F1 指标如公式(4)所示：

$$F1_c = \frac{2P_c R_c}{P_c + R_c} \left(P_c = \frac{n_{cc}}{\sum_{i=1}^{21} n_{ci}}, R_c = \frac{n_{cc}}{\sum_{i=1}^{21} n_{ic}} \right) \quad (4)$$

4.3. 结果分析

本文在 `pytorch` 框架下，以 SGD 算法训练优化前后 BiSeNet 网络，采用交叉熵相似度结合的损失函数，并结合上下文分支中的 $1/32$ 特征图和 1×1 特征图，分别增加两个等比重的辅助损失。训练中，将辅助损失系数设为 1，批量设为 512，训练轮数设为 230 轮，最初学习率与其衰减率分别设为 $2.5e^{-2}$ 和 0.97，动量为 0.9。

4.3.1. 整体指标对比分析

优化前后 BiSeNet 模型整体指标对比如表 4 所示。

Table 4. Overall indicator comparison

表 4. 整体指标对比

模型	Time	PA (%)	$mIOU$ (%)	$mF1$ (%)
原 BiSeNet 模型	52 ms	85.36	54.23	67.65
改进后 BiSeNet 模型	78 ms	87.42	63.30	75.81

表中将单张分割时长(Time)作为模型的整体轻量级指标，将像素准确率(PA)、平均交并比($mIOU$)、平均 F1 分数($mF1$)作为整体性拟合指标。首先，改进后 BiSeNet 模型单张分割用由 52 ms 增加至 78 ms，虽然增加了 26 ms，但其仍在轻量级范围内，说明改进后 BiSeNet 模型具有良好轻量内嵌性能。其次，PA 值、 $mIOU$ 值及 $mF1$ 值分别达到了 87.42%、63.30%、75.81%，均超过 50%，其整体拟合指标均达到到置信要求，优化结构将 PA 值、 $mIOU$ 值及 $mF1$ 值分别提升提升了 2.06%、9.07%、8.16%，说明优化后 BiSeNet 模型具有更强的整体拟合性。伴随训练过程，优化前后 BiSeNet 模型 PA、 $mIOU$ 、 $mF1$ 走势图如图 8 所示，绿色曲线描述优化 BiSeNet 指标走势，红色曲线描述原 BiSeNet 指标走势，绿线均高于红线，可直观观察到改进后 BiSeNet 模型整体拟合性更强。

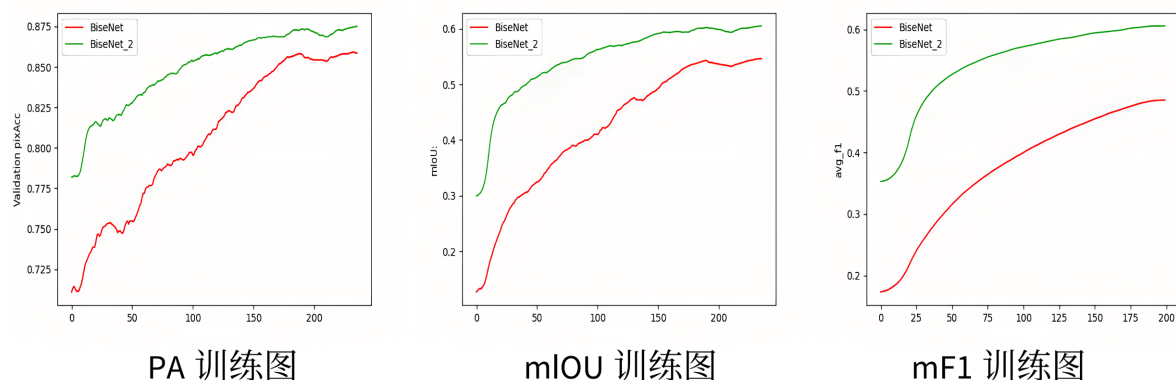


Figure 8. Overall fitting index trend chart

图 8. 整体拟合指标走势图

4.3.2. 各类别指标对比分析

优化前后 BiSeNet 模型在 VOC2012 的 21 个具体类别上得出结果，各类别指标对比如表 5 所示。

Table 5. System resulting data of standard experiment

表 5. 各类别结果指标数据

类别	IOU (%)	IOU' (%)	$F1$ (%)	$F1'$ (%)
背景	86.54	88.21	92.37	93.97
飞机	66.80	74.16	78.99	85.20
自行车	35.62	34.19	53.42	52.67
鸟类	46.39	71.29	62.53	84.76
船	46.35	50.06	61.43	64.98
瓶子	48.32	48.65	65.83	65.91
公交车	67.31	80.95	77.93	89.47
汽车	64.92	76.29	78.00	87.68
猫	67.25	78.43	78.39	87.55
椅子	25.64	33.62	35.82	38.45
牛	50.66	68.17	68.51	79.86
餐桌	42.67	53.47	53.79	68.88
狗	59.45	67.39	74.77	83.62
马	49.75	65.20	65.98	78.43
自行车	63.52	67.49	77.48	81.13
人物	67.92	73.46	83.61	86.76
盆栽	34.37	46.83	42.29	60.38
绵羊	60.24	69.93	76.34	84.16
沙发	39.10	42.03	48.36	55.72
火车	61.83	76.19	77.15	86.58
显示器	48.55	57.32	67.38	70.04

表中 IOU 和 $F1$ 对应原 BiSeNet 交并比与 $F1$ 指标, IOU' 和 $F1'$ 对应优化后模型指标。由表可知, 不考虑背景类, 对 20 个实物类别的分割结果中, 单独考虑优化 BiSeNet 模型, 其对飞机、鸟类、公交车、汽车、猫、人物、火车等 7 个类别分割结果优秀, 其交并比和 $F1$ 指标均达到 70% 以上, 其中对公交车类指标最高, 交并比和 $F1$ 指标分别达到了 80.95%、89.47%; 对船、牛、餐车、狗、马、自行车、绵羊、显示器等 8 个类别分割良好, 双指标均达到 50% 以上, 说明优化 BiSeNet 模型具有可信性。对比结果显示, 优化后 BiSeNet 模型增强了大多数像素类别的预测准确性, 优化后 BiSeNet 模型对飞机、鸟类、公交车、汽车、猫、牛、餐桌、狗、马、盆栽、绵羊、火车等 12 个类别提升效果显著, 双指标均提升 5% 以上, 对应交并比指标分别提升 7.36%、24.9%、13.64%、11.37%、11.18%、17.51%、10.8%、7.94%、15.45%、12.46%、9.69%、14.36%, 对应 $F1$ 分数指标分别提升 6.21%、22.23%、11.54%、9.68%、9.16%、11.35%、15.09%、8.85%、12.45%、18.09%、7.82%、9.43%, 其中鸟类分割优化最显著, 双指标均提升了 20% 以上, 验证了优化结构的合理性。

4.3.3. 分割效果对比分析

在测试集中随机选取两张图片, 优化前后 BiSeNet 模型对比效果图如图 9 所示。

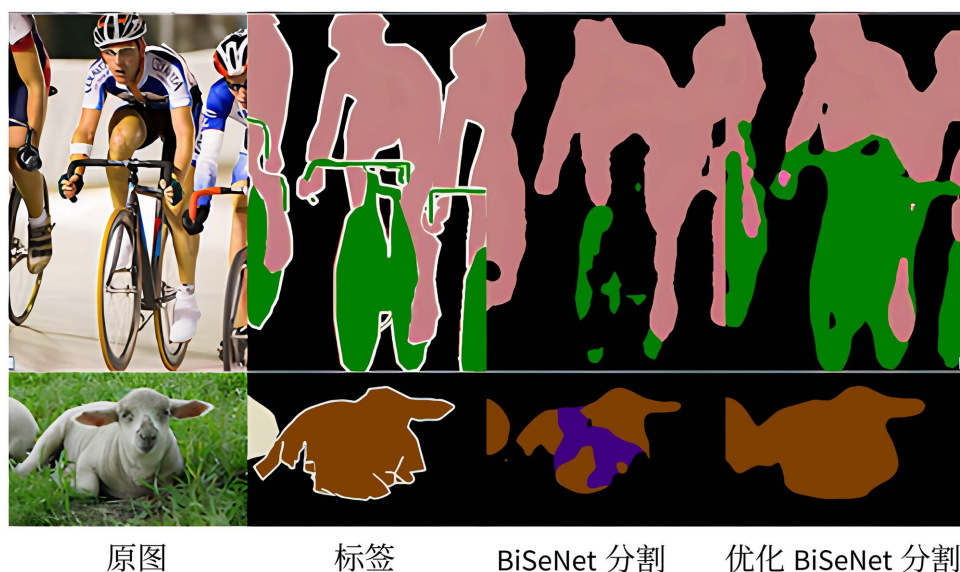


Figure 9. Comparison of segmentation effect
图 9. 分割效果对比图

分析图例举的效果对比图, 可观察到原 BiSeNet 模型将图例中的自行车类别像素大面积归类为背景, 且将羊内部分像素错误预测为狗类; 优化后 BiSeNet 模型则解决了原模型中出现的问题, 还原了大面积的自行车区域, 矫正了对羊内部像素的错误分类, 可直观对比出优化后 BiSeNet 模型分割结果更加贴合真实值标签, 优化模型具有增强效果。

5. 总结

本文基于原 BiSeNet 模型, 提出优化结构, 并验证其合理性。优化改进包括: (1) 在原 BiSeNet 辅助分支表层化卷积中, 以空洞卷积升级普通卷积, 提高图片像素整体相关特征提取。(2) 对 BiSeNet 辅助分支表层卷积输出, 进行轻量 SAM 空间加权操作, 增强辅助分支数据质量。(3) 以金字塔结构完成 BiSeNet 上下文主分支及辅助空间保留分支的输出融合, 提高融合复杂度。VOC2012 验证结果为: (1) 79 ms 的分

割速度验证了优化后 BiSeNet 模型的轻量性。(2) 优化后 BiSeNet 模型绝大多数单类别、类平均指标及整体分割评价指标均优于原模型, 优化后 BiSeNet 分割图直观上比原模型更贴合标签, 验证了优化结构的合理性。

参考文献

- [1] Wang, Z., Tang, C., Sima, X., *et al.* (2021) Research on Application of Deep Learning Algorithm in Image Classification. 2021 *IEEE Asia-Pacific Conference on Image Proceedings, Electronics and Computers (IPEC)*, Dalian, 14-16 April 2021, 1122-1125. <https://doi.org/10.1109/IPEC51340.2021.9421185>
- [2] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [3] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [4] Howard, A.G., Zhu, M., Chen, B., *et al.* (2017) Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv: 1704.04861.
- [5] Chen, L.C., Papandreou, G., Kokkinos, I., *et al.* (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- [6] Yu, C.Q., Wang, J.B., Chao, P., *et al.* (2018) BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Proceedings of the 15th European Conference on Computer Vision*, 334-349. https://doi.org/10.1007/978-3-030-01261-8_20
- [7] Zhao, H.S., Qi, X.J., Shen, X.Y., *et al.* (2018) ICNet for Real-Time Semantic Segmentation on High-Resolution Images. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Proceedings of the 15th European Conference on Computer Vision*, 418-434. https://doi.org/10.1007/978-3-030-01219-9_25
- [8] Tang, X.Y., Tu, W.X., Li, K.Q., *et al.* (2021) DFFNet: An IoT-Perceptive Dual Feature Fusion Network for General Real-Time Semantic Segmentation. *Information Sciences*, **565**, 326-343. <https://doi.org/10.1016/j.ins.2021.02.004>
- [9] 张铮, 钱勤建, 周嘉政, 等. 基于改进 Fast-SCNN 的裂缝实时分割算法[J]. 应用光学, 2023, 44(3): 539-547.
- [10] 方家吉, 赖一波, 唐正涛, 等. 基于 DeepLabV3+ 的轻量级电力线语义分割方法[J]. 计算机时代, 2023(9): 19-23, 28.
- [11] 肖哲璇, 陈辉, 王硕. 基于双分支多尺度特征融合的道路场景语义分割[J]. 宁夏师范学院学报, 2024, 45(1): 81-92.
- [12] 谢刚, 王荃毅, 谢新林, 等. 融合多尺度深度卷积的轻量级 Transformer 交通场景语义分割算法[J]. 通信学报, 2023, 44(10): 212-225.