

融合空间细节与语义特征的指针式仪表 双边图像分割网络

朱耀晖, 吴志刚*, 陈敏

江西理工大学能源与机械工程学院, 江西 南昌 330013

摘要 针对指针式仪表的小目标图像分割特点以及现有方法存在的局限性, 提出融合空间细节与语义特征的双边路深度学习主干网络 BiUnet, 用于指针式仪表图像分割。该网络以 BiSeNet V2 算法作为出发点, 重新设计其语义支路、细节支路以及双边融合层。首先, 利用 ConvNeXt 卷积模块对细节支路进行调整和优化, 提升算法对指针及刻度线边界细节的特征提取能力; 其次, 结合编码器-解码器 U 型结构能够融合不同尺度语义信息的优势对语义支路进行重新设计, 提高语义支路对指针和刻度这类小目标的分割能力; 最后, 提出双引导拼接聚合层, 用于更好地融合细节支路和语义支路特征。在自制仪表图像分割数据集上的消融实验证实了所提出的网络设计方案的有效性和可行性。在仪表数据集上与不同主干网络进行对比实验, 实验结果表明, BiUnet 对仪表的平均分割精度 mIoU (mean intersection of union) 达到了 88.66%, 相比于 BiSeNet V2 网络的 80.02% 提升 8.64 个百分点, 并且这两种网络的分割精度均优于基于 Transformer 和纯卷积的常见主干网络。

关键词 指针式仪表; 双边路主干网络; 深度学习; 图像分割

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP231461

A Pointer Meter Bilateral Image Segmentation Network Integrating Spatial Details and Semantic Features

Zhu Yaohui, Wu Zhigang*, Chen Min

*School of Energy and Mechanical Engineering, Jiangxi University of Science and Technology,
Nanchang 330013, Jiangxi, China*

Abstract Aiming at the characteristics of small target image segmentation of pointer meter and the limitations of existing methods, a bilateral deep learning backbone network called BiUnet is proposed for pointer meter image segmentation, which combines spatial details and semantic features. Starting from BiSeNet V2 algorithm, the semantic branch, detail branch and bilateral fusion layer are redesigned in this network. First, the ConvNeXt convolution block is used to adjust and optimize the detail branch to improve the feature extraction ability of the algorithm for pointer and scale line boundary details. Second, the semantic branch is redesigned based on the advantages of the U-shape structure of encoder and decoder to integrate different scales of semantic information, which improves the special segmentation ability of the semantic branch for small objects such as pointer and scale. Finally, a bilateral-guide splicing aggregation layer is proposed to fuse the detail branch and the semantic branch features. The ablation experiments on the self-made instrument image segmentation dataset confirm the validity and feasibility of the proposed network design scheme. Comparative experiments with different backbone networks are carried out on the instrument dataset, the experimental results show that the mIoU (mean intersection of union) of BiUnet's instrument segmentation accuracy reaches 88.66%, which is 8.64 percentage points higher than the BiSeNet V2 network (80.02%). Both of them have better segmentation accuracy than common backbone networks based on Transformer and pure convolution.

Key words pointer meter; bilateral road backbone network; deep learning; image segmentation

收稿日期: 2023-06-05; 修回日期: 2023-06-24; 录用日期: 2023-07-24; 网络首发日期: 2023-08-18

基金项目: 江西省教育厅科技项目(GJJ200833)、江西理工大学博士启动基金(A08)

通信作者: *wzgang2017@163.com

1 引言

在工业制造过程中,利用人工读取仪表数据信息效率低、成本高且易出错,因此,针对自动指针式仪表图像分割的计算机视觉方案得到了广泛的关注^[1]。

传统视觉算法方面,基于图像处理的利用霍夫变换、图像减法和图像细化的指针式仪表图像采集技术^[2]可以得到指针的轮廓信息,但该方法依赖于对某个场景下的仪表信息的先验知识,需要根据不同的环境信息动态调整参数。结合轮廓特征筛选和线性检测的指针识别方法在一定程度上提高了指针识别的准确性^[3],然而,它要求极高的图像质量。将 RGB 空间转换为 HSV 空间,利用颜色特征来检测仪表中的指针^[4],指针的颜色特征受光照影响比较明显,使用时间长可能会导致指针颜色褪色,因此该方法的可靠性相对较低。一方面,基于机器视觉的算法都受到特定环境和位置的限制;另一方面,对于不可控因素,如不均匀的光照、光照变化范围大、复杂的背景、图像模糊和尺度变化等,这些算法需要人为调整参数,不具有较好的鲁棒性。因此,研究者提出了一种非局部信息和反投影聚类算法,在图像分割中用于平衡噪声抑制和细节保留之间的关系^[5]。针对应用环境为水下的视觉任务,分解网络和物理成像模型图像恢复技术被提出^[6]。这类技术为更多场景下指针式仪表的图像分割提供了实现方案。

基于上述挑战,利用神经网络中端到端的模式被广泛研究。如:通过 Yolov4 跟踪仪表位置信息,并利用改进后的 U-net 对仪表指针及刻度信息进行分割^[7]。但该方法在 U-net^[8]的各个阶段增加特征通道使得模型的参数量增加,降低了分割的实时性。全卷积网络(FCN)对输入图像^[9]的各个像素进行分类,没有充分考虑像素与像素之间的关系,忽略了在通常的基于像素分类的分割方法中使用的空间规整步骤,缺乏空间一致性。万吉林等^[10]提出了一种利用 Faster R-CNN (region convolution neural network)进行表盘检测,然后利用 U-net 对表盘进行刻度和指针分割的双阶段算法。此外,研究者提出的带有关键点检测的指针仪表自动读取方法的主要思想是首先应用改进的 mask RCNN 检测仪表上刻度和指针的关键点^[11]。虽然这些方法可以达到实时的推理速度,但它们极大地牺牲了准确性和效率,损失了底层的细节。

总体来说,利用深度学习分割指针式仪表图像的方式相比于传统的视觉算法来说具有更强的泛化能力和鲁棒性,且具有较高的精度。但上述利用深度学习的方法存在精度与实时性无法达到较好的平衡的问题。BiSeNet V2^[12]网络利用语义支路、细节支路及双边路主干网络更好地学习目标的语义特征和空间细节特征,同时考虑了底层细节和高层语义,以达到高精度和高效率。此外,该网络可通过调整细节支路和语义

支路的网络结构来调整网络模型的大小。然而,针对指针和刻度线这类小目标的特征边缘细节分割及深层语义特征的目标位置信息提取,BiSeNet V2 仍有很大的提升空间。

本文基于 BiSeNet V2 提出了一种新型的双边路 U-net 主干网络并将其命名为“BiUnet”,用于更好地处理指针式仪表小目标分割任务,以达到较高的准确性和效率。首先,该主干网络保留了 BiSeNet V2 所提出的细节支路和语义支路这种双边路思想,细节支路用来捕捉宽通道和浅层的空间细节;语义支路采用窄通道、深层次的分类语义提取方法;双边路将浅层级空间细节和深层次的语义信息分开处理,以实现高精度和高效率的实时语义分割。其次,利用 ConvNeXt^[13]卷积模块重新设计了细节支路,并调整每个阶段卷积模块的数量使模型达到更高的精度。接着,结合 U-net 的编码器-解码器思想对语义支路进行重新设计,利用 U 型网络能够融合不同尺度和上下文语义特征的优势使语义支路更好地融合高级语义信息,提升模型对指针及刻度这类小目标的分割精度。同时,调整了编码器-解码器的每个阶段的通道数量,使模型在保持较高的分割精度的同时具有较为轻量化的特性,降低网络模型的计算量。最后,对原始 BiSeNet V2 提出的聚合层进行重新设计,简化原有聚合层,使细节支路与语义支路在特征图尺寸和通道数量上保持一致,并利用通道维度的拼接操作来融合双支路特征信息。通过消融实验及算法对比实验和可视化分析验证了所提 BiUnet 的有效性和可靠性。

2 BiUnet 搭建

BiUnet 的网络框架如图 1 所示,主要由三个部分组成:细节支路主干、语义支路主干以及双支路聚合模块层。细节支路有三个阶段,每个阶段都是基于 7×7 卷积核和倒置瓶颈结构组成的残差卷积模块,每个阶段的卷积模块数量比为 1:3:3,每个阶段的通道数分别为 384、192、128。在每个阶段开始时会对图像进行二倍下采样和归一化处理。语义支路为轻量化 U-net,网络的 5 个阶段通道数量依次为 16、32、64、128、256。BiUnet 放弃了 BiSeNet V2 的双引导聚合层结构,引入了一种简化且精度更高的双引导拼接聚合层结构。

以 BiSeNet V2 作为参照,对细节支路、语义支路和聚合层依次进行优化和重新设计,逐步实现 BiUnet 的搭建。

2.1 细节支路构建

细节支路主要针对指针及刻度的空间细节,即低级信息。浅层的高分辨率图像蕴含了丰富的空间细节信息,例如:目标的大小、纹理、边界等,因此,该支路需要丰富的通道容量和高性能的卷积核来保留空间细节信息。BiSeNet V2 中使用的传统黄金标准内核大小 (3×3) 的卷积在现代图形处理器(GPU)上具有很好

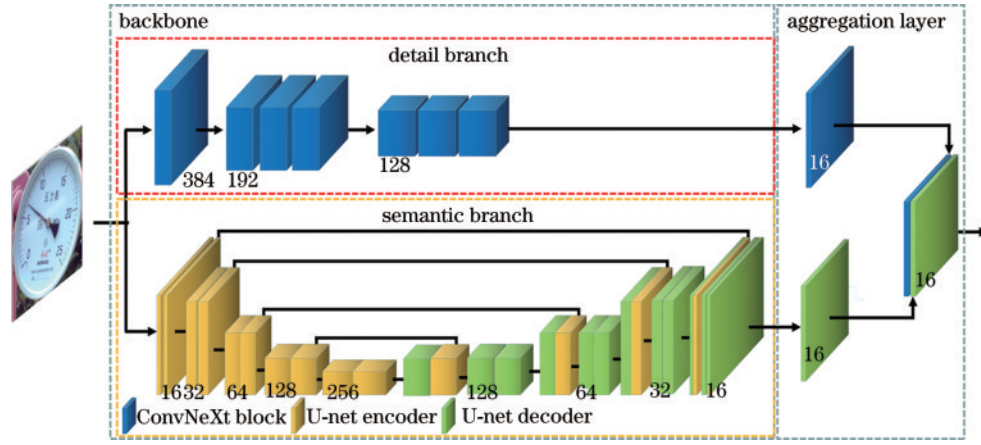


图 1 BiUnet 框架
Fig. 1 BiUnet framework

的效果^[14],但 3×3 传统卷积模块生成指针及刻度密集特征的效果还欠佳。在 ConvNeXt 实验中,通过将 ResNet-50 中 3×3 的卷积核替换为 7×7 的卷积核,算法的精度从 79.9% 提升到 80.6%;另外,在 ResNet 中应用倒瓶颈结构机制,算法的精度从 81.9% 提高至 82.6%^[13]。将 BiSeNet V2 细节支路原有的 3×3 卷积替换为 ConvNeXt 卷积模块,调整前后卷积模块的对比如图 2 所示。在 BiSeNet V2 细节支路中,每个阶段的卷积模块都是 3×3 且具有相同通道数量的卷积核;而 ConvNeXt 卷积模块是一个 7×7 和两个 1×1 卷积核组成的残差卷积模块,其中隐藏通道是输入输出通道 C 的 4 倍。

除了对卷积模块进行调整之外,每个阶段开始前都要经过一次 Conv2d 下采样和归一化,以减少计算量,防止过拟合。每个阶段的输出特征通道由原先的 64、64、128 调整为 384、192、128。将高通道容量卷积核用于高分辨率的浅层特征学习来丰富空间细节,详细参数如表 1 所示。其中包含每个阶段(stage)的一个或多个操作(opr),如 ConvNeXt、Conv2d 等,每个操作

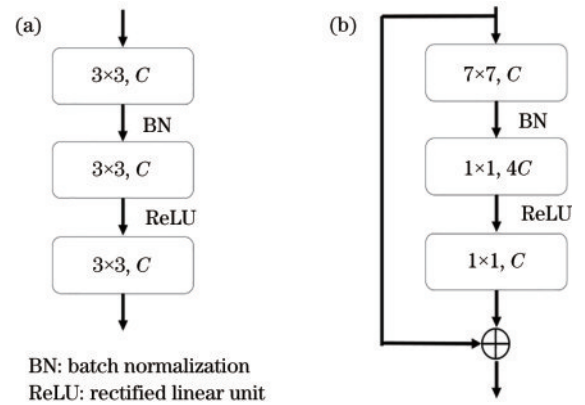


图 2 细节支路卷积模块调整前后对比。(a)BiSeNet V2 细节支路卷积模块;(b)ConvNeXt 卷积模块
Fig. 2 Comparison of the detail branch convolution block before and after adjustment. (a) Detail branch convolution block of BiSeNet V2; (b) ConvNeXt convolution block

都有填充量 p 、内核大小 k 、步幅 s 、输出通道 C 和重复次数 r 参数。

表 1 原始与改进 BiSeNet V2 的细节支路参数对比
Table 1 Comparison of detail branch parameters of original and improved BiSeNet V2

Stage	BiSeNet V2 detail branch						Improved detail branch						Output size
	Opr	p	k	C	s	r	Opr	p	k	C	s	r	
Input													512×512
1	Conv2d	1	3	64	2	1	Conv2d	1	2	384	2	1	256×256
	Conv2d	1	3	64	1	1	ConvNeXt	3	7	384	1	1	
2	Conv2d	1	3	64	2	1	Conv2d	1	2	384	2	1	128×128
	Conv2d	1	3	64	1	2	ConvNeXt	3	7	192	1	3	
3	Conv2d	1	3	128	2	1	Conv2d	1	2	192	2	1	64×64
	Conv2d	1	3	128	1	2	ConvNeXt	3	7	128	1	3	

通过对细节支路的重构建使算法的平均分割精度 mIoU (mean intersection of union) 由 80.02% 提高到 84.96%。至此,对于细节支路的构建是有效的。接着以此作为新的起点继续构建语义支路和双引导拼接聚

合层。

2.2 语义支路构建

语义支路采用窄通道、深层次的分类语义提取方法,旨在捕获更深层指针及刻度的高级语义。如图 3(a)

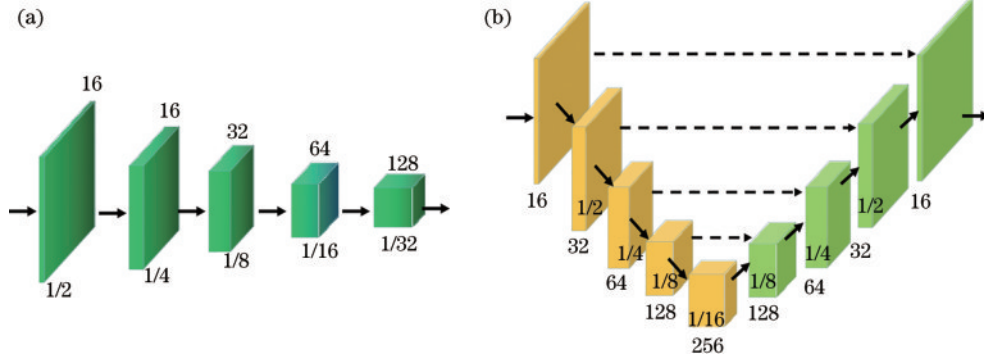


图 3 语义支路结构对比。(a)BiSeNet V2原始语义支路;(b)改进后的语义支路

Fig. 3 Comparison of semantic branch structures. (a) BiSeNet V2 primitive semantic branch; (b) improved semantic branch

所示, BiSeNet V2 通过 5 个阶段的卷积模块提取特征语义信息, 输出最后阶段的深层次高级语义特征。每个阶段图像的尺寸是上一个阶段图像尺寸的 1/2, 每个阶段的通道数依次为 16、16、32、64、128。对不同尺度的高级语义进行融合更有助于提高算法对指针及刻度这类小目标的分割能力。参考 U-net 改进的语义支路如图 3(b) 所示。该支路中的具体操作为: 通过第 5 个阶段中由最大值池化和卷积组成的编码器将输入图像编码为不同级别的语义特征表示; 再通过第 4 个阶段中上采样和卷积组成的解码器, 将编码器学习到的判别性特征语义逐步投影至与原始空间分辨率一致的像素空间以进行密集预测分类; 接着通过“跳跃连接”操作, 在网络进行上采样的同时将来自编码器的更高分辨率的特征图与上采样的特征图拼接起来, 以更好地学习后续卷积的特征。为减小网络模型、降低计算量, 调整每个阶段的通道数为原 U-net 的 1/4, 依次为 16、32、64、128、256。相比 BiSeNet V2 语义支路, U-net 可以更好地融合不同尺度和不同深度的语义信息。

通过对语义支路的改进, 算法的平均分割精度 mIoU 由 84.96% 提升到 85.47%。至此, 对于 BiSeNet V2 语义支路的改进是有效的。后续以此作为新的起点继续构建双引导拼接聚合层。

2.3 双引导拼接聚合层构建

细节支路和语义支路的特征表示是互补的, 它们互相不知道另一个支路的信息, 两个支路的输出具有不同的特征表示水平。因此, 聚合层旨在合并两种类型的特征表示。如图 4 所示, BiSeNet V2 在细节支路和语义支路上分别利用卷积、深度卷积、平均池化、归一化等操作使图像的尺寸和通道对齐, 再令逐像素点特征相乘后相加, 以融合两个支路的互补信息。然而该方法对语义支路中 U-net 提取的结果进行高度下采样和逐像素点融合, 这破坏了双支路网络对指针和刻度特征的提取效果。改进后的聚合层通过将细节支路的 128 通道降为 16 通道来将提取的图像上采样至 256 pixel × 256 pixel, 将语义支路下采样至 16 通道输出图像尺寸为 256 pixel × 256 pixel 的特征图。这里的聚合使用的是拼接操作而非 BiSeNet V2 中的逐点相乘再相加操作。

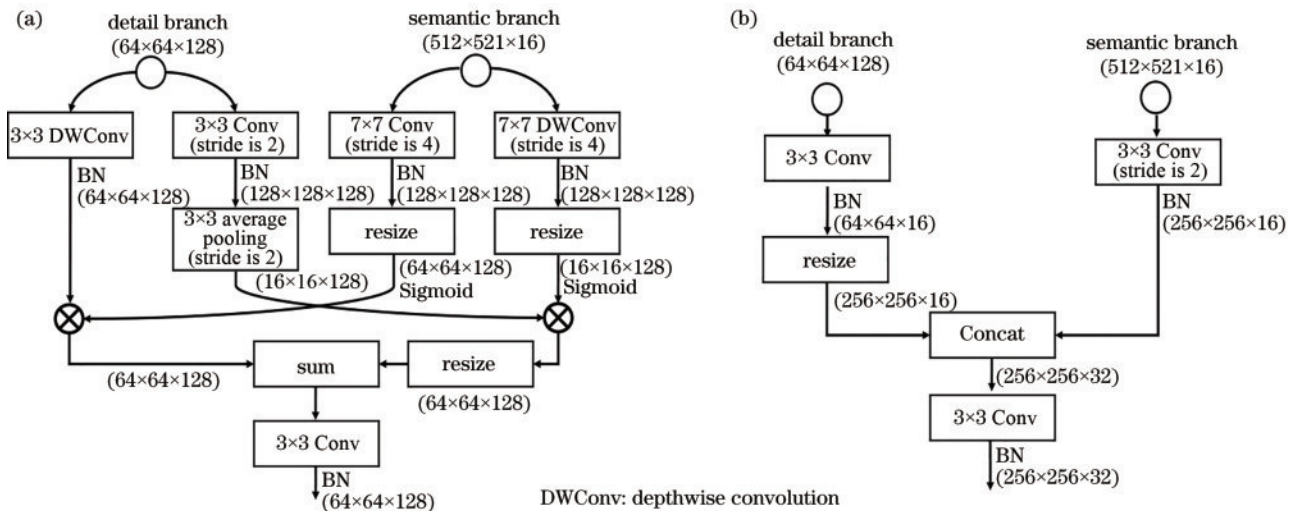


图 4 双引导聚合层与双引导拼接聚合层的结构对比。(a) BiSeNet V2 的双引导聚合层;(b) 双引导拼接聚合层

Fig. 4 Comparison of the structure of bilateral-guided aggregation layer and bilateral-guided splicing aggregation layer. (a) Bilateral-guided aggregation layer of BiSeNet V2; (b) bilateral-guided splicing aggregation layer

3 实验研究

接下来,首先介绍所用的仪表数据集和算法的实现细节;然后,研究所提方法中每个组件对算法性能的影响;最后,对比仪表数据集上所提算法与其他算法的分割精度、模型参数数量和计算量(flops)指标。

3.1 数据集及实验环境

实验所使用的数据集是人工标注的 350 张小批量仪表分割数据集,其中包含 4 个类别:背景(background)、仪表刻度线(line)、仪表指针中心点(pointer_mid)、指针(pointer)。训练与验证数据集的比例为 7:3,即其中 245 张用于训练,105 张用于验证。训练过程中 batchsize 为 2,实际输入图像的尺寸为 512 pixel×512 pixel,通过 400 次迭代消除误差波动,

取迭代过程中的最高精度为最终精度。同时,将平均联合交集的标准度量 mIoU、像素分类正确率的平均值 aAcc(average accuracy)和类别的像素分类平均正确率 mAcc(mean accuracy)作为评价指标。此外,基于 PyTorch 1.13 和 MMLab 框架进行实验。通过单张型号为 NVIDIA GeForce RTX 3060,具有 12 GB 显存的 GPU 在 CUDA 11.6 环境下进行推理。

3.2 仪表数据集消融实验

通过消融实验验证所提 BiUnet 算法的每个组件的有效性。实验结果如表 2 所示,其中:baseline 为 BiSeNet V2 算法中原有的各个模块;improved 为所提对于语义支路及细节支路的改进方案;BGA 为 BiSeNet V2 算法提出的双边引导聚合;BGCA 为所提双引导拼接聚合。

表 2 消融实验结果

Table 2 Ablation experimental results

Detail branch	Semantic branch	Aggregation layer	IoU / %				mIoU / %	#Param. / 10 ⁶
			Background	Line	Pointer_mid	Pointer		
Baseline	Baseline	BGA	98.54	58.15	85.91	77.49	80.02	3.34
Improved	Baseline	BGA	98.88	65.68	89.35	85.93	84.96	7.19
Baseline	Improved	BGA	98.96	67.98	89.55	86.01	85.62	3.26
Improved	Improved	BGA	98.94	68.25	89.79	84.88	85.47	5.76
Improved	Improved	BGCA	99.29	78.02	89.67	87.67	88.66	4.93

首先,以 BiSeNet V2 作为基线在仪表数据集上进行训练,其平均分割精度 mIoU 达到了 80.02%。改进细节支路后,算法的 mIoU 达到了 84.96%。算法在 line、pointer_mid、pointer 三个类别上的分割精度 IoU 均有明显提升,验证了该方案的可行性,证明改进细节支路提升了算法对刻度线、指针中心点以及指针小目标图像的分割能力。保留原始细节支路改进语义支路后,算法的 mIoU 达到了 85.62%,相较于原始 BiSeNet V2,提升明显,验证了语义支路改进的有效性。所提细节支路及语义支路共同构建的双支路主干

网络的精度有极小幅度的下降,主要原因在于算法在指针类别上的分割精度有所下降。然而,通过重新设计双支路聚合方案,即令特征图卷积后对齐拼接的方式替换原有 BGA 模块,使算法的平均分割精度提升至 88.66%,在仪表刻度线(line)分类上的分割精度比 BGA 方案提升了近 10 个百分点。该结果验证了与 BGA 方案相比,所提 BGCA 双支路融合方案对不同支路语义信息的互补更充分,减少了双支路融合后的特征丢失和受到的干扰。

对分割结果进行可视化,效果如图 5 所示。其中:

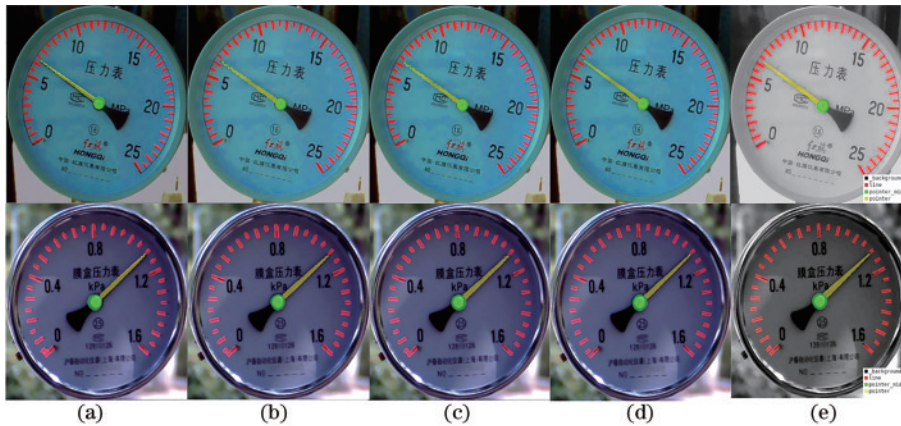


图 5 可视化仪表分割结果。(a) BiSeNet V2;(b)细节支路改进;(c)语义支路改进;(d)聚合层改进;(e)参照组

Fig. 5 Visualization of instrument segmentation results. (a) BiSeNet V2; (b) improved detail branch; (c) improved semantic branch; (d) improved aggregation layer; (e) ground truth

图 5(a)是 BiSeNet V2 算法的仪表图像分割结果；图 5(b)是以 BiSeNet V2 算法作为基线对语义支路进行改进后的仪表图像分割结果；图 5(c)是在图 5(b)算法基础上对细节支路进行改进后的仪表图像分割结果；图 5(d)是在图 5(c)算法基础上用双引导拼接聚合层替换原有双引导聚合层后的仪表图像分割结果；图 5(e)是作为仪表图像分割参照组的人工标注数据。通过可视化结果能够看到,所提 BiUnet 仪表图像分割算法在 line、pointer_mid、pointer 三个类别上的分割精度显著提升。

3.3 算法对比实验

对所提 BiUnet 主干网络与现有基于 Transformer 和纯卷积等主流主干网络在仪表数据集上的表现进行对比。实验方案是:在 BiUnet 算法的基础上仅替换现有主干网络,head 层以及损失函数等参数与 BiUnet 保持一致,以展示 BiUnet 与其他主干网络在仪表数据集 background、line、pointer_mid、pointer 类别上的分割精度、mIoU、mAcc、aAcc、模型参数量以及计算量。

实验结果如表 3 所示,从表 3 中可以发现,ViT^[15]、PoolFormer^[16]、BEiT^[17]、Swin-T^[18]、Twins^[19] 此类基于 Transformer 的主干网络在仪表刻度线和指针这类小目标上的分割表现欠佳,并且其模型参数量较大,网络较为复杂。尽管 ConvNeXt-T^[13] 将纯卷积网络的效果发挥到最大,但缺少不同尺度下的语义特征融合,因此,在指针和刻度线这类小目标上的分割精度 IoU 较低。BiSeNet V2 通过双支路学习特征,尽管在指针和刻度线上的分割精度优于 ConvNeXt-T,但其同样缺少语义支路的特征融合从而使算法在指针和刻度线上的分割精度较低。而所提 BiUnet 主干网络相比于其他主干网络在指针及刻度线上的分割精度高,且模型参数量较小,计算量较低,实现了高精度、轻量化的指针仪表图像分割。

为了更直观地区别不同算法在仪表数据集上的表现,对表 3 中不同算法的实际仪表图像分割结果进行可视化,如图 6 所示。其中:图 6(a)是作为测试标准的人工标注数据;第一行是基于 Transformer 结构主干网

表 3 BiUnet 算法与其他算法在仪表数据集上的表现

Table 3 Performance of BiUnet algorithm and other algorithms on the instrumentation dataset

Backbone	IoU / %				mIoU / %	mAcc / %	aAcc / %	#Param. / 10 ⁶	Flops / 10 ⁹
	Background	Line	Pointer_mid	Pointer					
PoolFormer ^[16]	97.11	16.19	81.63	44.11	59.76	65.60	97.13	12.59	9.82
Twins ^[19]	97.21	15.06	83.14	46.00	60.35	66.39	97.23	25.95	16.44
Swin-T ^[18]	97.20	15.28	83.46	46.77	60.68	66.44	97.22	31.06	25.59
BEiT ^[17]	97.74	29.47	88.72	73.51	72.36	78.05	97.77	23.96	18.46
ViT ^[15]	97.75	30.30	88.85	74.90	72.95	78.87	97.77	23.76	16.65
ConvNeXt-T ^[13]	97.26	17.14	84.68	48.97	62.01	67.91	97.28	34.92	25.11
BiSeNet V2 ^[12]	98.54	58.15	85.91	77.49	80.02	87.60	98.57	3.34	12.31
CGNet ^[20]	98.74	61.25	86.01	81.98	81.99	87.93	98.76	16.64	25.95
Fast-SCNN ^[21]	98.82	64.05	87.72	84.43	83.75	89.49	98.84	1.95	3.19
MobileNetV3 ^[22]	98.88	65.39	87.68	84.81	84.19	89.77	98.90	11.83	42.26
BiUnet (ours)	99.29	78.02	89.67	87.67	88.66	93.72	99.30	4.93	20.82

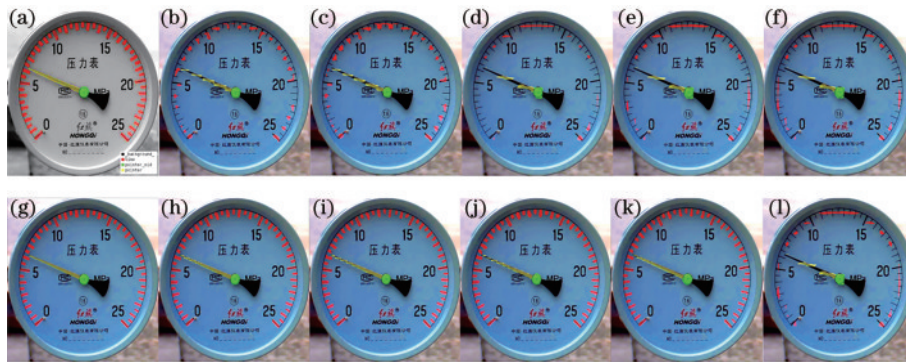


图 6 不同主干网络仪表图像分割结果对比。(a)测试标准;(b) ViT;(c) BEiT;(d) Swin-T;(e) Twins;(f) PoolFormer;(g) BiUnet (ours);(h) MobileNetV3;(i) Fast-SCNN;(j) CGNet;(k) BiSeNet V2;(l) ConvNeXt

Fig. 6 Comparison of the results of different backbone networks in instrument image segmentation. (a) Ground truth; (b) ViT; (c) BEiT; (d) Swin-T; (e) Twins; (f) PoolFormer; (g) BiUnet (ours); (h) MobileNetV3; (i) Fast-SCNN; (j) CGNet; (k) BiSeNet V2; (l) ConvNeXt

网络的分割效果图;第二行是基于纯卷积主干网络的分割效果图(从左至右按照精度降序排列)。可以直观地看到:Transformer 主干网络对刻度线和仪表指针的语义分割较为困难[图 6(b)~(f)];ConvNeXt 主干网络无法完成对仪表图像的分割[图 6(l)];MobileNetV3、Fast-SCNN、CGNet、BiSeNet V2 等主干网络对刻度线的分割大部分呈不规则掩码,与人工标注的矩形掩码差距较大,对指针针头的分割效果较弱,且存在不完全分割等情况,导致下游任务即数值读取存在误差;而所提 BiUnet 主干网络对刻度线的分割掩码均接近“ground truth”中的矩形掩码,相较于其他主干网络,在指针刻度线及针头上的分割效果更优。

4 结 论

由于深度学习比传统视觉算法在指针仪表图像分割上更有优势,针对现有深度学习仪表图像分割存在的局限,利用双支路主干网络来平衡准确性及速率,并差异化地设计了空间细节和深度语义特征的提取支路。以 BiSeNet V2 为出发点,提出 BiUnet 主干网络,用于指针式仪表的图像分割。通过 ConvNeXt 卷积模块替代细节支路中传统的 3×3 卷积模块提升了网络对指针及刻度等空间细节的学习能力,并通过实验验证了该方案可以提高指针式仪表图像分割的整体分割精度。通过 U-net 的编码器-解码器结构改进语义支路,融合不同尺度下的语义信息,提高了网络对小目标的分割能力。实验表明,该方案在刻度线这类小目标上的分割精度有显著提升。同时,提出的双引导拼接聚合层更好地融合了双支路的特征结果,进一步提升了网络对仪表图像的分割效果。与近年常见的图像分割网络进行了对比实验,实验结果表明,所提 BiUnet 主干网络优于 BiSeNet V2 等常见主干网络,证实了 BiUnet 主干网络在仪表图像分割中具有较小的网络模型和较高的精度。同时也为 BiSeNet V2 主干网络的后续研究以及指针式仪表图像分割领域的研究作出了贡献。

参 考 文 献

- [1] Ueda S, Suzuki K, Kanno J, et al. A two-stage deep learning-based approach for automatic reading of analog meters[C]//2020 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS-ISIS), December 5-8, 2020, Hachijo Island, Japan. New York: IEEE Press, 2021.
- [2] Ye X F, Xie D L, Tao S. Automatic value identification of pointer-type pressure gauge based on machine vision[J]. Journal of Computers, 2013, 8(5): 1309-1314.
- [3] Zhang L, Fang B, Zhao X, et al. Pointer-type meter automatic reading from complex environment based on visual saliency[C]//2016 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), July 10-13, 2016, Jeju, Korea (South). New York: IEEE Press, 2016: 264-269.
- [4] Chi J N, Liu L, Liu J W, et al. Machine vision based automatic detection method of indicating values of a pointer gauge[J]. Mathematical Problems in Engineering, 2015, 2015: 1-19.
- [5] Zhang X F, Sun Y J, Liu H, et al. Improved clustering algorithms for image segmentation based on non-local information and back projection[J]. Information Sciences, 2021, 550: 129-144.
- [6] Cui Y F, Sun Y J, Jian M W, et al. A novel underwater image restoration method based on decomposition network and physical imaging model[J]. International Journal of Intelligent Systems, 2022, 37(9): 5672-5690.
- [7] Peng J S, Xu M Y, Yan Y F. Automatic recognition of pointer meter reading based on Yolov4 and improved U-net algorithm[C]//2021 IEEE International Conference on Electronic Technology, Communication and Information (ICETCI), August 27-29, 2021, Changchun, China. New York: IEEE Press, 2021: 52-57.
- [8] 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.
- [9] Zheng Y, Chen M L, Peng J S, et al. An automatic reading method for pointer meter based on one-stage detector[J]. Proceedings of SPIE, 2020, 1651: 165104.
- [10] 万吉林, 王慧芳, 管敏渊, 等. 基于 Faster R-CNN 和 U-net 的变电站指针式仪表读数自动识别方法[J]. 电网技术, 2020, 44(8): 3097-3105.
- [11] Wan J L, Wang H F, Guan M Y, et al. An automatic identification for reading of substation pointer-type meters using faster R-CNN and U-net[J]. Power System Technology, 2020, 44(8): 3097-3105.
- [12] Fang Y X, Dai Y, He G L, et al. A mask RCNN based automatic reading method for pointer meter[C]//2019 Chinese Control Conference (CCC), July 27-30, 2019, Guangzhou, China. New York: IEEE Press, 2019: 8466-8471.
- [13] Yu C Q, Gao C X, Wang J B, et al. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129(11): 3051-3068.
- [14] Liu Z, Mao H Z, Wu C Y, et al. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 11966-11976.
- [15] Lavin A, Gray S. Fast algorithms for convolutional neural networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 4013-4021.
- [16] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: transformers for image recognition at scale[C]//International Conference on Learning

- Representations, May 3-7, 2021, Virtual, Online. La Jolla: ICLR Press, 2021: 1-22.
- [16] Yu W H, Luo M, Zhou P, et al. MetaFormer is actually what you need for vision[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 10809-10819.
- [17] Bao H B, Dong L, Piao S H, et al. BEiT: BERT pre-training of image transformers[C]//10th International Conference on Learning Representations, April 25-29, 2022, Virtual, Online. La Jolla: ICLR Press: 2022: 1-18.
- [18] Liu Z, Lin Y T, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 9992-10002.
- [19] Chu X, Tian Z, Wang Y, et al. Twins: revisiting the design of spatial attention in vision transformers[C]//Advances in Neural Information Processing Systems, December 6-14, 2021, Virtual, Online. Canada: NIPS Press, 2021: 9355-9366.
- [20] Wu T Y, Tang S, Zhang R, et al. CGNet: a light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [21] Poudel R K, Liwicki S, Cipolla R. Fast-SCNN: fast semantic segmentation network[C]//30th British Machine Vision Conference, September 9-12, 2019, Cardiff, United Kingdom. London: BMVA Press, 2019: 1-9.
- [22] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27 - November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 1314-1324.