Multi-scale dense and attention mechanism for image semantic segmentation based on improved DeepLabv3+

一.成果完成人及所在单位： 王佐帅，光电与通信工程学院 指导教师：张宏怡

二.成果简介：

语义场景分割已经成为计算机视觉中的一个重要应用，是智能交通系统中对周围环境进行完整场景理解的一个重要部分。一些基于卷积神经网络的方法已经出现，但也存在一些问题，包括小范围的目标损失、不准确的细节区域分割、边界类别混淆等。在本文中，我们基于金字塔理论，通过不同的浅层特征来利用全局背景信息的能力。构建了一个加权金字塔特征融合模块来融合骨干网络生成的不同尺度的特征图，并通过可训练的参数动态更新特征融合的比例。之后，引入注意力机制，以发现空间通道相互依赖的信息。最后，DeepLabv3+网络的 ASPP 模块通过空洞卷积不同空洞率之间进行密集连接从而进一步改进。实验结果表明，与 DeepLabv3+相比，所提出的方法有 4.1%的 mPA 和 3.92%的 mIOU 改进，而且语义分割的结果更加准确。

三. 主要研究内容及创新点：

(1）图像语义分割通常被认为是一个逐个像素的分类任务，因此不同尺度的浅层特征，如边缘和纹理，所包含的信息对网络最终输出的分割结果起着重要作用。但是，很多网络没有注意到这一点，或者没有充分利用骨干网络提取的不同尺度的浅层特征，这些浅层特征中包含的信息被浪费了，这就导致网络模型在区分前景信息较弱的像素或者属于空间结构较小的物体的像素时，分类效果往往不理想。

(2)与传统的卷积算子相比，空洞卷积可以实现更大的感受野大小。为此，DeepLabv3+提出了 ASPP，使输出特征图中的神经元包含多个不同大小的感受野。然而，ASPP 仍然受到另一个限制。具体来说，如果输入图像的分辨率很高，这就要求神经元有更大的感受野。然而，随着扩张率的增加，空洞卷积变得越来越无效，并逐渐失去其能力。

受金字塔理论的启发，本文提出了一个密集金字塔特征融合模块。除了传统的自上而下的路径和横向连接策略外，加权金字塔特征融合（WPFFM）模块引入了自下而上的路径，并优化了 WPFFM 的连接性，从而形成了一个简化而密集的双向连接模块，可以更充分地利用三个不同尺度的特征图。同时，参考 ASFF，我们在 WPFFM 进行特征融合的同时，为每个特征图引入了一个可学习的参数，其范围在（0，1）之间，可以参与网络的正常训练，最终作为不同特征图融合的权重。我们重新考虑了 ASPP 模块的优势。也就是说，它能使输出特征图中的神经元包含不同大小的多个感受野。为了解决从不同分支获得的感受野的稀疏性问题，我们提出了 Multi76 Scale ASPP（M-ASPP），它以密集的方式连接一组腹卷积层，这样它产生的多尺度特征不仅覆盖较大的尺度范围，而且密集地覆盖该范围。实验结果表明，与 DeepLabv3+相比，所提出的方法有 4.1%的 mPA 和 3.92%的 mIOU 改进，而且语义分割的结果更加准确。

最终提出方法与 DeepLabv3+网络在 cityscapes 数据集上的可视化对比图 1 所示。

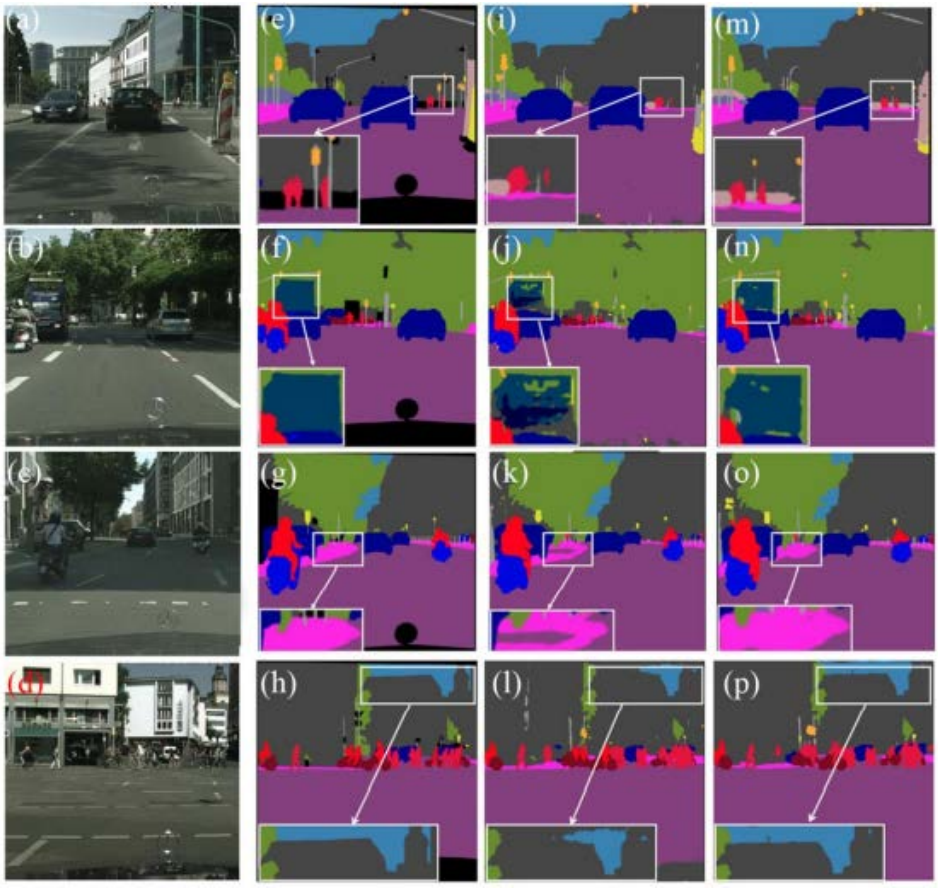


图 1 Cityscapes 数据集可视化结果. (a-d) 原始图像，(e-h) 标记的图像，(i-l) DeepLabv3+，(m-p) 提出的算法

四．主要研究成果及应用情况

在 Journal of Electronic Imaging 期刊上发表论文 1 篇。(Zuoshuai Wang, Hongyi Zhang, Zhiquan Huang, Zhibin Lin, and Hangxing Wu "Multi-scale dense and attention mechanism for image semantic segmentation based on improved DeepLabv3+," Journal of Electronic Imaging 31(5), 053006 (8 September 2022)).

联系人：王佐帅，　　电　话：18796617197　　Email: 841473215@qq.com