摘要

图像分割是计算机视觉中的一个基本任务，其目的是将图像划分成若干区域。图像语义分割根据所属于的物体类别划分区域，实例分割则更加进一步地精细划分以区别不同物体。语义分割和实例分割算法被广泛的应用到手机 APP、视频监控、自动驾驶、医疗图像、机器人感知、三维重建等场景中。

当前，语义分割以及实例分割利用深度网络对图像进行像素级别的理解。尽管深度学习在语义分割和实例分割取得了显著的进展，但依旧存在着许多问题。本文主要对语义分割中存在的像素特征之间的非关联性和多尺度问题、实例分割中的多任务问题进行研究，并提出以下创新模型：

1）提出了一种基于上下文感知的图像语义分割模型

建立像素特征之间的关联对于获得整体的分割结果至关重要，针对当前全卷积网络中相邻像素特征无关联的现象，本文提出了 Relation Module，建模空间中相邻像素特征之间的关系。与此同时，为了有效地区分以及约束相邻像素特征的距离，提出了空间约束损失函数对相邻像素特征进行优化。实验表明，本文所提出的网络结构 RelationNet 相对于基准网络在 Cityscapes 有 1.5% mIoU 的提升。

2）提出了一种新型多尺度特征金字塔的图像语义分割模型

多尺度特征表达对于不同尺度物体的语义分割至关重要，为此，本文提出了 Dense Relation Module，该模块采取跳跃循环神经网络构建特征金字塔，有效增大了网络的感受野，遏制了尺度大的物体欠分割的现象。我们所提出来的 Dense Relation Network 在数据集 Cityscapes 上达到 82.8% mIoU、Pascal Context 数据集上达到 49.0% mIoU 的结果。

3）提出了一种针对多任务的实例分割模型

本文在实例分割算法 Mask RCNN 的基础上，提出了多尺度感兴趣特征提取以及边缘优化损失函数。探讨出实例分割中的分类、回归、分割子网络，设计出一种适应于实例分割的头部设计。实验表明，其能够在数据集 COCO 实例分割上达到 1.7% 以及在 COCO 目标检测上达到 1.2% 的性能增益。

关键词：图像分割，语义分割，实例分割，深度学习

Study on Deep Learning Based Image Semantic Segmentation and Instance Segmentation

Yueqing Zhuang (Computer Applied Technology)

Directed by Prof. Xiaodong Xie

ABSTRACT

Image segmentation is one of basic tasks in computer vision, which aims to divide image into lots of similar regions. Image semantic segmentation divides image into regions according to the object class which they belong to, while instance segmentation further finely divides image to distinguish different objects. Semantic segmentation and instance segmentation have a wide range of applied scenarios, such as mobile application, video surveillance, autopilot, medical image process, robot perception, 3D reconstruction and so on.

Recently, semantic segmentation and instance segmentation use deep learning to have a pixel-level understanding for image. Obviously, deep learning has achieved significant performance on semantic segmentation and instance segmentation. However, there are still many problems that need to be studied and solved. In this paper, we study the problem of the non-associativity between pixels' feature and multi-scale problems and propose the innovative network as following:

1) A context-based network for image semantic segmentation is proposed.

It is critical to establish the association of pixels' feature representations to achieve consistent segmentation results. There is a problem that the adjacent pixels' feature representations of the current image semantic segmentation network are not associated. In this paper, we propose Relation Module to model the relationship between the adjacent pixels. At the same time, in order to distinguish the features of adjacent different types of pixel more effectively, a spatial correlation loss is proposed to retrict the feature representation. The proposed RelationNet has improved 1.5% mIoU in Cityscapes.

2) An new multi-scale pyramid network for image semantic segmentation is designed.

The multi-scale feature representation is very important for the semantic segmentation on objects of different scales. In this paper, we proposes Dense Relation Module, which uses the skipping recurrent neural network to construct feature pyramid, which effectively increases the network's receptive field and prevent under-segmentation for large objects . The proposed

Dense Relation Network achieves the remarkable results of 49.0% mIoU, and 82.8% mIoU on Pascal Context, Cityscapes dataset respectively.

3) An instance segmentation network for multitasking is presented.

Based on the instance segmentation algorithm Mask RCNN, multi-scale region of interest feature extraction and edge loss function are proposed. The correlation of classification, regression and segmentation in the instance segmentation is discussed in detail, and a head structure adapted to the instance segmentation is designed. Experiments show that the proposed instance segmentation network can obtain the performance gain of 1.7% mmAP on the COCO instance segmentation and the performance gain of 1.2% mmAP on the COCO detection dataset.

KEY WORDS: Image Segmentation, Semantic Segmentation, Instance Segmentation, Deep Learning