Deep Learning Assignment4

RE6111024 葉嘉宏

Abstract—This assignment is to use a model that can learn object detection and semantic segmentation capabilities from two different datasets respectively, then produce the two outputs. But in this assignment, we only do the part of object detection. The code can be found at https://github.com/chyeh1126/ DeepLearning-2023/tree/main/HW4.

Index Terms—Deep Learning, Neural Network, Object Detection, Semantic Segmentation

I. Introduction

In the field of computer vision, common tasks include image classification, object detection, and so on. The goal of object detection is to identify and localize objects of interest within an image or a video and the goal of semantic segmentation is to assign a class label to each pixel in an image.

II. Method

Single Shot Multibox Detector (SSD) [1] is a object detection algorithm that is widely used for real-time object dectection tasks. The key idea behind SSD is to combine the efficiency of one-shot detectors with the accuracy of multi-stage detectors.
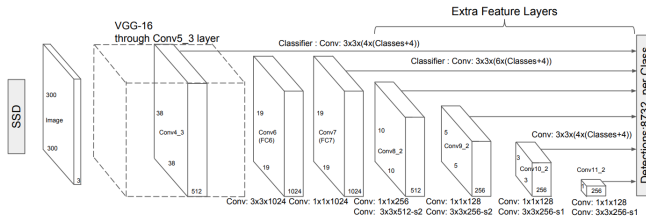

Fig. 1. The architecture of SSD

The architecture of SSD is shown in Fig 1. SSD use pre-trained VGG16 model on ImageNet as a base model and truncate the last classification layer to extract image features. Then, SSD adds extra convolution layers on top of the base model to create a feature pyramid. It allows the model to capture features at multiple scales and different levels of abstraction.

III. Experiment

A. Dataset

VOC2007 is a benchmark dataset for object detection, classification, and semantic segmentation. It contains 9963 images and 20 categories. In this assignment, we use this dataset on object detection task.

ADE20K contains 2000 images and 150 categories. This dataset is used for semantic segmentation task. Since we fail to build the part of semantic segmentation, this dataset has been not utilized.

B. Experiment Result

Table 1. shows the performance of SSD on testing set.

TABLE I
SSD performance on testing set

Model	Metric
	Mean AP
SSD	0.74

IV. Conclusion

In this assignment, we only do the part of object detection and fail to combine the part of semantic segmentation. Modifying the source code directly is challenging, maybe we should try other approaches.

References

References

[1] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. (2016). Ssd: Single shot multibox detector. In Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11—14, 2016, Proceedings, Part I 14 (pp. 21-37). Springer International Publishing.