論文の要旨

題　　　目　Enhanced Object Detection and Instance Segmentation Through Advanced Prior Information Integration in Deeplearning
（ディープラーニングにおける高度な事前情報統合による物体検出とインスタンスセグメンテーションの強化）

氏　　名 内之浦　伸治

1. Improved head and data augmentation to reduce artifacts at grid boundaries in object detection

We investigated the influence of horizontal shifts of the input images for one stage object detection method. We found that the object detector class scores drop when the target object center is at the grid boundary.Many approaches have focused on reducing the aliasing effect of down-sampling to achieve shift-invariance.

However, down-sampling does not completely solve this problem at the grid boundary; it is necessary to suppress the dispersion of features in pixels close to the grid boundary into adjacent grid cells.

Therefore, this paper proposes two approaches focused on the grid boundary to improve this weak point of current object detection methods.One is the Sub-Grid Feature Extraction Module, in which the sub-grid features are added to the input of the classification head.

The other is Grid-Aware Data Augmentation, where augmented data are generated by the grid-level shifts and are used in training.The effectiveness of the proposed approaches is demonstrated using the COCO validation set after applying the proposed method to the FCOS architecture.

2. An Object Detection Method Using Probability Maps for Instance Segmentation to Mask Background

This paper proposes a two-step detector called Segmented Object Detection (SODet), whose performance is improved by masking the background region. Previous single-stage object detection methods suffer from the problem of imbalance between foreground and background classes, where the background occupies more regions in the image than the foreground.

Thus, the loss from the background is firmly incorporated into the training. RetinaNet addresses this problem with Focal Loss, which focuses on foreground loss.

Therefore, we propose a method that generates probability maps using instance segmentation in the first step and feeds back the generated maps as background masks in the second step as prior knowledge to reduce the influence of the background and enhance foreground training. We confirm that the detector can improve the accuracy by adding instance segmentation information to both the input and output rather than only to the output results. On the CityScapes dataset, our method outperforms the state-of-the-art methods.

3. Graph Laplacian Regularization based on the Differences of Neighboring Pixels for Conditional Convolutions for Instance Segmentation

We propose a simple and effective regularization method for instance segmentation, GLRDN-L2 (Graph Laplacian Regularization based on Differences of Neighboring Pixels).
Instance segmentation is a challenging task in computer vision.
For many years, ROI-based methods such as Mask R-CNN have dominantly presented the top performances; however, the recently proposed CondInst, which employs dynamic FCNs as a mask head and performs instance-aware mask prediction, outperforms Mask R-CNN.
To our best knowledge, all methods optimize a model based on pixel-wise losses such as Dice Loss.

Even with the results of high-resolution masks, there are problems such as blurred boundaries and hollows in the instances.
We assume that these problems are due to the spatial structure and contextual information contained in the relationships between neighboring pixels not being incorporated well in the model. To address these problems, we propose a regularization that penalizes the errors in the spatial structure with a graph composed of the differences between neighboring pixels.