Deep Occlusion-Aware Instance Segmentation with Overlapping BiLayer

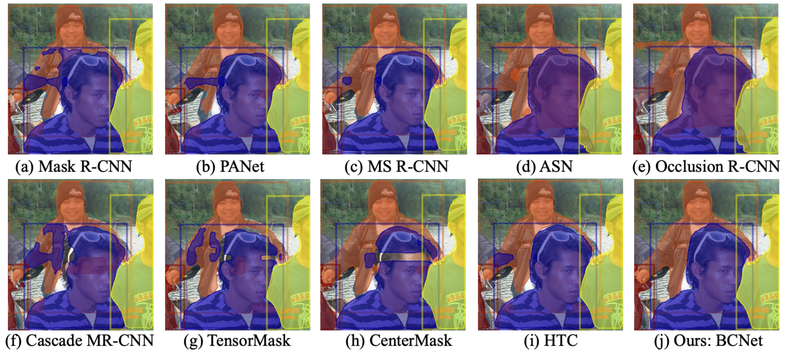
Lei Ke[1], Yu-Wing Tai[2], Chi-Keung Tang[1]
[1] HKUST, [2] Kuaishou Technology
https://github.com/lkeab/BCNet

Highlight

- **BCNet:** Two/one-stage (detect-then-segment) instance segmentation with state-of-the-art performance on three large-scale benchmarks (COCO, KINS, COCOA).
- **Novelty:** Explicit occlusion modeling in RoI with **bilayer decoupling** for the occluder and occludee.
- **Efficacy:** Large improvement using the FCOS (anchor-free) and Faster R-CNN (anchor-based) detectors.
- **Simple**: Small additional computation burden and easy to use.

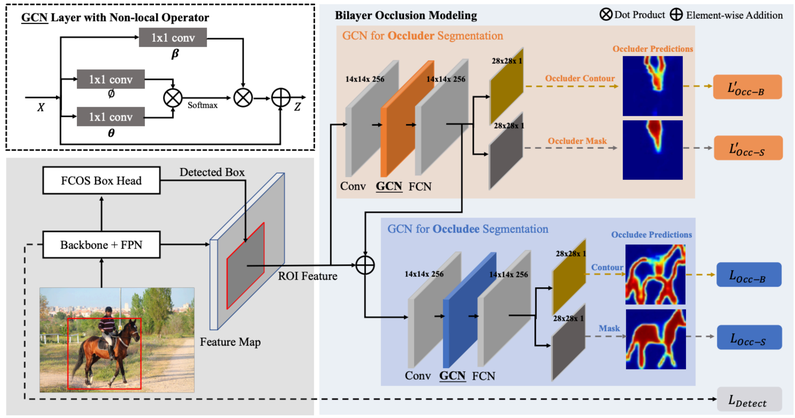
Introduction

- **Motivation**: **1)** A lot of segmentation errors are caused by overlapping objects, especially for object instances belonging to the same class. **2)** Most existing improvements come from better backbone designs, with little attention paid in the instance mask regression after RoI extraction.
- **Solution**: We propose Bilayer Convolutional Network (**BCNet**) for object contour and mask regression, which explicitly decouples overlapping objects in the same RoI into **two separate image layers** - top layer handles the **occluders** and bottom layer for the **occludee**.



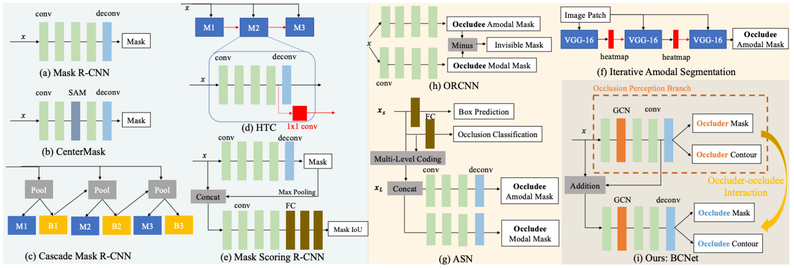
BCNet Framework

Bilayer GCN structure: For cropped ROI feature, the first GCN explicitly models occluders by simultaneously detecting occlusion contours and masks which distills essential shape and position information to guide the second GCN in mask prediction for the occludee (target object).



Comparison of Mask Head Design

The previous occlusion-aware mask heads only regress both modal and amodal masks from the occludee **w/o occluder-occludee interaction**.



Experimental Results

- Ablation experiments on bilayer structure

Table 1. Effect of the first GCN for occlusion modeling by predicting contours and masks on COCO with ResNet-50-FPN model.

Occlusion (Occluder) Modeling		COCO-OCC		COCO	
Contour	Mask	AP	AP_{50}	AP	AP_{50}
		29.04	49.22	32.65	52.39
	✓	29.65	49.42	33.25	52.82
✓		30.18	49.94	33.41	53.02
✓	✓	**30.37**	**50.40**	**33.43**	**53.12**

Table 2. Effect of the second GCN for detecting occludee contours for final mask prediction ***guided*** by the output of first GCN.

Target (Occludee) Modeling			COCO-OCC		COCO	
Guidance	Contour	Mask	AP	AP_{50}	AP	AP_{50}
		✓	29.45	49.73	32.56	52.21
✓		✓	30.37	50.40	33.43	53.12
✓	✓	✓	**30.68**	**50.62**	**33.62**	**53.26**

- Comparison with SOTA methods on COCO test-dev set, KINS, COCOA.

Method	Backbone	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Mask R-CNN [21]	ResNet-50	35.6	57.6	38.1	18.7	38.3	46.6
PANet [42]	ResNet-50	36.6	58.0	39.3	16.3	38.1	**52.4**
BCNet + Faster R-CNN [48]	ResNet-50	**38.4**	**59.6**	**41.5**	**21.9**	**40.9**	49.3
Mask R-CNN [21]	ResNet-101	37.0	59.2	39.5	17.1	39.3	52.9
MaskLab [8]	ResNet-101	37.3	59.8	39.6	19.1	40.5	50.6
Mask Scoring R-CNN [25]	ResNet-101	38.3	58.8	41.5	17.8	40.4	54.4
BMask R-CNN [13]	ResNet-101	37.7	59.3	40.6	16.8	39.9	**54.6**
HTC [7]	ResNet-101	39.7	**61.8**	43.1	21.0	42.2	53.5
BCNet + Faster R-CNN [48]	ResNet-101	**39.8**	61.5	**43.1**	**22.7**	**42.4**	51.1
YOLACT [4]	ResNet-101	31.2	50.6	32.8	12.1	33.3	47.1
TensorMask [9]	ResNet-101	37.1	59.3	39.4	17.4	39.1	51.6
ShapeMask [31]	ResNet-101	37.4	58.1	40.0	16.1	40.1	53.8
CenterMask [33]	ResNet-101	38.3	-	-	17.7	40.8	**54.5**
BlendMask [6]	ResNet-101	38.4	60.7	41.3	18.2	41.5	53.3
BCNet + FCOS [51]	ResNet-101	**39.6**	**61.2**	**42.7**	**22.3**	**42.3**	51.0

Coco-test-dev

Model	AP_{Det}	AP_{Seg}
Mask R-CNN [16]	26.97	24.93
Mask R-CNN + ASN [46]	27.86	25.62
PANet [42]	27.39	25.99
PANet + ASN [46]	28.41	26.81
BCNet	**28.87**	**27.30**

KINS

Model	AP_{all}	AP_t	AP_s
AmodalMask [65]	5.7	5.9	0.8
AmodalMRCNN [16]	21.51	21.09	9.0
ORCNN [16]	20.32	20.63	7.8
BCNet	**23.09**	**22.72**	**9.53**

COCOA

- Visual comparison. Heatmaps denote contour and mask predictions respectively.