

SOLO: Segmenting Objects by Locations

Appendix

Xinlong Wang¹, Tao Kong², Chunhua Shen¹, Yuning Jiang², and Lei Li²

¹ The University of Adelaide, Australia

² ByteDance AI Lab

A More Method Details

A.1 Multi-level Prediction

We use five FPN pyramids to segment objects of different scales (Table 1). Scales of ground-truth masks are explicitly used to assign them to the levels of the pyramid. Multi-level prediction gives 6.8 AP gains on the single-scale SOLO.

Table 1. we use five **FPN pyramids** to segment objects of different scales. The grid number increases for smaller instances due to larger existence space.

pyramid	P2	P3	P4	P5	P6
re-scaled stride	8	8	16	32	32
grid number	40	36	24	16	12
instance scale	<96 48~192 96~384 192~768 ≥384				

B More Experimental Results

B.1 Single-scale 1× Training

We list the 1x single-scale results in Table 2.

Table 2. Results with single-scale training. The models are trained with “1×” schedule and evaluated on val2017.

	AP	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L
SOLO	32.9	53.2	34.8	12.7	36.2	50.5
D-SOLO	33.9	54.0	35.7	13.8	36.9	51.0

B.2 Dice Loss

To make a fair comparison, we show the results of Mask R-CNN with Dice loss in Table 3. It shows that Dice loss is not suitable for Mask R-CNN, as it performs worse (-0.9AP) than original BCE loss. It is because the 'detect-then-segment' methods do not have the fg/bg imbalance issue, as they segment the foreground pixels in a local bounding box.

Table 3. Mask R-CNN with Dice loss. The models are trained with “3×” schedule and evaluated on `val2017`.

	AP	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L
BCE	36.2	58.0	38.9	20.1	39.5	49.0
DL	35.3	57.8	37.4	19.3	39.1	47.8

B.3 SOLO for Instance Contour Detection

Our framework can easily be extended to instance contour detection. We first convert the ground-truth masks in MS COCO into instance contours using OpenCV’s `findContours` function, and then use the binary contours to optimize the mask branch in parallel with the semantic category branch. Here we use Focal Loss to optimize the contour detection, other settings are the same with the instance segmentation baseline. Figure 1 shows some contour detection examples generated by our model. We provide these results as a proof of concept that SOLO can be used in contour detection.



Fig. 1. Visualization of SOLO for **instance contour detection**. The model is trained on COCO `train2017` dataset with ResNet-50-FPN. Each instance contour is shown in a different color.

B.4 Qualitative Results

We show more visualization results in Figure 2.

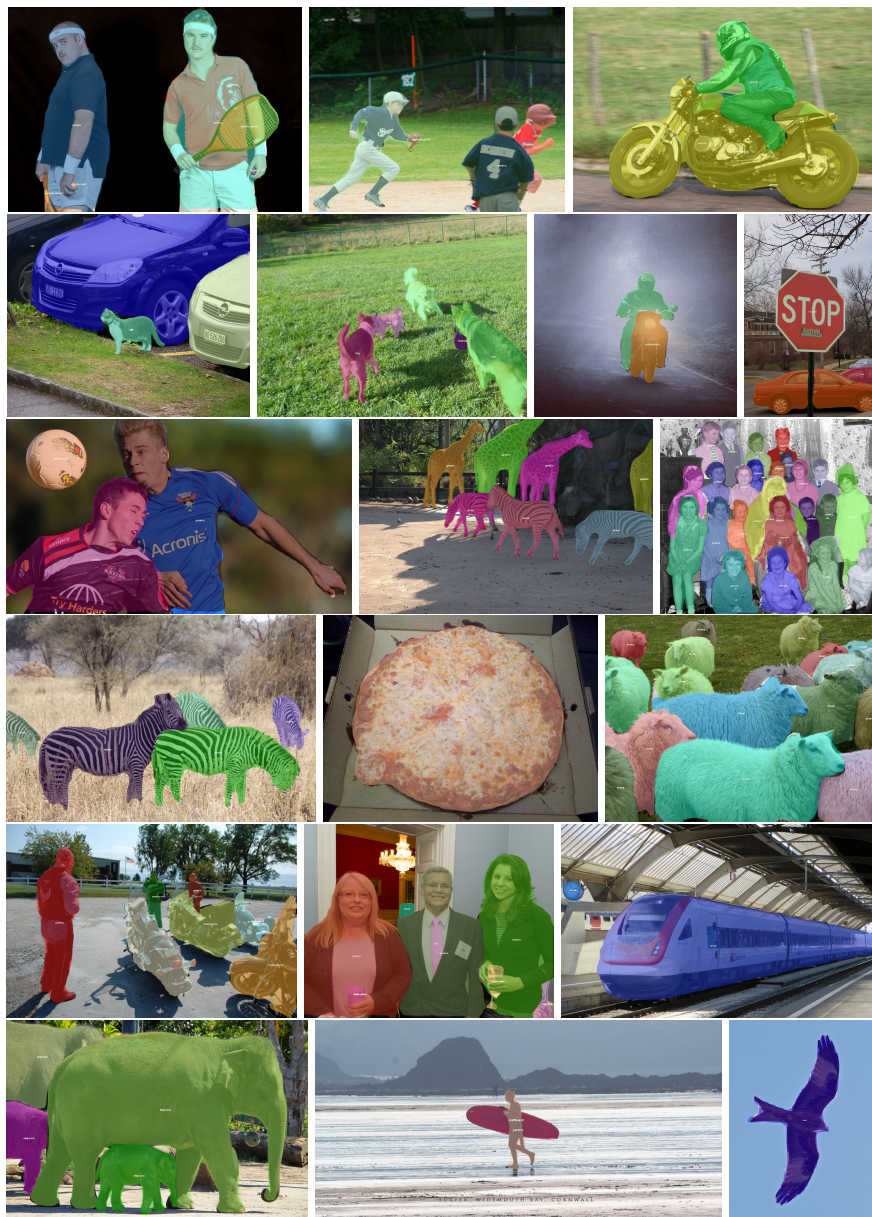


Fig. 2. Visualization of instance segmentation results using the Res-101-FPN backbone. The model is trained on the COCO train2017 dataset, achieving a mask AP of 37.8 on the COCO test-dev.