

Coarse-to-fine Image Co-segmentation with Intra and Inter Rank Constraints

-supplementary materials-

Submission ID: 2627

1 Qualitative results on iCoseg dataset

In this document, we provide more qualitative results (Fig. 1 - Fig. 16) on iCoseg dataset. The first column shows the original images, the second and third columns are the proposed coarse background (marked with cyan) and foreground (marked with cyan), and the last column are the segmentation results. From these results, we can see that our method can generate good performance in most cases.



Figure 1: Results on ‘Alaskan bear’.

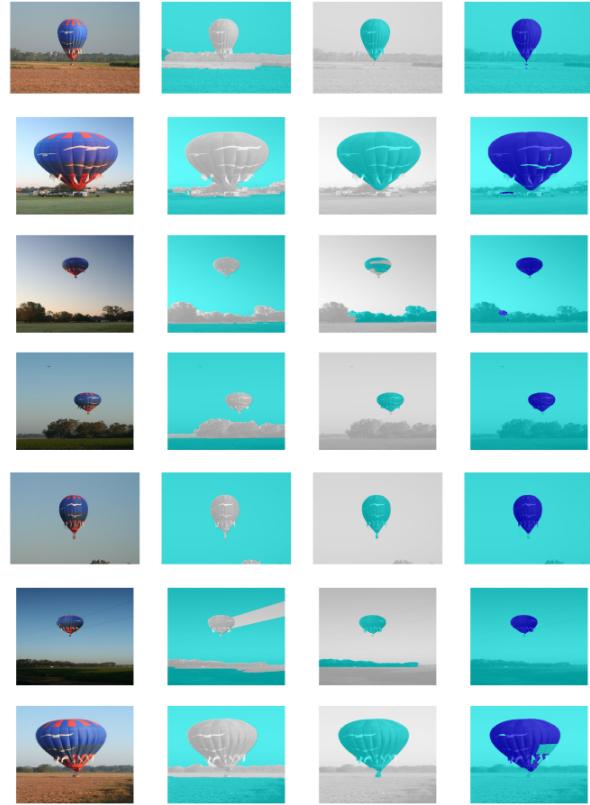


Figure 2: Results on ‘Ballon’.

2 Failure cases

We also plot some failure cases of our algorithm, and the results are shown in Fig. 17. There are several possible reasons for the failure of our algorithm. First, when the object itself has significant visual changes, e.g., the wheels and body of the car are significantly different, our used HSV feature may not be robust enough to segment them into foreground. Second, bad foreground and background proposals (e.g., rows 3 and 4) will also decrease the performance of our method. Thirdly, complex background (e.g., row 2) may introduce some noise.

Table 1: Accuracy of our image co-segmentation algorithms using different features

Dataset	HSV	LAB	Texture	SIFT	Deep Feature
iCoseg	94.1	90.4	88.3	82.6	81.5
MSRC	86.2	84.6	78.8	70.4	73.5

2.1 Effect of the features

In this subexperiment, we study the effects of different features, and show the results in Table 1. In both datasets, HSV achieved the best average performance compared with other four features. In general, color features (i.e., HSV and LAB) are better than Texture features and SIFT. However, different features have different advantages, and some of the features perform better than HSV for some class, e.g., LAB achieves the accuracy of 91.8% for car(front) compared to 88.7% of HSV. The unsatisfactory performance for deep features is probably due to that the original size of the feature maps is very small. Even after the resizing operation, the feature maps are still very coarse. Two different superpixels might be covered by one response point in the feature map. Therefore, they might not be well separated.

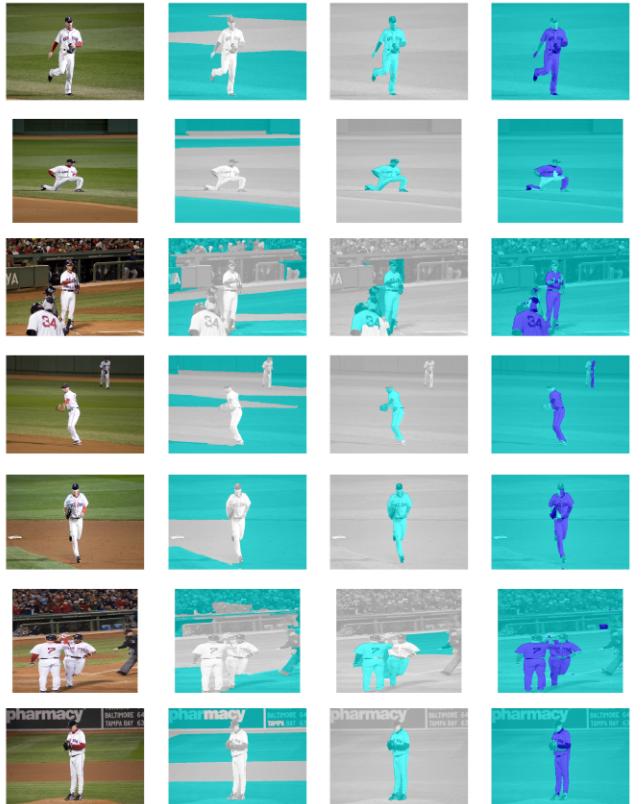


Figure 3: Results on ‘Baseball’.

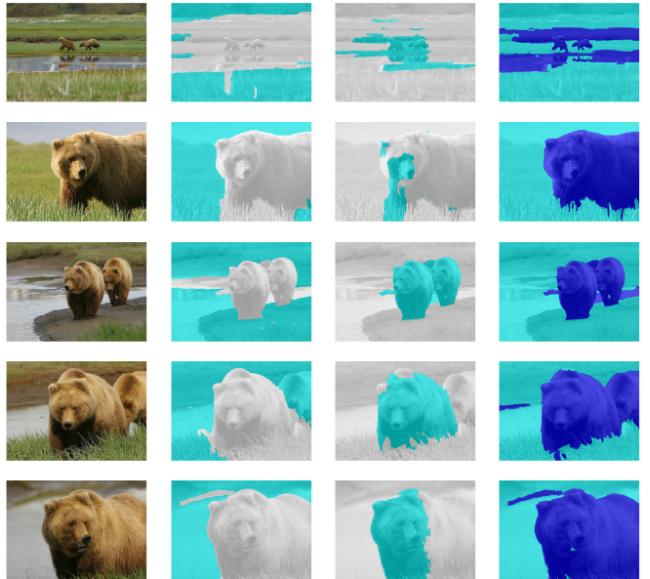


Figure 4: Results on ‘Bear’.

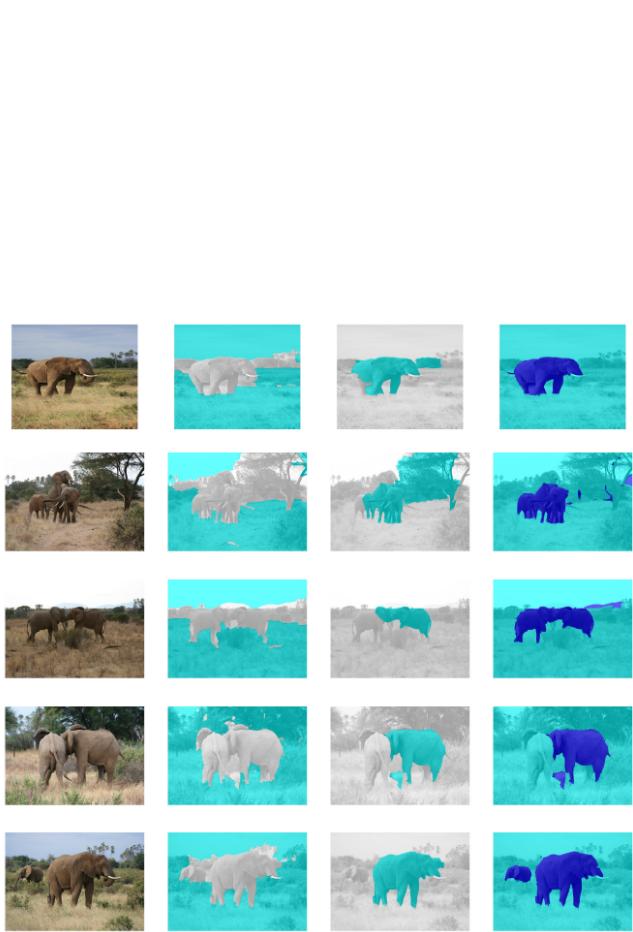


Figure 5: Results on ‘Elephant’.

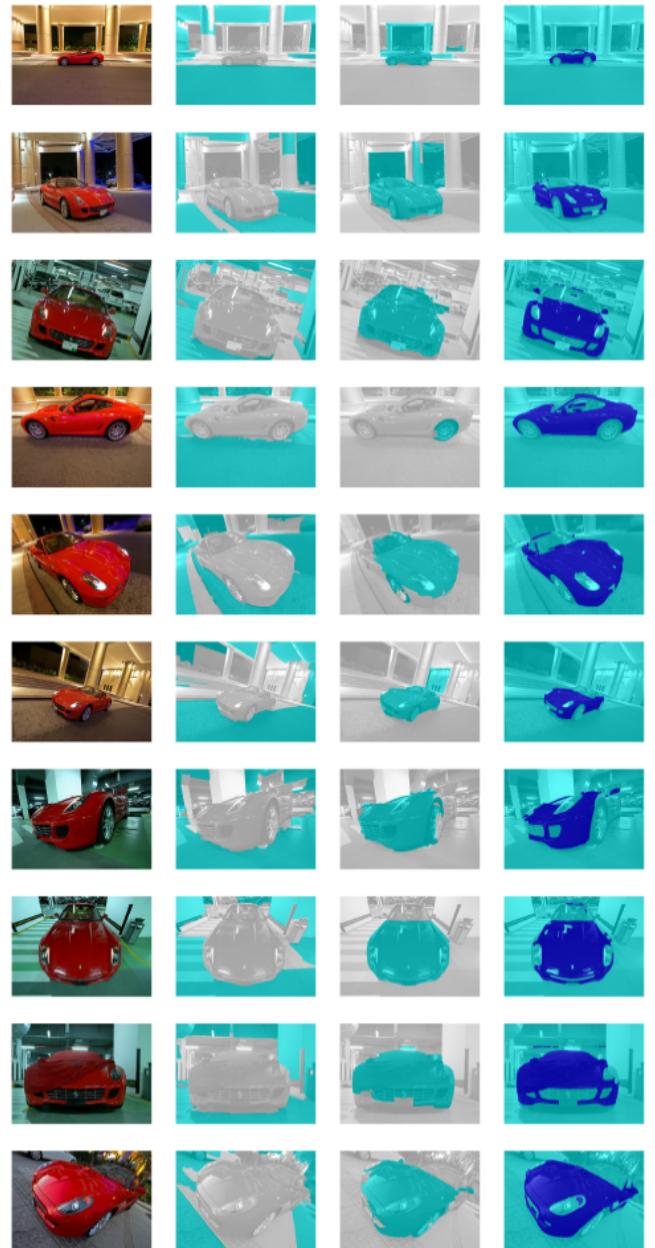


Figure 6: Results on ‘Ferrari’.

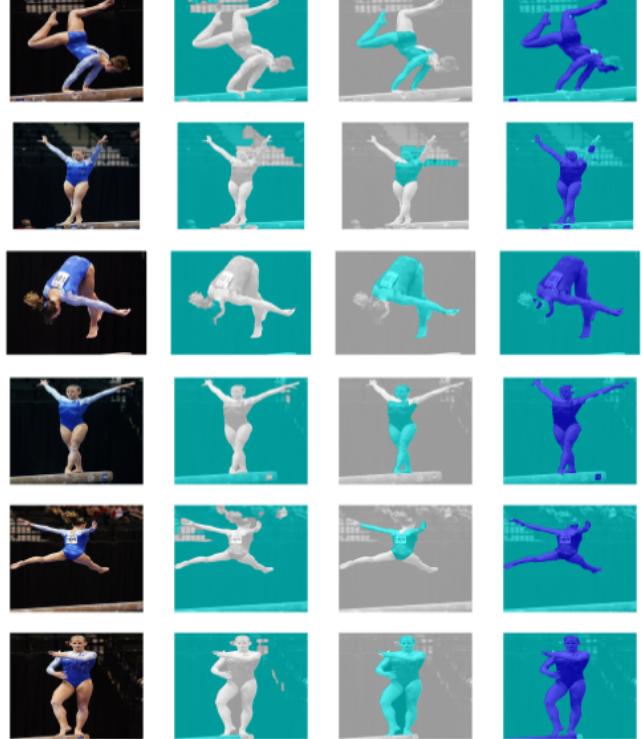


Figure 7: Results on ‘Gymnastics’.



Figure 8: Results on ‘Kite’.

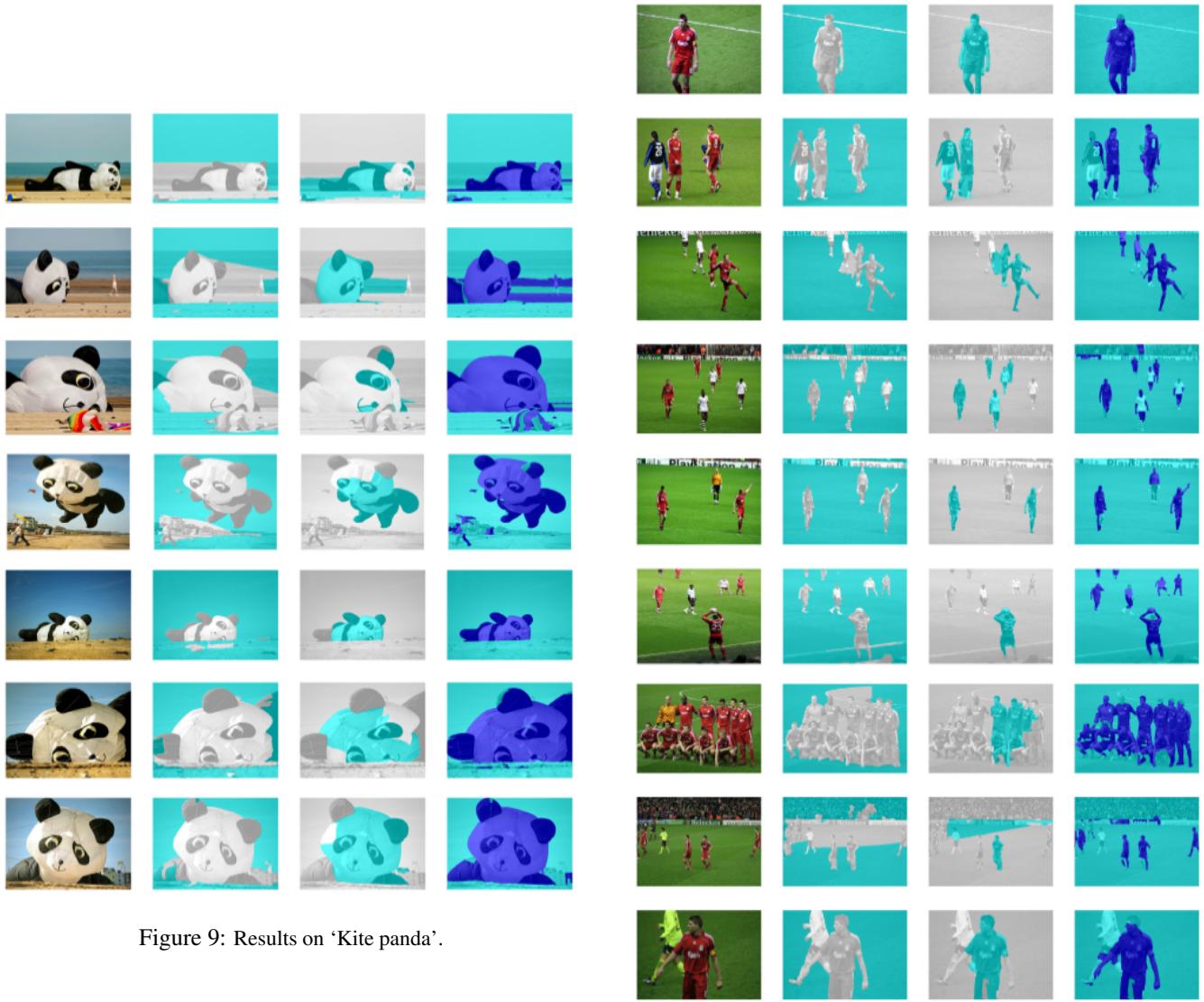


Figure 9: Results on ‘Kite panda’.

Figure 10: Results on ‘Liverpool’.

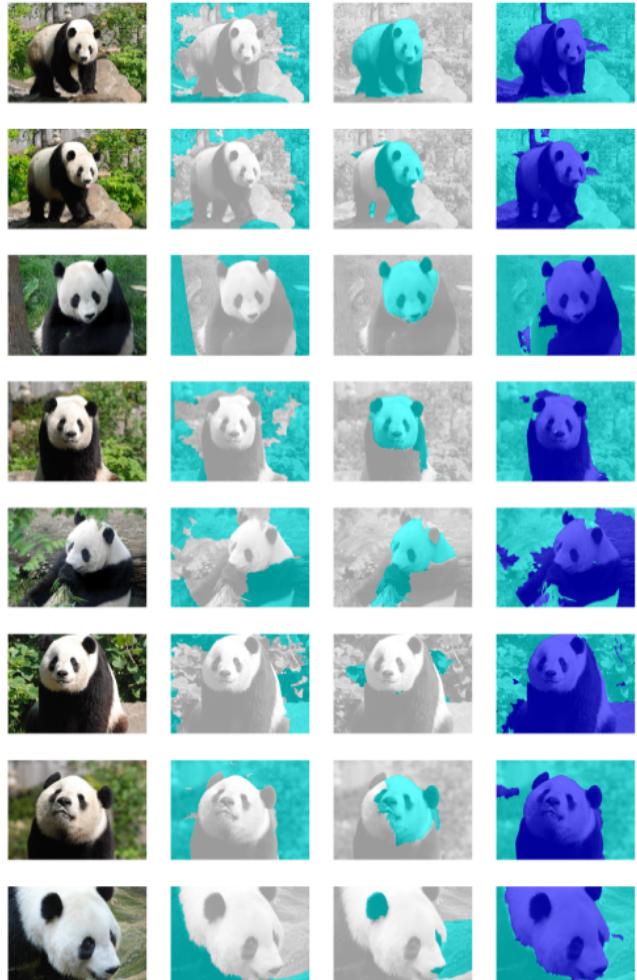


Figure 11: Results on ‘Panda’.



Figure 12: Results on ‘Skating’.

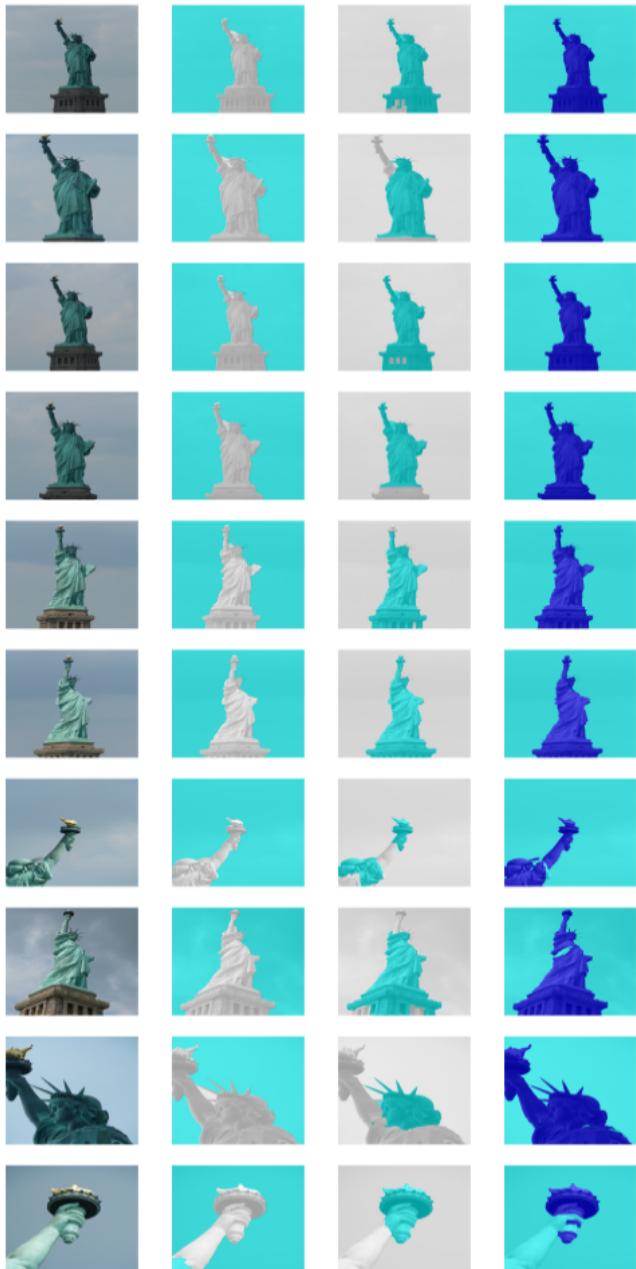


Figure 13: Results on ‘Statue’.



Figure 14: Results on ‘Stonehenge’.

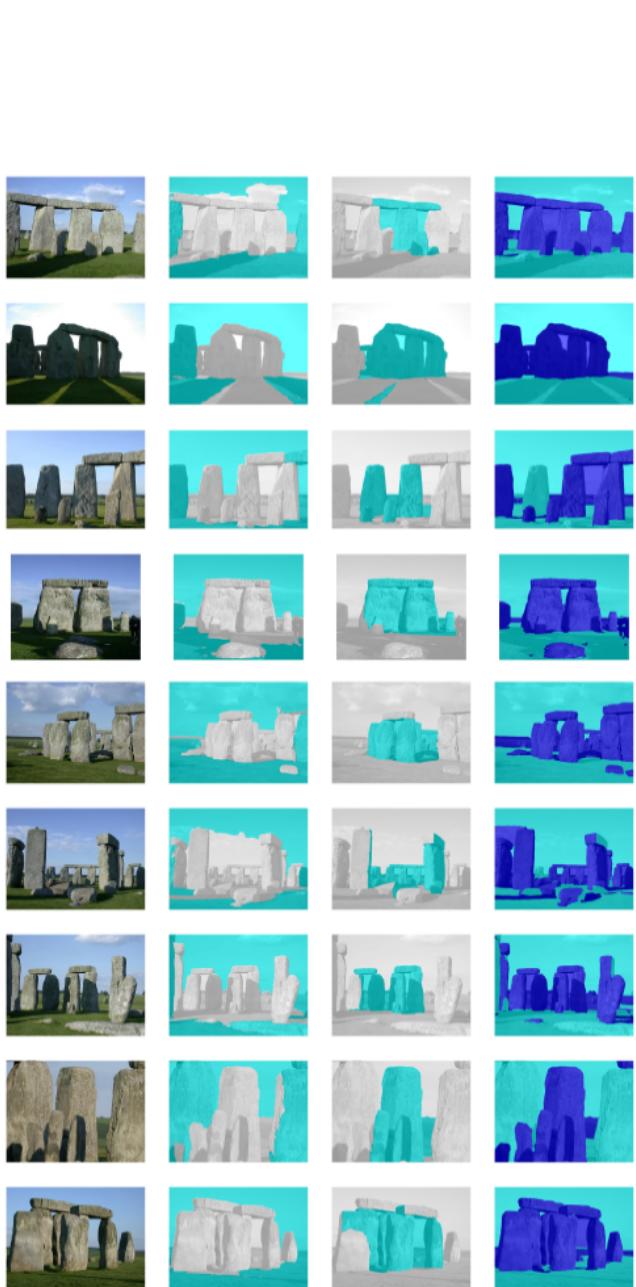


Figure 15: Results on ‘Stonehenge2’.



Figure 16: Results on ‘Taj Mahal’.



Figure 17: Failure cases.