

SIGN: Spatial-information Incorporated Generative Network for Generalized Zero-shot Semantic Segmentation

Supplementary Material

A. Detailed SIM Structure

Fig. 6 shows the reparameterization module in Spatial Information Module. Fig. 7 shows the self-attention module and Fig. 8 shows the details of four architectures tested for SIM in Section 4.4.

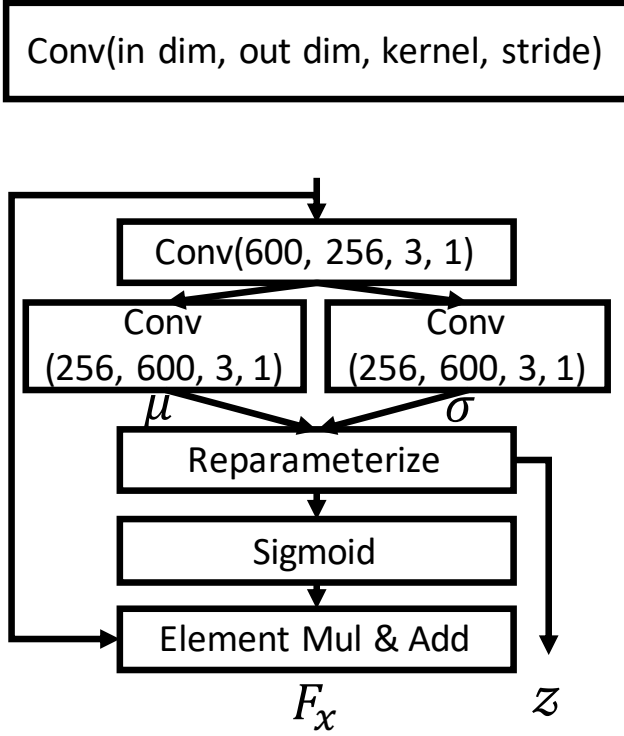


Figure 6. Reparameterization module used in SIM.

B. Additional Qualitative Results

We show additional qualitative results of our SIGN model in Figs. 9 to 11.

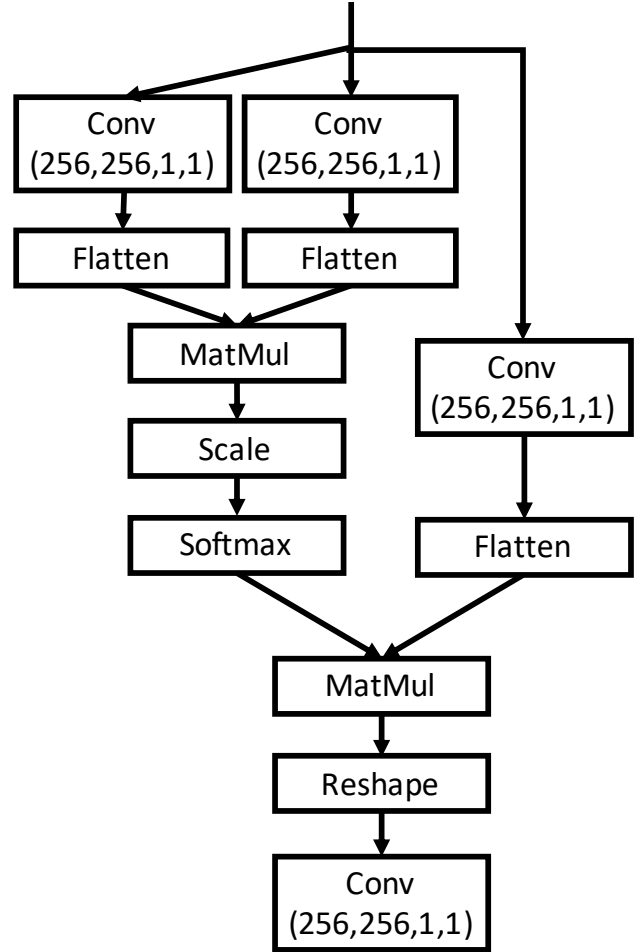


Figure 7. Self-attention module used in SIM

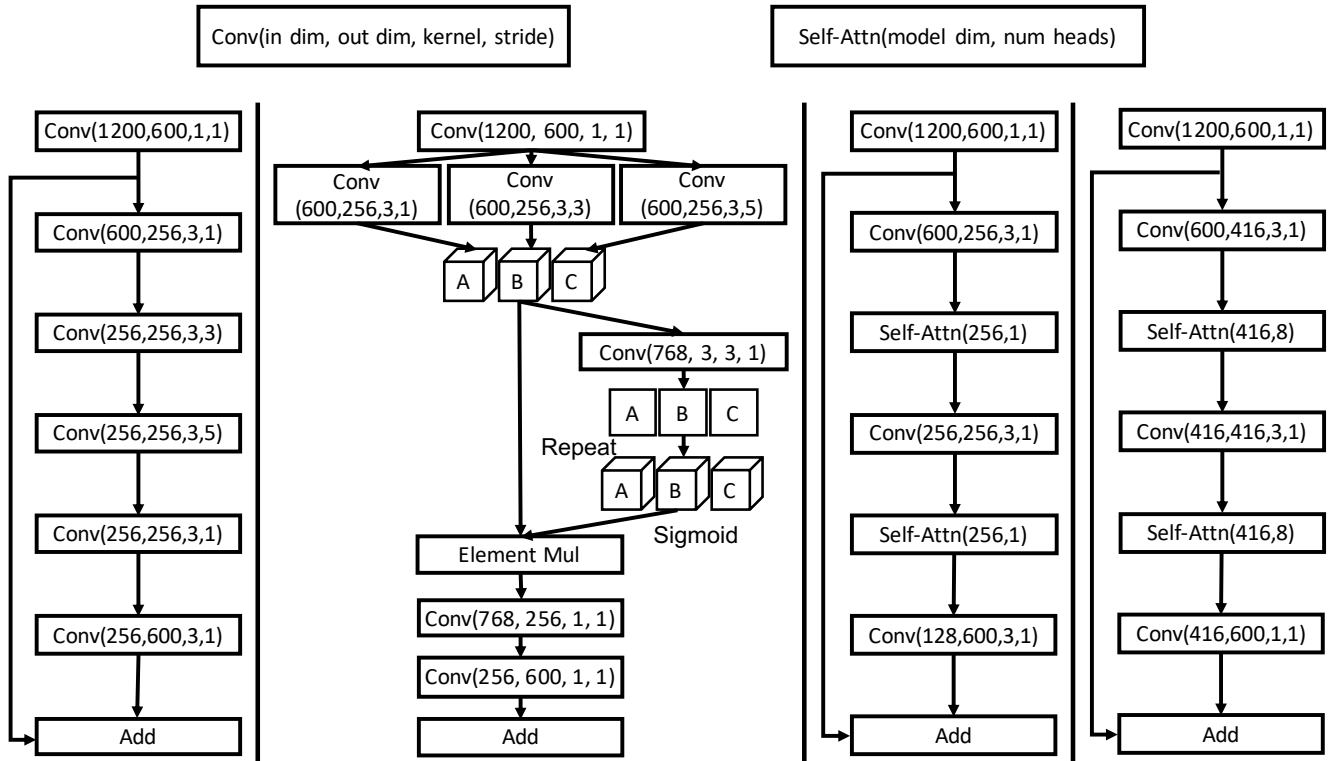


Figure 8. Four architectures tested in SIM . a) Convolution-based[21]. b) Attention-based[18]. c) Self-attention-based[50]. d) Multihead self-attention-based[50]

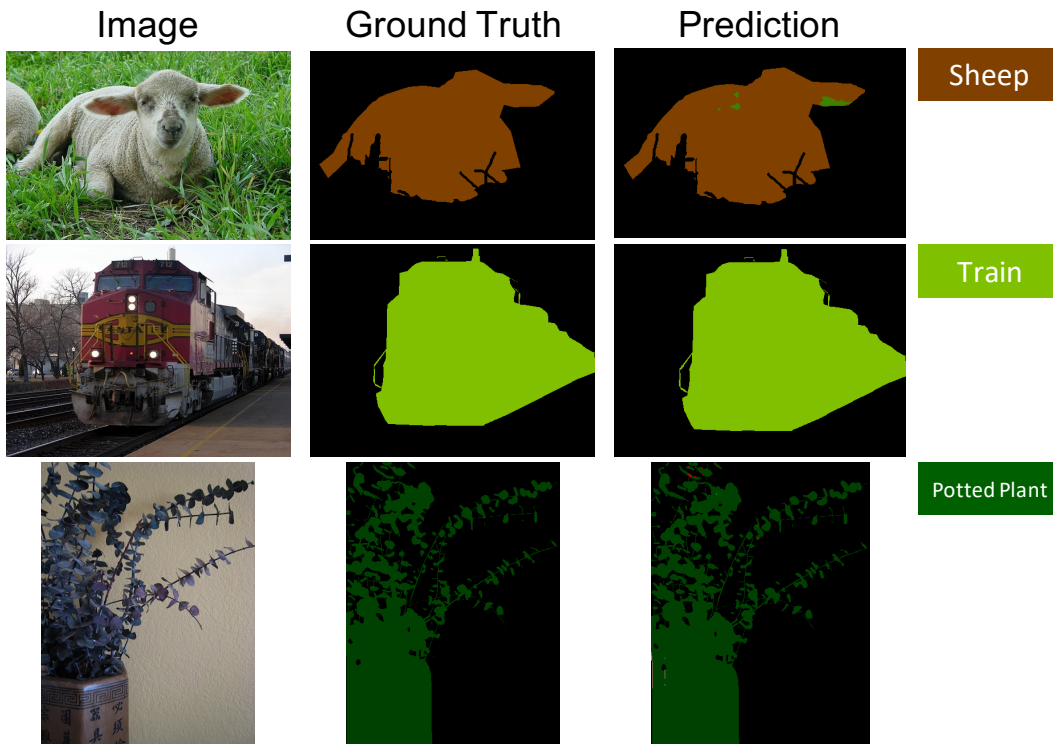


Figure 9. Additional qualitative results on Pascal VOC. Rectangle on the right side indicates the unseen class in the sample on the left

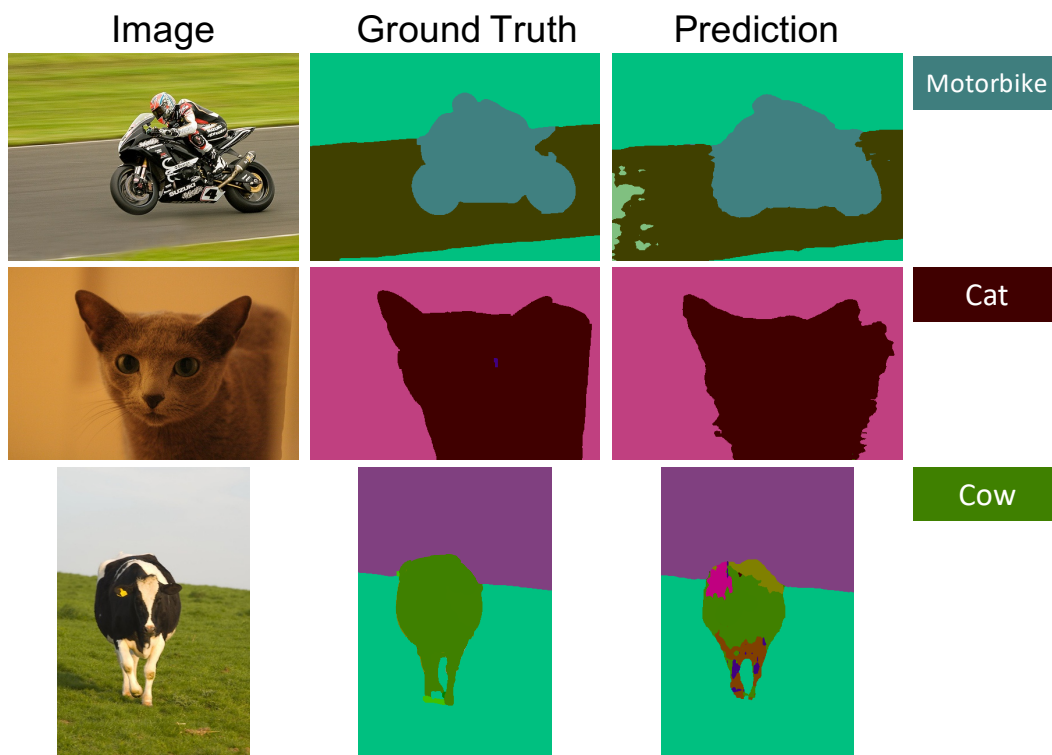


Figure 10. Additional qualitative results on Pascal Context. Rectangle on the right side indicates the unseen class in the sample on the left

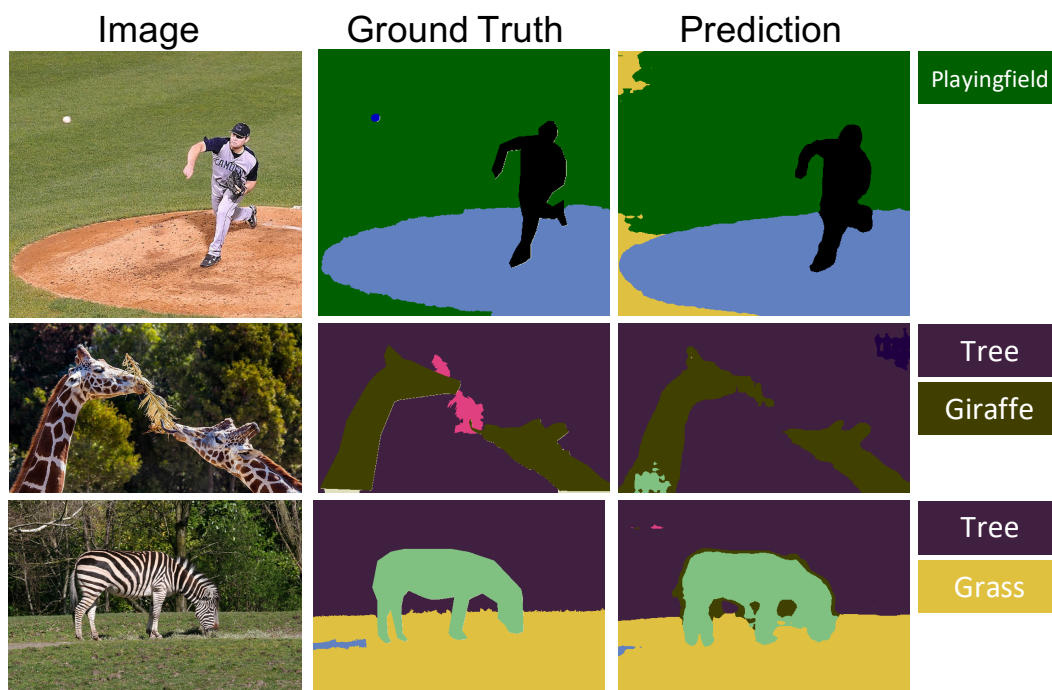


Figure 11. Additional qualitative results on COCO Stuff. Rectangle(s) on the right side indicates the unseen class(es) in the sample on the left