Supplementary Materials

Haoxiang Ma[1,2]
mahaoxiang822@buaa.edu.cn

Hongyu Yang[†3]
hongyuyang@buaa.edu.cn

Di Huang[1,2]
dhuang@buaa.edu.cn

State Key Laboratory of Software Development Environment[1]
Beihang University, Beijing, China

School of Computer Science and Engineering[2]
Beihang University, Beijing, China

Institute of Artificial Intelligence[3]
Beihang University, Beijing, China

Ablation study on MSB input. In Table 1, we report the ablation study results on MSB input to verify its impact on segmentation performance. As we can see, when connecting the features extracted from different stages of the backbone (ResNet101), both low-level details and high-level semantics can be utilized by the proposed MSB module, which brings better performance. By concatenating the features of all four stages, the mIoU value improves by 0.97%. The results above clearly validate the effectiveness of our model design.

MSB Input	mIoU%
{S4}	79.06
{S3,S4}	79.93
{S2,S3,S4}	79.47
{S1,S2,S3,S4}	**80.03**

Table 1: Effect of MSB input. S1, S2, S3 and S4 represent the output stages of the backbone.

†Corresponding author