Supplementary Material for:
TB-Net: A Three-Stream Boundary-Aware Network for Fine-Grained Pavement Disease Segmentation

Yujia Zhang[1], Qianzhong Li[1,2], Xiaoguang Zhao[1], and Min Tan[1]
[1]Institute of Automation, Chinese Academy of Sciences
[2]School of Artificial Intelligence, University of Chinese Academy of Sciences
{zhangyujia2014, liqianzhong2017, xiaoguang.zhao, min.tan}@ia.ac.cn

1. More Ablation Analysis

We provide a few visual examples of the ablation studies in Figure 1 and 2, in order to further analyze the effects of 1) the context-aware attention module, 2) incorporating the boundary information and 3) the weighting mechanism in the loss function.

Note that *TB-Net w/o attn*, *TB-Net w/o BS* and *TB-Net w/o w* denote the models that do not make use of the attention module, the boundary stream and the weighting mech-

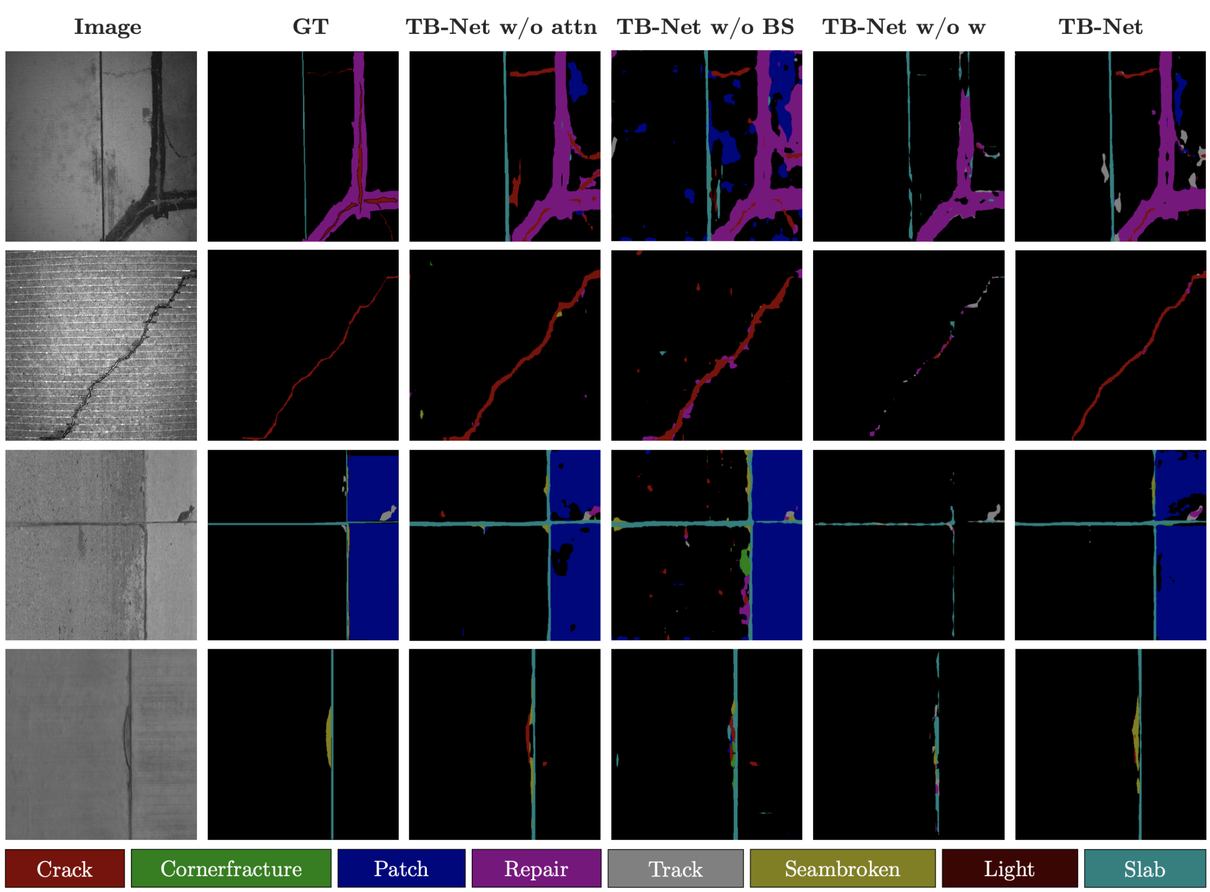

Figure 1. Visual examples of the proposed method and the three models that does not make use of the attention module (*TB-Net w/o attn*), the boundary stream (*TB-Net w/o BS*) and the weighting mechanism (*TB-Net w/o w*), respectively.

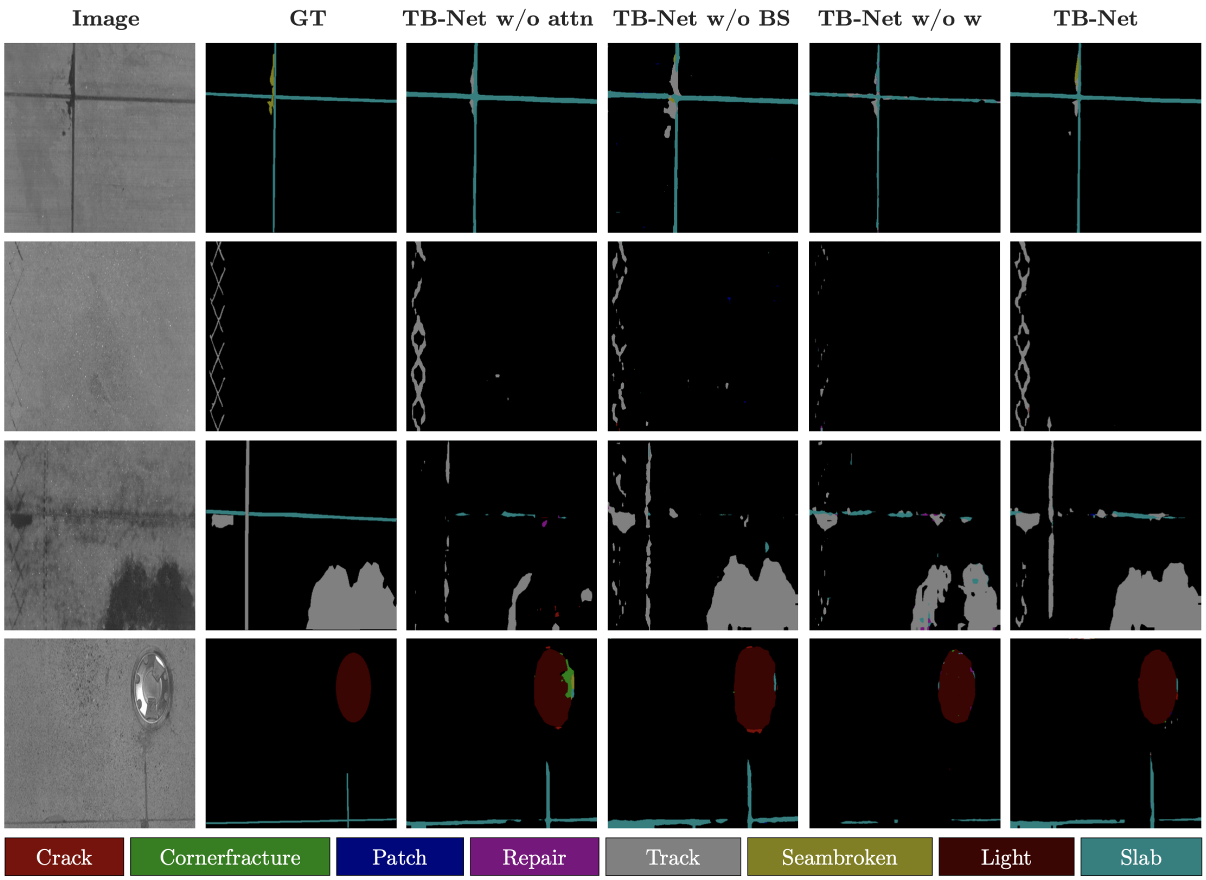

Figure 2. Visual examples of the proposed method and the three models that does not make use of the attention module (*TB-Net w/o attn*), the boundary stream (*TB-Net w/o BS*) and the weighting mechanism (*TB-Net w/o w*), respectively.

anism, respectively. It can be seen that our method generally achieves better segmentation results than all its variants. Specifically, the full model performs well on small *Seambroken* and *Track* while the variants fail to correctly segment most disease areas. Besides, for the images that have relatively large disease areas, our TB-Net that fuses three different feature representations achieves more favorable results. This demonstrates the effectiveness of our proposed context-aware attention module, the boundary stream and the weighting mechanism.

2. More Qualitative Results

Figure 3 provides additional qualitative results of our TB-Net and a state-of-the-art method on semantic segmentation, BiSeNet [2]. Note, the ground-truth boundary is obtained by using a recent edge detection method [1] given the annotated segmentation maps. We can observe that our model performs better both on large diseases (such as *Repair*, *Patch* and *Light*) and small diseases (such as *Crack* and *Seambroken*, while BiSeNet tends to struggle with large *Patch* and *Repair*. In general, our model can better discriminate the disease areas and the background.

References

[1] Jianzhong He, Shiliang Zhang, Ming Yang, Yanhu Shan, and Tiejun Huang. Bi-directional cascade network for perceptual edge detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3828–3837, 2019.

[2] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In *Proceedings of the European Conference on Computer Vision*, pages 325–341, 2018.

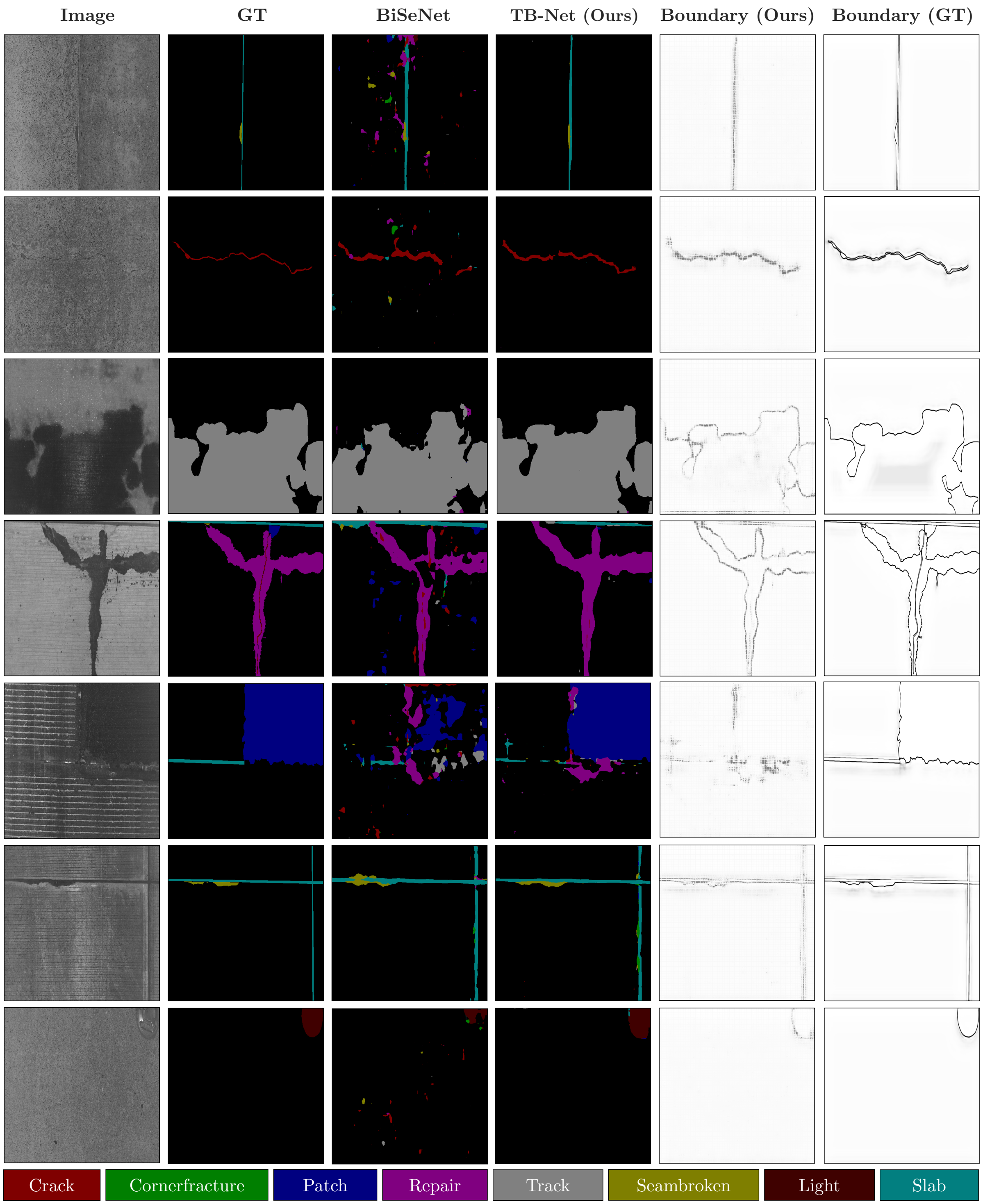

Figure 3. More qualitative results of our proposed TB-Net and one of the competing models, BiSeNet [2]. From left to right: image, ground-truth, predictions of BiSeNet and our TB-Net, boundary prediction of TB-Net and boundary ground-truth obtained using [1].