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 mechanism, respectively.

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 do 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.

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


<table>
<thead>
<tr>
<th>Image</th>
<th>GT</th>
<th>BiSeNet</th>
<th>TB-Net (Ours)</th>
<th>Boundary (Ours)</th>
<th>Boundary (GT)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="gt1.png" alt="GT" /></td>
<td><img src="bi-se1.png" alt="BiSeNet" /></td>
<td><img src="tb-net1.png" alt="TB-Net (Ours)" /></td>
<td><img src="boundary1.png" alt="Boundary (Ours)" /></td>
<td><img src="boundary-gt1.png" alt="Boundary (GT)" /></td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="gt2.png" alt="GT" /></td>
<td><img src="bi-se2.png" alt="BiSeNet" /></td>
<td><img src="tb-net2.png" alt="TB-Net (Ours)" /></td>
<td><img src="boundary2.png" alt="Boundary (Ours)" /></td>
<td><img src="boundary-gt2.png" alt="Boundary (GT)" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="gt3.png" alt="GT" /></td>
<td><img src="bi-se3.png" alt="BiSeNet" /></td>
<td><img src="tb-net3.png" alt="TB-Net (Ours)" /></td>
<td><img src="boundary3.png" alt="Boundary (Ours)" /></td>
<td><img src="boundary-gt3.png" alt="Boundary (GT)" /></td>
</tr>
</tbody>
</table>

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].