Supplementary Materials

Figure S.1 shows examples of 2-class segmentation results. In the examples, we can observe that the characteristics of tumor regions in these two examples are very different; Upper: the cancer cells sparsely distributed; Bottom: the cancer cells densely distributed. Though the color of tumor regions tends to be darker than that of normal regions, it is difficult to segment these only using the color information, (pathologists also use texture information). In this comparison, the proposed results were slightly better than that of SegNet (d) and DeepLabv3+ (f), as same as the experiments in the paper. At the boundary of the tumor regions in Figure S.1, it is qualitatively observed that the proposed methods smoothly segmented regions, and the noises in the tumor of our results were less than those of the compared methods. These are few pixels, which did not contribute much to the quantitative improvement of accuracy, but they are improving the discrimination when practically examining the cancer distribution.

Figure S.2 shows examples of multi-class segmentation results in comparison. These were generated by overlaying multi-subtype segmentation results in the tumor region of the two-class segmentation. Compared to a single input comparison method, the proposed method using multiple views seems to be more accurate. In particular, compared with U-net (c) that was used as the expert network of the proposed method, our results were better than those of U-net in all images, both of qualitatively and quantitatively. Moreover, in the bottom example, the result of the Adaptive-Weighting-Multi-Field-of-View (i) was better than those of the fixed weighting multi-filed-of-view method (h), though the performance of results in the top example was no much different.

The Figure S.3 shows examples of enlarged images. The index corresponds to Figure S.2. In the top image, the difference between the characteristics in Lepidic (red) and Acinar/Papillary (yellow) regions is not clear to the non-professional eye. Expert pathologists commented that the sparseness and the texture were different between the subtypes, but the boundaries of subtypes were not clear. That is a reason why the boundary area was not labeled as any subtypes (black). It is difficult to segment subtypes accurately than the semantic segmentation in a natural image that objects have clear boundaries. In our results (h)(i), the center area of Acinar/Papillary (yellow) regions was correctly segmented and the boundary area of that was incorrectly identified as Lepidic. In the results of U-net (c), the almost area of Acinar/Papillary regions were incorrectly identified as Lepidic. Moreover, the result of the Adaptive-Weighting-Multi-Field-of-View (i) was better than that of the fixed weighting multi-filed-of-view method (h). Middle images also show the difficulties in segmenting the subtypes between Solid and Micro-papillary. In this example, the results also have the same trends as the top example. In the bottom, segmentation is not performed correctly in any methods. In this way, there were some images with partial ambiguity and difficult to divide, but as a whole, by combining multiple fields of view the performance of semantic segmentation improved as shown in Table 2 in the paper.

References

[1] V. Badrinarayanan, A. Kendall, and R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *TPAMI*, 39(12):2481–2495, 2017. 2, 3

[2] L. Chen, G. Papandreou, and F. Schroff. Rethinking atrous convolution for semantic image segmentation. In *arXiv*, 2017. 2, 3

[3] bonlime/keras-deeplab-v3-plus: Keras implementation of deeplab v3+ with pretrained weights. https://github.com/bonlime/keras-deeplab-v3-plus. (Accessed on 11/16/2018). 2, 3

[4] Segnet model implemented using keras framework. https://github.com/imlab-uiip/keras-segnet. (Accessed on 11/16/2018). 2, 3

[5] Keras implementation of u-net. https://github.com/aymanshams07/Ultra/blob/master/unet.py. (Accessed on 11/16/2018). 2, 3, 4

[6] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015. 2, 3, 4

[7] D. B. Sam, S. Surya, and R. V. Babu. Switching convolutional neural network for crowd counting. In *CVPR*, 2017. 2, 3

[8] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. In *ICLR*, 2016. 2, 3

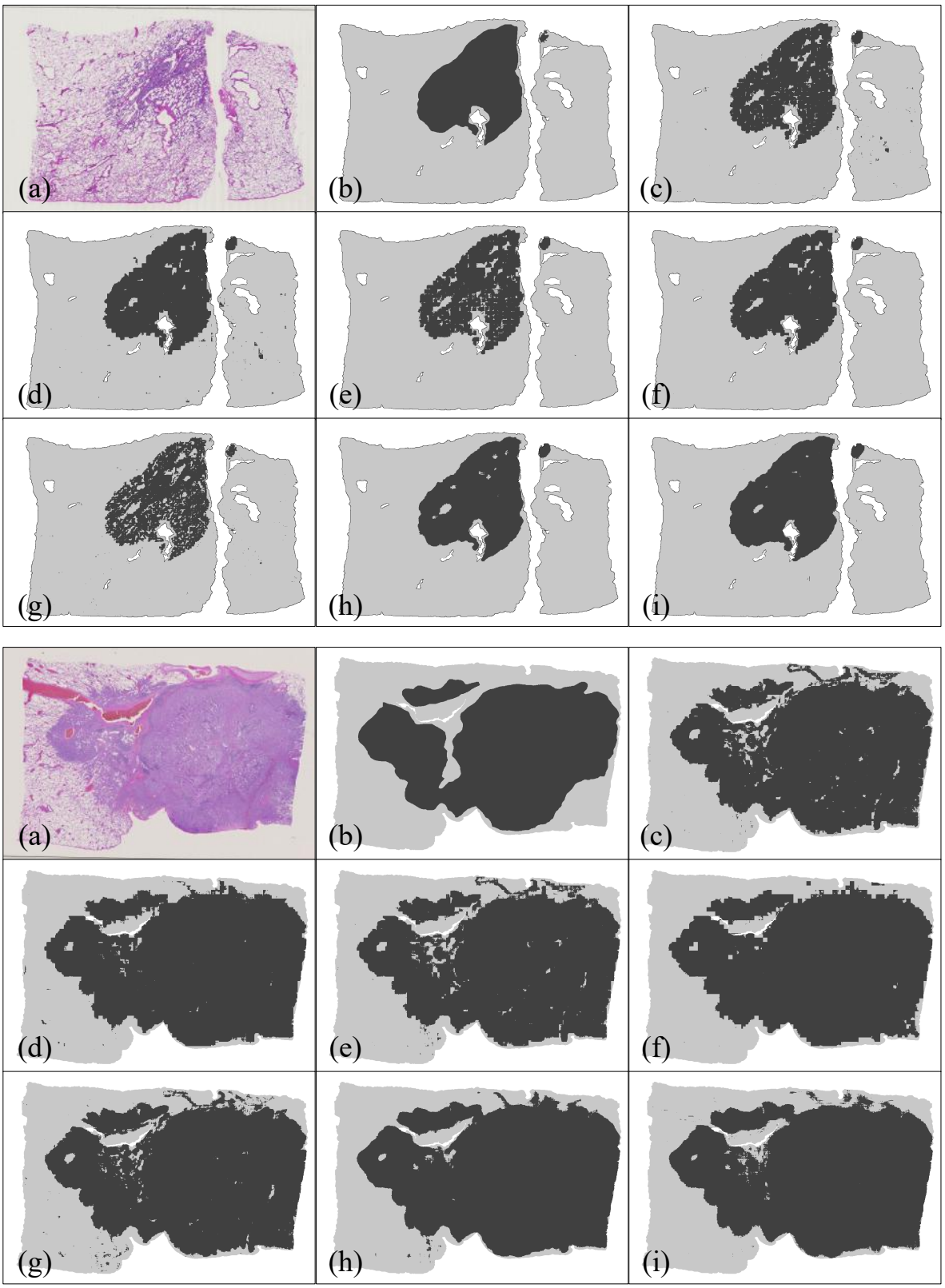

Figure S.1. Examples of 2-class normal or tumor segmentation results. (a) original images, segmentation images from (b) manual annotation, (c) U-net [5][6], (d) SegNet [1][4], (e) Dilated-net [8], (f) DeepLabv3+ [2][3], (g) Hard-Switch-CNN [7], (h) Ours (Fixed), and (i) Ours (Adaptive). The color of the region indicates the class (see Figure 4 in the paper).

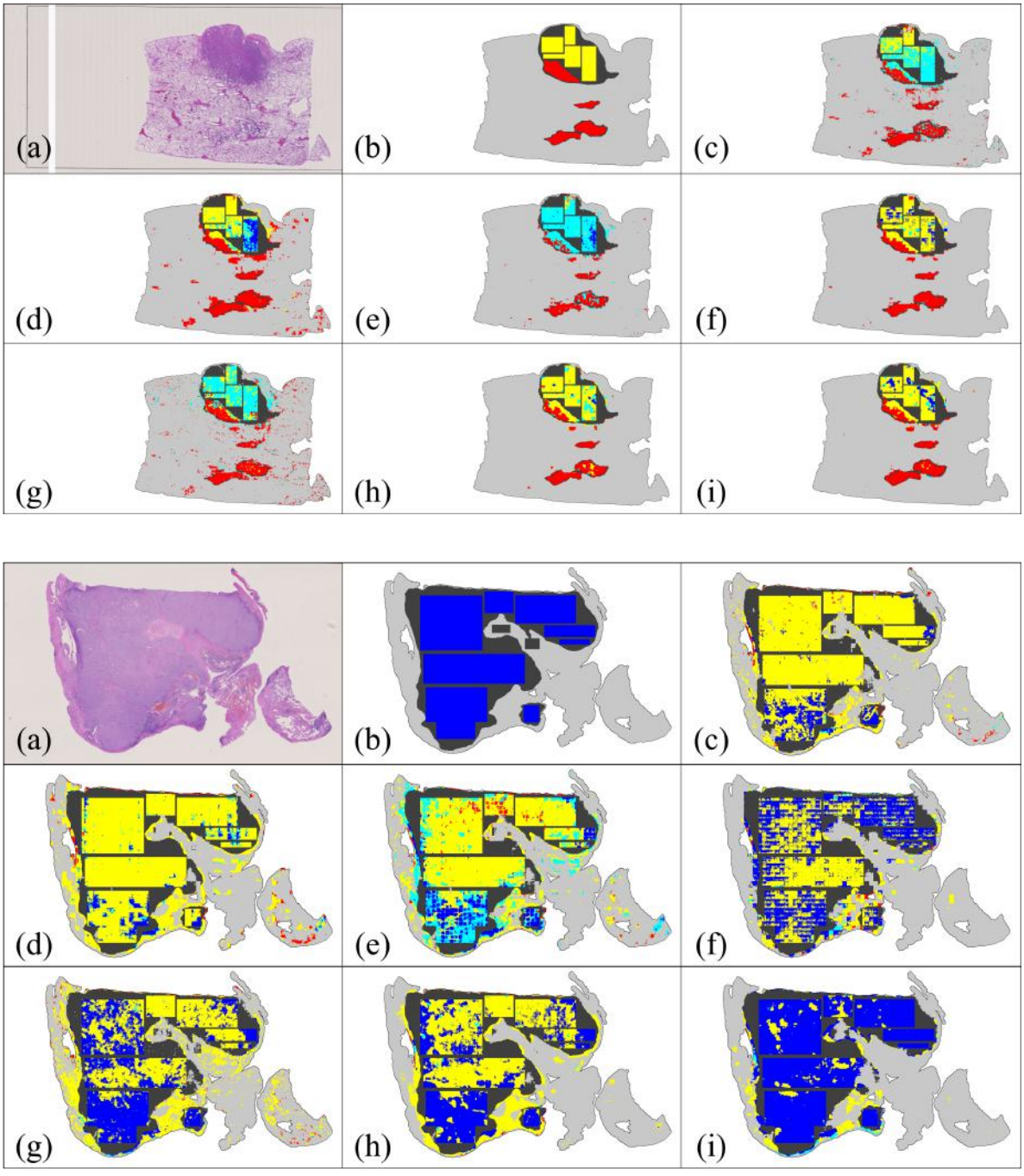

Figure S.2. Examples of multi-class segmentation results. (a) original images, segmentation images from (b) manual annotation, (c) U-net [5][6], (d) SegNet [1][4], (e) Dilated-net [8], (f) DeepLabv3+ [2][3], (g) Hard-Switch-CNN [7], (h) Ours (Fixed), and (i) Ours (Adaptive). The color of the region indicates the class (see Figure 4 in the paper).

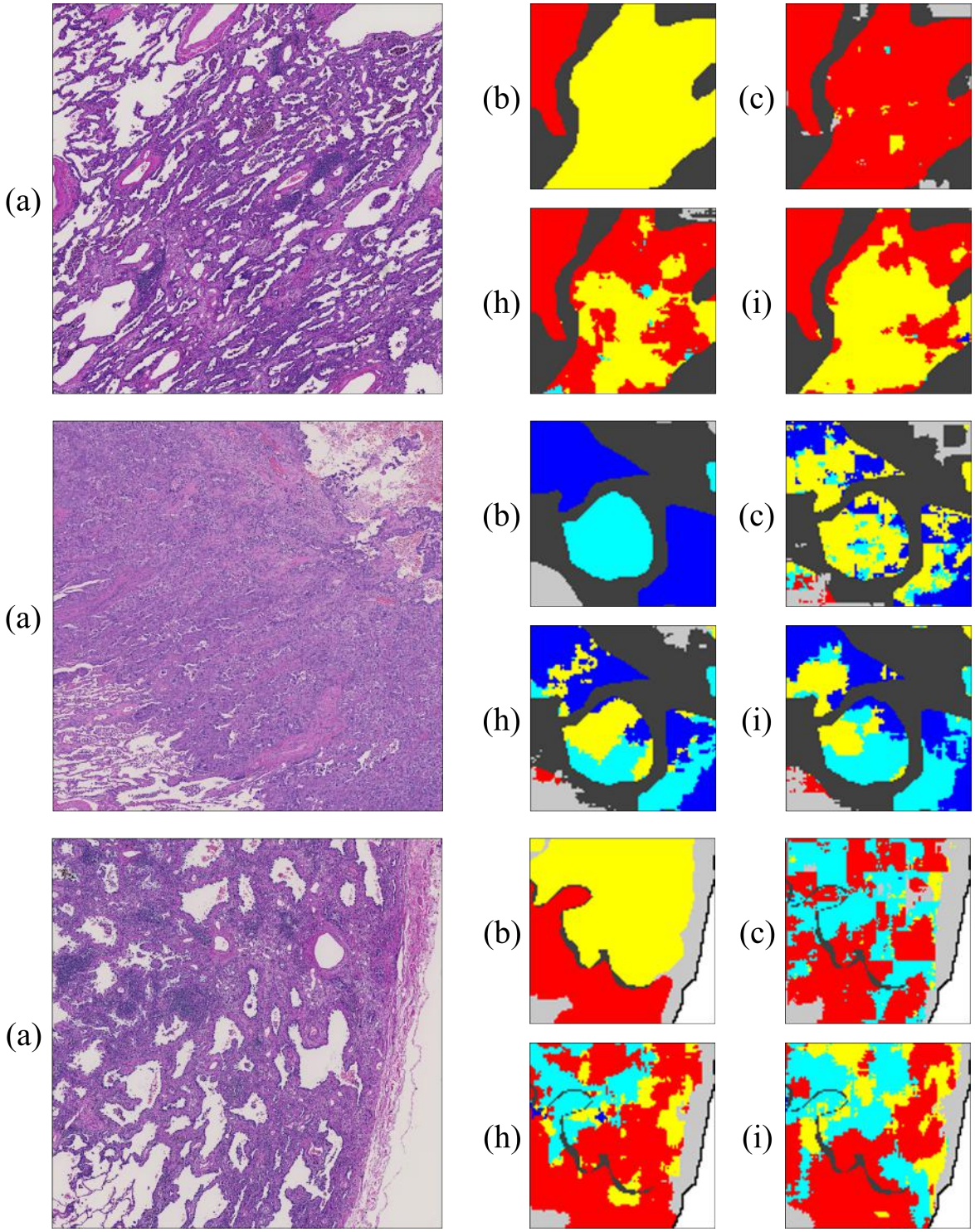

Figure S.3. Examples of enlarged multi-class segmentation results. (a) original images, segmentation images from (b) manual annotation, (c) U-net [5][6], (h) Ours (Fixed), and (i) Ours (Adaptive). The color of the region indicates the class (see Figure 4 in the paper).