Supplementary Material for *HistoSegNet: Weakly-Supervised Semantic Segmentation of Histological Tissue Type in Digital Pathology*

A Detailed Methodology Description in Mathematical Notation

In this section, we describe in detail the constituent operations of the four-stage HistoSegNet methodology in mathematical notation (as shown in Figure 1). Each black-box stage is investigated in further detail in the following subsections.

Our intention is to clarify the precise operations used in our methodology with a consistent notation (hence the slightly different notation from that used in the main paper) and we direct the reader to more thorough works written by others for more complete treatment of CNNs, Grad-CAM, and fully-connected CRFs. For higher-level explanation of the methodology and rationales behind the design choices, we refer the reader to the main paper.

A.1 Patch-level HTT classification CNN

Our CNN is a shortened, multi-label variant of the VGG16 architecture, with 7 convolutional layers, a global max pooling layer, and a single fully-connected layer. It takes as input an RGB patch image $\mathbf{X} = \mathbf{A}^{(0)}$ of size $N \times N \times 3$, outputting (1) a continuous confidence score vector $\mathbf{y}$ and (2) a boolean prediction $\mathbf{p}$, both of size C, the number of HTTs.

Each of the $\ell \in \{1, \cdots, 7\}$ convolutional layers take an input feature map $\mathbf{A}^{(\ell-1)}$ sized $N_{\ell-1} \times N_{\ell-1} \times D_{\ell-1}$ and convolves it with the convolutional kernels $\mathbf{W}^{(\ell)}$ sized $M_\ell \times M_\ell \times D_{\ell-1} \times D_\ell$ to produce an output feature map $\mathbf{A}^{(\ell)}$ of size $N_\ell \times N_\ell \times D_\ell$. This is followed by a ReLU activation, batch normalization, and either a dropout or max pooling layer. We provide the equations for each type of layer below.

Convolutional Layer.

$$\hat{\mathbf{A}}_d^{(\ell)} \leftarrow \sum_{d'=1}^{D_{\ell-1}} \mathbf{A}_{d'}^{(\ell-1)} * \mathbf{W}_{d',d}^{(\ell)}, \quad \forall d \in \{1, \cdots, D_\ell\} \tag{1}$$

ReLU Activation Layer (Feature Extractor).

$$\hat{\mathbf{A}}^{(\ell)} \leftarrow \max(\hat{\mathbf{A}}^{(\ell)}, 0) \tag{2}$$

Batch Normalization Layer. For the b-th batch,

$$\{\hat{\mathbf{A}}^{(\ell)}\}_b \leftarrow \gamma_\ell \frac{\{\hat{\mathbf{A}}^{(\ell)}\}_b - \mathbb{E}[\{\hat{\mathbf{A}}^{(\ell)}\}_b]}{\sqrt{\text{Var}[\{\hat{\mathbf{A}}^{(\ell)}\}_b] + \epsilon}} + \beta_\ell, \quad \forall b \in \{1, \cdots, B\}. \tag{3}$$

Dropout Layer.

$$\hat{\mathbf{A}}_d^{(\ell)} \leftarrow \begin{cases} \hat{\mathbf{A}}_d^{(\ell)}, & \text{with probability } 1 - P_\ell \\ 0, & \text{with probability } P_\ell \end{cases}, \quad \forall d \in \{1, \cdots, D_\ell\}. \tag{4}$$

Max Pooling Layer.

$$\mathbf{A}^{(\ell)}(i, j) \leftarrow \max_{(i',j') \in \mathcal{P}_{(i,j)}^{(\ell)}} \left(\hat{\mathbf{A}}^{(\ell)}(i', j') \right), \tag{5}$$

where $\mathcal{P}_{(i,j)}^{(\ell)} \in \mathbb{R}^{2 \times 2}$ is a 2×2 window around the pixel (i, j).

Global Max Pooling Layer.

$$a_d \leftarrow \max_{(i',j') \in \mathcal{P}^{(7)}} \left(\mathbf{A}_d^{(7)}(i', j') \right), \tag{6}$$

where $\mathcal{P}^{(7)} \in \mathbb{R}^{N_7 \times N_7}$ is the entire receptive field of the final convolutional layer's output.

Fully Connected Layer.

$$\hat{y}_c \leftarrow \sum_{d=1}^{D^{(7)}} a_d w_{d,c}, \quad c \in \{1, \cdots, C\} \tag{7}$$

ReLU Activation Layer (Classifier).

$$\hat{\mathbf{y}} \leftarrow \max(\hat{\mathbf{y}}, 0) \tag{8}$$

Sigmoid Layer.

$$\mathbf{y} \leftarrow \frac{1}{1 + \exp(-\hat{\mathbf{y}})} \tag{9}$$

Thresholding Layer.

$$p_c \leftarrow [y_c \geq \theta_c], \quad c \in \{1, \cdots, C\} \tag{10}$$

A.2 Pixel-level HTT Segmentation

Grad-CAM.

In the Grad-CAM method proposed by [1] to infer pixel-level class activation in a CNN, a partial backpropagation is conducted from the confidence score y_c to the final activation output $\hat{\mathbf{A}}_d^{(7)}$ (before pooling) to obtain the gradient $\frac{\partial y_c}{\partial \hat{\mathbf{A}}_d^{(7)}}$. Then, we obtain the "class feature weight" between c and d as follows:

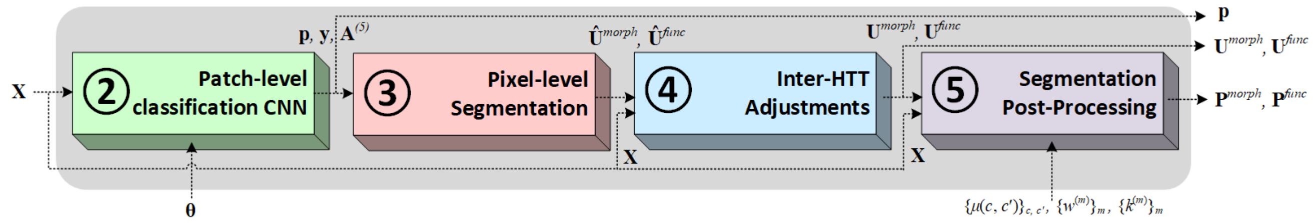

X
p, y, $\mathbf{A}^{(5)}$
2 Patch-level classification CNN
$\hat{\mathbf{U}}^{morph}$, $\hat{\mathbf{U}}^{func}$
3 Pixel-level Segmentation
$\mathbf{U}^{morph}$, $\mathbf{U}^{func}$
4 Inter-HTT Adjustments
5 Segmentation Post-Processing
p
$\mathbf{U}^{morph}$, $\mathbf{U}^{func}$
$\mathbf{P}^{morph}$, $\mathbf{P}^{func}$
X
X
θ
$\{\mu(c,c')\}_{c,c'}$, $\{w^{(m)}\}_m$, $\{k^{(m)}\}_m$

Figure 1. *The overall patch processing pipeline of HistoSegNet for a given patch, expressed as a sequence of black-box operations: the inputs to the pipeline are the input patch* $\mathbf{X}$ *and the class thresholds* $\boldsymbol{\theta}$*, the outputs are the class confidence scores* $\mathbf{p}$*, the Class-Specific Grad-CAMs* $\mathbf{U}^{morph}$, $\mathbf{U}^{func}$*, and the predicted segmentations* $\mathbf{P}^{morph}$, $\mathbf{P}^{func}$.

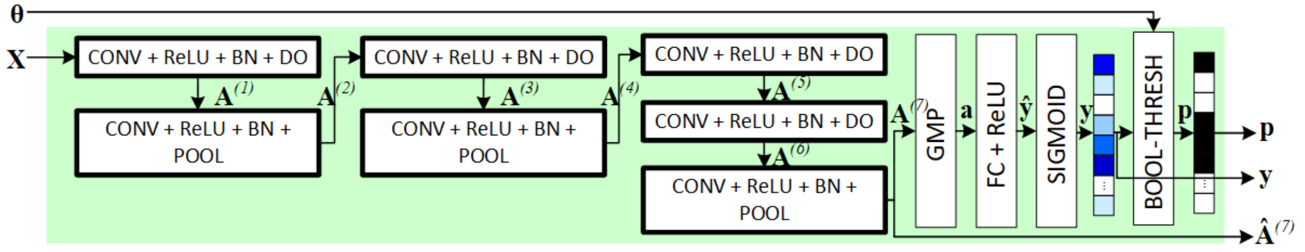


Figure 2. *Expanded view of the CNN's constituent operations.*

$$\alpha_{c,d} \leftarrow \frac{1}{N_7^2} \sum_{i=1}^{N_7} \sum_{j=1}^{N_7} \frac{\partial y_c}{\partial \hat{\mathbf{A}}_d^{(7)}(i,j)} \tag{11}$$

And the corresponding Grad-CAM (sized $N_7 \times N_7$) is:

$$\tilde{\mathbf{U}}_c \leftarrow \text{ReLU}\left(\sum_{d=1}^{D_7} \alpha_{c,d} \hat{\mathbf{A}}_d^{(7)}\right) \tag{12}$$

Finally, the Grad-CAM is upsampled back to the original image size $N \times N$ using bilinear interpolation and normalized by its 2D maximum as follows:

$$\tilde{\mathbf{U}}_c \leftarrow \frac{\tilde{\mathbf{U}}_c}{\max(\tilde{\mathbf{U}}_c)} \tag{13}$$

.

Scaling by HTT Confidence Scores.

The Grad-CAM is then scaled by the boolean-thresholded patch-level HTT confidence scores $y_c \cdot p_c$ in order to ignore non-confident HTT predictions and boost confident predictions relative to less-confident predictions:

$$\hat{\mathbf{U}}_c \leftarrow \begin{cases} y_c \hat{\tilde{U}}_c, & \text{if } p_c = 1 \\ 0, & \text{if } p_c = 0 \end{cases} \tag{14}$$

Then, the Grad-CAMs are split into morphological ($\hat{\mathbf{U}}^{\text{morph}}$) and functional ($\hat{\mathbf{U}}^{\text{func}}$) types for separate processing.

A.3 Inter-HTT Adjustments

Background Activation.

We produce our background activation map two steps: first, the smooth white-illumination image $\hat{\mathbf{U}}_B$ is obtained by applying a scaled-and-shifted sigmoid to the mean-RGB image $\overline{\mathbf{X}}$; then, we subtract the appropriate transparent-staining class activations, and finally we filter with a 2D Gaussian blur $\mathbf{U}_{\mu,\sigma}$ to reduce the prediction resolution:

$$\begin{aligned} \hat{\mathbf{U}}_B &\leftarrow \frac{0.75}{1+\exp\left[-4(\overline{\mathbf{X}}-240)\right]} \\ \hat{\mathbf{U}}_B^{\text{morph}} &\leftarrow (\hat{\mathbf{U}}_B - \max(\hat{\mathbf{U}}^{\text{A.W}}, \hat{\mathbf{U}}^{\text{A.B}})) * \mathbf{U}_{0,2} \\ \hat{\mathbf{U}}_B^{\text{func}} &\leftarrow (\hat{\mathbf{U}}_B - \max(\hat{\mathbf{U}}^{\text{G.O}}, \hat{\mathbf{U}}^{\text{G.N}}, \hat{U}^{\text{T}})) * \mathbf{U}_{0,2}. \end{aligned} \tag{15}$$

"Other" Activation.

To generate our non-functional tissue activation map, we first take the maximum of: (1) other functional type activations $\hat{\mathbf{U}}_c^{\text{func}}$, (2) adipose activations $\hat{\mathbf{U}}_A = \max(\hat{\mathbf{U}}^{\text{A.W}}, \hat{\mathbf{U}}^{\text{A.B}})$, and (3) background activation $\mathbf{U}_B^{\text{func}}$. Then, we subtract this activation from one, scale it:

$$\hat{\mathbf{U}}_O^{\text{func}} \leftarrow 0.05\left[1 - \max\left(\{\hat{\mathbf{U}}_c^{\text{func}}\}_{c=1}^{C}, \hat{\mathbf{U}}_B^{\text{func}}, \hat{\mathbf{U}}_A\right)\right]. \tag{16}$$

Depth Concatenation.

$$\mathbf{U}^{\text{morph}} \leftarrow \{\hat{\mathbf{U}}_B^{\text{morph}}, \{\hat{\mathbf{U}}_c^{\text{morph}}\}_{c=1}^{C^{\text{morph}}}\} \tag{17}$$

$$C^{\text{morph}} \leftarrow C^{\text{morph}} + 1 \tag{18}$$

$$\mathbf{U}^{\text{func}} \leftarrow \{\hat{\mathbf{U}}_B^{\text{func}}, \hat{\mathbf{U}}_O^{\text{func}}, \{\hat{\mathbf{U}}_c^{\text{func}}\}_{c=1}^{C^{\text{func}}}\} \tag{19}$$

$$C^{\text{func}} \leftarrow C^{\text{func}} + 2 \tag{20}$$

A.4 Post-process HTT Segmentation

Borrowing some notation from [2], we can express the fully-connected CRF as an iteratively-updated mean field approximation. Note that, for the rest of this section, we adopt the notations $(\mathbf{U}, \mathbf{P}, C)$ as short-hand for $(\mathbf{U}^{\text{morph}}, \mathbf{P}^{\text{morph}}, C^{\text{morph}})$ and $(\mathbf{U}^{\text{func}}, \mathbf{P}^{\text{func}}, C^{\text{func}})$, since the morphological and functional types are post-processed in

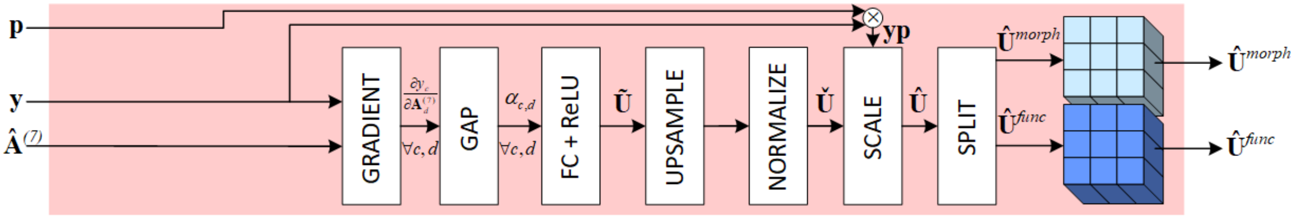


Figure 3. *Expanded view of the pixel-level HTT segmentation's (i.e. Grad-CAM) constituent operations.*

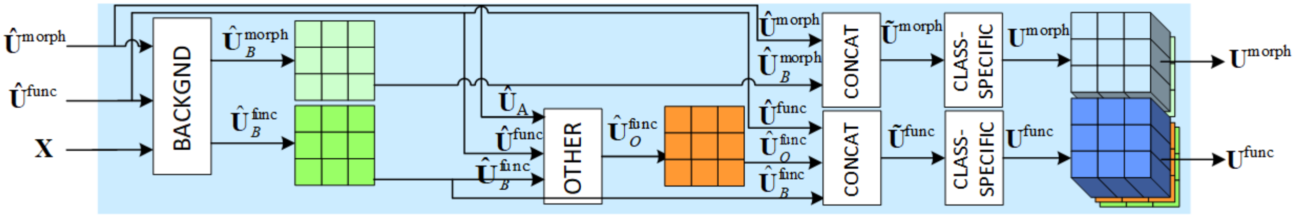

$\hat{\mathbf{U}}^{\mathrm{morph}}$
$\hat{\mathbf{U}}^{\mathrm{func}}$
X
BACKGND
$\hat{\mathbf{U}}_B^{\mathrm{morph}}$
$\hat{\mathbf{U}}_B^{\mathrm{func}}$
$\hat{\mathbf{U}}_A$
$\hat{\mathbf{U}}^{\mathrm{func}}$
$\hat{\mathbf{U}}_B^{\mathrm{func}}$
OTHER
$\hat{\mathbf{U}}_O^{\mathrm{func}}$
$\hat{\mathbf{U}}^{\mathrm{morph}}$
$\hat{\mathbf{U}}_B^{\mathrm{morph}}$
$\hat{\mathbf{U}}^{\mathrm{func}}$
$\hat{\mathbf{U}}_O^{\mathrm{func}}$
$\hat{\mathbf{U}}_B^{\mathrm{func}}$
CONCAT
CONCAT
$\tilde{\mathbf{U}}^{\mathrm{morph}}$
$\tilde{\mathbf{U}}^{\mathrm{func}}$
CLASS-SPECIFIC
CLASS-SPECIFIC
$\mathbf{U}^{\mathrm{morph}}$
$\mathbf{U}^{\mathrm{func}}$
$\mathbf{U}^{\mathrm{morph}}$
$\mathbf{U}^{\mathrm{func}}$

Figure 4. *Expanded view of the Inter-HTT Adjustment's constituent operations.*

the same way. Further, we chose to run CRF for a fixed number of iterations as the convergence criterion.

```
Data: $\mathbf{U}, \mathbf{X}, \{\mu(c,c')\}_{c,c'}, \{w^{(m)}\}_m, \{k^{(m)}\}_m$
Result: $\mathbf{P}$
for $(i,j) \in \mathbb{N}^{N \times N}, c \in \{1, \cdots, C\}$ do
    $Z(i,j) \leftarrow \sum_{c=1}^{C} \exp\left(-U_c(i,j)\right)$ ;                    // (1)
    $Q_c(i,j) \leftarrow \frac{1}{Z(i,j)} \exp\left(-U_c(i,j)\right)$
end
while not converged do
    for $(i,j) \in \mathbb{N}^{N \times N}, c \in \{1, \cdots, C\}$ do
        $\tilde{Q}_c^{(m)}(i,j) \leftarrow$
          $\sum_{(i,j) \neq (i',j')} k^{(m)}\left(\mathbf{f}(i,j), \mathbf{f}(i',j')\right) Q_c(i',j')$
          ;                                                   // (2)
        $\check{Q}_c(i,j) \leftarrow \sum_{m=1}^{M} w^{(m)} \tilde{Q}_c^{(m)}(i,j)$ ;
          // (3)
        $\hat{Q}_c(i,j) \leftarrow \sum_{c'=1}^{C} \mu(c,c') \check{Q}_c(i,j)$ ;
          // (4)
        $\breve{Q}_c(i,j) \leftarrow -U_c(i,j) - \hat{Q}_c(i,j)$ ;  // (5)
        $Z(i,j) \leftarrow \sum_{c=1}^{C} \exp\left(\breve{Q}_c(i,j)\right)$ ;  // (6)
        $Q_c(i,j) \leftarrow \frac{1}{Z(i,j)} \exp\left(\breve{Q}_c(i,j)\right)$
    end
end
$P(i,j) \leftarrow \underset{c \in \{1, \cdots, C\}}{\operatorname{argmax}} \left(Q_c(i,j)\right)$ ;                   // (7)
```

Algorithm 1: *Fully-connected CRF, expressed in algorithmic form*

(1) Initialization (Softmax). In the first step of the CRF (1), the unary potentials are initialized from $-\mathbf{U}$, the negative Grad-CAM activation, at all pixel locations (i,j) for each of the c-th HTTs, by applying the softmax function. This ensures that all the potentials at each pixel location sum up to one across all HTTs.

(2) Message Passing ($N^2 \times 5$-Convolution). Next, in order to encode the local feature relations between pixels and pass the "message" from all pixels (i',j') in a pairwise fashion to all pixels (i,j), we do the following. First, the image features are extracted at each pixel by concatenating the pixel locations and RGB values: $\mathbf{f}(i,j) = [i, j, X_R(i,j), X_G(i,j), X_B(i,j)]$. Then, the $m \in \{1, \cdots, M\}$-th Gaussian kernel $k^{(m)} \in \mathbb{R}^{N^2 \times 5}$ is applied and the unary potentials' feature distances are used to weigh all other pixels' unary potential values $Q_c(i',j')$ and are summed up for all (i',j').

(3) Weighting Filter Outputs ($1 \times 1 \times M \times 1$-Convolution). Then the M filter outputs are weighted and summed with fixed weights $w^{(m)}$.

(4) Compatibility Transform ($1 \times 1 \times C \times 1$-Convolution). To encode the compatibility relations between HTTs and transform the message from all HTTs c' in a pairwise fashion to all HTTs c, we do the following. First, we weight all the incoming features for a given HTT c at a fixed location (i,j) by its compatibility weights with all other HTTs c', $\mu(c,c')$. Then, we sum up across the other HTTs.

(5) Local Update (Addition). The potential at each pixel location $U_c(i,j)$ is updated by adding its update value $\hat{Q}_c(i,j)$.

(6) Normalizing (Softmax). And then, in order to ensure that the potentials at each pixel location sum up to one across all HTTs before the next iteration of the algorithm, we apply the softmax function.

(7) Argmax. Once the algorithm has converged, we apply the argmax function to the potentials at each pixel location across all HTTs $c \in \{1, \cdots, C\}$ in order to obtain the most-confident HTT prediction at each pixel.

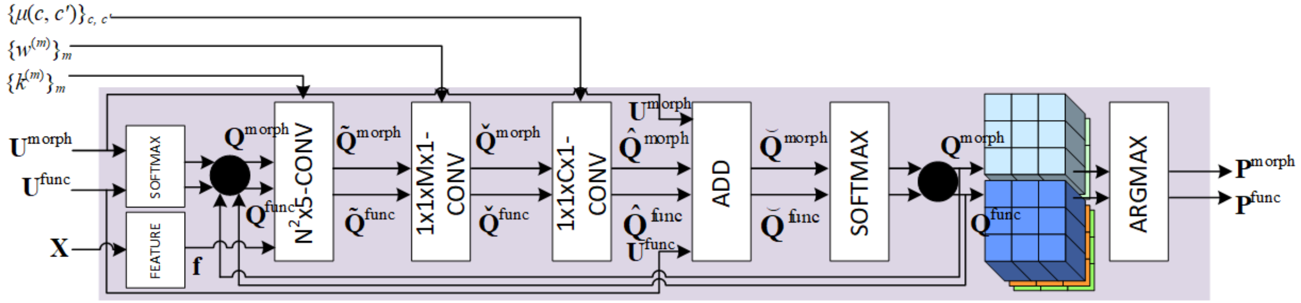


Figure 5. *Expanded view of the Segmentation Post-Processing's constituent operations.*

B Additional Pathologist Slide Segmentation Validation

In Figure B, we present additional segmentation masks of morphological and functional types for additional gastrointestinal slides and in Figure B, we do the same for other organ systems (e.g. thyroid, lymphatic, breast). For the rest of this section, we present additional comments and feedback from the validating pathologist on five select slides that we omitted from the main paper for reasons of space.

C Comparison with Competitive GlaS Gland Segmentation Algorithm

In this section, we compare the performance of HistoSegNet against that of another gland segmentation algorithm on the GlaS Challenge dataset by Gaudet *et al.*[1], who train a U-Net-like Encoder-Decoder CNN with skip connections between different convolutional layers. Their model consists of five convolutional blocks of two convolutional layers each, with skip connections between the second and third blocks, and between the first and fourth blocks. Leaky ReLU activation and spatial dropout regularization are also used. Training is conducted with binary cross-entropy loss on the provided binary gland segmentation masks provided by the GlaS challenge.

The GlaS data consists of five different combinations of Tumor classes i.e. Healthy (H), Adenomatous (A), Moderately differentiated (M), Poorly-to-Moderately differentiated (PM), and Poorly differentiated (P). Each image patch is annotated at the pixel level (done by pathologists) by creating a binary mask image separating between glandular and non-glandular areas. Note that Gaudet *et al.*'s method has been trained only on two different classes introduced by the binary masks ('0' for non-gland, and '1' for gland), where all five tumor classes are treated the same, since they belong to the glandular areas. In contrast, our proposed method HistoSegNet is trained on the ADP database to classify between different healthy tissue classes and is applied to segment out the glands in GlaS without retraining.

As the range of healthiness of glands in GlaS dataset varies from healthy to poorly differentiated, the prediction of such cancerous grades becomes less confident for HistoSegNet. This is shown visually in Figure 14 for five different grades, where Gaudet *et al.*'s method is well capable of segmenting out the glands regardless of their tumor grade but the prediction of HistSegNet becomes visibly less confident with worsening tumor grade. Furthermore, we analyze demonstrate the evaluation performance of both methods numerically in terms of all five different tumor grades in Figure 15. HistoSegNet yields better differentiability between healthy and all four different cancer grades than Gaudet *et al.*'s method. For instance, HistoSegNet yields 18.57% drop in Dice score from healthy 'H' to the next best performing class 'M', while this is only 5.23% drop from healthy to the second best Dice score i.e. 'MP'. Moreover, HistoSegNet provides consistent differentiability across two different scores of Dice and Hausdorff, while Gaudet *et al.*'s method confuses the rank order between healthy and the other four cancer classes across Hausdorff score (intuitively, Dice and Hausdorff should be worsening with worsening tumor grades. Overall, HistoSegNet's performance deterioration from healthy to poorly differentiated tumorous glands is relatively more distinguishable than with a comparable gland segmentation algorithm.

References

[1] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 618–626, 2017. 4321

[2] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr. Conditional random fields as recurrent neural networks. In *Proceedings of the IEEE international conference on computer vision*, pages 1529–1537, 2015. 4322

[1] https://github.com/gaudetcj/GlandSegmentation

Morphological Types

□ Background

🟦 Simple Squamous Epithelium (E.M.S)
🟩 Simple Cuboidal Epithelium (E.M.U)
🟧 Simple Columnar Epithelium (E.M.O)
⬜ Stratified Squamous Epithelium (E.T.S)
🟥 Stratified Cuboidal Epithelium (E.T.U)
🟥 Stratified Columnar Epithelium (E.T.O)
🟫 Pseudostratified Epithelium (E.P)

🟦 Dense Irregular Connective (C.D.I)
🟪 Dense Regular Connective (C.D.R)
🟥 Loose Connective (C.L)

🟧 Erythrocytes (H.E)
🟩 Leukocytes (H.K)
🟪 Lymphocytes (H.Y)

🟧 Compact Bone (S.M.C)
🟪 Spongy Bone (S.M.S)
🟫 Endochondral Bone (S.E)
🟪 Hyaline Cartilage (S.C.H)
🟪 Marrow (S.R)

🟦 White Adipose (A.W)
⬛ Brown Adipose (A.B)
🟫 Marrow Adipose (A.M)

🟨 Smooth Muscle (M.M)
🟥 Skeletal Muscle (M.K)

🟦 Neuropil (N.P)
🟩 Nerve Cell Bodies (N.R.B)
🟨 Nerve Axons (N.R.A)
🟧 Microglial Cells (N.G.M)
🟪 Schwann Cells (N.G.W)

Functional Types

□ Background
🟦 Other

🟦 Exocrine Gland (G.O)
🟩 Endocrine Gland (G.N)

🟥 Transport Vessel (T)

Figure 6. *HTT Segmentation Mask Legend, for both Morphological and Functional types.*

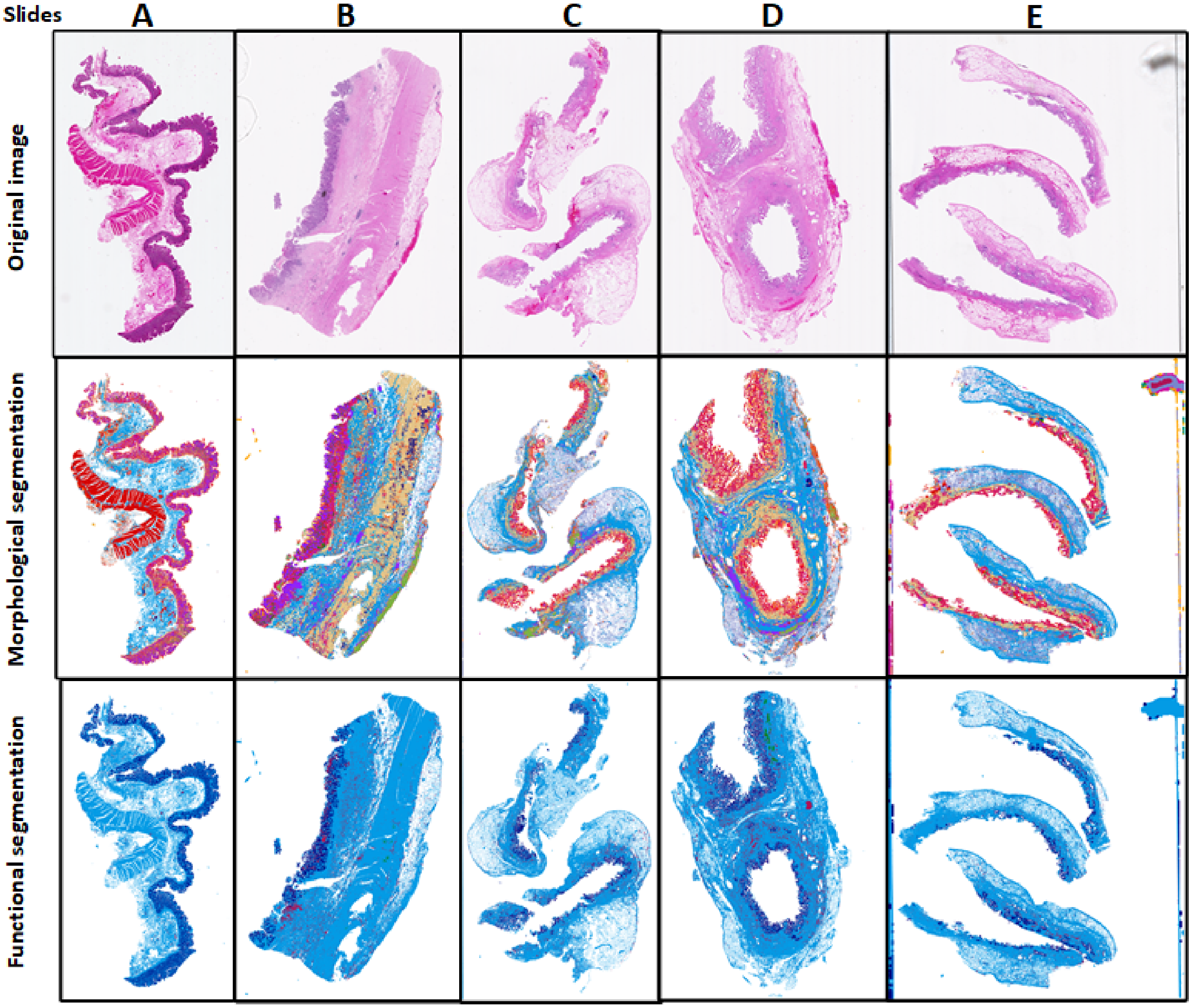


Figure 7. *Morphological and Functional Segmentation on Gastrointestinal Tissues: (A) colon, (B) pylorus, (C) jejunum/ileum, (D) jejunum/ileum, and (E) jejunum/ileum.*

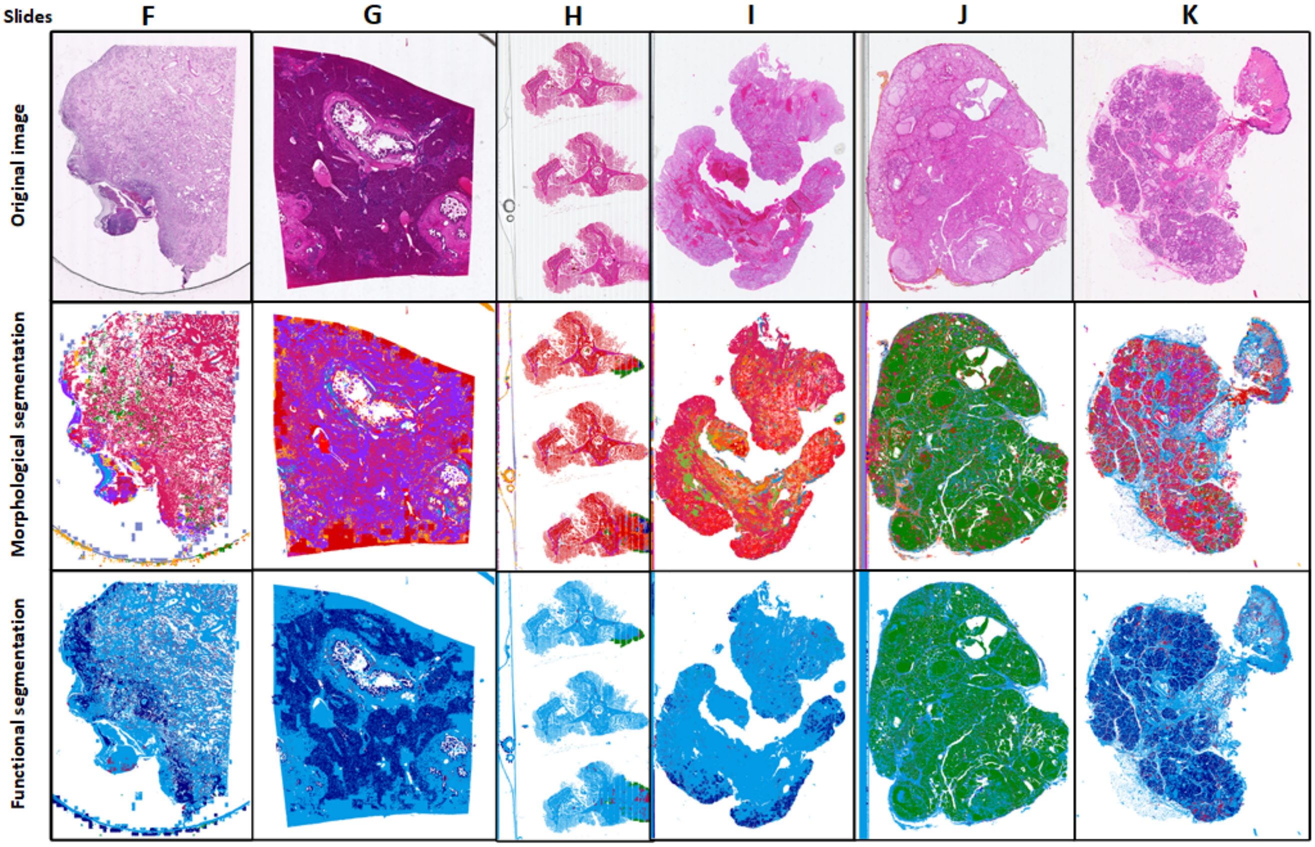
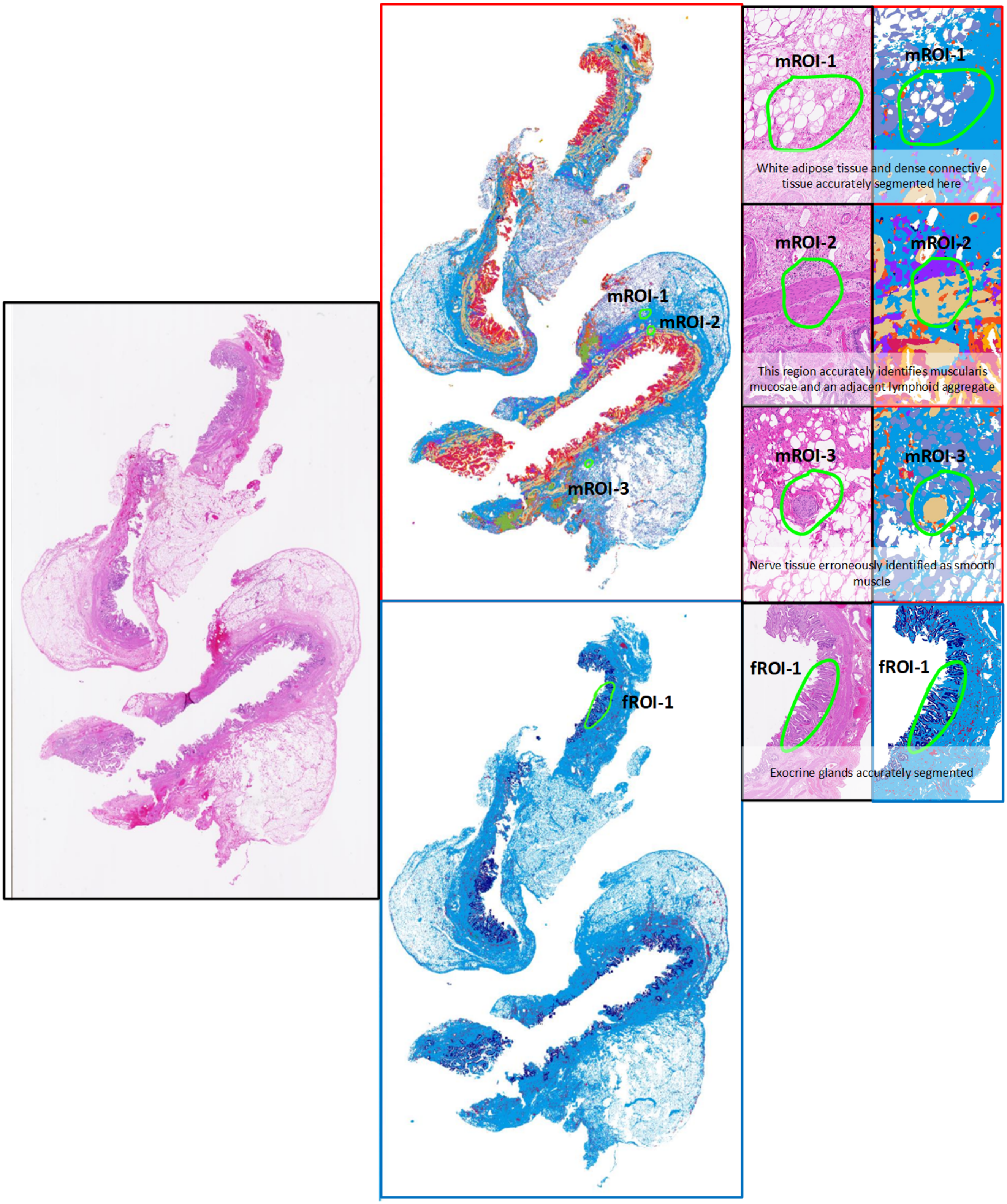


Figure 8. *Morphological and Functional Segmentation on other tissues: (F) lymphatic tissue, (G) liver, (H) vertebrae, (I) glandular (?), (J) thyroid, and (K) breast.*



White adipose tissue and dense connective tissue accurately segmented here

This region accurately identifies muscularis mucosae and an adjacent lymphoid aggregate

Nerve tissue erroneously identified as smooth muscle

Exocrine glands accurately segmented

Figure 9. *Pathologist's Validation Notes on Slide C: original image(s) are shown at left, morphological segmentation shown at top, functional segmentation shown at bottom, and annotated regions shown at right.*

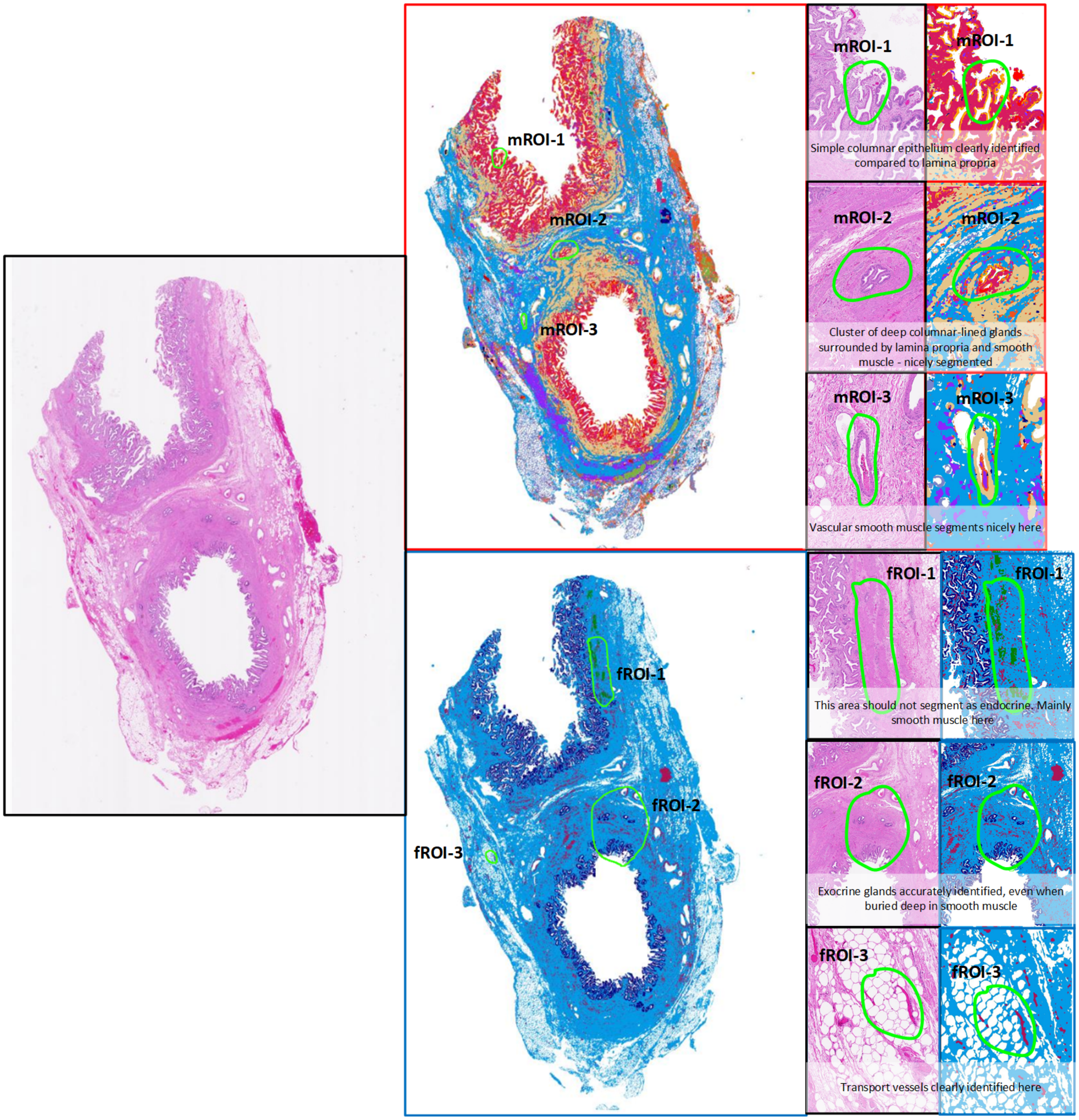
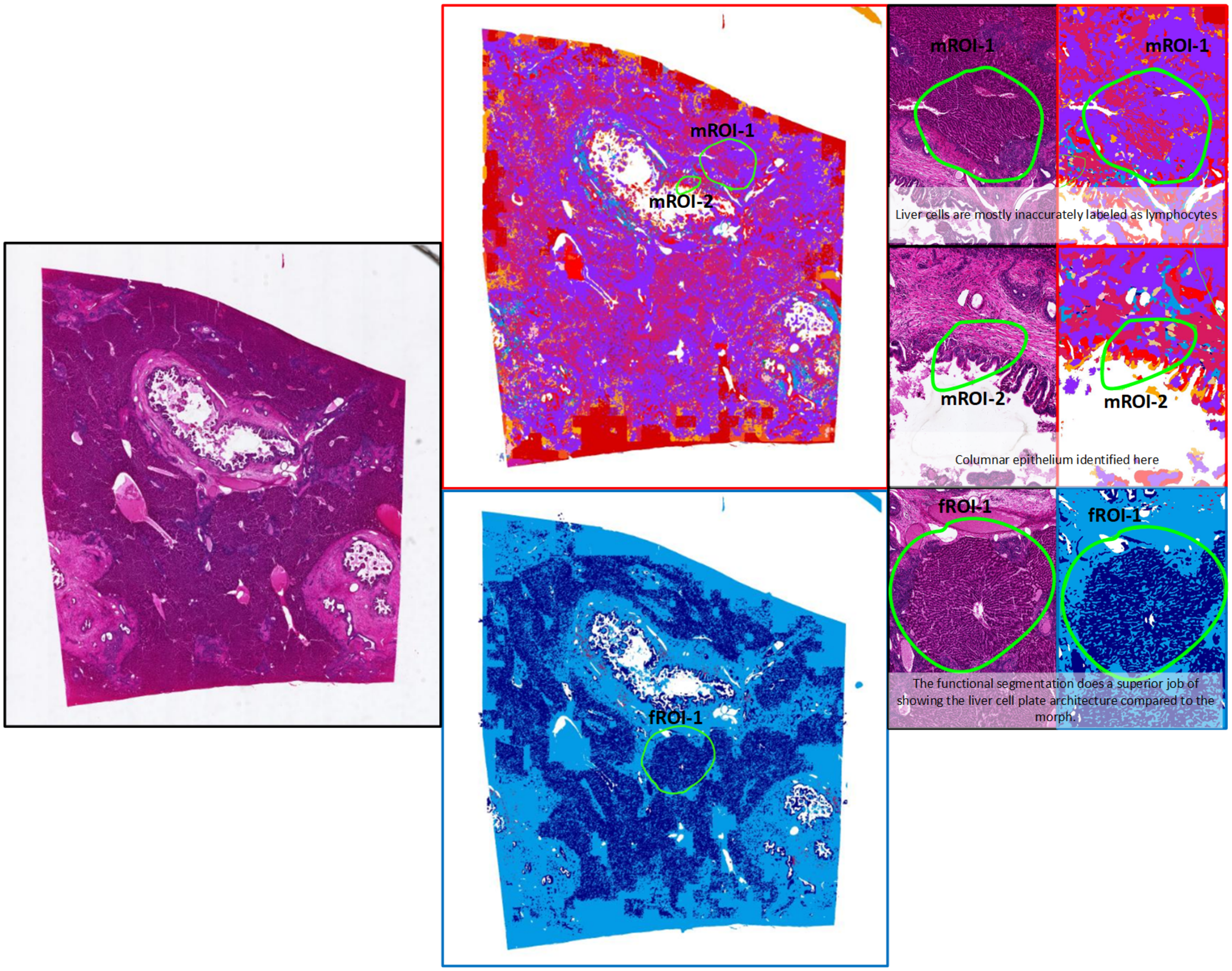

mROI-1
mROI-2
mROI-3
mROI-1
mROI-1
Simple columnar epithelium clearly identified compared to lamina propria
mROI-2
mROI-2
Cluster of deep columnar-lined glands surrounded by lamina propria and smooth muscle - nicely segmented
mROI-3
mROI-3
Vascular smooth muscle segments nicely here
fROI-1
fROI-2
fROI-3
fROI-1
fROI-1
This area should not segment as endocrine. Mainly smooth muscle here
fROI-2
fROI-2
Exocrine glands accurately identified, even when buried deep in smooth muscle
fROI-3
fROI-3
Transport vessels clearly identified here

Figure 10. *Pathologist's Validation Notes on Slide D: original image(s) are shown at left, morphological segmentation shown at top, functional segmentation shown at bottom, and annotated regions shown at right.*



Liver cells are mostly inaccurately labeled as lymphocytes

Columnar epithelium identified here

The functional segmentation does a superior job of showing the liver cell plate architecture compared to the morph.

Figure 11. *Pathologist's Validation Notes on Slide G: original image(s) are shown at left, morphological segmentation shown at top, functional segmentation shown at bottom, and annotated regions shown at right.*

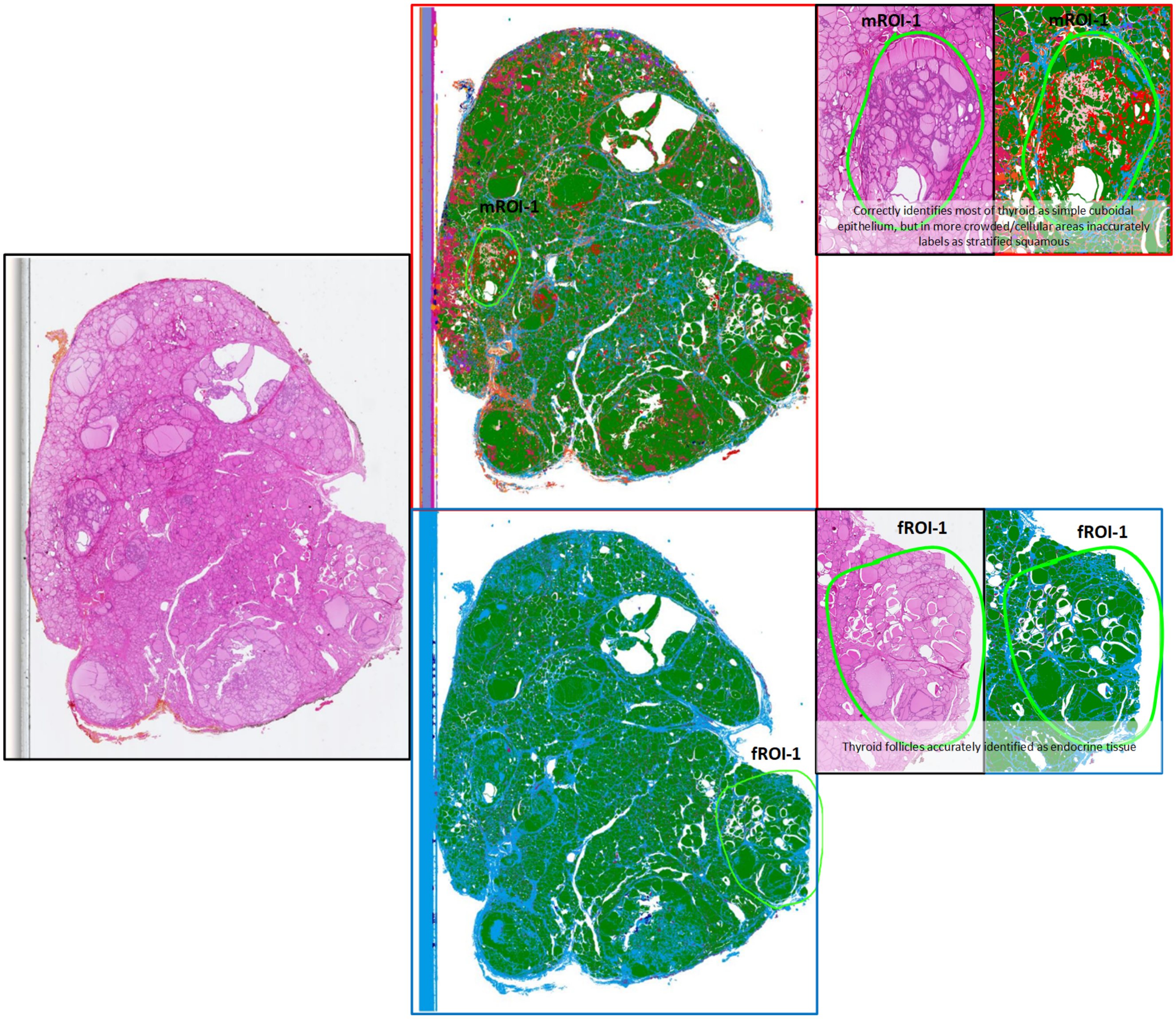
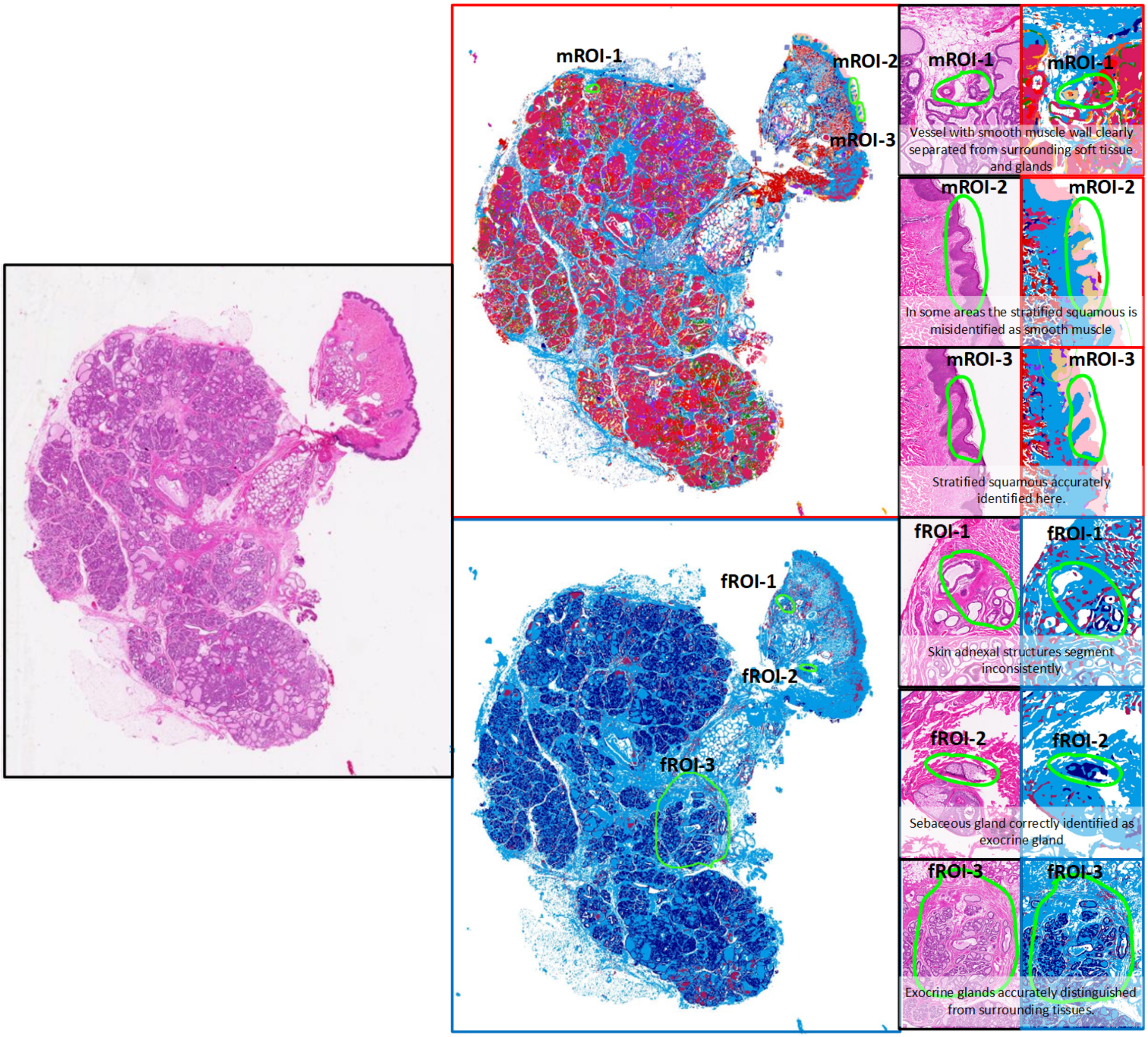

mROI-1
mROI-1
mROI-1
Correctly identifies most of thyroid as simple cuboidal epithelium, but in more crowded/cellular areas inaccurately labels as stratified squamous
fROI-1
fROI-1
fROI-1
Thyroid follicles accurately identified as endocrine tissue

Figure 12. *Pathologist's Validation Notes on Slide J: original image(s) are shown at left, morphological segmentation shown at top, functional segmentation shown at bottom, and annotated regions shown at right.*



Vessel with smooth muscle wall clearly separated from surrounding soft tissue and glands

In some areas the stratified squamous is misidentified as smooth muscle

Stratified squamous accurately identified here.

Skin adnexal structures segment inconsistently

Sebaceous gland correctly identified as exocrine gland

Exocrine glands accurately distinguished from surrounding tissues.

Figure 13. *Pathologist's Validation Notes on Slide K: original image(s) are shown at left, morphological segmentation shown at top, functional segmentation shown at bottom, and annotated regions shown at right.*

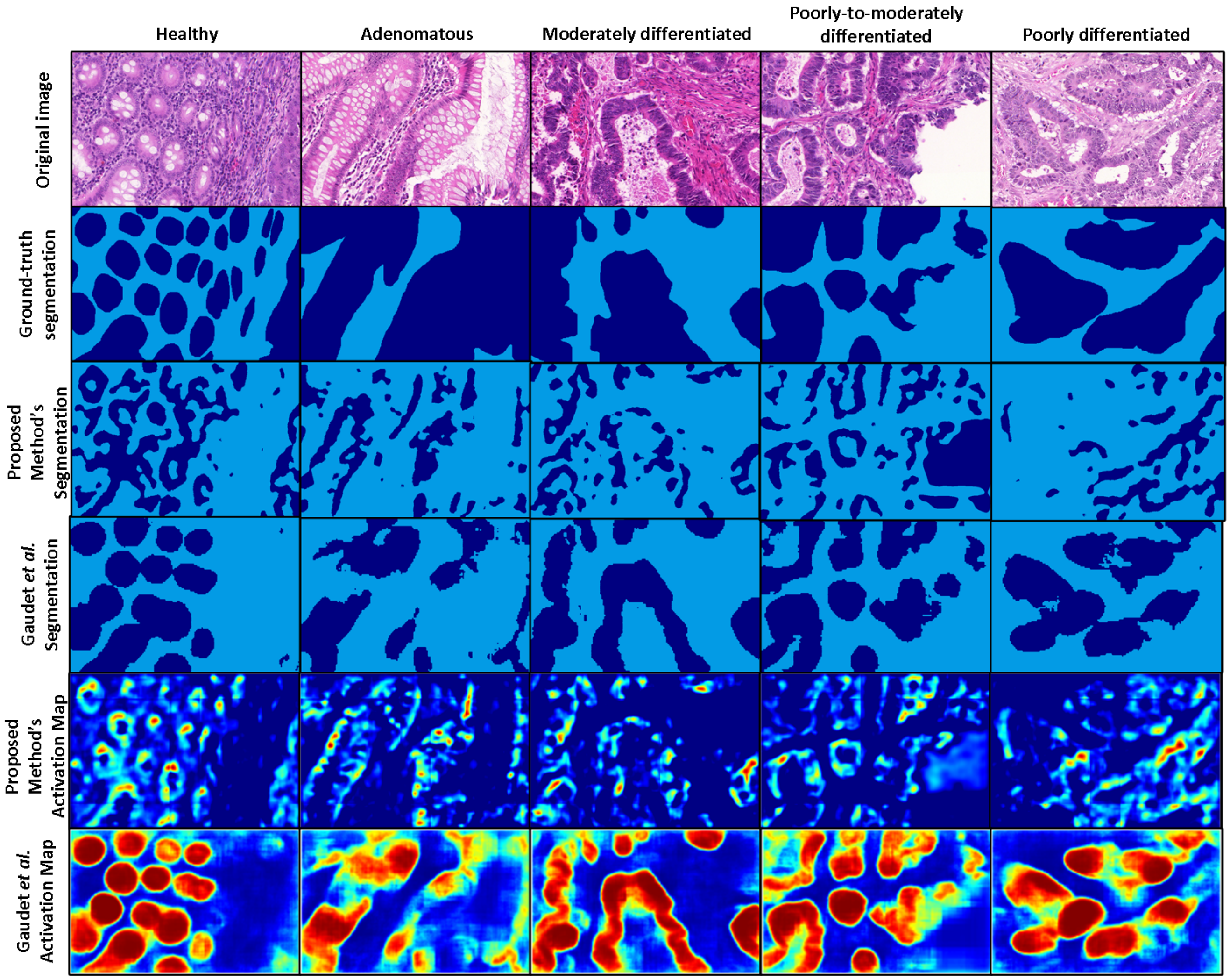
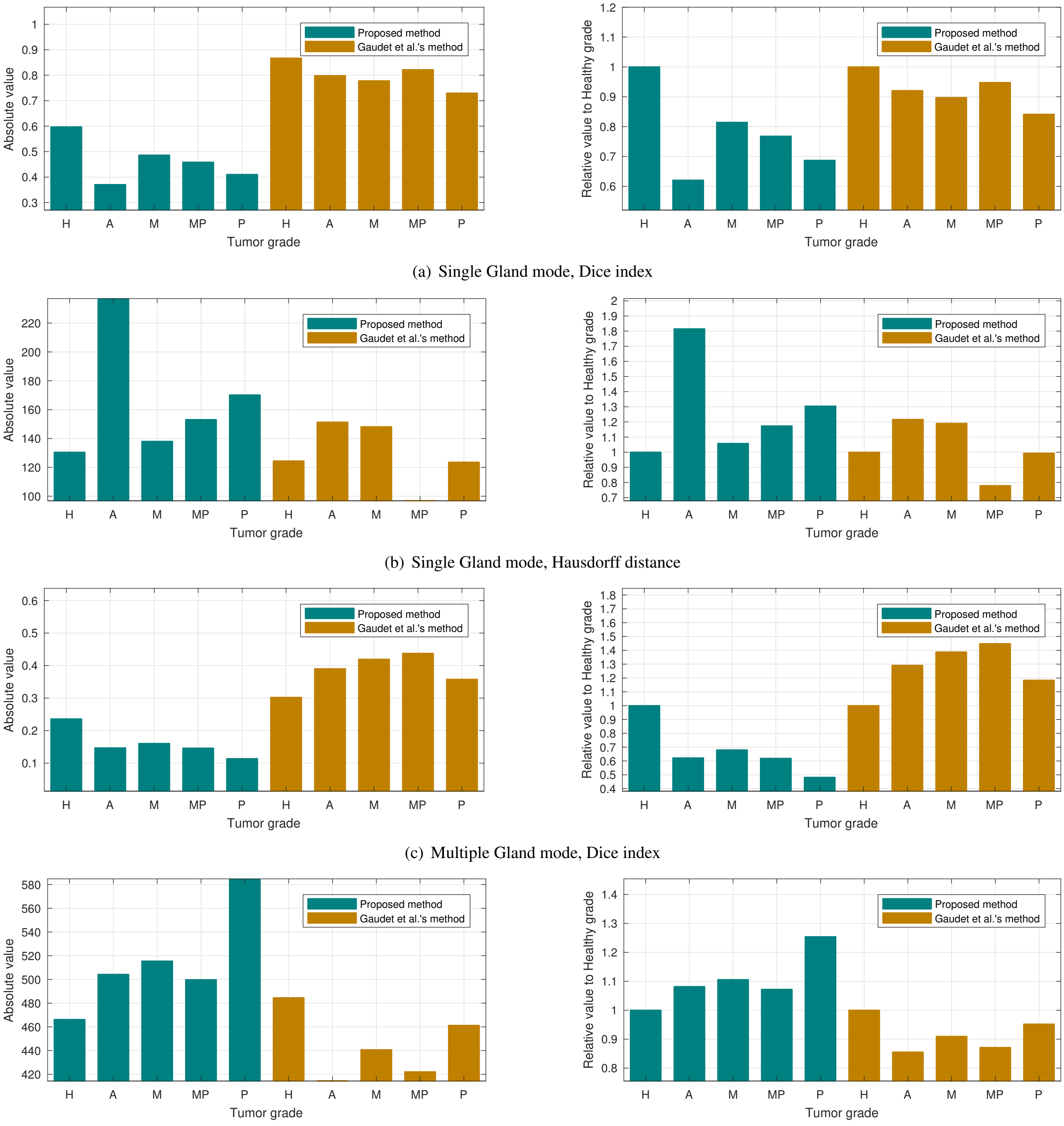

Healthy
Adenomatous
Moderately differentiated
Poorly-to-moderately differentiated
Poorly differentiated
Original image
Ground-truth segmentation
Proposed Method's Segmentation
Gaudet et al. Segmentation
Proposed Method's Activation Map
Gaudet et al. Activation Map

Figure 14. *HistoSegNet vs. Gaudet* et al.*'s gland segmentation algorithm on GlaS challenge images, visually compared at different tumor grades: note how the performance of HistoSegNet steadily degrades with worsening tumor grade while Gaudet* et al.*'s method remains relatively stable throughout.*



(a) Single Gland mode, Dice index

(b) Single Gland mode, Hausdorff distance

(c) Multiple Gland mode, Dice index

(d) Multiple Gland mode, Hausdorff distance

Figure 15. *HistoSegNet vs. Gaudet* et al.*'s gland segmentation algorithm, assessed quantitatively at different tumor grades in four modes: (1) Single gland and Dice index, (2) Single gland and Hausdorff distance, (3) Multiple gland and Dice index, and (4) Multiple gland and Hausdorff distance. The tumor grades are abbreviated as follows: Healthy (H), Adenomatous (A), Moderately differentiated (M), Moderately-to-Poorly differentiated (MP), and Poorly differentiated (P).*