A Bayesian Approach to Weakly-supervised Laparoscopic Image Segmentation - Supplementary Material -

Proof. As for $p(\mathbf{x},\mathbf{y})$, we have:

$$\begin{aligned}\log p(\mathbf{x},\mathbf{y}) &= \log \int_{\mathbf{z}} p(\mathbf{x},\mathbf{y},\mathbf{z})\, d\mathbf{z} = \log \int_{\mathbf{z}} \frac{p(\mathbf{x},\mathbf{y},\mathbf{z})\, q(\mathbf{z}|\mathbf{x})}{q(\mathbf{z}|\mathbf{x})}\, d\mathbf{z} \\ &\geq \mathbb{E}_{\mathbf{z}\sim q}\left[\log \frac{p(\mathbf{x},\mathbf{y},\mathbf{z})}{q(\mathbf{z}|\mathbf{x})}\right] = \mathbb{E}_{\mathbf{z}\sim q}\left[\log \frac{p(\mathbf{y}|\mathbf{x},\mathbf{z})\, p(\mathbf{x}|\mathbf{z})\, p(\mathbf{z})}{q(\mathbf{z}|\mathbf{x})}\right] \\ &= \mathbb{E}_{\mathbf{z}\sim q}\left[\log p(\mathbf{y}|\mathbf{x},\mathbf{z}) + \log p(\mathbf{x}|\mathbf{z})\right] - \mathbb{E}_{\mathbf{z}\sim q}\left[\log \frac{q(\mathbf{z}|\mathbf{x})}{p(\mathbf{z})}\right],\end{aligned} \tag{10}$$

where $q(\mathbf{z}|\mathbf{x})$ is a variational distribution, and $\mathbb{E}_{\mathbf{z}\sim q}$ denotes the expectation over $q(\mathbf{z}|\mathbf{x})$. We finish the proof by deriving the ELBO in Eq. (5). ∎

Table 4. Implementation details. Experiments were performed on PyTorch.

Dataset	CholecSeg8k	AutoLaparo	ACDC
Backbone	U-Net	U-Net	U-Net
Preprocessing	Resized each image to 432×240 pixels and normalized the intensities to [0,1]	Resized each image to 480×240 pixels and normalized the intensities to [0,1]	Resized each slice to 256×256 pixels and normalized the intensities to [0,1]
Input size	432×240	480×240	256×256
Optimizer	Adam with a weight decay of 10^{-4}	Adam with a weight decay of 10^{-4}	SGD with a weight decay of 10^{-4} and a momentum of 0.9
Batch size	8	8	8
Training epochs or iterations	1st stage: $\mathbf{e}_1$, $\mathbf{d}_1$, $\mathbf{e}_2$, and $\mathbf{d}_2$ were jointly trained for 100 epochs, 2nd stage: $\mathbf{w}$ was trained for 100 epochs	1st stage: $\mathbf{e}_1$, $\mathbf{d}_1$, $\mathbf{e}_2$, and $\mathbf{d}_2$ were jointly trained for 200 epochs, 2nd stage: $\mathbf{w}$ was trained for 200 epochs	1st stage: $\mathbf{e}_1$, $\mathbf{d}_1$, $\mathbf{e}_2$, and $\mathbf{d}_2$ were jointly trained for 90000 iterations, 2nd stage: $\mathbf{w}$ was trained for 90000 iterations
Learning rate	1st stage: 10^{-4}, 2nd stage: 10^{-4}	1st stage: 10^{-4}, 2nd stage: 10^{-4}	1st stage: $10^{-2} \times (1 - \eta/90000)^{0.9}$, 2nd stage: $10^{-2} \times (1 - \eta/90000)^{0.9}$, η is the current iteration
Dimension of $\mathbf{z}$	256	256	256
α	10^{-3}	10^{-3}	10^{-3}
β	10^{-1}	10^{-1}	10^{-1}
γ	10^{-8}	10^{-8}	10^{-8}
N	3	3	3
T	15	15	15
Execution manner	5-fold cross validation	5-trial repeats	5-fold cross validation

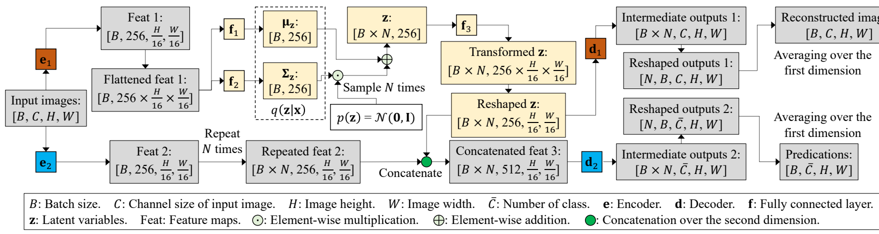

Fig. 2. Network configuration for modeling $p(\mathbf{x}, \mathbf{y}|\mathbf{z})$. For simplicity, specifics of the encoder and decoder layers are excluded, and skip connections are omitted.

Table 5. Illustration of the CholecSeg8k.

Class ID	Object	Color
Class 0	Black background	
Class 1	Abdominal wall	
Class 2	Liver	
Class 3	Gastrointestinal tract	
Class 4	Fat	
Class 5	Grasper	
Class 6	Connective tissue	
Class 7	Blood	
Class 8	Cystic duct	
Class 9	L-hook electrocautery	
Class 10	Gallbladder	
Class 11	Hepatic vein	
Class 12	Liver ligament	

Table 6. Illustration of the AutoLaparo. I: Instrument

Class ID	Object	Color
Class 0	Background	
Class 1	Manipulation of I-1	
Class 2	Shaft of I-1	
Class 3	Manipulation of I-2	
Class 4	Shaft of I-2	
Class 5	Manipulation of I-3	
Class 6	Shaft of I-3	
Class 7	Manipulation of I-4	
Class 8	Shaft of I-4	
Class 9	Uterus	

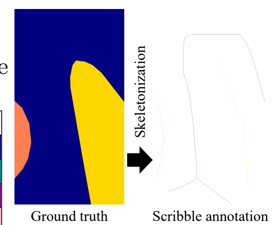

Fig. 3. An example of weak annotation simulation with skeletonization. The white area indicates unlabeled region.

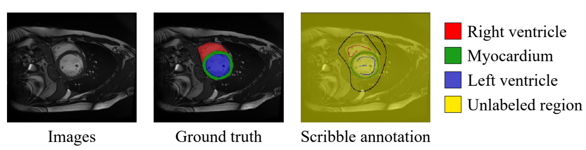

Fig. 4. An example slice of the ACDC dataset.

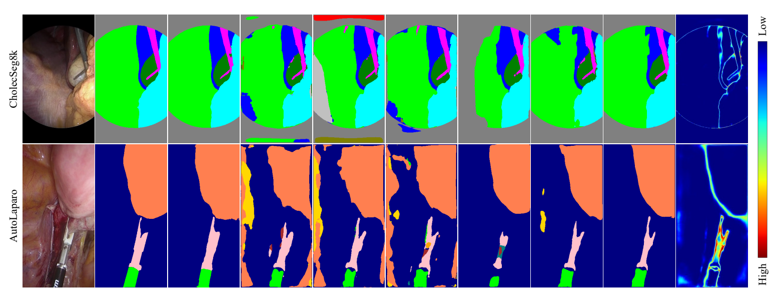

Fig. 5. Visualization results of various methods.