Supplementary: Weakly supervised segmentation with point annotations for histopathology images via contrast-based variational model

Hongrun Zhang†, Liam Burrows†, Yanda Meng, Declan Sculthorpe, Abhik Mukherjee, Sarah E Coupland, Ke Chen, Yalin Zheng*

1. The impact of deep features to CVM

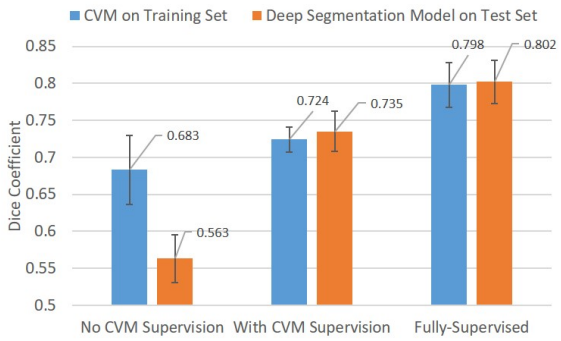

Figure 1. Dice coefficient values of the contrast-based variational model (CVM) on the training set of Camelyon-16 using different deep features. This figure suggests the performance of CVM is related to the quality of the deep features, and the deep segmentation model and CVM are mutually reliant.

We further explore the impact of the learned deep features on the segmentation performance of the contrast-based variational model (CVM). Figure.1 shows the corresponding results on Camelyon-16, in which the blue bar represents the CVM model on the training sets, while the orange bar represents the deep segmentation model on the test set. Note that the CVM is only applied to the training set, and is not applicable to the testing set as there are no input annotated points during the inference period. The three groups (from left to right) represent the deep segmentation models trained without CVM as the complementary supervision, with CVM as the complementary supervision, and the fully-supervised, respectively. From the first two groups it is clear that if the learned features are with better quality that the deep segmentation model achieves higher performance on the test set when complementary supervision from CVM is applied, the performance of CVM on the training set can be accordingly improved (from 0.683 to 0.724 in Dice coefficient). More notably, when the CVM utilizes the deep features from the deep segmentation model trained in a fully-supervised fashion, the corresponding Dice coefficient can increase to 0.798. This observation indicates that the segmentation performance of CVM relies on the quality of the learned deep features, although it can still achieve promising performance on lower-quality deep features.

†: Equal contribution; ∗: Corresponding author

2. Extension to region annotation

As mentioned, point annotation is a special case of region or brushstroke annotation, but is usually faster. In fact, region annotation can be obtained through expanding point click in which an annotated region is centered at the annotated point. The proposed method can be adapted for such an extension, whereas only the CVM needs minor modification. One possible solution is to use the mean value of region features to compute the correlation. Random and center selections of region features are another two alternatives. To further evaluate, extra experiments on Camelyon-16 have been conducted with ground-truths being expanded from points to regions of various sizes (at output dimension,i.e.,256x256) and using center selection. Results in Table.1 show that the region annotation has limited improvement over the point annotation, probably because the tissue within a small region has high morphological similarity, so involving surrounding region labels introduces limited new knowledge.

3. Correlation maps and contrast maps

Figure.2 presents examples of the correlation maps generated based on the pairs of annotated in-target and out-of-target points, and the corresponding contrast maps. As can be seen, except for target regions, there are still comparably strong activations outside the target regions in the in-target correlation maps (third column in Figure.2), which also appear in the same locations in the out-of-target correlation maps. With the subtraction operation to generate the contrast maps, these undesired activations can be removed or alleviated in the resulting contrast maps.

Table 1. Dice coefficient values on Camelyon-16

label region size	1x1 (point)	3x3	5x5
Partial CE	0.563±0.032	0.571±0.058	0.590±0.086
Ours	0.735±0.027	0.736±0.038	0.740±0.035

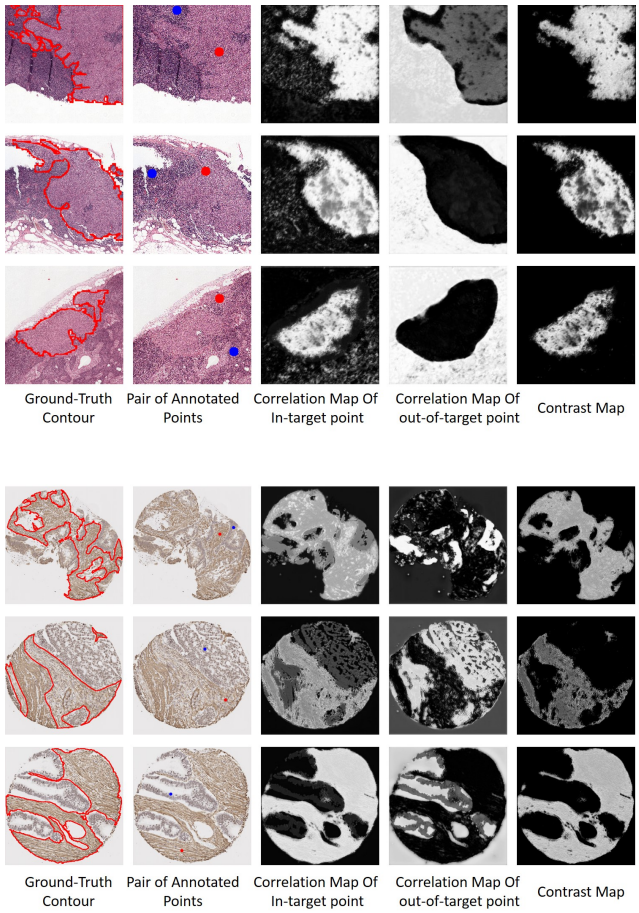

Figure 2. The pair of annotated in-target (red dot) and out-of-target (blue dot) points, and their correlation maps and the corresponding contrast map. The contrast maps will have fewer false activations than the correlation maps.

4. Numerical scheme

The non-linear partial differential equations (PDE) of the main paper is solved using an Additive Operator Splitting scheme (AOS). To do this, we rewrite the PDE as:

$$\frac{\partial u_p}{\partial t} = \Big(\partial_{x_1}(G\partial_{x_1}u) + \partial_{x_2}(G\partial_{x_2}u)\Big) - f_p,$$

where $G = G(u) = \frac{g(|\nabla z_p|}{|\nabla u|_\epsilon}$, with $g(s) = \frac{1}{1+\iota s^2}$ as an edge detector, $|\nabla u|_\epsilon = \sqrt{(\partial_{x_1}u)^2 + (\partial_{x_2}u)^2 + \epsilon}$ with ϵ a small parameter to ensure numerical stability, and $f_p = \lambda\big((z_p - c_{1,p})^2 - (z_p - c_{2,p})^2\big)$. Here we use x_1 and x_2 to denote each spatial direction. Discretizing this system results in the following:

$$\frac{u^{k+1} - u^k}{\tau} = \sum_{\ell=1}^{2} A_\ell(u^k)u^{k+1} - f_p,$$

where τ is a time-step parameter and A_ℓ is a discretized version of $\partial_{x_\ell}(G(u^k)\partial_{x_\ell}u^{k+1})$. This can be rearranged to:

$$u^{k+1} = \big(I - \tau\sum_{\ell=1}^{2} A_\ell(u^k)\big)^{-1}\big(u^k - \tau f_p\big),$$

which can be rearranged to give the final AOS scheme [5]:

$$u^{k+1} = \frac{1}{2}\sum_{\ell=1}^{2}\big(I - 2\tau A_\ell(u^k)\big)^{-1}\big(u^k - \tau f_p\big),$$

where $k = 0, ..., N$ denotes iterations. More information can be found in [5].

5. Contrast maps vs CVM results

Figure.3 shows examples of the mean contrast map and the CVM result, where a mean contrast map is defined as the mean value of all contrast maps of an image w.r.t. the annotated point pairs, i.e., $\frac{\sum_p z_p}{|P|}$. We can see that the mean contrast maps can already present the rough shapes of the target regions, yet they are not very regionally consistent and have rugged boundaries. On the contrary, the CVM results are highly consistent in regions and have smoother boundaries. The phenomenon is also in accordance with the performance results shown in Table.1 in the main paper, where the proposed method, which utilizes CVM results as the complementary supervision, outperforms the method that uses the mean contrast maps as the complementary supervision (the 'Supervised By Contrast Map').

6. Comparisons with conventional variational selective segmentation models

We also evaluate the proposed CVM with recent selective variational segmentation methods on the training sets, namely models composed of: Chan-Vese with Euclidean distance (CVE) [4], Chan-Vese with geodesic distance (CVG) [2], reformulated Chan-Vese with geodesic distance [3], and Mumford-Shah based with Euclidean distance [1] (MSE).

Table 2. Performance comparison with conventional variational methods on the training sets.

	Camelyon16			
	Dice Coefficient	Accuracy	Cohen's Kappa	AUC
CVE [4]	0.158 ± 0.235	0.699 ± 0.256	0.115 ± 0.196	0.560 ± 0.102
CVG [2]	0.071 ± 0.111	0.610 ± 0.250	-0.004 ± 0.126	0.493 ± 0.076
RCVG [3]	0.681 ± 0.187	0.825 ± 0.193	0.575 ± 0.251	0.812 ± 0.112
MSE [1]	0.501 ± 0.247	0.597 ± 0.254	0.130 ± 0.248	0.594 ± 0.151
CVM	0.724 ± 0.017	0.896 ± 0.019	0.625 ± 0.042	0.923 ± 0.023
	Colorectal tissue cores			
	Dice Coefficient	Accuracy	Cohen's Kappa	AUC
CVE [4]	0.315 ± 0.123	0.757 ± 0.125	0.218 ± 0.139	0.590 ± 0.056
CVG [2]	0.249 ± 0.112	0.734 ± 0.136	0.154 ± 0.141	0.564 ± 0.060
RCVG [3]	0.193 ± 0.136	0.693 ± 0.139	0.051 ± 0.092	0.524 ± 0.046
MSE [1]	0.426 ± 0.136	0.532 ± 0.279	0.100 ± 0.189	0.564 ± 0.060
CVM	0.680 ± 0.035	0.815 ± 0.047	0.557 ± 0.043	0.891 ± 0.016

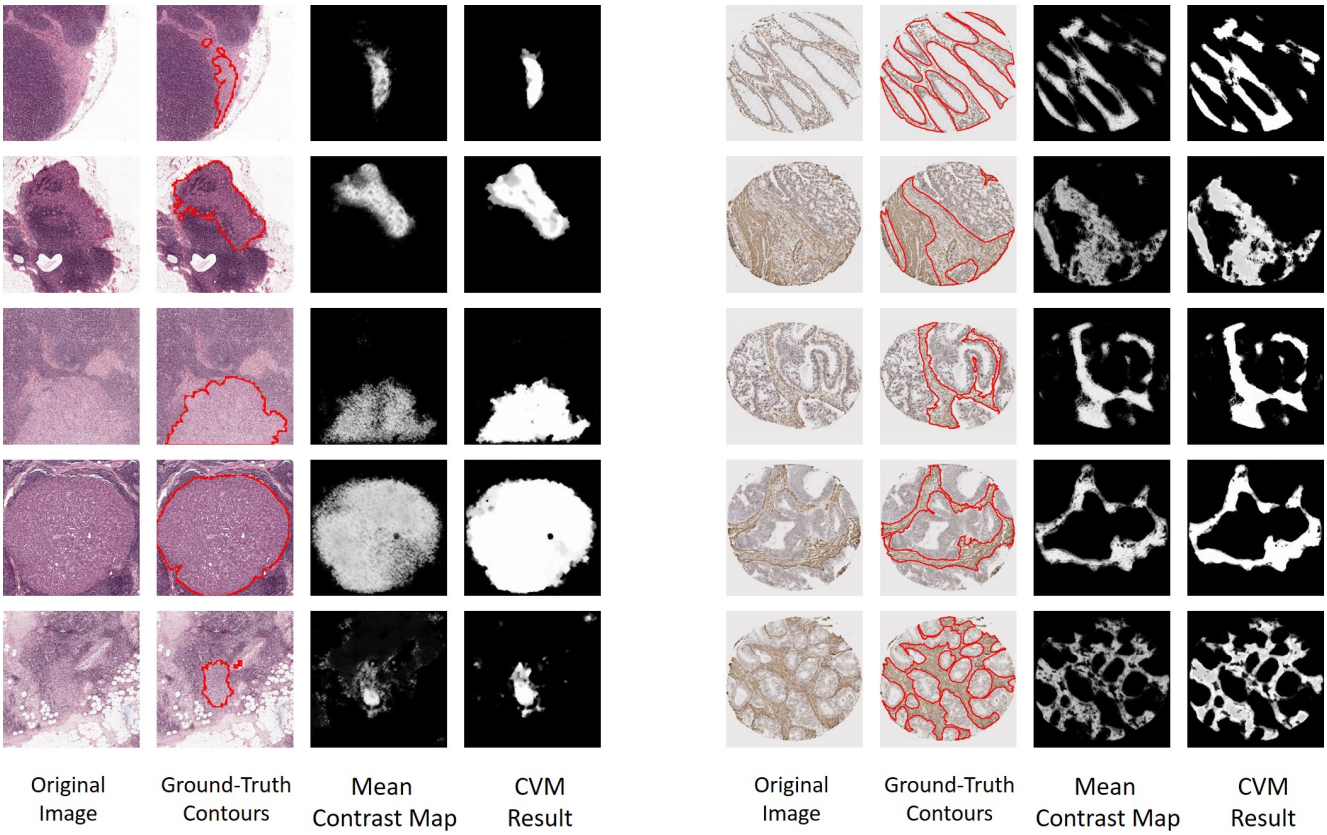

Figure 3. Examples of contrast maps and the CVM results from Camelyon-16 (left) and from colorectal tissue cores (right).

In Table.2 detailed performance comparisons of the CVM with the selective segmentation of variational methods are presented. Clearly, the proposed CVM outperforms the traditional variational methods in all performance metrics. This is due to a number of reasons. The Chan-Vese fidelity used in [2, 4] is insufficient to capture details in the images used, and in all four of the competitors, the geometric constraints are not sophisticated enough to effectively restrict the domain to the region of interest only. Qualitative examples (heatmaps) are provided in Figure.5, which show the proposed CVM can produce more regional consistent segmentation results with smoother boundaries in

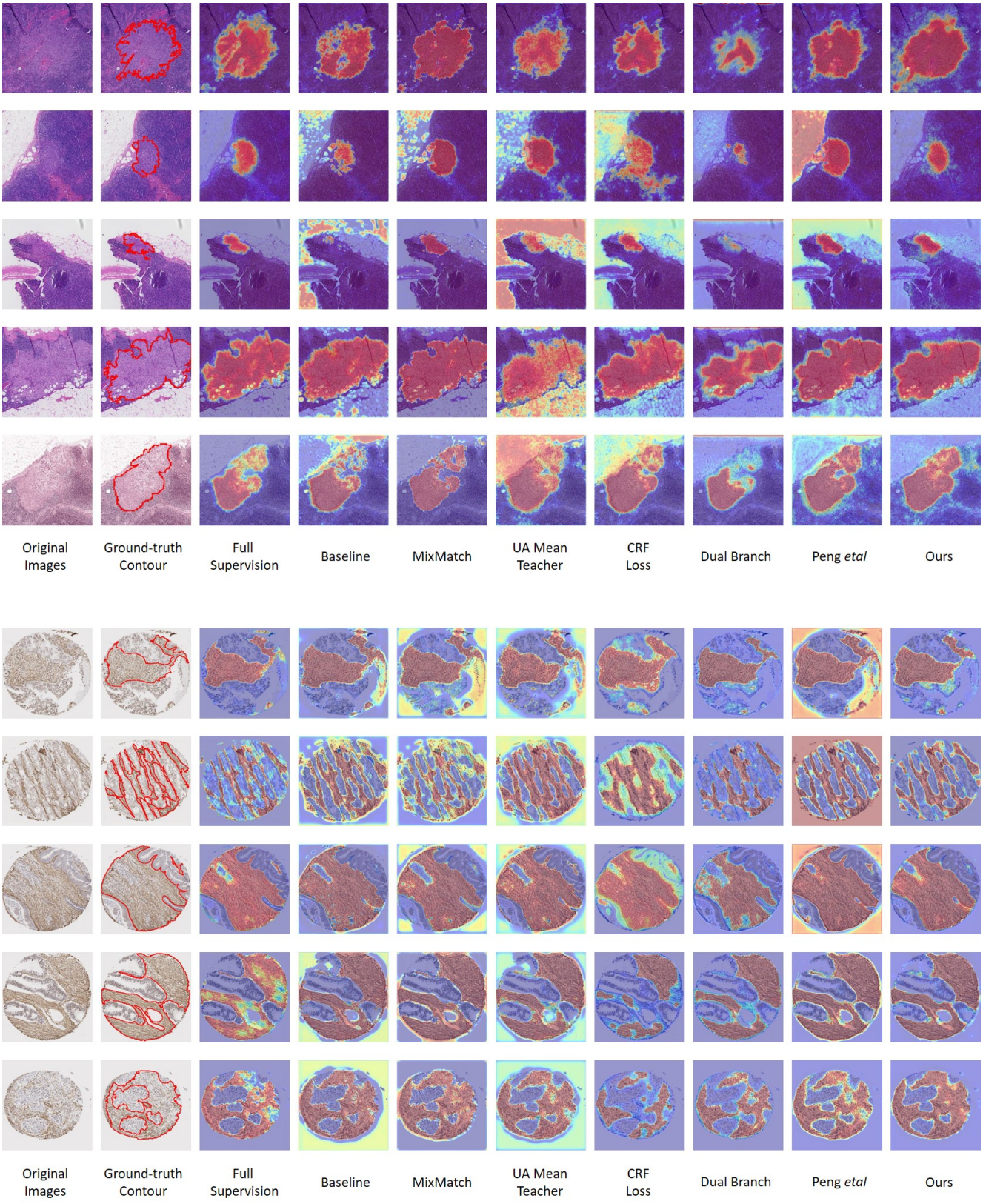

Figure 4. Heatmaps of different methods. Top: examples from the Camelyon-16; Bottom: examples of the colorectal tissue cores.

comparison to other conventional variational methods. Notably, RCVG [3] also achieves promising performances on the Camelyon-16 datasets, and the corresponding qualitative results also show its high segmentation capability, although its performances on the colorectal tissue cores are much more inferior to the proposed CVM.

7. More qualitative results

More qualitative results of heatmaps in addition to the Figure.3 of the main paper are shown in Figure.4. As can be seen, the dual branch method also presents promising segmentation ability in the colorectal tissue core images, yet it tends to shrink the segmentation.

8. Limitation

The time consumption is considerable for the solvers of variational methods, as mentioned in the main paper, and at present there is no GPU version solver to speed up the calculation process. However, the calculation process of the proposed CVM is conducted only during the training, and the trained deep segmentation model can be used as a regular deep learning model in the inference period. In addition, images in the training set can be processed by multiple CPU cores in parallel, which substantially reduces the computational time for CVM.

As mentioned in Section.1, the segmentation performance of CVM is related to the quality of the learned deep features from the deep segmentation model, and the CVM may fail in some cases if the corresponding learned deep features are not relevant enough for rare cases. Figure.6 shows some failed examples of CVM. Particularly, in some cases of colorectal tissue cores, some target regions (stroma) have sparse tissue and light colors, being similar to the white background, and the CVM model cannot well recognize them as the target regions. Besides, for some target regions with very rugged boundaries, the CVM may generate segmentation results with over-smooth boundaries.

It is also notable that in the shown failed cases on the Camelyon-16, some other conventional variational methods, especially RCVG, can still generate comparably good segmentation results, which inspires us in future work to incorporate the strengths of these conventional variational methods into the proposed framework for better segmentation performance.

References

[1] Chunxiao Liu, Michael Kwok-Po Ng, and Tieyong Zeng. Weighted variational model for selective image segmentation with application to medical images. *Pattern Recognition*, 76:367–379, 2018. 2, 3

[2] Michael Roberts, Ke Chen, and Klaus L Irion. A convex geodesic selective model for image segmentation. *Journal of Mathematical Imaging and Vision*, 61(4):482–503, 2019. 2, 3

[3] Michael Roberts and Jack Spencer. Chan–vese reformulation for selective image segmentation. *Journal of Mathematical Imaging and Vision*, 61(8):1173–1196, 2019. 2, 3, 5

[4] Jack Spencer and Ke Chen. A convex and selective variational model for image segmentation. *Communications in Mathematical Sciences*, 13(6):1453–1472, 2015. 2, 3

[5] Joachim Weickert, BM Ter Haar Romeny, and Max A Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998. 2

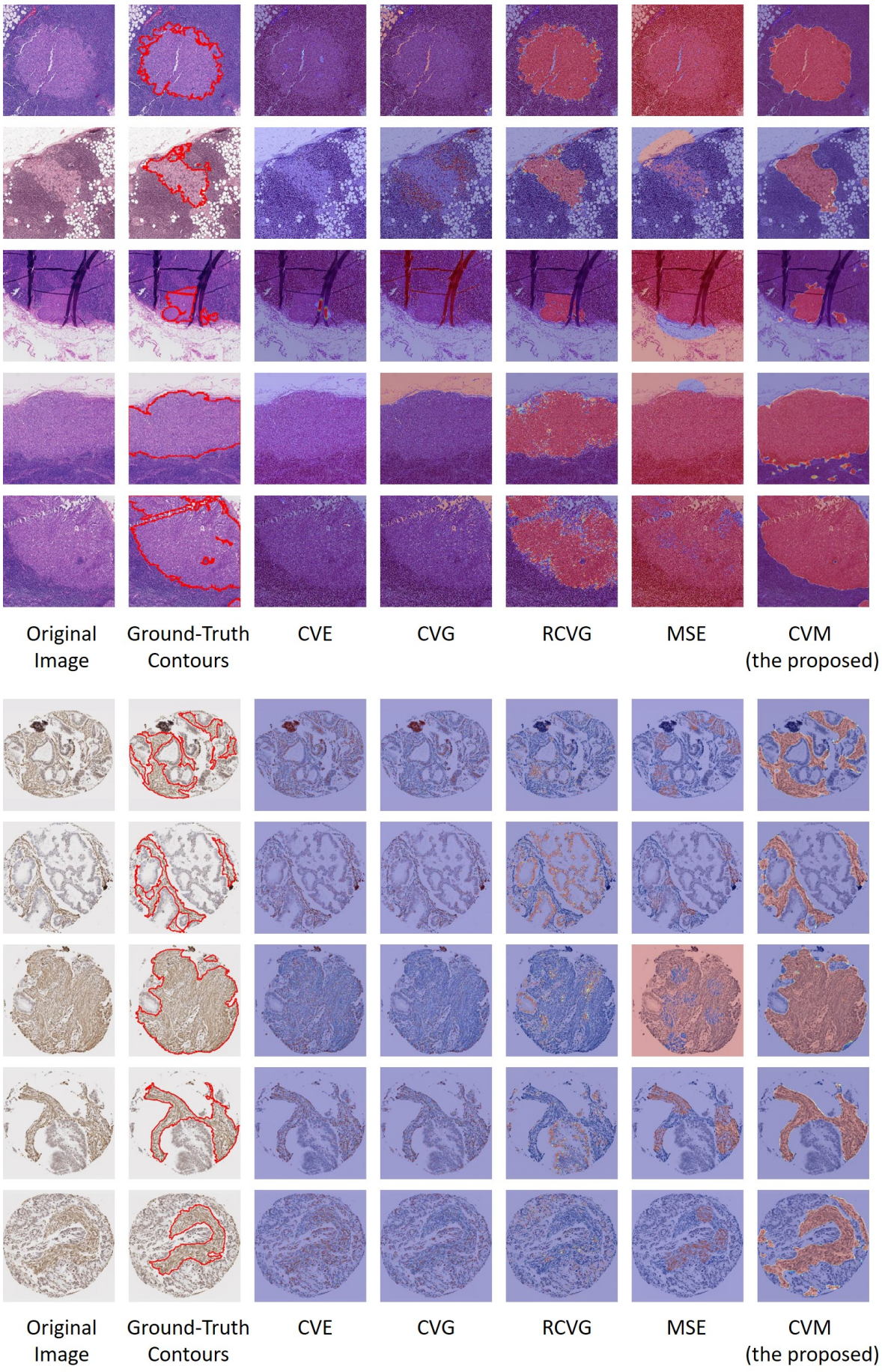

Figure 5. Heatmaps of different variational selective segmentation models on the training sets.

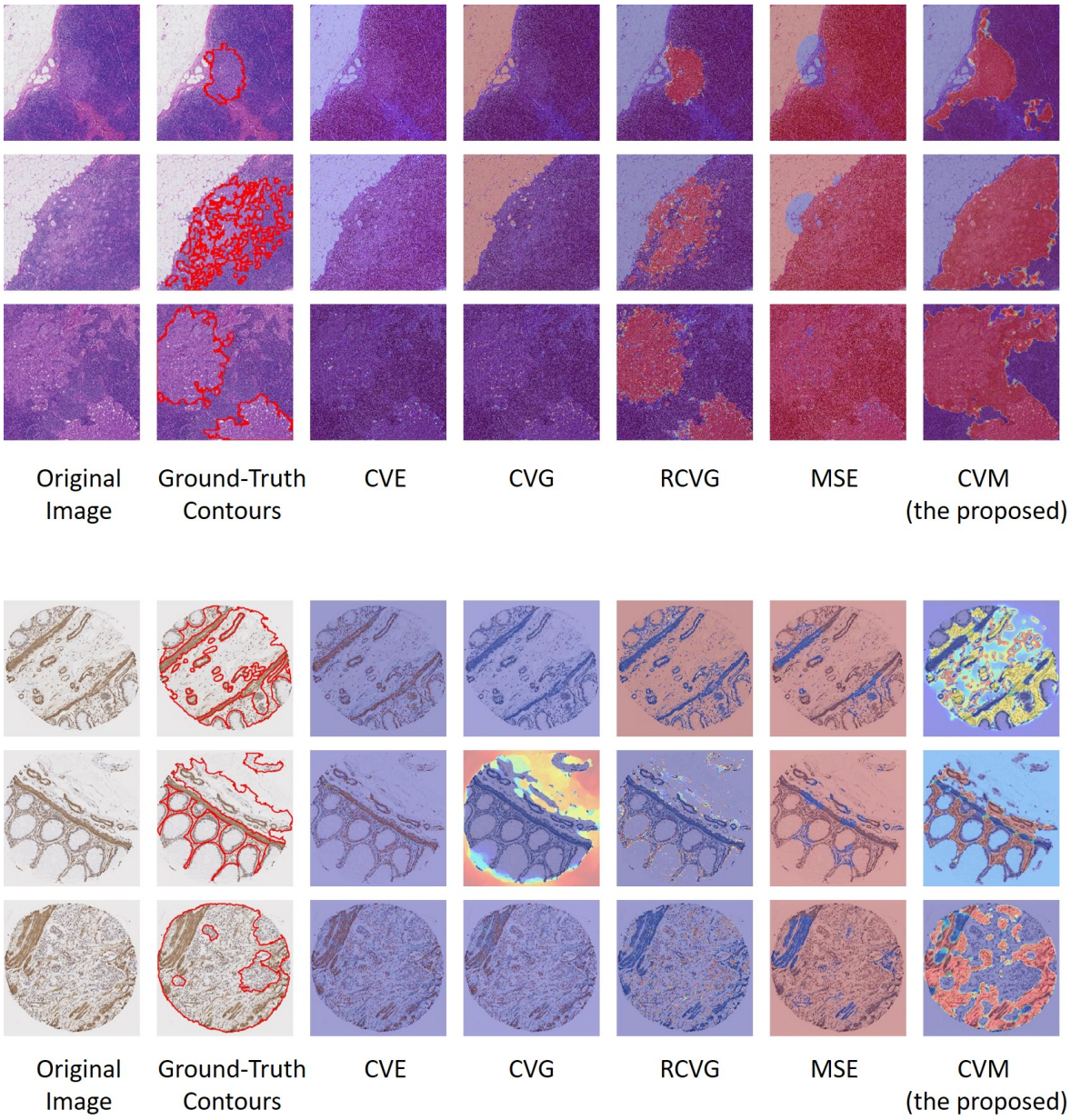

Figure 6. Examples of failed cases of CVM.