Leveraging Uncertainty for Deep Interpretable Classification and Weakly-Supervised Segmentation of Histology Images

Soufiane Belharbi[1] SOUFIANE.BELHARBI.1@ENS.ETSMTL.CA
Jérôme Rony[1] JEROME.RONY.1@ENS.ETSMTL.CA
Jose Dolz[2] JOSE.DOLZ@ETSMTL.CA
Ismail Ben Ayed[1] ISMAIL.BENAYED@ETSMTL.CA
Luke McCaffrey[3] LUKE.MCCAFFREY@MCGILL.CA
Eric Granger[1] ERIC.GRANGER@ETSMTL.CA
[1] *LIVIA, Dept. of Systems Engineering, ÉTS, Montreal, Canada*
[2] *LIVIA, Dept. of Computer Engineering, ÉTS, Montreal, Canada*
[3] *Goodman Cancer Research Centre, Dept. of Oncology, McGill University, Montreal, Canada*

Editors: Under Review for MIDL 2022

Abstract

Trained using only image class label, deep weakly supervised methods allow image classification and ROI segmentation for interpretability. Despite their success on natural images, they face several challenges over histology data where ROI are visually similar to background making models vulnerable to high pixel-wise false positives. These methods lack mechanisms for modeling explicitly non-discriminative regions which raises false-positive rates. We propose novel regularization terms, which enable the model to seek both non-discriminative and discriminative regions, while discouraging unbalanced segmentations and using only image class label. Our method is composed of two networks: a localizer that yields segmentation mask, followed by a classifier. The training loss pushes the localizer to build a segmentation mask that holds most discrimiantive regions while simultaneously modeling background regions. Comprehensive experiments[1] over two histology datasets showed the merits of our method in reducing false positives and accurately segmenting ROI.

Keywords: Deep Weakly-Supervised Learning, Image Classification, Semantic Segmentation, Histology Images, Interpretability

1. Vulnerability of weakly-supervised methods to pixel-wise false positives over histology data

Weakly-supervised learning (WSL) has seen a lot success in different applications mainly over natural images. However, recent study over histology data (Rony et al., 2019) showed that these methods yield high pixel-wise false positives highlighting the difficulty of transferring global labels to pixel-level in this data. This is mainly caused by two factors: 1) strong visual similarity of ROI and background in histology images making it difficult to spot ROI. 2) since WSL methods are trained to maximize their class confident, ROI segmentation is left free. Without any constraints, the model can yield an arbitrary segmentation that maximizes classification. Both factors combined lead to over-segmentation, hence high false positive. In this work, we propose to constrain ROI to be most discriminative while simultaneously modeling background regions. This increases the model awareness to the presence of background and reduces over segmentation, therefore reducing false positives.

1. Publicly available code at: https://github.com/sbelharbi/deep-wsl-histo-min-max-uncertainty.

2. Proposed method

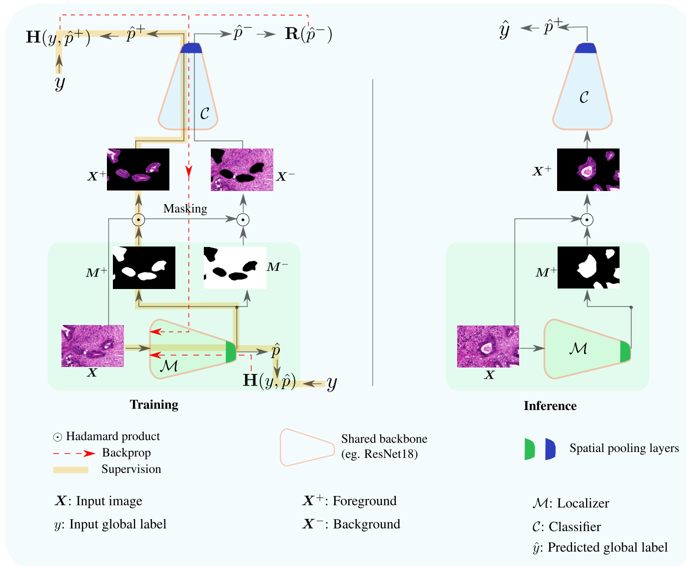

Figure 1: Illustration of our approach.

Fig.1 depicts training and evaluation of our proposal. It aims to simultaneously segment foreground (FG) and background (BG) regions. To do so, we constrain the classifier response p^+ over FG to be more certain with respect to the true class p by minimizing cross-entropy $\min \mathbf{H}(p, \hat{p}^+)$. The BG is constrained to have high uncertainty with respect to the class caused by the expected absence of discriminative regions. This is achieved either by explicitly maximizing entropy $\max -\mathbf{H}(\hat{p}^-)$, or, by matching a surrogate distribution with the highest uncertainty that is uniform distribution q, $\min \mathbf{H}(q, \hat{p}^-)$. To avoid that one region dominates the other, we constrain the model to find the *largest* FG and BG regions by imposing size constraints formulated through a log-barrier method.

$$\min_{\boldsymbol{\theta}_\mathcal{C}} \quad \mathbf{H}(p, \hat{p}^+) + \lambda\, \mathbf{R}(\hat{p}^-) - \frac{1}{t}\left[\log \boldsymbol{s}^+ + \log \boldsymbol{s}^-\right], \tag{1}$$

where $\mathbf{R}(\hat{p}^-) = -\mathbf{H}(\hat{p}^-)$; or $\mathbf{R}(\hat{p}^-) = \mathbf{H}(q, \hat{p}^-)$, λ is a balancing positive scalar, and $t > 0$ is a parameter that determines the accuracy of the approximation of the barrier method. We define the size of each mask as: $\boldsymbol{s}^+ = \sum_{z\in\Omega} \boldsymbol{M}^+(z)$, $\boldsymbol{s}^- = \sum_{z\in\Omega} \boldsymbol{M}^-(z)$, where Ω is the spatial image domain. We use SGD to optimize Eq.1 which trains simultaneously the localizer and classifier in an end-to-end manner (Fig.1).

3. Experiments

We evaluated our method over two public histology benchmarks: GlaS[2] dataset for colon cancer, and patch based Camelyon16[3] for breast cancer. For evaluation, we consider classification error for image class, and Dice index (F1) over foreground ($\mathrm{F1}^+$) and background

2. GlaS dataset: https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest
3. Camelyon16 dataset: https://camelyon16.grand-challenge.org. Patches: (Rony et al., 2019).

(F1$^-$). Quantitative results are presented in Tab.1 which show that our method has better segmentation of ROI and low false positives. Visual results are presented in Fig2.

Table 1: Image classification and pixel-level segmentation performances on the GlaS and Camelyon16 test sets. Cl: classification. The best performance is shown in bold.

	GlaS			Camelyon16-P512		
Method	**Image level** Cl. error (%)	**Pixel level** F1$^+$ (%)	F1$^-$ (%)	**Image level** Cl. error (%)	**Pixel level** F1$^+$ (%)	F1$^-$ (%)
All-ones (Lower-bound)	−−	66.01	00.00	−−	59.44	00.00
PN	−−	65.52	24.08	−−	31.15	37.36
ERASE	7.50	65.60	25.01	8.61	31.30	42.48
Max-pooling	1.25	66.00	26.32	10.06	48.28	81.92
CAM-LSE	1.25	66.05	27.93	1.51	64.31	63.78
Grad-CAM	**0.00**	66.30	21.30	2.40	62.78	79.05
GAP	**0.00**	66.90	17.88	2.40	62.75	79.05
WILDCAT	1.25	67.21	22.96	**1.48**	62.73	72.59
Deep MIL	**2.50**	68.52	41.34	1.93	59.01	36.94
Ours (EEM)	**0.00**	**72.11**	69.07	6.26	67.98	**88.80**
Ours (SEM)	**0.00**	71.94	**69.23**	6.95	**68.26**	88.55
U-Net (Upper-bound)	−−	90.19	88.52	−−	71.11	89.68

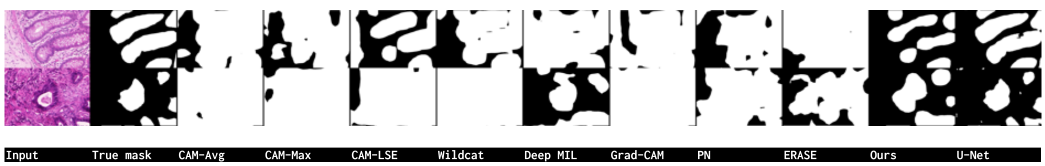

Figure 2: **GlaS dataset**: Predicted binary mask on GlaS test images.

4. Conclusion

Explicitly incorporating background prior in deep WSL methods has shown to reduce false positives with large margin over histology data. This was achieved by reducing over-segmentation, a common issue in WSL techniques over histology images, and in turn helps yielding accurate ROI segmentation.

Acknowledgments

This research was supported in part by Canadian Institute of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC) and Compute Canada.

References

J. Rony, S. Belharbi, J. Dolz, I. Ben Ayed, L. McCaffrey, and E. Granger. Deep weakly-supervised learning methods for classification and localization in histology images: a survey. *CoRR*, abs/1909.03354, 2019.