Uncertainty Driven Multi-Loss Fully Convolutional Networks For Histopathology

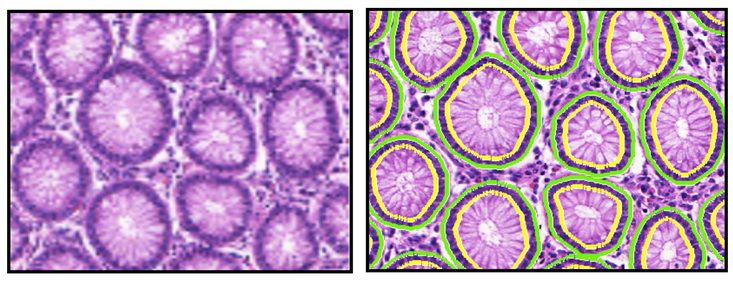
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Aïcha BenTaieb and Ghassan Hamarneh
abentaie@sfu.ca hamarneh@sfu.ca

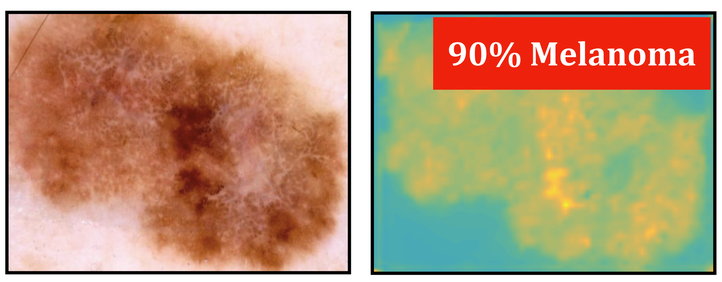
MICCAI Quebec City CONFERENCE 10.-14. SEPTEMBER 2017
LABELS 2017 MICCAI Workshop on Large-scale Annotation of Biomedical data and Expert Label Synthesis

Uncertainty In Medical Image Analysis

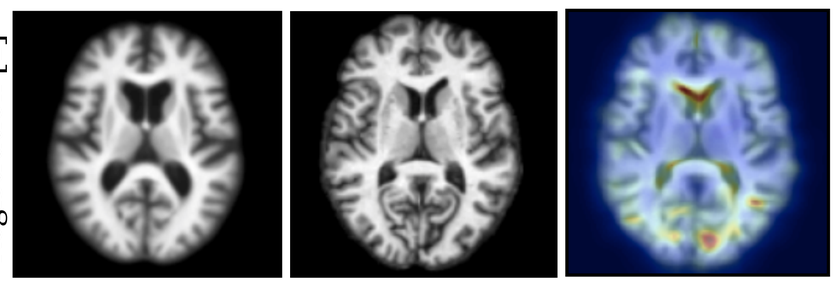
Segmentation



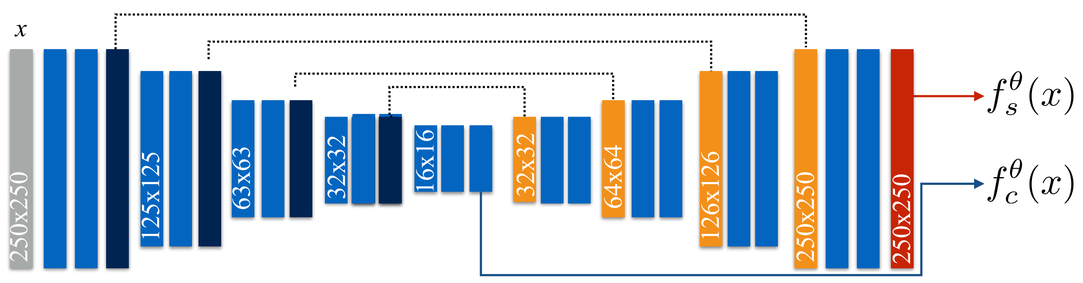
Classification



Registration [1]



State of the art deep learning models .... can encode flexible priors as multi loss functions [2]



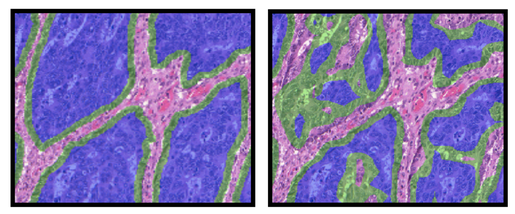
$$\mathcal{L}_{total}(x;\theta) = \sum_{i=1}^{T} \lambda_i \mathcal{L}_i(x;\theta)$$

Generally, set via grid search

Could be set based on model's uncertainty [3]?

Proposed Method

$\mathcal{L}_i$= Segmentation Loss



Uncertainty λ_i Uncertainty λ_i

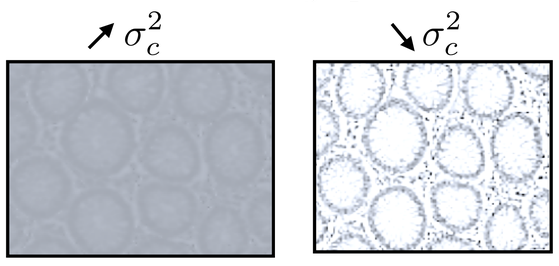
Uncertainty Driven Multi Loss Function

$$\mathcal{L}_{total}(x;\theta,\sigma_c,\sigma_s,\sigma_t,\sigma_g) = \mathcal{L}_c(x;\theta,\sigma_c) + \mathcal{L}_s(x;\theta,\sigma_s) + \mathcal{L}_t(x;\theta,\sigma_t) + \mathcal{L}_g(x;\theta,\sigma_g)$$

Per-Term Uncertainty | Classification | Segmentation | Topology | Geometry

Uncertainty-Based Predictions

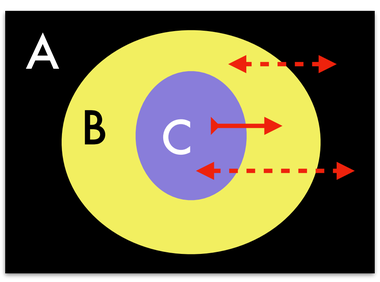
$$P(C_k = 1|x,\theta,\sigma_c) = \frac{\exp(\frac{1}{\sigma_c^2} f_{c_k}^\theta(x))}{\sum_{k'=1}^{K} \exp(\frac{1}{\sigma_c^2} f_{c_{k'}}^\theta(x))}$$



Uniform $f_{c_k}^\theta(x)$ Non-Uniform $f_{c_k}^\theta(x)$

Multi-Region Interactions

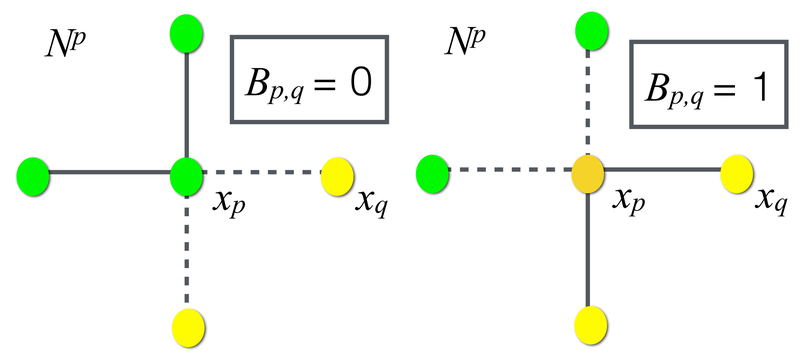
A ⇠ - - Excludes - - ⇢ B B ⟵ Contains ⟶ C



A	1	0	0
B	0	1	1
C	0	0	1
$V(S_p)$	1	1	1
S_p	**A**	**B**	**C**

$$P_t(S_p^r|x_p;\theta) = \frac{1}{Z} V(S_p) \prod_{r=1}^{L} \exp\left(\frac{1}{\sigma_t^2} f_{s_r}^\theta(x_p)\right) \times S_p^r$$

Smooth Object Boundaries



Connected neighbours with different labels.

Connected neighbours sharing a label.

$$\mathcal{L}_c(x;\theta,\sigma_c) = \frac{1}{\sigma_c^2} \sum_{k=1}^{K} -C_k \log P(C_k = 1|x_p;\theta) + \log \sigma_c^2$$

$$\mathcal{L}_t(x;\theta,\sigma_t) = \frac{1}{\sigma_t^2} \sum_{p\in\Omega} \sum_{r=1}^{L} -S_p^r \log P_t(S_p^r = 1|x,\theta) + \log \sigma_t^2$$

$$\mathcal{L}_s(x;\theta,\sigma_s) = \frac{1}{\sigma_s^2} \sum_{p\in\Omega} \sum_{r=1}^{L} -S_p^r \log P(S_p^r = 1|x,\theta) + \log \sigma_s^2$$

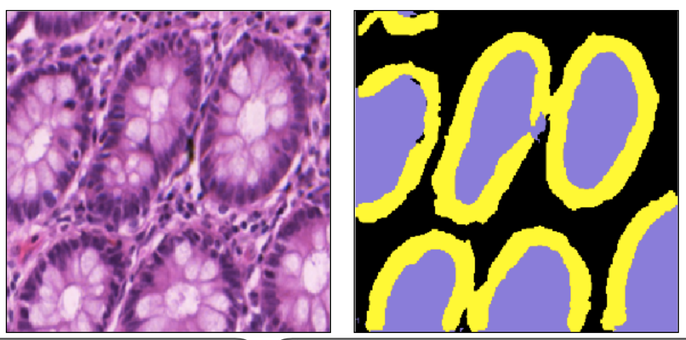
$$\mathcal{L}_g(x;\theta,\sigma_g) = \frac{1}{\sigma_g^2} \sum_{p\in\Omega} \sum_{r=1}^{L} \sum_{q\in\mathcal{N}^p} S_p^r \left| \log \frac{P_t(S_p^r|x_p;\theta)}{P_t(S_q^r|x_q;\theta)} \right| B_{p,q} + \log \sigma_g^2$$

Validation And Performance

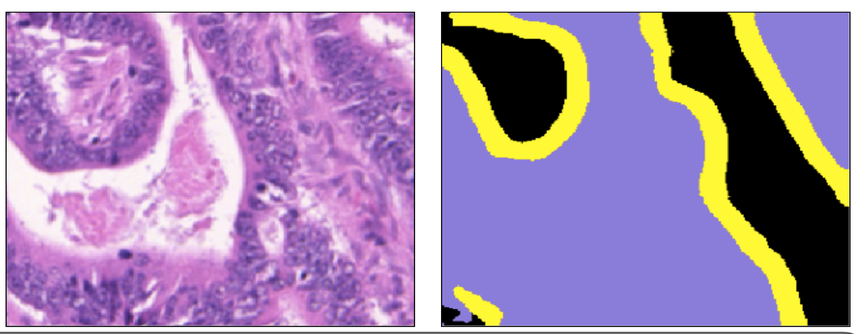
Dataset

- Warwick-QU Colon Adenocarcinoma Dataset [4]
- 70 Training / 15 Validation / 80 Test Images
- Image and Pixel-Level Annotation
- All models were randomly initialized

Benign

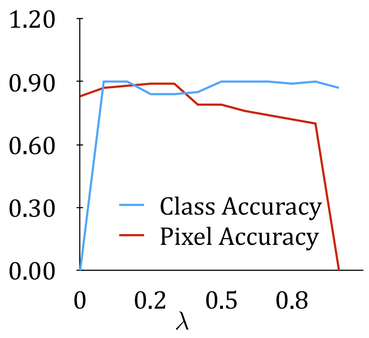


Malignant



Multi Loss vs Single Loss

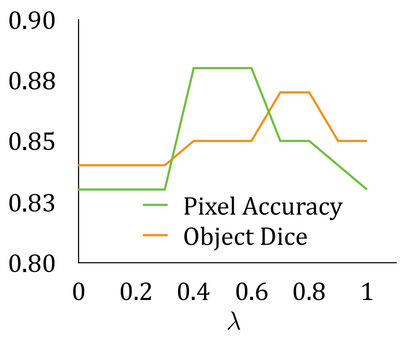
$\mathcal{L}_{\text{total}} = \lambda\mathcal{L}_c + (1-\lambda)\mathcal{L}_s$



+3% classification accuracy
+6% pixel accuracy

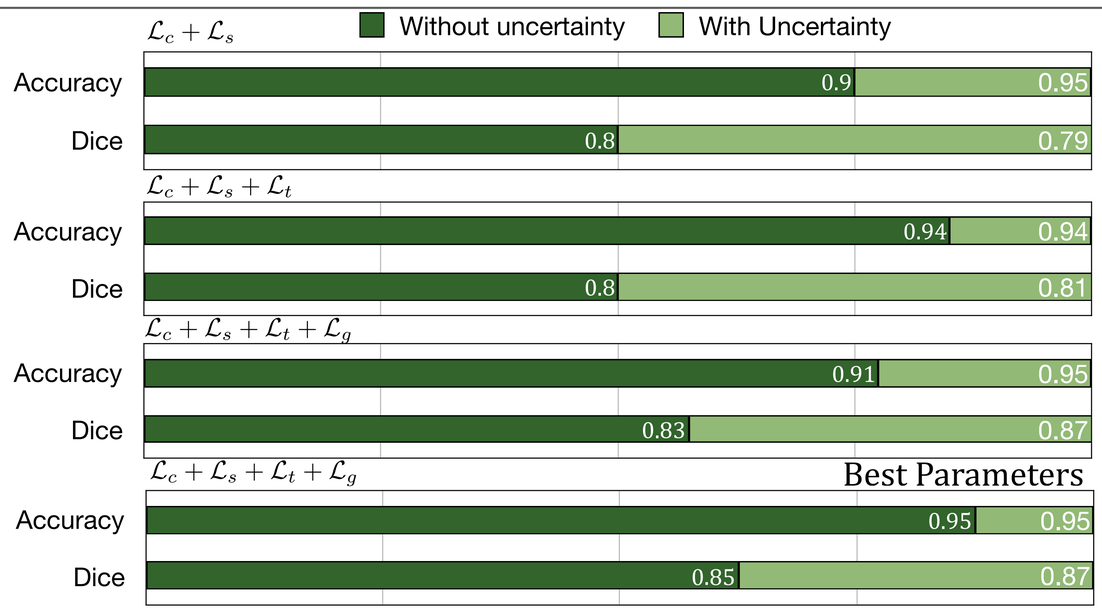
Penalty Terms Trade-Off

$\mathcal{L}_{\text{total}} = \mathcal{L}_s + \lambda\mathcal{L}_t + (1-\lambda)\mathcal{L}_g$



Minimal change in performance when λ varies by +/- 20%

Uncertainty Driven Trade-Off



+2% Object Dice when using uncertainty driven loss vs extensive grid search

[1] Yang et al. Fast Predictive Image Registration, LABELS 2016
[2] BenTaieb et al. Topology Aware Fully Convolutional Networks for Histology gland Segmentation, MICCAI'16
[3] Kendall et al. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, arXiv 2017
[4] Sirinukunwattana et al. Gland Segmentation in colon histology images: the GlaS contest, arXiv 2016

Acknowledgements: We thank NSERC for funding and NVIDIA Corporation for the donation of a TITAN X GPU.