Stain Based Contrastive Co-training for Histopathological Image Analysis

Bodong Zhang[1,2], Beatrice Knudsen[3], Deepika Sirohi[3], Alessandro Ferrero[2], Tolga Tasdizen[1,2]

1 Department of Electrical and Computer Engineering, 2 Scientific Computing and Imaging Institute, 3 Huntsman Cancer Institute
University of Utah, Salt Lake City, UT, USA

MICCAI2022 Singapore
MILLanD workshop

Introduction

- Deep learning models are widely used in histopathology image classification.
- Expert annotation at tile level for providing labels in model training has very high cost and is infeasible beyond a small number of whole slide images (WSIs).
- Semi-supervised learning (SSL) utilizes unlabeled data in training combined with limited amount of labeled data.
- Co-training approach to SSL achieves excellent results when multiple conditionally independent views of each sample are available, and each view is able to support accurate classification on its own.
- We explored if H&E slides can be color deconvoluted to Hematoxylin (H) and Eosin (E) stain images to fulfill co-training's view requirements.
- A novel contrastive co-training model with H and E views was tested on a clear cell renal cell carcinoma (ccRCC) dataset and a prostate cancer dataset.
- Our co-training model always had the best performance over other state-of-the art SSL methods in both datasets, including consistency regularization, MixMatch and FixMatch.

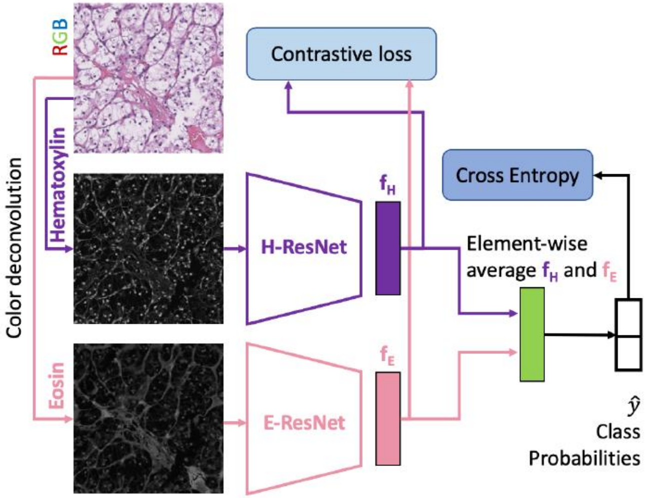
Model design

Stain separation: $\begin{bmatrix} H \\ E \end{bmatrix} = \begin{bmatrix} 1.838 & 0.034 & -0.760 \\ -1.373 & 0.772 & 1.215 \end{bmatrix} \begin{bmatrix} \log_{10} 255/R \\ \log_{10} 255/G \\ \log_{10} 255/B \end{bmatrix}$

Contrastive loss:

$$\mathcal{L}_{c.t.}(x_i) = \max\left(\| f_H(x_i) - f_E(x_i) \|_2 - \| f_H(x_i) - f_E(x_k) \|_2 + m, 0\right)$$

Total loss: $\mathcal{L} = \sum_{x_j \in L} y_j \log \hat{y}_j + \lambda \sum_{x_i \in L \cup U} \mathcal{L}_{c.t.}(x_i)$



Experiment settings

- In clear cell renal cell carcinoma (ccRCC) dataset*:
  - ✓ Histologic growth pattern (HGP) tiles were cropped from expert annotated polygons from 53 WSIs.
  - ✓ Tiles from same patient are only in same set.
  - ✓ 10% tiles set as labeled data from training polygons in SSL.
  - ✓ Performed nested vs. diffuse (non-nested) classification and compared with other SSL models.
- In prostate cancer dataset*:
  - ✓ 5% tiles set as labeled data from training polygons in SSL.
  - ✓ Performed benign vs. cancer classification and compared with other SSL models.
- Randomly selected labeled data 5 times to calculate mean and standard deviation of classification accuracies.

*Experiments and research supported by Computational Oncology Research Initiative (CORI) at the Huntsman Cancer Institute, ARUP Laboratories and the Department of Pathology at University of Utah.

Experiment results

ccRCC Model	Test Accuracy	Prostate Model	Test Accuracy
100% label RGB ResNet	$84.8 \pm 2.4\%$	100% label RGB ResNet	$77.5 \pm 2.5\%$
100% label H/E co-train	$\mathbf{92.0 \pm 2.6\%}$	100% label H/E co-train	$\mathbf{79.1 \pm 2.0\%}$
10% label RGB ResNet	$76.9 \pm 5.9\%$	5% label RGB ResNet	$73.4 \pm 1.0\%$
10% label RGB consis	$86.8 \pm 3.3\%$	5% label RGB consis	$74.7 \pm 1.3\%$
10% label RGB MixMatch	$85.9 \pm 5.7\%$	5% label RGB MixMatch	$73.7 \pm 5.0\%$
10% label RGB FixMatch	$88.3 \pm 3.8\%$	5% label RGB FixMatch	$78.2 \pm 3.8\%$
10% label H/E co-train	$\mathbf{92.3 \pm 2.1\%}$	5% label H/E co-train	$\mathbf{78.7 \pm 1.9\%}$

H-only and E-only models' test accuracy for the ccRCC dataset to show each view (H or E) can do classification alone:

Model	Accuracy	Model	Accuracy
100% label H ResNet	$79.4 \pm 3.7\%$	10% label H ResNet	$73.5 \pm 4.0\%$
100% label E ResNet	$94.0 \pm 1.4\%$	10% label E ResNet	$82.3 \pm 7.0\%$

Coefficient of determination(R^2) of image mapping between various channels on ccRCC validation set to show independence between channels:

Experiments	R^2 value	Experiments	R^2 value
$H \Rightarrow E$	0.5223	$E \Rightarrow H$	0.4613
$R \Rightarrow G$	0.8464	$G \Rightarrow R$	0.7833
$R \Rightarrow B$	0.8207	$B \Rightarrow R$	0.7713
$G \Rightarrow B$	0.8522	$B \Rightarrow G$	0.8824

Ablation study on ccRCC (Test set accuracy of ResNet and co-training models taking only 2 channels from RGB as input with 10% labeled data in training):

Model	Accuracy	Model	Accuracy	Model	Accuracy
RB ResNet	$77.5 \pm 6.6\%$	RG ResNet	$80.2 \pm 6.4\%$	GB ResNet	$78.4 \pm 9.7\%$
R/B co-train	$78.2 \pm 4.5\%$	R/G co-train	$79.8 \pm 5.6\%$	G/B co-train	$76.6 \pm 7.3\%$

Conclusion

- The experiments show that our model outperforms all other tested state-of-the-art semi-supervised learning methods in both datasets.
- The experiments prove that the benefit of our model is due to contrastive co-training loss on hidden features combined with more independent H and E view selection.