MISSISSIPPI STATE UNIVERSITY
JAMES WORTH BAGLEY COLLEGE OF ENGINEERING
Department of Industrial and Systems Engineering

Unsupervised Domain Adaptation for Cell Detection on Medical Images

Yibin Wang and Haifeng Wang
Department of Industrial and Systems Engineering, Mississippi State University

Abstract
Stimulated Ramen scattering (SRS) imaging can generate virtually stained mapping images in cells level within live tissues rapidly, which provides a high potential for real-time intelligent medical diagnosis. The objective of this research is to detect the cells on SRS images by transferring the knowledge learned from standard hematoxylin and eosin (H&E) staining images using an unsupervised domain adaption model. The outcomes of this research can be used for cell counting as well as calculating the cell density which could classify tissues in order to assist the surgical decision making.

Methodology
The architecture includes three components:
1) feature extractor, G_f
2) label predictor, G_y
3) domain classifier, G_d

A gradient reversal layer (GRL) multiplies the gradient by a certain negative constant during the backpropagation-based training. The GRL can be defined as R(x) by describing its backpropagation behavior:

$$\frac{dR(x)}{d\boldsymbol{x}} = -\lambda \boldsymbol{I}$$

where $\boldsymbol{I}$ is an identity matrix.
An adversarial loss optimization process aims to:
1) minimize the loss of the label classifier
2) maximize the loss of domain classifier

The loss function of this network can be represented as:

$$E(\theta_f, \theta_y, \theta_d) = \sum_{\substack{i=1\ldots N \\ \boldsymbol{d}_i=0}} L_y\big(G_y\big(G_f(\boldsymbol{x}_i;\theta_f);\theta_y\big), \boldsymbol{y}_i\big) - \lambda \sum_{i=1\ldots N} L_d\big(G_d\big(G_f(\boldsymbol{x}_i;\theta_f);\theta_d\big), \boldsymbol{d}_i\big)$$

$$= \sum_{\substack{i=1\ldots N \\ \boldsymbol{d}_i=0}} L_y^i(\theta_f, \theta_y) - \lambda \sum_{i=1..N} L_d^i(\theta_f, \theta_d)$$

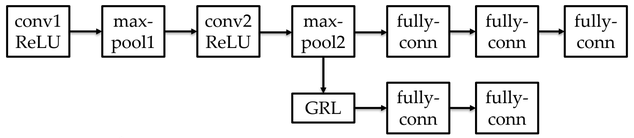
where λ controls the balance between the two objectives in the model, L_y refers to the loss of cell label predictor and L_d refers to the loss of the domain classifier.

$$(\hat{\theta}_f, \hat{\theta}_y) = arg \min_{\theta_f, \theta_y} E(\theta_f, \theta_y, \hat{\theta}_d)$$

$$\hat{\theta}_d = arg \max_{\theta_d} E(\hat{\theta}_f, \hat{\theta}_y, \theta_d)$$

L_y^i and L_d^i represent the corresponding loss functions evaluated at the i-th training example. The parameters $\hat{\theta}_f, \hat{\theta}_y, \hat{\theta}_d$ are optimized using stochastic gradient descent (SGD) solver.

CNN Architecture



Data Preprocessing
A public medical imaging dataset is applied with paired H&E and SRS images of a frozen sectioned specimen.

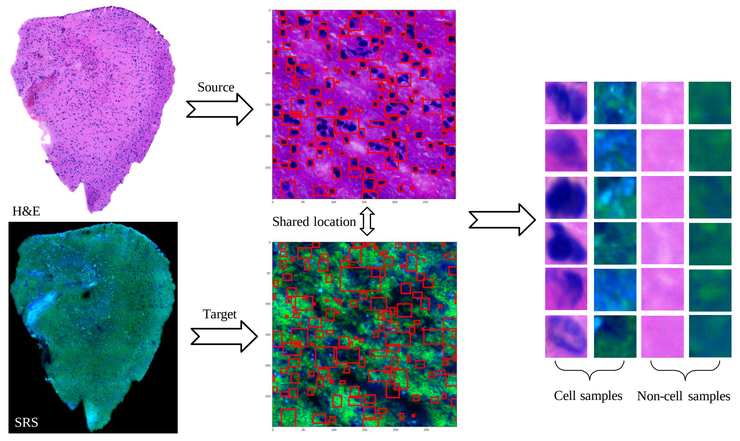


Figure 1: Source and domain samples generated through preprocess

During data preparation, both cell and non-cell samples are extracted from H&E and paired SRS images. Otsu thresholding method is applied to detect the cells in H&E images.

Results
The performance of the model is evaluated with different batch size, i.e., 2, 5, 10, 15, and 20. The results are gathered and summarized in Table 1.

Table 1: Average accuracy and loss calculated with different batch size

Batch size		2	5	10	15	20
Accuracy(%)	Training	72	97.5	99	99.73	**99.75**
	Testing on source	94.17	95	95.83	**99.17**	**99.17**
	Testing on target	37.5	**91.67**	64.1	50.83	50.83
Loss	Label prediction	0.434	0.0874	**0.0047**	0.0056	0.0055
	Source domain	**0.6925**	0.6781	0.6833	0.6872	0.6908
	Target domain	0.6932	**0.7049**	0.7032	0.6829	0.6842
	Total	1.8197	1.4704	1.3912	**1.3757**	1.3805

The best transfer performance on test target domain has been achieved when batch size equals to 5. In this situation, the domain loss remains a relatively high level, and the label loss is low, which both contribute to the learning performance.

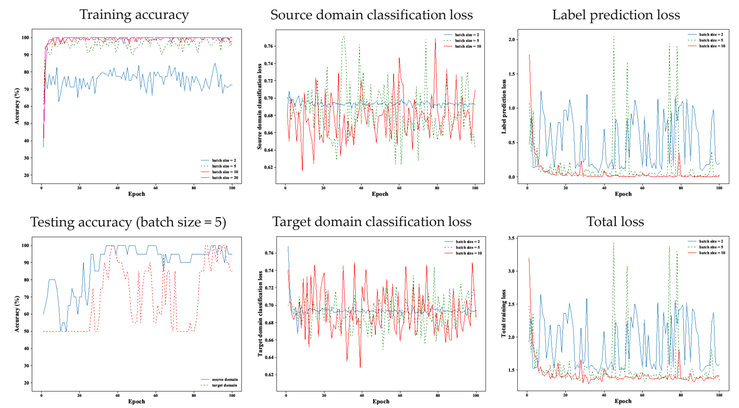


Figure 2: Classification accuracy and loss with different batch size

Conclusions and Future Work
This research studied an unsupervised deep domain adaptation-based model for a cell detection problem in medical images. The model results are tested based on collected samples from different domains, and results suggest that the model is a powerful methodology to classify cell and non-cell samples using transferred knowledge.

Further studies will focus on comparing with other state-of-the-arts domain adaption methods. From model implementation prospective, the performance of the proposed image classification model needs to be further tested and verified with more datasets.

Reference
1. Wang, H., D. Won, and S. W. Yoon. "A deep separable neural network for human tissue identification in three-dimensional optical coherence tomography images." IISE Transactions on Healthcare Systems Engineering 9 (3), 250-271, 2019.
2. Zhang, Q., H. Wang, H. Lu, D. Won, and S. W. Yoon. "Medical image synthesis with generative adversarial networks for tissue recognition." In 2018 IEEE International Conference on Healthcare Informatics (ICHI), 199-207. IEEE, 2018.