Histopathologic Cancer Detection of Lymph Node Patches for Breast Cancer

Gowtham Gundu (gowthamg@stanford.edu), Jad Aboudiab (jadab@stanford.edu)
Department of Computer Science, Stanford University (SCPD)

Stanford Computer Science

Abstract

Manual pathology classification is a time consuming task that requires deep domain expertise by medical board certified pathologists. It can also be an expensive task to have pathologists spend time on classifying even the simplest positive cases, but also the extra time required to spot the difficult cases. Assistive image classification for histopathologic scans that achieves human like or better quality has become an area of deep interest within the medical industry, in the last several years.

This work explores various deep neural network based techniques to come up with a robust way to build a high recall assistive system to detect metastatic cancer by analyzing small hematoxylin and eosin (H&E) stained patch images of lymph node sections. We use a publicly available Kaggle dataset to train, evaluate, and compare various techniques.

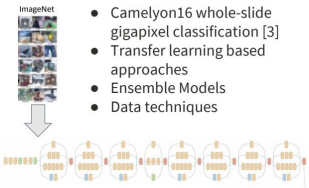
Stanford University

Background

Manual Pathological Assessments

- Time consuming, expensive, and error prone.
- 25% of classifications change upon secondary review [1].
- 38% Precision under time constraint [2].
- **Assistive detection systems have started to become areas of wide interest.**

Prior Work



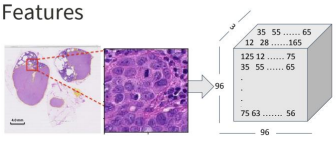
- Camelyon16 whole-slide gigapixel classification [3]
- Transfer learning based approaches
- Ensemble Models
- Data techniques

Data and Features

Dataset

Training Instances	220,025 (40% positives)
Size	96x96 (32x32 has target content)
Channels	3 (8 bits per channel)

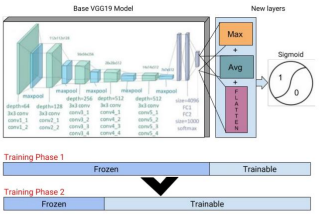
Features



Data Augmentation

- Elastic deformations
- Flipping
- Rotations
- Brightness
- Gaussian noise
- Cropping

Modeling



Architecture

- Transfer learning from VGG19 for imagenet
- 28 layers, with 17 trainable layers
- 20,030,017 trainable parameters

Training

- Phase 1: Train last 5 layers using Adam optimizer
- Phase 2: Train last 17 layers using SGD with smaller learning rate

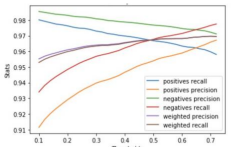
Tuning

Bayesian Hyperparameter optimization

- Surrogate model
- AUC for ROC to guide parameter prediction

Threshold Tuning

- Recall not optimizable
- Tune threshold on validation set and verify on test set. It works well!



Picking threshold of 0.12 based on the plot also gives 98% recall for positives on test set as well.

Results Overview

Threshold Tuned Model

98% recall on positive regions

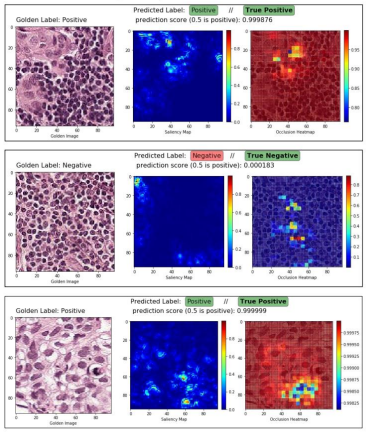
Label	Precision	Recall	F1-score	instances
0	0.98	0.94	0.96	13,179
1	0.92	**0.98**	0.95	8,824
overall	0.96	0.96	0.96	22,003

Predicting Metastasis

Top models for each architecture trained, with data augmentation, hyperparameter tuning, threshold tuning, etc.

Base architecture	Train acc.	Validation acc.	Test acc.
VGG19	**98.5**	**97.2**	**97.1**
VGG16	98.5	96.3	96
Resnet50	99.6	96.7	96.8
InceptionV3	90	89.7	89.8
MLP	69	68	68.2

Random samples of predictions with saliency maps and occlusions



Discussion

- Our best model, transfer learned from VGG19 (with tuning, augmentation, etc.) was able to distinguish areas of large abnormal lymphocytes, enlarged nuclei, & pink ample cytoplasm with high accuracy.
- 10 random samples were shown to a Board Certified Anatomic and Clinical Pathologist, Aida Rechdouni M.D.
- For 5/7 TPs, model picked exact regions as Dr. Rechdouni. Remaining 2 were correct, though not very confident in identifying every metastasis in the slide.
- For 3/3 TNs, model predicted accurately.
- Connections between base pre-trained and new FC layers, threshold tuning, hyperparameter tuning, and data augmentation are key!

Future Work

- Explore active learning from false negatives on validation set to further improve model.
- Address any sampling bias with respect to representative slides or patients.
- Reduce model size with smaller architectures like MobileNet and DNN compression techniques.

References

- [1] Vestjens JH, Pepels MJ, De boer M, et al. Relevant impact of central pathology review on nodal classification in individual breast cancer patients. Ann Oncol. 2012 https://www.ncbi.nlm.nih.gov/pubmed/22495317
- [2] Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer: https://jamanetwork.com/journals/jama/fullarticle/2665774
- [3] 270 Whole-Slide-Images of lymph node sections for breast cancer: https://camelyon16.grand-challenge.org/Data/