

Computer Vision for Cancer Detection

Trevor Adriaanse, Jack Hart, Norman Hong

Departments of Linguistics and Analytics

Georgetown University

Washington, D.C.

{tva7, csh87, nh530}@georgetown.edu

Abstract

Breast cancer is the most common form of cancer in women.^[1] Clinical management of breast cancer is predicated on classifying the extent of the cancer. One of the central components in this staging process is determining lymph node involvement.^[2] Histological assessment is a diagnostic procedure performed by expert pathologists to determine whether the tissue has been afflicted by cancer. Automating approaches to detect lymph node metastasis is an attractive avenue of exploration to aid the pathologist in their workload.

We develop neural network models to perform binary classification on histopathologic image data, apply transfer learning, and inspect predictions through an analysis of saliency and occlusion maps.

1 Introduction

The tumor, node, metastases (TNM) Classification system is the most widely-accepted standard guiding cancer stage description and categorization. Lymph node metastases is a major factor in cancer staging and is often assessed with the sentinel lymph node procedure. This process involves excising the lymph node most likely to contain metastasized cancer, histopathologically processing the specimen, and having a pathologist perform an examination. By coupling these whole-slide images (WSIs) of sentinel lymph nodes with a binary label (presence of

cancer), researchers created the Camelyon16 dataset to bring deep learning techniques to bear on an image classification problem of particular importance for the medical community and countless people affected by cancer.^[3]

As a computer vision problem, a natural architectural choice is the convolutional neural network (convnet/CNN). Convnets rely on kernel convolutions to take advantage of the structural properties inherent in image data. Specifically, convolutional neural networks home in on local correlations while ignoring distant ones. As a baseline, we build a convolutional neural network, trained end-to-end. To see how previous task knowledge can be leveraged for potential performance boons, transfer learning using a variety of pre-trained networks is also explored.

Convnets have state of the art performance for various image recognition benchmarks^[4] but at the cost of greater abstraction. For problem domains such as cancer detection, where the cost of incorrectly classifying an image as benign is high, model interpretability and explainability are exceedingly critical. Visualizations that provide insight into how and why a model is classifying a certain image as cancerous or benign would serve as useful tools for pathologists.

2 Related Work

Applying deep neural networks to histopathologic cancer detection has enjoyed past success. Ragab et al. employed AlexNet to obtain a 71% accuracy on a few breast mammography datasets.^[5] Kassani et al. used ensembles of pre-trained networks to obtain

an accuracy of 95% on the PCam dataset.^[6] Other approaches include bringing CNNs to bear on related problems such as gland segmentation and nuclei segmentation.^[7,8]

Understanding classification decisions is paramount when the decision potentially carries life-changing consequences. To this end, visualizations will be used to shed light on network behavior. For instance, occluding parts of the lymph node scan and inspecting activations in certain layers or the output probabilities of the network for a particular classification decision can inform which regions of the image were most important in making the decision.^[9] In addition, we explore leveraging saliency maps, which generate visualizations using the gradient of the unit of maximum activation in the network’s output layer.^[10]

3 Dataset

The Camelyon16 dataset contains 400 WSIs of sentinel lymph nodes, together with a label denoting whether there is any presence of cancer in the image. Building off of this dataset, Veeling et al. constructed the PatchCamelyon (PCam) dataset. The PCam dataset consists of 327,680 patches extracted from Camelyon16, at a size of 96×96 pixels at $10\times$ magnification, with a $75 - 12.5 - 12.5$ train/validate/test split. A positive label means that the image contains at least one pixel of cancer tissue in the middle 32×32 region of the slide. Therefore, outside of the 32×32 center of the image, none of the image data plays a role in labeling. For instance, it is possible for cancerous tissue to be present outside of the 32×32 region but still have a negative label correctly associated with the image.

The sentinel lymph node scans are hematoxylin and eosin–stained (H&E). This technique is the most widely used staining method in medical diagnosis, and it yields slides primarily with purple and pink colors.^[11] Figure 1 shows two examples of images from the PCam dataset; (a) is a negative example and (b) is a positive one.

4 Methods

All models were trained on five epochs over the test data with Adam as the optimizer. A learning rate of 0.001 was used for all experiments. Results are

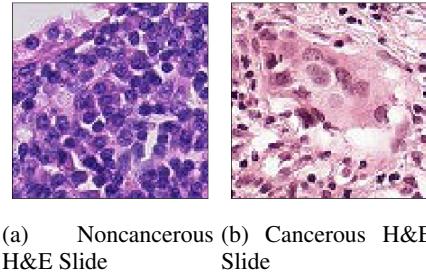


Figure 1: H&E Slides

based on the performance of models on the test set, which consisted of 32,768 patches (16,377 cancerous and 16,391 non-cancerous).

4.1 Baseline

The baseline model consists of three pairs of convolutional layers, with filter sizes set to 16, 32, and 64, respectively. The kernel size was set to 3×3 in each convolutional layer, and no padding was used. Batch normalization was applied after each pair of convolutions, to be followed by a maxpooling layer with pool size 2×2 and strides of 2×2 . A few Dense layers bridged the final convolutional block with the classification layer.

4.2 Transfer Learning

Transfer learning relies on the idea that representations learned by deep CNNs trained on large amounts of image data can be useful for specific image tasks beyond just the ones they were trained for. We explore transfer learning using the ResNet50, InceptionResNetV2, and Xception neural networks.^[12,13,14]

5 Results and Discussion

5.1 Baseline

The baseline Convnet obtained an accuracy of 0.82 and F1 score of 0.81. Figure 2 the probabilities generated by the model for each class. The model had an easier time classifying non-cancerous data, evidenced by the high density of non-cancerous examples with probabilities close to zero.

5.2 Transfer Learning Models

Fine-tuning was not performed on any of the transfer models. Rather, features generated from these

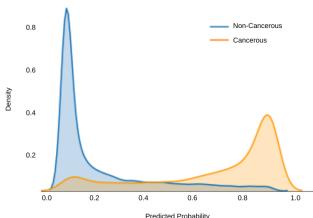
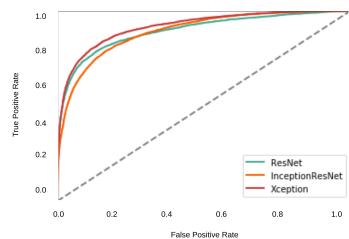


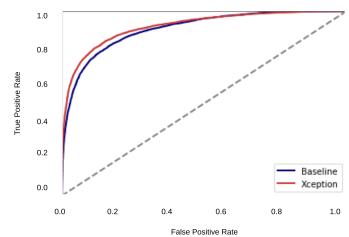
Figure 2: Baseline Class Probability Distributions

models were treated as fixed inputs. Performances between each of the transfer models varied slightly, with improved performances for more complicated model architectures. For instance, the Xception architecture demonstrated the strongest performance among the transfer models, with an accuracy of 0.81 and an F1 score of 0.78. Xception’s performance was followed by InceptionResNetV2 with an accuracy of 0.79, and then ResNet with an accuracy of 0.78.

No transfer model attained better accuracy than the baseline Convnet. Even though Xception had a lower accuracy than the baseline, it’s performance was comparable given the model has a slightly larger AUC (refer to Figure 3(b)). These results are indicative that pre-trained convolutional layers are too general to pick up on specific features important for this dataset. Transfer models will likely benefit from supervised fine-tuning.



(a) Transfer Models’ ROC Curves



(b) Xception and Baseline ROC Curves

Figure 3: ROC Curves

Architecture	Accuracy	F1
Baseline	0.82	0.81
ResNet50	0.78	0.73
InceptionResNetV2	0.79	0.77
Xception	0.81	0.78

Table 1: Model Results Summary

5.3 Saliency and Occlusion Maps

Saliency and occlusion maps are a visualization technique applicable to neural networks performing image classification. Specifically, an image is fed through the network many times with specific groups of pixels greyed out. The map is created by plotting the output of the model as a color for each greyed out patch. Since we are performing a binary classification task, the occlusion map plots the probability of cancer predicted by the model.

In Figure 4, red indicates high predicted probabilities and blue indicates low probabilities. For cancerous images, the saliency maps would ideally be majority red. For instances where the model correctly classified these images as cancerous (i.e. true positives), there appear to be some areas where the performance of the model is hurt by occlusion, thus the map shows blue regions. This suggests that the areas highlighted in blue may be cancerous patterns that the model has correctly picked up on.

Saliency maps can also be useful when trying to determine what features a model is incorrectly using for classification. For example, the saliency maps of cancerous images that were incorrectly classified as non-cancerous contain regions that, when occluded, improve model performance. These are patterns the model believes are indicative of cancer.

6 Future Work

We explored transfer learning with deep pre-trained networks with the assumption that the representations learned by the networks would be sufficient to perform well at the cancer detection task at hand. Evidently, WSIs look vastly different from, for instance, images contained in the ImageNet dataset. Therefore, fine-tuning the borrowed model parameters to be specific to the WSIs contained in PCam

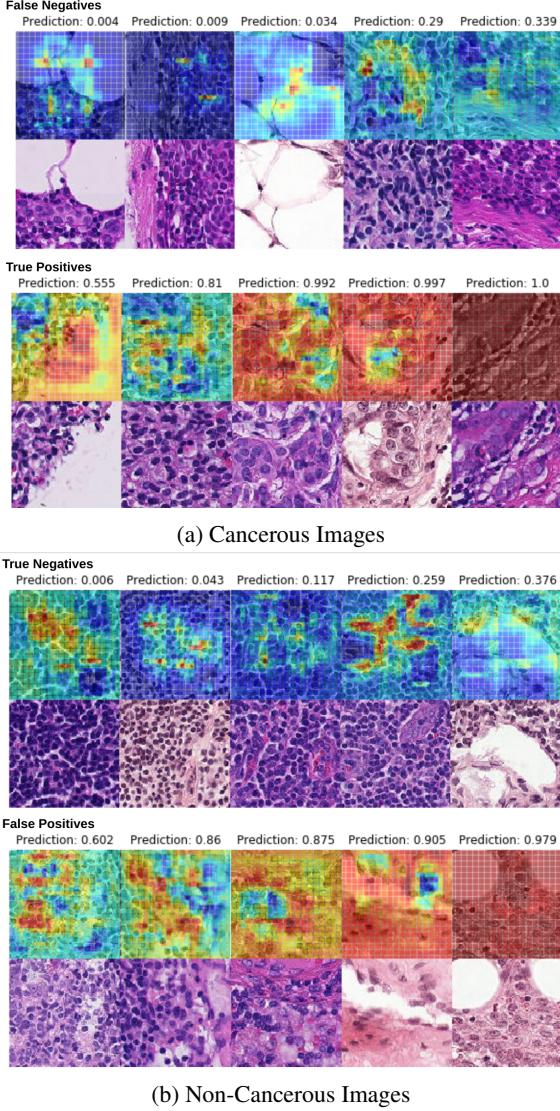


Figure 4: Occlusion and Saliency Maps

may offer stronger results than training a classification layer alone.

For a binary classification task with high-stakes decisions, preventing false negatives is paramount. Whereas a false positive decision is sure to induce stress and sorrow in a patient, a false negative can lead to death if treatment is forestalled because of an incorrect diagnosis. So, exploring techniques to optimize for recall is another continuing objective.

7 Conclusion

CNNs offer a promising avenue to assist pathologists in inspecting lymph node scans for metastasis.

In addition to constructing a baseline and exploring transfer learning, various visualization techniques were effective for determining important sections of images and provided insight into the predictions made by these models.

References

- [1] F. Bray, et al. *Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. Wiley Online Library, American Cancer Society, 2018.
- [2] L. Sabin, et al. *TNM Classification of Malignant Tumours*. Wiley-Blackwell, 2011.
- [3] B. Bejnordi, et al. *Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer*. The Journal of the American Medical Association, 2017.
- [4] Z. Alom, et al. *The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches*. arXiv: 1803.01164, 2018.
- [5] D. Ragab, et al. *Breast Cancer Detection using Deep Convolutional Neural Networks and Support Vector Machines*. PeerJ 7:e6201 <http://doi.org/10.7717/peerj.6201>, 2019.
- [6] S. Kassani, et al. *Classification of Histopathological Biopsy Images Using Ensemble of Deep Learning Networks*. arXiv:1909.11870, 2019.
- [7] Yang, Lin, et al. *BoxNet: Deep Learning Based Biomedical Image Segmentation Using Boxes Only Annotation*. arXiv:1806.00593, 2018.
- [8] M. Salvi, et al. *Multi-tissue and multi-scale approach for nuclei segmentation in HE stained images*. Biomedical Engineering Online 17.1: 89., 2018.
- [9] Zeiler et al. *Visualizing and Understanding Convolutional Networks*. <http://arxiv.org/abs/1311.2901>, 2018.
- [10] Simonyan et al. *Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps*. <http://arxiv.org/abs/1312.6034>, 2013.
- [11] M. Titford. *The Long History of Hematoxylin*. Biotechnic Histochemistry. Europe PMC, 2005.
- [12] K. He, et al. *Deep Residual Learning for Image Recognition*. arXiv:1512.03385, 2015.
- [13] C. Szegedy, et al. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*. arXiv:1602.07261, 2016.
- [14] F. Chollet. *Xception: Deep Learning with Depthwise Separable Convolutions*. arXiv:1610.02357, 2017.