Computer Vision for Cancer Detection

Trevor Adriaanse, Jack Hart, Norman Hong

Introduction and Convolutional Neural Network Architectures

Abstract

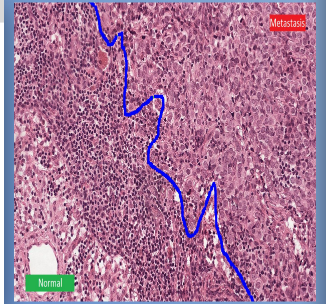
We use a convolutional neural network (CNN) to perform binary classification on the PCam dataset.[1]

Background

Lymph node metastases is a major factor in cancer staging.[2] Thus, detecting cancer presence in lymph nodes affects clinical management and treatment. Augmenting pathologist review with neural models has the potential to benefit pathologist workflows.

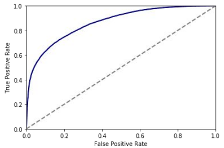
Dataset

PCam consists of 327,680 WSIs, at a size of 96x96 pixels, at 10x magnification. A positive label means that the image contains at least one pixel of cancer tissue in the middle 32x32 region of the slide.



Network Choices

Baseline: Three pairs of convolutional layers, filter sizes set to 16, 32, and 64, respectively. Batch normalization applied after each pair of convolutions.



ResNet50: A 50-layer model trained on ImageNet that uses Skip connections to circumvent the vanishing gradient issue. This means that the network can be deep without having gradient updates be ineffective.

InceptionResNetV2: Continuing with the use of Skip connections, this neural architecture is substantially larger than ResNet50, with more than 30 million parameters than that architecture. Inception networks use filters of multiple sizes within the same layer of a network, so a single layer can take advantage of information that is local with smaller kernels and global with larger ones.

XCeption: In this architecture, Inception modules are replaced with depthwise separable convolutions. An “extreme” form of Inception, as noted by XCeption’s author, applies a 1x1 convolution to map cross-channel correlations and then uses 3x3 convolutions on each resulting output channel.[3]

Model Comparison at a Glance: Accuracy

Baseline	ResNet50	InceptionR esNetV2	XCep -tion
76.20	78.50	78.57	82.12

Interpretability and Data Visualizations

Background

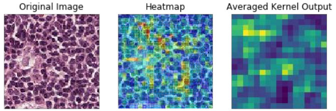
Deep learning models have state of the art performance for various image recognition benchmarks[4] but at the cost of greater abstraction. For problem domains such as cancer detection, where the cost of incorrectly classifying an image a benign is high, model interpretability and explainability are exceedingly important.

There has been recent and frequent research on Explainable AI (XAI) and model interpretability[5]. In the case of CNNs, various visualization techniques such as occlusion mappings and deconvolutional networks have been proposed. We implement various visualizations for our CNNs with the goal to gain better insight into how and why the model is classifying a certain image as cancerous or benign.

Kernel Output Averaging

Kernels closer to the end of CNNs can pick up on higher level patterns in the data, making them useful for interpreting how the model is classifying a certain image. The last convolutional layer in our baseline model had 64 kernels each of size 24x24. To see if information could be gained from these kernels, we took an average of all their outputs for a given image.

The oscillation map for the following example appeared to be very similar to the averaged kernel output. This is indicative that the averaged output of kernels may be useful for determining where a model is “paying attention”.



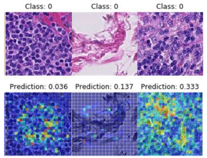
Conclusion

CNNs offer a promising avenue to assist pathologists in performing rote inspection for metastasis in lymph nodes. In addition to constructing a baseline and exploring transfer learning, various visualization techniques were found to be effective tools for determining important sections of images. These visualizations used in conjunction with deep learning models can provide better explainability of predictions to pathologists, while also maintaining the high accuracies from CNNs.

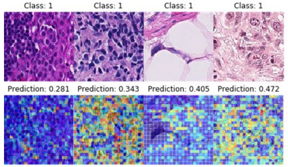
Occlusion Maps

Occlusion maps[6] are a visualization technique used on neural networks performing image classification. Specifically, an image is fed through the network many times with specific groups of pixels greyed out. A heatmap can then be 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.

The occlusion maps for the non-cancerous images to the left show majority blue (low probabilities). All of these examples were correctly classified as non-cancerous (e.g. true negatives) given the threshold was set to 0.5.



The occlusion maps for the cancerous images below would ideally show majority red (high probabilities). However, all of these examples were incorrectly classified as non-cancerous (e.g. false negatives) given the threshold was set to 0.5. For all the examples, there are some sections of the images that significantly change the probability predicted by the model.



References

[1] B. S. Veeling, et al. “Rotation Equivariant CNNs for Digital Pathology”
[2] L. Sobin, et al. “TNM Classification of Malignant Tumours”
[3] F. Chollet, “XCeption: Deep Learning with Depthwise Separable Convolutions”
[4] Z. Alom, et al. “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches”
[5] L. Gilpin, et al. “Explaining Explanations: An Overview of Interpretability of Machine Learning”
[6] M. Zeiler, R. Fergus. "Visualizing and Understanding Convolutional Networks"