Histopathological Image Analysis using Multiple Instance Learning

October 19th, 2020

By

Team Members

Nupur Yadav

Lokesh Vadlamudi

Chetan Kulkarni

Abstract:

Deep learning is being used in every field nowadays, but it has played a very crucial role in healthcare and pathology. The advanced equipment, such as specialized scanning machines and strides in storage capabilities has made it easy to store and process microscopic glass slides in digital forms on the computer. This has resulted in remote diagnosis, faster analysis, improved accuracy, and safe storage of pathology information.

Though fully supervised techniques such as CNN have proved to be state-of-the-art for many image classification datasets such as ImageNet, automating the classification of histopathological slides is challenging because of the following reasons. Firstly, the images that CNN works well on are low-resolution images while histopathological images are high-resolution images that are meant to be examined under the microscope by the pathologists. The image can contain billion of pixels while the area of interest can be just a few thousand pixels. So, to apply a fully supervised technique, the Whole Slide Image (WSI) has to be divided into several small tiles, and the classifier is to be applied individually to each tile. The output is then aggregated to obtain a final classification for the WSI. Secondly, it is extremely difficult to obtain a large dataset of WSIs with detailed pixel-wise annotations that make it challenging to use a supervised deep learning classifier.

Therefore, in this project, we aim to use weakly supervised techniques such as Multiple Instance Learning (MIL) for histopathological image classification that addresses the above problems of fully supervised learning. MIL uses the whole-slide labels instead of using pixel-wise annotations. We will be performing histopathological classification on the prostate cancer dataset. In this project, we will also discuss the recent trends in performing histopathological image classification using MIL and future direction.

Goals:

1. Training the Model
2. Host the pickle file on an ec-2 instance in AWS.
3. Web application using Django to access the model’s prediction.

Role of the team members:

Nupur – Preparing Training data / Backend Infrastructure / Hosting application on AWS

Chetan – Model training / Web application

Lokesh – Model Finetuning / Visualization / Web application

Input:

The input for this project will be a medical image of prostate biopsy sample.

Output:

The model on the cloud analyses the image and send out a prediction on whether it is cancerous or benign.

Visualization Tools / Metrics:

Tensorboard metric