Deep Learning 2021

Lecturer:
Dr. Jonathan Rubin

Classification of histopathology images

Inbal Mishal & Or Ben-Aba

מכון טכנולוגי חולון
Holon Institute of Technology

Background

Early, fast and accurate diagnosis of diseases can often save lives. Diagnosis of diseases from tissue samples is done manually by pathologists. Usually, This process takes a lot of time and costs a lot of money.
Therefore, there is a need for an accurate and automated systems that can assist pathologists and help them decide quickly what type of tissue sample is in front of them. Those systems can give to the pathologists second opinion in critical cases and help them to save lives.
In this project, a deep-learning based transfer learning approach has been proposed to classify histopathology images automatically.

Dataset

- Dataset of images of Colorectal histology. This dataset is “tensorflow” dataset that included 5,000 images of f textures in colorectal cancer histology. Each example is a 150 x 150 x 3 RGB image of one of 8 classes.
- Dataset of large images of Colorectal histology. This dataset is “tensorflow” dataset that included 10 images of cells in Colorectal histology. Each example is a 5000 x 5000 x 3 RGB image.

Models

We tried several CNN models with different optimizers and different layers. Part of them was training with data augmentation, transfer learning and fine tuning. The best model built with transfer learning of VGG19 (not frozen weights), SGD optimizer and with data augmentation. This model achieved 92% of accuracy on the test set.

confusion matrix of the final model:

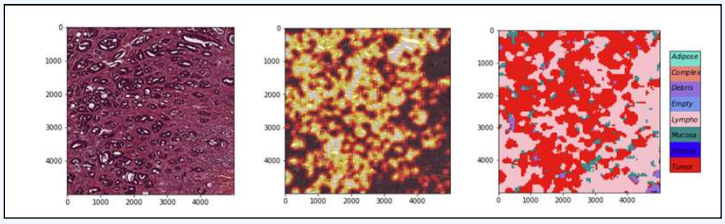
Results

After choosing the best model among 7 models, we continued to the second step.

We took the large images (5000 x 5000 x 3) and divided them into patches of (150 x 150 x 3, in order to adjust the images to our model) with 100px overlapping between the patches. Then, we put the patches on our model and got the classification of each patch, and then painted the large image by its patches’ classification.

The final step was to take each large image and predict the probability of each patch to be a cell with tumor.

Classification of the patches in the large image:



Conclusion

Detecting cancerous or abnormal cells automatically may lead to a revolutionary change that will save many lives. An automated system will help pathologists quickly and accurately diagnose the type of cell they are examining.
In this project, we built a deep learning model that identifies and manages to classify the cell type with accuracy of **91.68** percent.
Our project proves and illustrates that this system can help pathologists do their job better and save more lives.

Bibliography

- Tensorflow datasets
- VGG16 and VGG19 trained models
- Stack Overflow