Classification of histopathology images

Dr. Jonathan Rubin

Shay Shtern & Coral Ifergan 5781, Semester B – מסטר א, סמסטר



Objective

Train several neural network architectures to predict the tissue type of the Colorectal histology dataset, choose the best model among them and use it to classify the Colorectal histology large dataset

Specifications

Work Environment: Jupyter Notebook **Programming Language:** Python 3.8.3 **Libraries**: TensorFlow, Numpy, Matplotlib

Algorithm

Step I: Load the dataset and prepare it:

- Load the Colorectal histology dataset which consist of 5000 Tissue images of size 150x150x3. Each image has its own label (0-7).
- Split the dataset: 90% for train and 10% for test. Load it with shuffle ON so both the train and the test set will include all 8 classes.

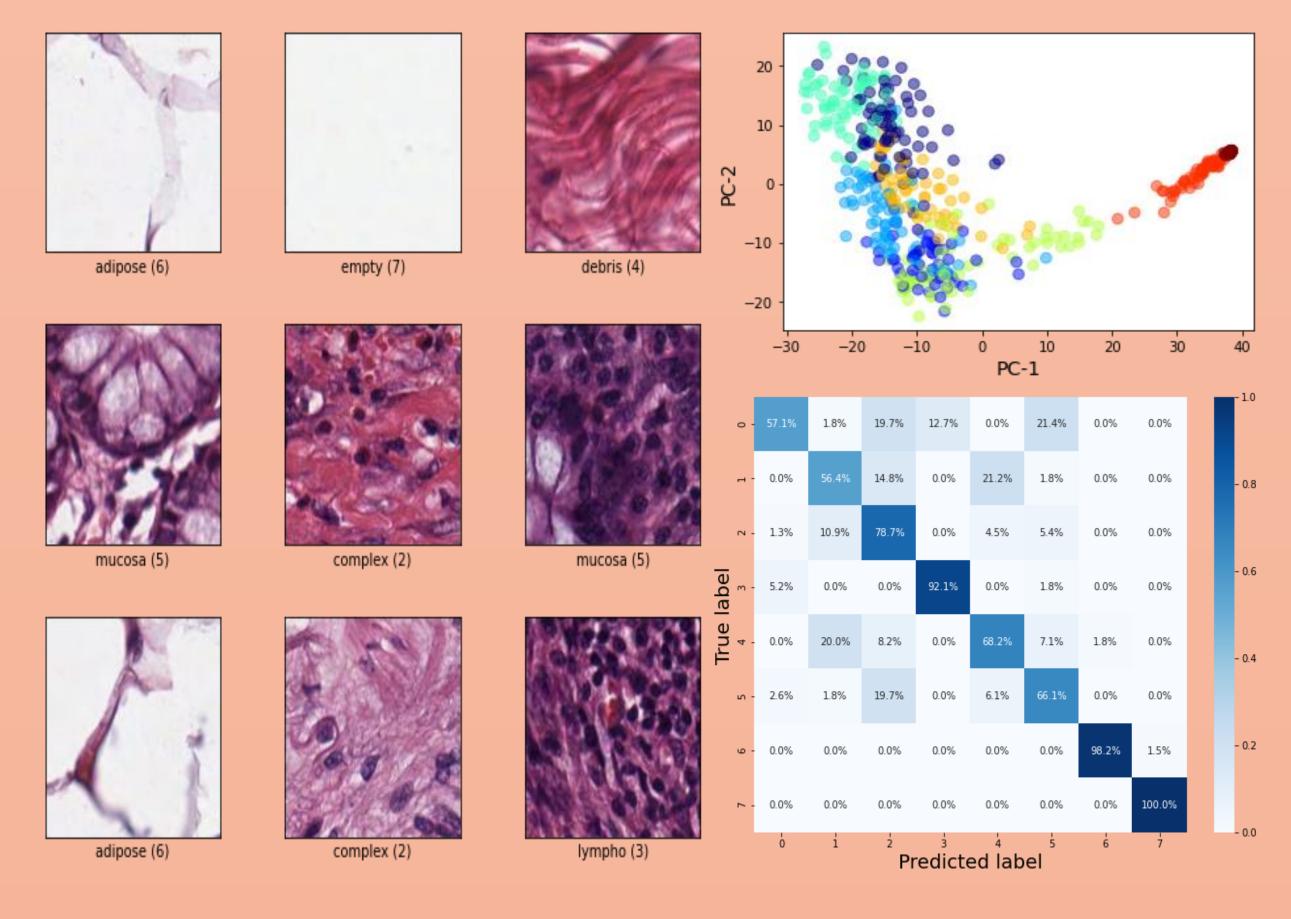
Step II: Create a customized CNN model* and use it to predict the Tissue type

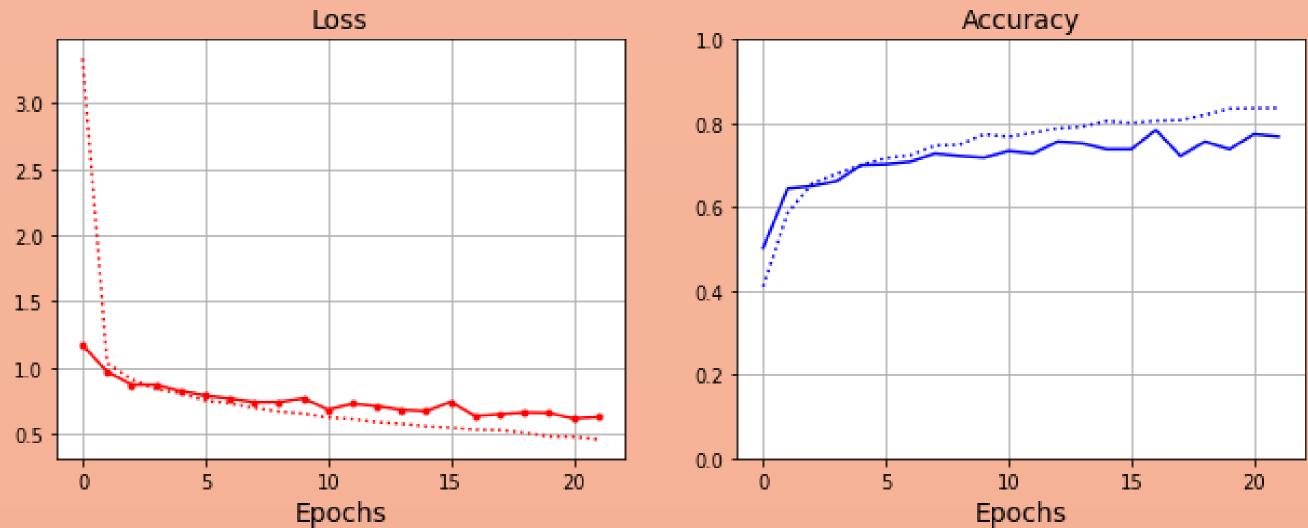
- The model consist of Conv2Dx4, MaxPoolingx4 A. between the Conv2D layers, Flattenx1, Dense(128) and Dense(8). The input images are of size 150x150x3 and the output is a 8 integer vector. Each of these integers corresponds to the probability of a given image belonging to a specific class.
- Compile the model with the SGD optimizer, small **B**. learning rate (0.0001) and a momentum of 0.5
- С. Train the model with 22 epoch and a batch size of 50.

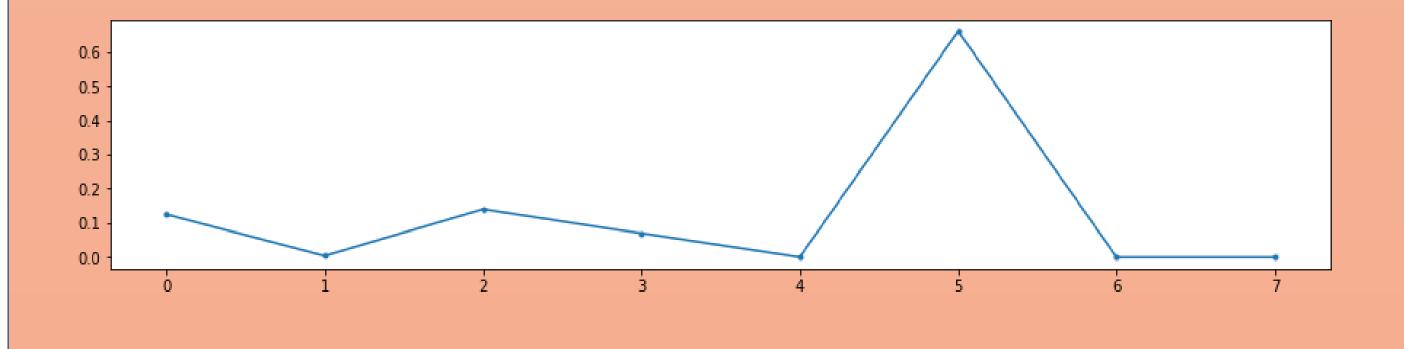
Step III: Test the model on unlabeled dataset:

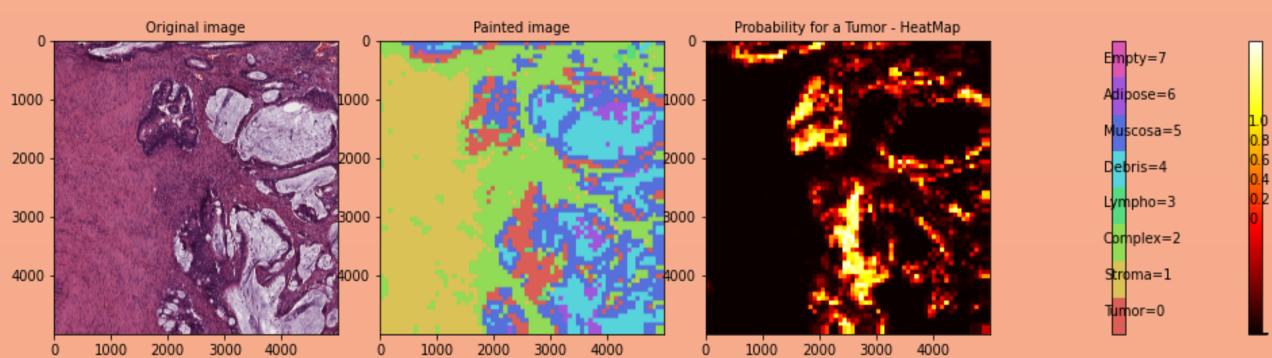
- Load 10 large images of size 5000x5000x3 that A. belongs to a textured colorectal cancer histology dataset.
- Split each image to smaller patches by using a floating window of size 150x150x3. Use the model to predict the class (0-7) on each of these patches. By using a 8-colormap palette, each of the classes is represented by a unique color.
- Plot both a colorful classification image output with respect to all classes and a heat-map with respect to the Tumor class (0)

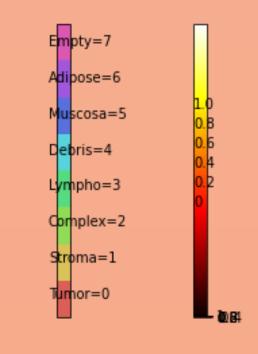
Pipeline











^{*}Compared several models and picked up the best one among them