Use Deep Learning to Predict Cancer Metastasis from Histopathologic Tissue Patches

Wenwen Ye[1], Zicheng Yuan[2] and Xue Zou[3]

[1]Department of Statistics, Duke University, [2]Department of Electrical and Computer Engineering , Duke University , [3]Computational Biology and Bioinformatics Program, Duke University

Abstract

Advances in the quality of clinical images have enabled the systematic digital recognition of metastatic cancer from pathologic scans [1], and deep learning algorithm has proved its effectiveness in identifying metastatic lymph nodes among MRI images [2]. To evaluate the accuracy and efficiency of digital assistance on identifying clinical-patches at different severity, we conducted a study utilizing deep learning algorithm for the detection of cancer metastasis in small image patches taken from larger digital pathology scans. Two hundred twenty thousand twenty-five images were used in the training and evaluation set, and two models were used to predict the binary labels and compared: 4-layer **Convolutional Neural Network (CNN)** and **ImageNet pre-trained model**. This study demonstrates the potential of a deep learning algorithm to improve the computer-assisted diagnostic in pathology and clinical care.

Data Description

- A large number of the small **pathology images** 96*96px
- Modified version of the **PatchCamelyon (PCam)** benchmark dataset (Original PCam contains duplicate images due to its probabilistic sampling)
- **Positive labels** indicate that the center 32*32px region of a patch contains at least one pixel of tumor tissue, tumor tissue in the outer region of the patch does not influence the label

Images Example

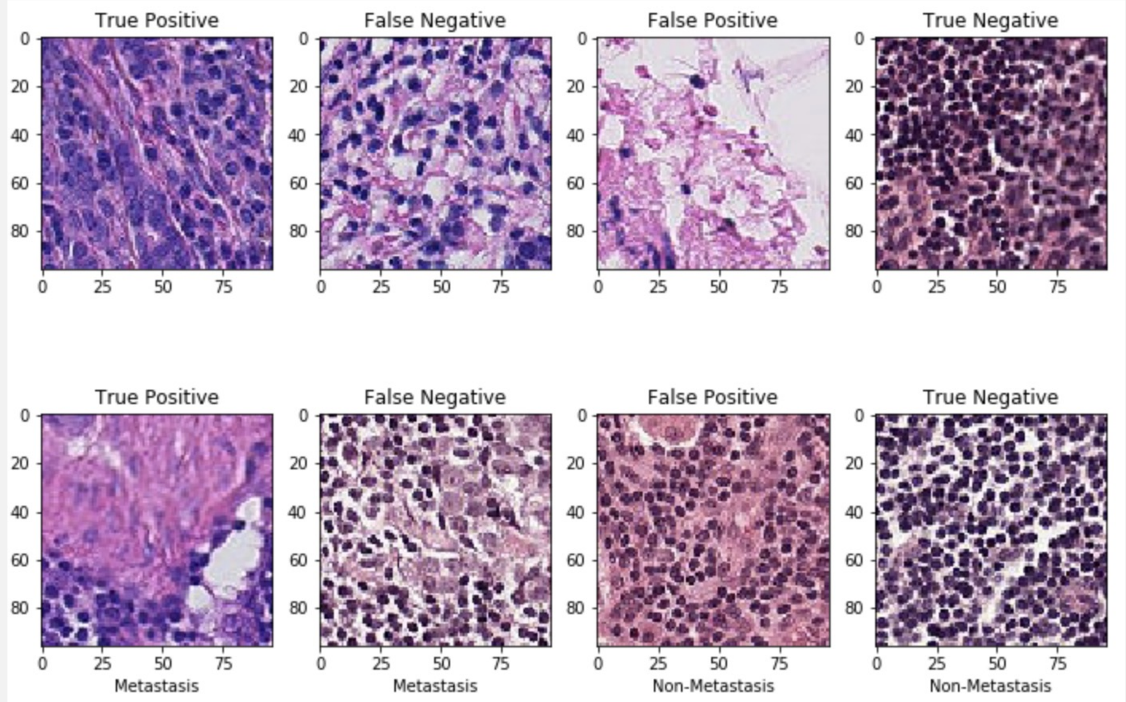

Figure: Two random selected examples for FP, FN, TP, TN, respectively

Contact Information

- Web: https://github.com/x811zou/STA561D-project
- Email: wenwen.ye@duke.edu zicheng.yuan@duke.edu xue.zou@duke.edu

Methods

Convolutional neural network (CNN) is the most popular neural network model used for feature identification and classification of the images. A typical CNN consists of multiple convolutional layers, ReLU layers, pooling layers as feature learning and fully connected layer and softmax as classification.

- **Convolutional Layer** extracts features from the input image by carrying out convolution operation with a Kernel/Filter, which has the same depth of input image. The first convolutional layer is typically used to capture low-level features such as edges. The additional layers then adapt to high-level features, giving a network of the overall understanding of the images in the data set.
- **Rectified Linear Unit (ReLU) Layer** enhances the training efficiency by only maintaining positive values of a convolved output and carrying those activated features into the next layer. The activation function is represented as:

$$f(x) = max(0, x)$$

- **Pooling Layer** performs nonlinear down-sampling layer that reduces the spatial size of a convolved output. Specifically, Max Pooling partitions the convolved output into a set of non-overlapping regions and outputs the maximum of each region.
- **Fully Connected Layer** performs high-level reasoning by learning a possibly non-linear function of the high-level features.
- **Softmax** produces a vector containing the probabilities of an image being classified into each image class. The softmax functin is written as:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{e^{z_1}+e^{z_2}} \text{ for i = 1, 2 and } \mathbf{z} = (z_1, z_2) \in \mathbb{R}^2$$

where z is the input probability vector.
- **Adam Optimization** is an extension to traditional stochastic gradient descent methods.It make use the moving average of first and second moments of the gradient to update weight as $w_t = w_{t-1} - \alpha\frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\epsilon}$, $\hat{m}_t = \frac{m_t}{1-\beta_1^t}$,$\hat{v}_t = \frac{v_t}{1-\beta_2^t}$,where m_t and v_t are the updated biased first and second moment.

4-layer CNN Model

The model in this paper has a 4-layer architecture. Each layer consists of one **convolutional-2d layer** with ReLu as the activation function, followed by **batch normalization**, one **maxpooling layer** and **dropout**. Those structures are designed to extract features in training images. At the end of the four layers, a flatten layer and several fully-connected layers are used to make final classifications.

Layer (type)	Output Shape	Param #
conv2d_9 (Conv2D)	(None, 94, 94, 32)	896
batch_normalization_13 (Batc	(None, 94, 94, 32)	128
conv2d_10 (Conv2D)	(None, 92, 92, 32)	9248
batch_normalization_14 (Batc	(None, 92, 92, 32)	128
max_pooling2d_5 (MaxPooling2	(None, 46, 46, 32)	0
dropout_11 (Dropout)	(None, 46, 46, 32)	0
conv2d_11 (Conv2D)	(None, 44, 44, 64)	18496
batch_normalization_15 (Batc	(None, 44, 44, 64)	256
dropout_12 (Dropout)	(None, 44, 44, 64)	0
conv2d_12 (Conv2D)	(None, 42, 42, 128)	73856

Figure: model's summary

batch_normalization_16 (Batc	(None, 42, 42, 128)	512
max_pooling2d_6 (MaxPooling2	(None, 21, 21, 128)	0
dropout_13 (Dropout)	(None, 21, 21, 128)	0
flatten_3 (Flatten)	(None, 56448)	0
dense_7 (Dense)	(None, 512)	28901888
batch_normalization_17 (Batc	(None, 512)	2048
dropout_14 (Dropout)	(None, 512)	0
dense_8 (Dense)	(None, 128)	65664
batch_normalization_18 (Batc	(None, 128)	512
dropout_15 (Dropout)	(None, 128)	0
dense_9 (Dense)	(None, 2)	258

Figure: models's summary(continued)

Illustration for the 4-layer CNN model

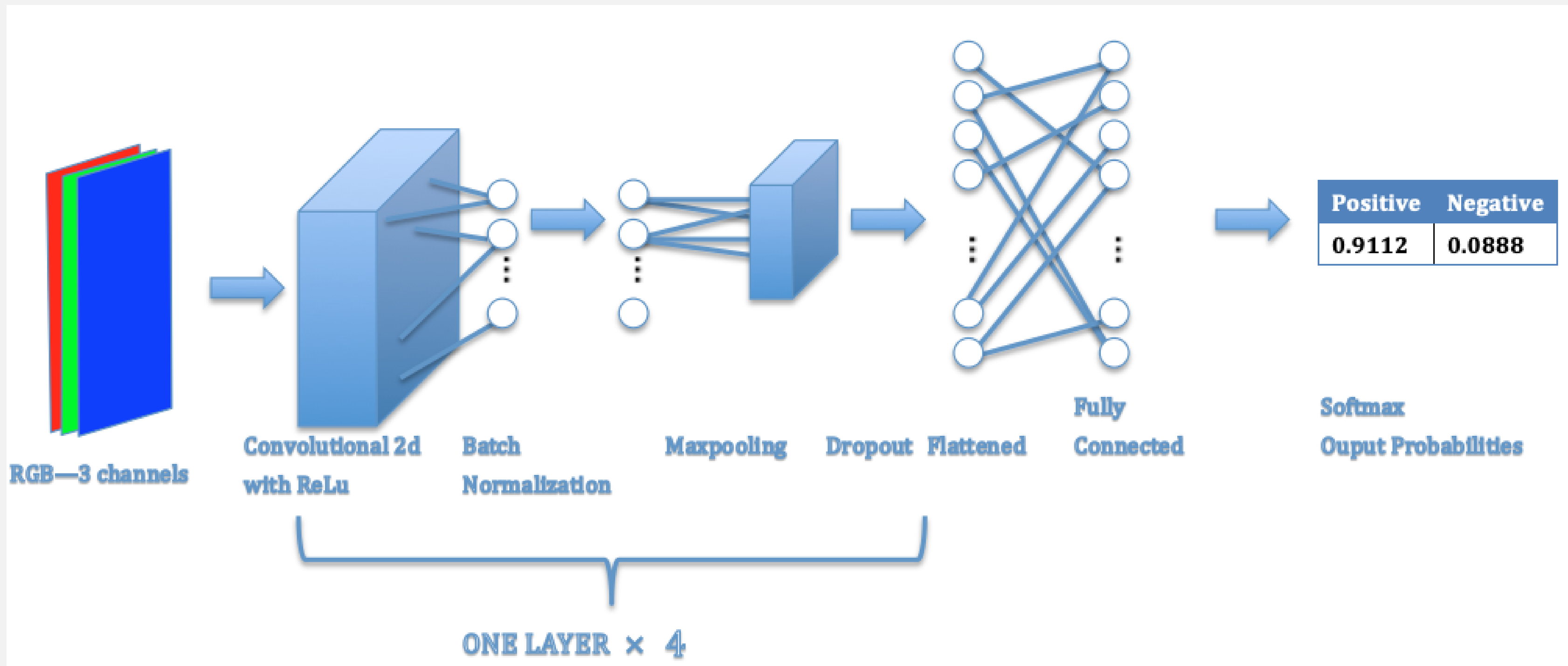

Figure: 4-layer CNN represented by one layer shown in the figure for four times. A fully connected layer and the softmax decide the final classification of the image.

Results

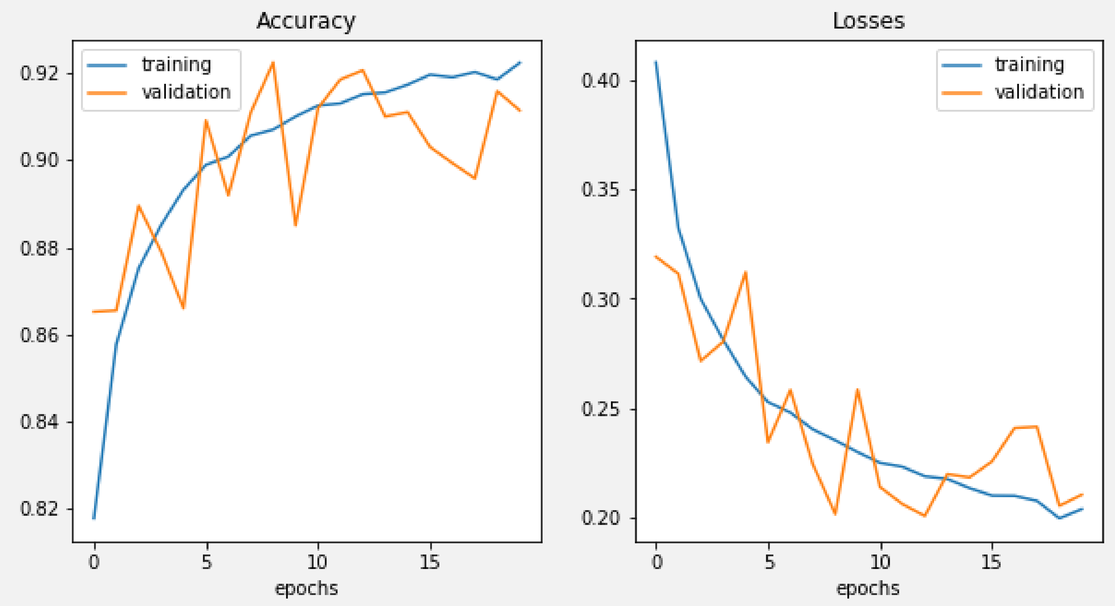

Figure: Accuracy and Losses of 20 epochs

Accuracy	Losses
0.9224	0.2015

Table: Accuracy and Loss of the Best model

The image above shows how accuracy and losses for training(80%) and validation(20%) datasets change as the number of epochs increases. The training accuracy as the fitting becomes more complicated, while the validation accuracy shows tendency to drop after the peak.

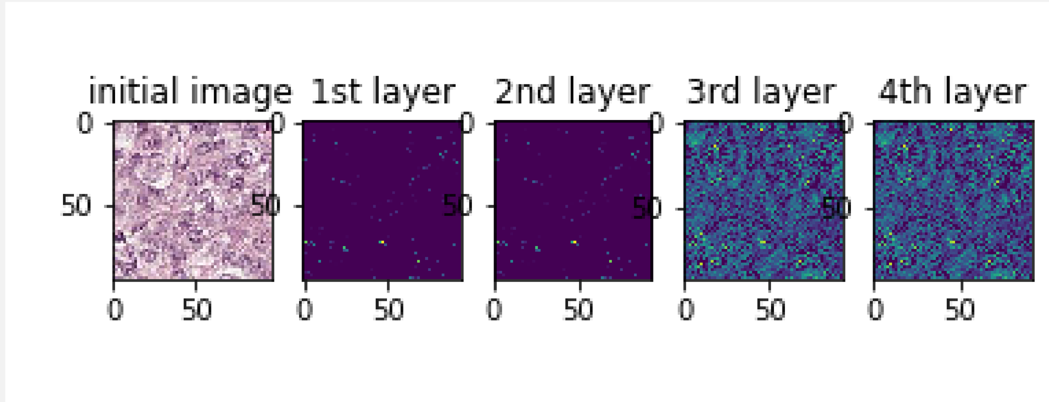

Figure: Activations of four layers

This image shows the four layers of the best model predicting on a sample image. The 1st and 2nd layer are only picking out rough features comparing to the much fine-grid 3rd and 4th layer.

Discussion and Future Direction

The overall accuracy shows that the convolutional network works decently on classifying our image dataset. However, the learning rates needs to be tuned and other optimizers might be tried for a better performance. In the meantime, we are working on improving our 4-layer CNN model and compare it with pre-trained model from ImageNet benchmarks such as VGG16.

References

[1] Steiner, F.D., et al., Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. Am J Surg Pathol, 2018.
[2] Lu, Y., et al., Identification of Metastatic Lymph Nodes in MR Imaging with Faster Region-Based Convolutional Neural Networks. Cancer Res, 2018. 78(17): p. 5135-5143.