Breast Cancer Detection by Patch-based Convolutional Neural Network

MSBD 6000B Project

NAME	ITSC	STU NUM
GAN Weisen	wgan	20446377
KANG Jiayue	jkangad	20475029
LIU Yang	yliufh	20475562
QIAN Zhiying	zqianab	20447591

Introduction

I. Background

Convolution Neural Network(CNN) is one of the most advanced image classification task. However, according to the X-ray images used in breast cancer detection, it costs large amount of space and time to train a CNN on gigapixel resolution of whole organ image. What's more, extensive image downsampling is required by which most of the discriminative details could be lost. In addition, it is possible that the CNN is able to learn from multiple discriminative modes in the image only, resulting in inefficient data. In this case, we split the image into several number of moderately sized patches. So, the challenge becomes how to predict the label of patches based on original image as well as how to determine the image label through patches label in test data.

II. Method

We propose using EM algorithm and logistic regression to classify patches and images respectively. Specifically, the first essential issue is to predict the label of individual patches rely on the ground truth label of image-level. Because tumours may have a mixture of structures and texture properties as well as blank black area, patch-level label are not necessarily consistent with the image-level label. In this part, we utilize Expectation Maximization(EM) method combining with CNN to output the patch-level label through iteration optimal parameters of neural network and the maximum probability by parameters above until converge. After that, we use logistic regression to predict the test image data using patch-level label of each image which is a binary classifier.

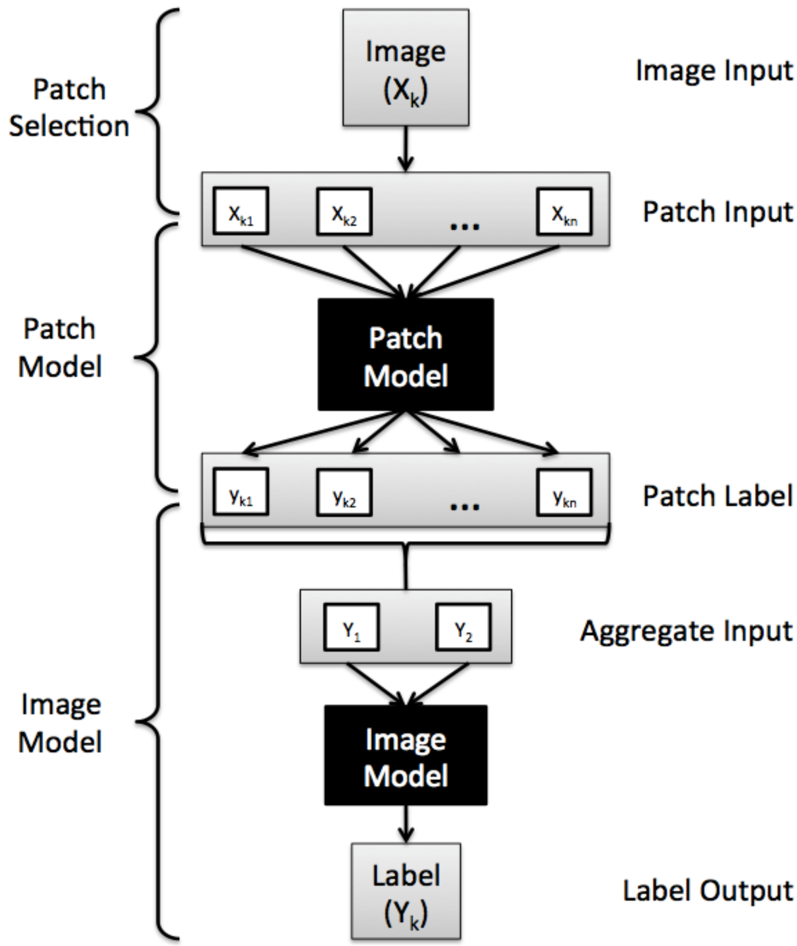
Model Design & Implementation

I. Overview

Our model will detect the breast cancer based on the input X-ray image. The input X is the RGB numerical array of the image and the output Y is defined as:

$$\mathrm{Y} = \begin{cases} 1 & \text{if no cancer detected} \\ 0 & \text{others} \end{cases}$$

The model's flow chart is as follows:



The model can be divided into three parts. The first part is Patch Selection that selects N discriminative patches from the input image. The second part is Patch Model that predicts N patch labels based on N patches inputs via a Convolutional Neural Network trained by EM algorithm. The third part is Image Model that predicts the image label based on the two aggregate inputs derived from patch labels.

II. Patch-selection

Before training the data using neural network, we decided to apply some pre-processing work to all images. First, we found out that there are two kinds of size for images X_k: $3328 * 4084$ and $2560 * 3328$. Because we want to process the images by cutting them into patches, we do some slightly modification to the size of images at the beginning. After that, we cut the images into 208 patches for image sized $3328 * 4084$ and 130 patches for images sized $2560 * 3328$. All the patches from all the images share the same size.

After that, we need to get rid of some patches containing only a small part of organ. We calculate mean grey scale for each patch and discard the patches having mean grey scale lower than a threshold we set. Then, for every image X_k, it generates n_k patches and finally there are $\sum n_k$ would be put into the model to train.

III. Patch-Model

To predict the patch label based on the patch input, we train a CNN model with EM algorithm. Firstly, we initialize the patch labels that match the image label Y, i.e. $y_{ik} = Y_k$ for image k. Then we train the CNN model with the training data $\{X_{ij}, y_{ij}\}$ and obtain the optimal parameter θ^* for the model (M-step).

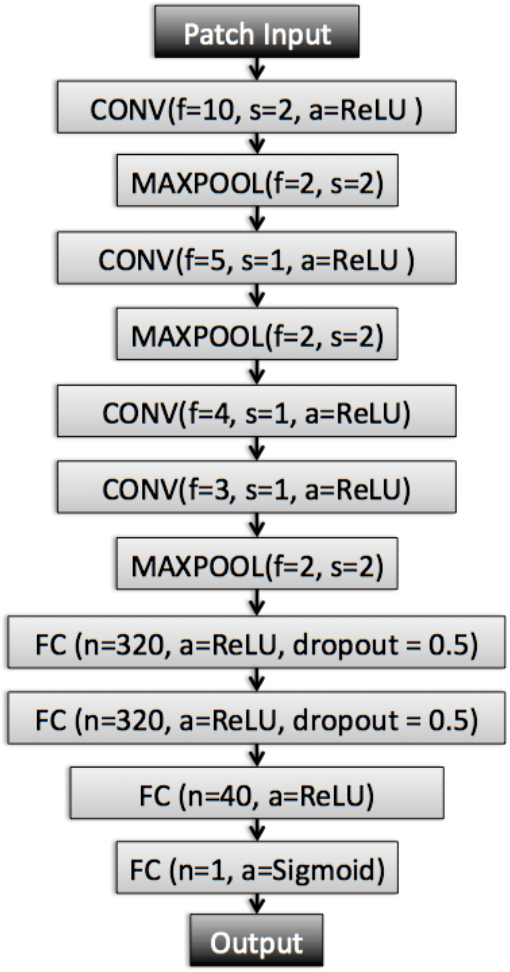
$$\theta^* = \arg\max_{\theta} \sum_{i=1}^{N} \sum_{j=1}^{n_i} P(y_{ij} \mid X_{ij}, \theta)$$

Afterwards we predict $\{X_{ij}\}$ based on our trained model and get the prediction list $\{y\_pred_{ij}\}$. We update the patch label y (E-step) based on rule

$$y_{ij} = \begin{cases} 1 & \text{if } y\_pred_{ij} \geq a \\ 0 & \text{others} \end{cases}$$

where a is a threshold (in our project we use 0.9). We then update $\{X_{ij}, y_{ij}\}$ with new values and repeat the EM training process until the parameter in the patch model converges.

For the patch model, we design a convolutional neural network with convolutional (CONV) layers followed by max-pooling (MAXPOOL) layers, and finally we add fully-connected (FC) layers and the final label output is a one-dimensional number. The details of our CNN model are shown in the graph below.



IV. Image-Model

After predicting the label for each patch for each image, all the patches have been classified into two classes, label '1' for abnormal patches while label '0' for normal patches.

For image denoted as X_k , it has n patches whose labels are denoted as $\{y_1, y_2, \dots \ , y_n\}$. Now, for each image, we have n patches to depict the image and summarize the situation. The patches for each image are classified into two classes. We compute the proportion of abnormal patches given all the patches and denoted as Y_1 while $Y_2 = 1 - Y_1$ for the proportion of normal patches. Then we add some nonlinearity into the result. Here we use sigmoid function to do that. And finally, we can classify all the images according to the probability generated by sigmoid function and threshold we set up.

V. Prediction

We will input test data to the model and predict the labels for each image. Because out model is well trained, therefore, I can get the predictions by just input the data. For every given test image, pre-processing would be applied on it and generated n patches, which would be inputted into the neural network to label the patches with ‘0’ and ‘1’. Then, a vector containing n labels would be outputted, given the proportions each label. Finally, a classifier would be used to classify the image into ‘0’and ‘1’.

Result & Conclusion

From the model above, the confusion matrix about training data is as follow:

	TRUE			
PREDICTION		1	0	**TOTAL**
	1	78	29	107
	0	22	281	303
	TOTAL	100	310	410

Accuracy=87.56%

Precision=72.90%

Recall=78.00%

We got the accuracy about 87.56%. According to the two types of error, we need to pay more attention to false negative rate which is same as recall. In other words, we would better decrease the misjudge about healthy person who has breast cancer indeed.

Improvement

1. We have used the probability of patches with label 1 as the input of the logistic regression. It is possible to choose a better decision fusion method to aggregate patch-level labels.
2. We could use Gaussian Smooth to reduce image noise and detail before processing, which is significant in eliminating the effect of edges.

Reference

[1] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz. Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. 2016.

[2] M. M. Kokar, J. A. Tomasik, and J. Weyman. Data vs. Decision Fusion in the Category Theory framework. 2001.

[3] D. C. Cires¸an, A. Giusti, L. M. Gambardella, and J. Schmidhuber. Mitosis detection in breast cancer histology images with deep neural networks. In MICCAI. 2013.

[4] 葛红，田联房 信息融合技术在模式识别中的应用 [A] 1001-3695 (2009) 01-0019-06.