A Novel Computer-aided Breast Cancer detection method based on Convolutional neural network

Xinkai Cai[1, a]

[1]Saint Anselm's Abbey School, Washington DC, USA

*Corresponding author e-mail: hzhc_wjdi@163.com

Abstract: Breast cancer is one of the main causes of cancer death worldwide. The diagnosis of cancer from eosin stained images is non-trivial and specialists often disagree on the final diagnosis. Computer-aided Diagnosis systems contribute to reduce the cost and increase the efficiency of this process. Conventional classification approaches rely on feature extraction methods designed for a specific problem based on field-knowledge. To overcome the many difficulties of the feature-based approaches, deep learning methods are becoming important alternatives. A method for the classification of hematoxylin and eosin stained breast biopsy images using Convolutional Neural Networks (CNNs) is proposed. Images are classified in four classes, normal tissue, benign lesion, in situ carcinoma and invasive carcinoma, and in two classes, carcinoma and non-carcinoma. The architecture of the network is designed to retrieve information at different scales, including both nuclei and overall tissue organization. This design allows the extension of the proposed system to whole-slide histology images. Accuracies of 77.8% for four classes is achieved. The sensitivity of our method for cancer cases is 95.6%.

1. Introduction

Breast cancer starts when cells in the breast begin to grow out of control. These cells usually form a tumor that can often be seen on an x-ray or felt as a lump. The tumor is malignant (cancer) if the cells can grow into (invade) surrounding tissues or spread (metastasize) to distant areas of the body. Breast cancer occurs almost entirely in women, but men can get breast cancer, too. Cells in nearly any part of the body can become cancer and can spread to other areas.

Different tests can be used to look for and diagnose breast cancer. Tranditional Tests including Chest x-ray,CT scan (computed tomography),MRI (magnetic resonance imaging),Ultrasound, PET scan, Bone scan[1,2],but these method of detecting breast cancer need doctors or Radiologists take a lot of time to interpreting images.

With the emergence of big data and the improvement of computer hardware computing power, deep learning has achieved great success in image processing, speech recognition, automatic driving, machine translation, etc[7].

Machine learning algorithms such as support vector machines are often used to detect and classify tumors. But they are often limited by the assumptions we make when we define features. This results in reduced sensitivity. However, deep learning could be ideal solution because these algorithms are able to learn features from raw image data.

2. Methods

2.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN), artificial neurons can cover a surrounding unit in a corresponding part, and have excellent performance for large image processing. This network includes a convolution layer and a pooling layer. This network consists of an input layer, a convolution layer, an activation function, a pooling layer, and a fully connected layer. Convolution

refers to the extraction of features on the original input, which extraction is simply a small fish in the original input to extract features.
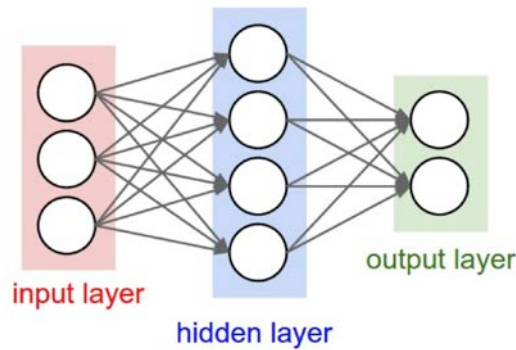

Fig.1. Fully connected neural network

2.2 Image Crops

Our raw images are 2048 * 1536 RGB Breast Cancer Images, are divided into Four kinds: Insuite, Benign, invasive, and normal. Since we only have 400 original images, we use image cutting to divide each image into 12 images of 512 * 512 color as input to our convolutional neural network. Four kinds of raw breast cance images are as follows:
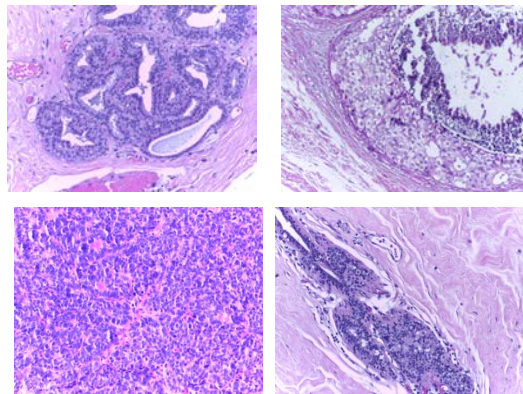

Fig.2. Four types of breast cancer, From left to right, top to bottom, the images are: Benign, Insitu, Invasive, Normal.

3. Experiment Design

3.1 DataSet Split

Our data set has a total of 4800 images and corresponding ground truth labels, divided into 3 data sets, namely: training set, verification set and test set, wherein the training set data is 4298 samples, used for model training, update Parameters such as weights and bias account for about 90% of the size of the entire data set, and the verification set is 478 samples. It is used to calculate the classification performance of the model after each round, and does not participate in training, accounting for about 10% of the data set. The remaining 24 samples are used to test the model effect after the model has completed all training. The data set format is as follows:

```
number of training examples = 4776
number of test examples = 24
X_train shape: (4776, 512, 512, 3)
Y_train shape: (4776, 4)
X_test shape: (24, 512, 512, 3)
Y_test shape: (24, 4)
```

Fig. 3. The dataset format as numpy array

3.2 Model Design

Our network hierarchy consists of 10 layers, including convolutional layer, max_pooling layer, and dropout layer, flatten layer and fully connected layer. The input layer receives 512*512*3 images, and the first layer is convolutional layer. The layer uses 16@3*3 convolution kernels, the step size is set to 1, the padding mode is 'valid', the activation function uses 'relu', followed by the max_pooling layer, and the pooling window is 3*3, step size Set to 3, followed by the second convolutional layer and the next max_pooling layer, with a dropout layer in the middle for regularization of the neural network to achieve random deactivation of the neural unit after convolution and pooling, The ratio is set to 0.5, and finally the flattening layer is used to change the multi-dimensional tensor into a 1-dimensional tensor, followed by 3 fully connected layers, and finally the softmax activation function is used to predict the probability of 4 types of breast cancer. The output gets the expected output. The model parameters and the output tensor for each layer are as follows.

Layer	Output Shape	Param #
input_1 (InputLayer)	(None, 512, 512, 3)	0
conv2d_1 (Conv2D)	(None, 510, 510, 16)	448
activation_1 (Activation)	(None, 510, 510, 16)	0
max_pooling2d_1 (MaxPooling2	(None, 170, 170, 16)	0
conv2d_2 (Conv2D)	(None, 168, 168, 32)	4640
activation_2 (Activation)	(None, 168, 168, 32)	0
max_pooling2d_2 (MaxPooling2	(None, 84, 84, 32)	0
conv2d_3 (Conv2D)	(None, 83, 83, 64)	8256
activation_3 (Activation)	(None, 83, 83, 64)	0
zero_padding2d_1 (ZeroPaddin	(None, 87, 87, 64)	0
max_pooling2d_3 (MaxPooling2	(None, 43, 43, 64)	0
dropout_1 (Dropout)	(None, 43, 43, 64)	0
conv2d_4 (Conv2D)	(None, 42, 42, 64)	16448
activation_4 (Activation)	(None, 42, 42, 64)	0
zero_padding2d_2 (ZeroPaddin	(None, 46, 46, 64)	0
max_pooling2d_4 (MaxPooling2	(None, 15, 15, 64)	0
conv2d_5 (Conv2D)	(None, 13, 13, 32)	18464
activation_5 (Activation)	(None, 13, 13, 32)	0
flatten_1 (Flatten)	(None, 5408)	0
dense_1 (Dense)	(None, 256)	1384704
dense_2 (Dense)	(None, 128)	32896
dense_3 (Dense)	(None, 4)	516

Fig. 4. Our designed network parameters and the output tensor for each layer

4. Data Analysis

4.1 Train&Validation Performace

When we experimented, we initially set the epochs to 30, and since we set the performance monitoring strategy in the model, that is, within 3 epoch, if the accuracy on the verification set is not improved, then we terminate the model training. The training process is as follows:

Fig. 5. The train process

The model performed best on the second epoch, with an accuracy of 0.9561 on the validation set, and the performance of the next 3, 4, and 5 epoch models did not exceed the second round, according to the set training rules. Our model will terminate training at the 5th epoch. At this point, we get the performance of the model on the training set, namely acc and loss, as shown below:
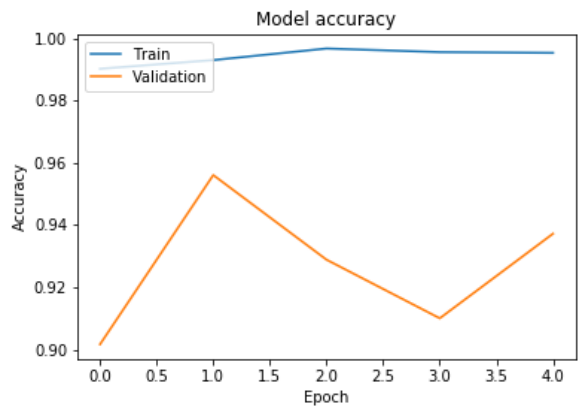

Fig. 6. Accuracy on the train set and validation set

4.2 Test Performance

After training the model, we fixed the parameters such as the weight and bias of the neural connection at each level of the network for the next model test. First we load our weight file and apply the model to the 24 test set images. To observe the effect of the model on the test set, the data is as follows:

Fig. 7. Performance on the test set

From the Figure 7, we can clearly see that the loss value on the test set is 2.826 and the acc value is 0.625. That is to say, 15 images in 24 images are correctly classified, and 9 images are misclassified. For each type of breast cancer, the model discrimination needs to be further improved. However, we found that the model distinguishes between benign and malignant breast cancers, which is 95.6%. It may be that the model is not very effective in distinguishing the characteristics of benign breast cancer. The image quality and early breast cancer itself are not obvious enough. The degree is not related to other factors.

5. Discussion And Conclusion

This study used convolutional neural networks to identify four types of breast cancer, achieving relatively high image classification accuracy and preliminary exploration of the feasibility of automated clinical detection of cancer. However, there are still many spaces for improvement in our

own convolutional neural network, including but not limited to: convolutional layer number, fully connected layer number, number of filters, filter size, number of hidden nodes, batch size, Learning rate, maximum size combined, discard rate, etc[8,9]. In addition, better image enhancement and consistent source images may perform better[8]. These are the direction of future research, providing better models and reliability for automated clinical testing of breast cancer.

References

[1] Bruening W, Uhl S, Fontanarosa J, Reston J, Treadwell J, Schoelles K. Noninvasive Diagnostic Tests for Breast Abnormalities: Update of a 2006 Review [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Feb. Accessed at www.ncbi.nlm.nih.gov/books/NBK84530/ on August 25, 2017.

[2] Caldarella C, Treglia G, Giordano A. Diagnostic performance of dedicated positron emission mammography using fluorine-18-fluorodeoxyglucose in women with suspicious breast lesions: A meta-analysis. Clin Breast Cancer. 2014;14(4):241-248.

[3] Jochelson MS. Chapter 12: Imaging Analysis: New Breast Imaging Techniques. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott Williams & Wilkins; 2014.

[4] Lee CI, Elmore JG. Chapter 10: Breast Cancer Screening. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed. Philadelphia, Pa: Lippincott Williams & Wilkins; 2014.

[5] Rhodes DJ, Hruska CB, Phillips SW, Whaley DH, O'Connor MK. Dedicated dual-head gamma imaging for breast cancer screening in women with mammographically dense breasts. Radiology. 2011;258(1):106-118.

[6] Weigert JM, Bertrand ML, Lanzkowsky L, Stern LH, Kieper DA. Results of a multicenter patient registry to determine the clinical impact of breast-specific gamma imaging, a molecular breast imaging technique. AJR Am J Roentgenol. 2012;198(1):W69-75.

[7] Hinton,G.E.,2007,Deep Belief Nets,2007 NIPS Tutorial Notes, Neural Information Processing Systems Conference, Vancouver, BC,December.

[8] Cheng, Jie-Zhi, et al. "Computer-Aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in CT scans." Scientific reports 6 (2016).

[9] Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." Nature (2017).

[10] MLA Krizhevsky, Alex, et al. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems (2012).