Breast Cancer Detection Using Transfer Learning in Convolutional Neural Networks

Shuyue Guan and Murray Loew
Department of Biomedical Engineering, George Washington University
Washington, D.C. USA

Motivation
For the U.S. women, breast cancer will be diagnosed among about 1 in 8 during their lifetime and it is the second leading reason for death [1].

Convolutional Neural Network (CNN) is a promising method for breast cancer detection, which can improve treatment outcomes for breast cancer and longer survival times for the patients [2].

To train the CNN from scratch needs a huge number of labeled images [4]. Such a requirement often is infeasible for mammographic tumor images. An alternative solution is to apply transfer learning.

Fig.3. Transfer learning: the features learned from natural images could be transferred to medical images

Training Sets for DDSM

Abnormal ROIs
Resize: 300x300
1300
benign (650) & malignant (650)

Normal ROIs
Resize: 300x300
1300

Validation: 10-fold cross
Classes: Binary

Fig.4. Mammography (BruceBlaus, 2014; Chung, 2008)

We firstly downloaded mammographic images from the DDSM database and cropped the Region of Interest images (ROIs) by given abnormal areas as ground-truth information.

Fig.5. (A) A mammographic image from DDSM rendered in grayscale; (B) Cropped ROI by the given truth abnormality boundary; (C) Convert Grey to RGB image by duplication.

DATA

METHODS

The structure of CNN in transfer learning was the combination of the 13 convolutional layers in pre-trained VGG-16 model with a simple FC layer. All the weights in 5 convolutional blocks were imported from the pre-trained VGG-16 model and not changed (“weights frozen”) during the training of this CNN.

Pre-trained VGG-16 model
Weights Frozen

Conv block 1
3x3 conv, 64
pre?
3x3 conv, 128
pre?
3x3 conv, 256
pre?
3x3 conv, 512
pre?
3x3 conv, 512
pre?

Conv block 2
3x3 conv, 64
pre?
3x3 conv, 128
pre?
3x3 conv, 256
pre?
3x3 conv, 512
pre?
3x3 conv, 512
pre?

Conv block 3
3x3 conv, 64
pre?
3x3 conv, 128
pre?
3x3 conv, 256
pre?
3x3 conv, 512
pre?
3x3 conv, 512
pre?

Conv block 4
3x3 conv, 64
pre?
3x3 conv, 128
pre?
3x3 conv, 256
pre?
3x3 conv, 512
pre?
3x3 conv, 512
pre?

Conv block 5
3x3 conv, 64
pre?
3x3 conv, 128
pre?
3x3 conv, 256
pre?
3x3 conv, 512
pre?
3x3 conv, 512
pre?

Flatten
Exten_1 layer
Input
0.016
0.081
0.163
0.245
0.327

Classifier:
Full-connected 256 neurons

Fig.6. Transfer learning CNN architecture: only weights in the FC layer were randomly initialized and updated by training.

RESULTS

Fig.7. Result: the red curve is training accuracy and blue is validation accuracy. The center line is average through 10-fold cross validation accuracy.

DISCUSSION

Such a training process can be seen as that the VGG-16 extracts features from input image and then these features were used to train a FC neural classifier.

Accuracy = 90.5
AUC=0.971

For best val-acc = 0.950

Pre-trained CNN on LSVRC datasets & Fine-tuning + Two-step decision (Jiao et al., 2016)
2-fold cross (600)
(Ben-Mal) 96.7 –

Pre-trained CNN with handcrafted features + RF (Dhungel et al., 2016)
5-fold cross (410)
(Ben-Mal) 91 ± 0.02 0.76

Pre-trained AlexNet + Sparse MIL (Zhu et al., 2016)
5-fold cross (410)
(Mal-norMal) 90.00 ± 0.02 0.85

Pre-trained VGG-16 + one FC layer (Curs)
10-fold cross (2600)
(Abnorm-Norm) 90.5 ± 3.2 0.96

Accuracy ± SD (Bootstrap k=10)

Fig.8. Receiver Operating Characteristic curve

Validity set (8 of images)

Main method
Validation
Accuracy %
AUC
Pre-trained CNN on LSVRC datasets & Fine-tuning + Two-step decision (Jiao et al., 2016)
2-fold cross (600)
(Ben-Mal) 96.7 –
Pre-trained CNN with handcrafted features + RF (Dhungel et al., 2016)
5-fold cross (410)
(Ben-Mal) 91 ± 0.02 0.76
Pre-trained AlexNet + Sparse MIL (Zhu et al., 2016)
5-fold cross (410)
(Mal-norMal) 90.00 ± 0.02 0.85
Pre-trained VGG-16 + one FC layer (Curs)
10-fold cross (2600)
(Abnorm-Norm) 90.5 ± 3.2 0.96