IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 5, Ver. II (Sep.- Oct. 2017), PP 50-55
www.iosrjournals.org

Transfer Learning on Pre-trained Deep Convolutional Neural Network for Classification of Masses in Mammograms

*Basanth M.S[1], Dr.Rajashree Shettar[2]
[1]Research Scholar, Department of CSE, R.V College of Engineering, Bengaluru, Karnataka, India
[2]Professor, Department of CSE, R.V College of Engineering,Bengaluru, Karnataka, India
Corresponding Author: Basanth M.S

***Abstract**: Computer-Aided Detection systems have developed as a support system for radiologists and physicians in lesions detection in mammograms, indicating the presence of breast cancer. Early detection of breast cancer could be facilitated through the use of computer technologies for identifying abnormalities such as masses, architectural distortion, and calcifications in mammograms. Efficient image classification and feature extraction will reduce the number of rescanning, the time required for diagnosis, accurate diagnosis, and quick treatment. Cost and waste reduction is important for making the diagnostics technologies reach the common people. Reading medical image need an expert radiologist or physician for the correct diagnosis. The minor human error can cause misdiagnosis and wrong treatment. The techniques in machine learning can support medical practitioners in reading the medical images. Masses in mammograms can be in any shape, but usually with convex outside borders. A mass found during a mammogram will typically be described according to its shape, density, and its margin. This paper presents classification of masses in mammograms (benign or malignant) using transfer learning on pre-trained Deep Convolutional Neural Network.*

***Keywords:** Computer-aided detection or diagnosis, Breast cancer, Convolutional NeuralNetwork, ConvNets, Transfer Learning, Deep Learning*

Date of Submission: 05-09-2017 Date of acceptance: 25-09-2017

I. Introduction

One of the leading reasons for cancer deaths is the Breast Cancer and it is one of the most common types of cancers in women worldwide Jemal, A., Siegel, R., & Ward, E. (2008). Diagnosis of malignant lesions during the initial stages in mammograms can reduce the death rate. The three assessments include clinical examination, pathological correlation, and radiological investigation. Diagnostic mammography is a standard procedure normally done for diagnosing breast cancer.

Radiologist looks for feature abnormalities in breast tissue that will aid in diagnostics. Occasionally they will notice simple changes in density, a ‘distortion‘, or the presence of micro-calcification, however, at other instances, the breast X-ray will surely monitor a ‘mass, which won't be clinically palpable. Most of the breast masses will be benign, however, if it gives suspicious characteristics of mass need further analyze for breast cancer. A mass is normally something a little more substantial and clearly visible than a ‘lesion’. It has volume and it occupies space. It could generally be in any shape, however, typically with convex outside borders. Mass normally has a tendency to be denser in the middle than in the direction of the edges. In comparison with a non-affected tissue, a mass will not be interspersed with fat cells to the same degree. A mass generally described according to its shape, density, and its margin.

Computer-Aided Detection (CAD) systems aid in increasing the accuracy and efficiency of the radiologist in identifying cancer cases. Jiang et al. Jiang, Y., Nishikawa, R. M., & Schmidt, R. A. (1999) and Chan et al. Chan, H. P., Doi, K., &Vybrony, C. J. (1990) concluded with their study that CAD could be used to support and improve radiologists in diagnosis of breast cancer.

In this paper, masses in a mammogram are detected and classified into benign or malignant using transfer learning of pre-trained convolutional neural network Lévy, D., & Jain, A. (Barcelona, Spain, Dec 2016). Mainly there are three phases. The first phase is fine tuning pre-trained Convolutional Neural Network (CNN) replacing layers to mammogram images. Second phase train the fine-tuned CNN using mammogram images create a model which can be used for classification. In the third phase, the trained model can be used for classification of input masses in mammograms into benign and malignant Jiao, Z., Gao, X., Wang, Y., & Li, J. (2016). Mammogram mass lesion classification using standard CNN is explored in Arevalo, J., Gonz´alez, F. A., Ramos-Poll´an, R., Oliveira, J. L., & Lopez, M. A. (aug 2015). In Liu, X., & Tang, J. (2014), authors discuss classification of mass in mammograms using geometry and texture features. In Chan, H. P., Doi, K., &Vybrony, C. J. (1990), authors describe a feature based framework for classification of breast masses.

This paper is based on designing a CAD system for semi-automatic mass classification in mammograms. The major objective is to classify original image to differentiate between benign and malignant masses in the breast. In Suzuki, S., Zhang, X., &Ichiji, K., et al. (2016), authors have used CNN to implement a classification system which efficiently categorizes images. CNNs are trained using a large number of mammogram images. This paper uses transfer learning to retrain AlexNet Suzuki, S., Zhang, X., Homma, N., &Ichiji, K. (2012), a pre-trained deep convolutional neural network for classifying mass in mammograms. From these large image databases, CNNs can learn rich features for a variety of mammogram images.

II. Methodology

Before going to the major phases, pre-processing techniques in mammograms for noise reduction and normalization is applied. This paperuses median filter for noise reduction and 'zero centre' for normalization. Images used for testing and validating the proposed solution are taken from "Digital Database for Screening Mammography (DDSM) database Digital Database for Screening Mammography (DDSM). (2004). A public database collaboratively maintained at the University of South Florida. This database has 2,620 studies. Images of each breast and image information in included in the study. Images which are having doubtful areas are associated with pixel-level "ground truth" information about the locations, and details of types of suspicious regions. The CBIS-DDSM dataset is a curated breast imaging subset of DDSM. It is a standardized version of DDSM. This database contains normal, benign, and malignant cases also provide verified pathology information. The size of the database along with detailed ground truth makes the DDSM a useful database for implementing and testing decision support systems. The CBIS-DDSM collection is a subset of the actual DDSM data, selected and curated by a trained mammographer. In this database, images are decompressed and converted to DICOM format from PNG format. The database also includes Region of Interest (ROI) segmentation and bounding boxes with the pathologic diagnosis of training data. During pre-processing, Noise reduction and contrast correction are carried out. Median filter for noise reduction and 'zero centre' for normalization has been used. Zero centre operation will subtract the image from its mean value. AlexNet, a Digital Database for Screening Mammography (DDSM). (2004) is a deep convolutional neural network trained to classify images into the 1000 different classes. AlexNet has multiple convolution layers before the pooling and activation layers, instead of alternating layers of convolution, activation, then pooling. AlexNet has dropout during training which reduces overfitting, and it also has data augmentation (rotations, translations, color variation) which improves robustness.

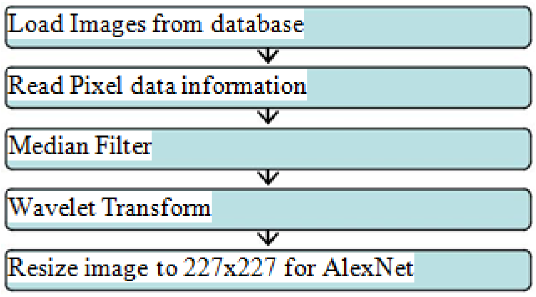

Figure 1: Preprocessing

Signal analysis is done using wavelet transforms, when signal frequency varies over time Arianna mencattini, Marcello Salmeri (2008), Chan, H. P., Doi, K., &Vybrony, C. J. (1990) and V, K., G., Ambalika, A.,S. (2010). Region of Interest (ROI) extraction is retrieved from the image using the truth information provided along with the DDSMdatabase. Resize the input image size to 227x227 as per the requirement of AlexNet.

III. Transfer Learning

A Convolutional Neural Network (CNN) is a powerful machine learning technique and it is also a deep learning technique. Convolutional neural networks are an important class of learnable representations to numerous computer vision related problems. Deep CCNs are composed of several layers of processing; each has linear and non-linear operators, to solve problems. These methods are used widely for retrieving features from audio, visual and text data.

Transfer Learning or Inductive learning stores knowledge attained from one problem solving and applies it to a completely different problem. In the neural network context, it is transferring learned features of a pre-trained network to a new problem. Training CNN from scratch is not an effective way especially with insufficient amount of training data. It is a general practice in deep learning to use a network that is trained on a large image dataset for a completely new problem. The network is fine-tuned by modifying the last few layers and keeping the initial layers of the pre-trained network as is, so as to learn the specific features of images in the new data set. Transfer learning generally results in faster training times than training from scratch the new CNN formed, reason is that the need to estimate all the parameters in the new network is not required.

The problem addressed in this paper is to categorize mammogram masses as either benign or malignant. The ROI images from DDSM dataset have been preprocessed and this is being used for training the pre-trained network. In this paper, a pre-trained neural network –AlexNet has been used for retraining our network with fresh data and task. The ROI images of benign and malignant categories are loaded into the CNN. The dataset is split into two different sets - training set and testing test. The pre-trained network, AlexNet which is a deep convolutional neural network has been pre-trained to recognize 1000 categories of objects and it has been trained on millions of images in the LSVRC-2010 ImageNet training set D., & J. (2009) and Russakovsky, O., Deng, J., Su, H., Krause, J., &Satheesh, S., et al. (2015). This network can be fine-tuned to transfer capability of the basic image pre-processing needed to classify the mammogram images.

AlexNet will be modified to recognize 2 classes instead of 1000 classes by modifying the corresponding layers. After this, the updated network learns weights for last few layers. To change the network, the learning rate of the modified layer is set to a small value (0.001).

Internally CNN extracts features from the ROI images which are fed as input. In this paper, images are classified into categories using a CNN after running transfer learning using ROI images. ConvNets are created using nodes or neurons with weights and biases used for learning. CNNarchitectures assume inputs are images, hence encoding of some characteristics into the architecture is possible. This makes efficient forward function and it reduces the number of parameters in the network. Convolutional Neural Networks allows constraining the architecture in a more sensible way as the inputs considered are images. ConvNet layers have neurons organized in three dimensions, i.e. width, height, and depth. Each layer in CNN transforms the 3D input to 3D output. The input layer has the image, so the dimensions of the image are the input layer width, height, and the depth which is three channels – Red, Green, and Blue.

A simple CNN uses a sequence of layers, and each layer has a differentiable function to transform one volume to another Z., D, M., & R.(2014). The three main types of layers used to build the CNN are: Convolutional Layer, Pooling Layer, and Fully Connected Layer. Stack of these layers forms a full CNN architecture.

Simplest CNN used for classification can have the architecture as shown below in the expression (1):

INPUT → CONVOLUTION → RELU → POOL → FC (1)

From expression (1), INPUT represents the Input layer which will hold the image pixel values which include height and width of the image along with Red, Green, and Blue representing the three color channels. CONVOLUTION represents the Convolution layer which calculates the neuron outputs which are connected to input local regions. The dot product of weights and a region they are connected to in the input volume is computed. The Fn filters will result in volume indicated as [Width x Height x Fn].

RELU layer represents the Rectified Linear Units which applies the activation function element-wise. One of the activation function that leaves the size of the volume unchanged is max(0, x), thresholding at zero.

POOL layer is for down sampling to reduce the volume size. Normally down sampling is done along the width and height of the image.

FC or Fully-Connected layer calculates the class or category scores which result in a volume of size [1x1xZ], where Z represents a class score. All the neurons of the previous layer are connected to each neuron in this layer.

CNN modifies the original image in each layer by considering the original image pixel values. The values are stored in to the final class or category scores. CONV/FC layers transformation function depends on the weights, input volume, and biases of the neurons. The RELU/POOL layers always implements a fixed function. The gradient descent is used to make sure the category scores which CNN calculated are matching with labels in the training set. In the classification phase, ROI images to be classified aregiven as input to the CNN machine model. The fine-tuned network will predict whether the new input image is malignant or benign.

IV. Result

The mammograms for experimentation are taken from Digital Database for Screening Mammography (DDSM) database. The overlay file from the DDSM dataset which has abnormality boundary information is read. Mammogram images with abnormality of type mass are only selected for training. Mammogram images are cropped as per the boundary value mentioned in the overlay file. Segregation of the mammogram images according to whether it is malignant or benign is carried out. Then apply data augmentation by preparing more images by rotating and flipping each image. Figure 2 shows the different operations part of experimental data preparation.

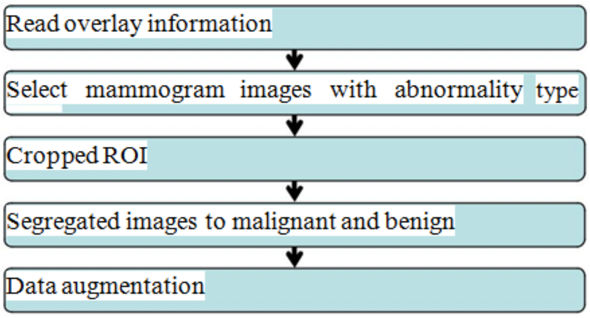

Figure 2: Process of Data Preparation

Data augmentation applied in cropped dataset which is a means by which add value to base data by adding information from other sources. This is useful to relatively small training sets which is normally the case with many medical image datasets.

V. Performance Analysis

Data augmentation can reduce the need for manually collecting huge data for training the model. It is also not easy to collect medical images huge in number for doing a deep CNN analysis. To increase the size of our dataset we prepared additional learning data by image rotation and mirroring Z., D, M., & R. (2014). For each training image, we perform 4 rotations and 2mirroring effectively augmenting training set size by a factor of 8. In the case of masses in mammograms these augmentations are justified since they have no inherent orientation and their diagnosis is invariant to rotation and mirroring. This paper used MatLab implementation of convolutional neural networks for performance analysis Roth, H. R., Lu, L., & Liu, J. (2015). With the data augmentation, system showed around 80% accuracy consistently. Standard CNN's are mainly invariant to transformations which are in the training data. So data augmentation makes the mass features rotation invariant/flip invariant. Table 1 shows the dataset size used for training and testing phases. Table 2 shows the confusion matrix for the test dataset.

Table 1: Dataset size with and without augmentation

Type	Training Set	Test Set	Total
Benign	4077	1019	5090
Malignant	4077	1019	5090

Table 2: Confusion matrix (Average number)

	Benign	Malignant	Total
Benign	TN = 878	FP = 141	1019
Malignant	FN = 213	TP = 806	1019
Total	1091	947	

Basic terms used in confusion matrix
(a) True Positives (TP) = 806
(b) True Negatives (TN) = 878
(c) False Positives (FP) = 141
(d) False Negatives (FN) = 213

Rates calculated from confusion matrix from binary classifier
(a) Accuracy = (TP+TN)/Total = (806+878)/2038 = 82.63
(b) Misclassification Rate = (FP+FN)/Total= (141+213)/2038 =17.37
(c) True Positive Rate = TP/Actual Yes= 806/1019 = 79.10
(d) False Positive Rate = FP/Actual No= 141/1019 = 13.84
(e) Specificity = TN/Actual No= 878/1019 = 86.16
(f) Precision = TP/Predicted Yes= 806/947 = 85
(g) Prevalence = Actual Yes/Total= 1019/2038 = 0.5

Table 3: Confusion Matrix (Percentage)

	Benign	Malignant
Benign	86.16% (Specificity)	13.84% (False Positive)
Malignant	20.90% (False Negative)	79.10% (True Positive)

The proposed system correctly predicted 86.16% malignant masses and 79.10% benign masses. System wrongly predicted 13.84% benign cases as malignant and 20.90% malignant cases as benign. This CAD system detects malignant masses with the sensitivity (true positive rate) of 79.10% which is good enough for using this system as a support system for radiologists.

VI. Discussion

This paper highlights the fact that the need for feature extraction V., A., &Lenc, K. (2014) is drastically reduced since in normal machine learning practices, the feature extraction phase is the most time-consuming procedure. CNN transfer learning is the most suited method for medical or mammogram imaging, as it is not practical to get huge number of mammogram images with truth information for training CNN. Also it is an uphill task to identify good number of useful features for classification.

Wavelet Transform D., Singh, N. A., &Selvi, S. T. (2014), which is one of the most popular time-frequency-transformations has also been used, which also improves model accuracy by 2-3%. The improvement shown is not very significant, but it helps in standardizing data for creating the model.

With the data augmentation, system shows around 80% accuracy. This shows importance of data augmentation in transfer learning using CNN. Data augmentations make the mass features in mammograms rotation and flip invariant. System is giving accuracy around 80% consistently with less number of false negatives and false positives for having used transfer learning technique for training the system.

VII. Conclusion

Mammography is the best available method for breast cancer detection, but there are cases in which experienced radiologists are unable to detect cancer despite their experience. Computer-aided methods presented in this paper can assist radiologists to improve the accuracy of detection. In this paper, classification of mammograms using CNN for feature detection has been presented. The proposed classification scheme has been applied on DDSM image dataset. Experimental results show that features detected using CNN show good classification accuracy with mammograms. The resulting high accuracy of classification of calcifications can reduce unnecessary biopsies.

Deep learning techniques such as CNN can give higher performance when properly tuned for classifying mammogram images. The major advantage of CNN has removed the need for feature engineering and theeffort to find the correct set of features which can be used for classification. Data augmentation is a method which will improve the learning accuracy especially when there is scarcity for training deep network or when using transfer learning.

Acknowledgment

The data used for experiment is from the “Digital Database for Screening Mammography” (DDSM) dataset which is in PNG format. **Courtesy - TM Deserno, Department of Medical Informatics, RWTH Aachen, Germany.**

References

[1]. Jemal, A., Siegel, R., &Ward, E. (2008). “Cancer statistics”, CA: A

[2]. Cancer Journal for Clinicians.“Cancer statistics”, CA: A Cancer Journalfor Clinicians, 58(2), 71-96.

[3]. Jiang, Y., Nishikawa, R. M., & Schmidt, R. A. (1999)."Improving breast cancer diagnosis with computer aided diagnosis”, Academy Radioliology, 6(1), 22-33.

[4]. Chan, H. P., Doi, K., &Vybrony, C. J. (1990).“Improvement in radiologist's detection of clustered microcalcifications on mammograms: the potential of computer aided diagnosis”, Investigative Radiology, 25(10), 1102-1110.

[5]. Lévy, D., & Jain, A. (Barcelona, Spain, Dec 2016). “Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks”,.Proceeding of the 30th Conference on Neural Information Processing Systems.

[6]. Jiao, Z., Gao, X., Wang, Y., & Li, J. (2016). "A deep feature based framework for breast masses classification”. Neurocomputing, 197, 221-231.

[7]. Arevalo, J., Gonz´alez, F. A., Ramos-Poll´an, R., Oliveira, J. L., & Lopez, M. A. (aug 2015). “Convolutional neural networks for mammography mass lesion classification”, Proceeding of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 797-800.

[8]. Liu, X., & Tang, J. (2014).“Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method”. IEEE Systems Journal.

[9]. Suzuki, S., Zhang, X., &Ichiji, K., et al. (2016).“Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis,".Proceeding of the 55th Annual Conference of the Society ofInstrument and Control Engineers of Japan (SICE), 1382 - 1386. DOI:10.1109/SICE.2016.7749265

[10]. Suzuki, S., Zhang, X., Homma, N., &Ichiji, K. (2012).“Imagenet classification with deep convolutional neural networks”, Advances inNeural Information Processing Systems., 1097 -1105.

[11]. Digital Database for Screening Mammography (DDSM). (2004).University of South Florida.

[12]. Arianna mencattini, Marcello Salmeri (2008), “Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing”.IEEE transactions.

[13]. V, K., G., Ambalika, A., & S. (2010).“Contrast Enhancement of Mammographic Images Using Wavelet Transform”, 3rdIEEE International Conference on Computer Science and Information Technology.

[14]. D., Singh, N. A., &Selvi, S. T. (2014).“Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach”, Journal of Biomedical Informatics, 49, 45-52.

[15]. D.,& J. (2009). "Imagenet: A large-scale hierarchical image database.", Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[16]. Russakovsky, O., Deng, J., Su, H., Krause, J., &Satheesh, S., et al. (2015).“ImageNet Large Scale Visual Recognition Challenge”, International Journal of Computer Vision (IJCV), 115(3), 211-252.

[17]. Z., D, M., & R. (2014). "Visualizing and understanding convolutional networks.", Computer Vision-ECCV by Springer International Publishing, 818-833.

[18]. Roth, H. R., Lu, L., & Liu, J. (2015).“Improving Computer aided detection using convolutional neural networks and random view aggregation”, IEEE Transactions Medical Imaging.

[19]. V., A., &Lenc, K. (2014)."MatConvNet-convolutional neural networks for MATLAB", ArXiv, 1412-1421.

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Basanth M.S. “Transfer Learning on Pretrained Deep Convolutional Neural Network for Classification of Masses in Mammograms.” IOSR Journal of Computer Engineering (IOSR-JCE), vol. 19, no. 5, 2017, pp. 50–55.