International Conference on Cancer Research 2019

Website: www.CancerConf.org

June 10-11, London, UK

Performance of a Transfer Learning Module with Pretrained Neural Network and Logistic Regression in Detection of Breast Carcinoma from Microphotographs of Fine Needle Aspiration Cytology Smears

Parikshit Sanyal [1], Anirudh Sharma [2], Prosenjit Ganguli [3*], Sanghita Barui [1]

[1] *Dept of Pathology, Military Hospital Jalandhar, Jalandhar Cantt, PIN 144005, India*

[2] *Undergraduate student, Vellore Institute of Technology, Katpadi,Vellore, Tamil Nadu, PIN 632014, India*

[3] *Dept of Pathology, Command Hospital, Alipore, Kolkata 700027, India*

** Corresponding author: prosenjitganguli@hotmail.com*

Abstract

Fine needle aspiration cytology (FNAC) is an useful modality for initial assesment of a palpable breast lump. The aim of this study was to develop a computer model for classification of microphotographs from FNAC smears of breast lesions, into two classes 'benign' and malignant'. We have used the transfer learning method, i.e. using a neural network which has been trained on a different dataset, to extract features from the present datatset. Apart from being of valuable diagnostic utility, the model will also provide key insights on machine learning and how a learner, human or machine, distinguishes benign from malignant.

A pretrained neural network (VGG16) which has been trained on the ImageNet database, was used for the study. A total of 2037 processed microphotographs from Romanowky stained FNAC smears were taken, all at 40x magnification. Images from two different microphotography systems in two different tertiary care centers of India was used. The images were then split into two sets, 'training' (1544 images) and 'validation' (493 images). During training, features were extracted with VGG16 and fit with original labels using logistic regression. After completion of training, images from the validation set was processed with the VGG16 network and the trained logistic regression model was used to generate predictions.

The model achieved 90.38% sensitivity, 87.12% specificity, 88.67% positive predictive value and 89.03% negative predictive value. A diagnostic accuracy of 89% was achieved. Receiver operating characteristic shows area under curve of 0.89, indicating good perforamance. 12.8% false positives and 9.6% false negatives were also reported by the model. The principal difficulties encountered were the distinction between the dark staining nuclei of myoepithelial cells and the hyperchromasia of a malignant epithelial cell. Also, hypocellular foci with single malignant epithelial cells were often reported to be falsely negative by the model.

Overall, the sensitivity, specificity, positive and negative predictive value of the model is close to FNAC reported by pathologists. It shows potential to be used as a screening tool, after validation on a larger dataset.

Keywords:

Pretrained Neural Network,

Logistic Regression,

Breast Carcinoma,

Microphotographs.

1. Introduction

Breast cancer is the most common cancer in women worldwide, with 2 million new cases reported in 2018.[1 2] Fine needle aspiration cytology (FNAC) is part of the triple assessment of a suspected breast lump as it is minimally invasive and provides cytological material for identifying malignant lesions. In a study from Italy of 210 patients, sensitivity, specificity, diagnostic accuracy, positive predictive value (PPV) and negative predictive value (NPV) of 97%, 94%, 95%, 91% and 98% were found for FNAC breast.[3] A large systematic review found that the sensitivity of core needle biopsy (CNB) is better than that of FNAC [87% (95% CI, 84%-88%, I2 = 88.5%) versus 74% (95% CI, 72%-77%, I2 =

88.3%)] and the specificity of CNB is similar to that of FNAC [98% (95% CI, 96%-99%, I2 = 76.2%) versus 96% (95% CI, 94%-98%, I2 = 39.0%)].[4]

The aim of the present study was to develop a machine learning model which will be able to scan Romanowsky stained smears, mark areas with features suggestive of malignancy and present them to the pathologist; in effect the model will serve as a screening tool for FNAC smears of breast lesions. The performance characteristics of the model will have to be comparable to the established sensitivity, specificity, PPV and NPV of manual FNAC by the pathologist.

Machine learning methods have evolved sufficiently in complexity so that such a task can be undertaken. In particular, Artificial Neural Networks (ANNs) have been found to be suitable for image recognition tasks. [5]Artificial neural networks (ANN) are mathematical models based on systems of linear algebra. An ANN constitutes of several layers of 'neurones', individual nodes which take a number as input, perform a linear operation on it and return an output. For each neurone, presented with an input *x*, the output *y* is:

$$y = f(wx + b)$$

Where 'w' and 'b' are two parameters of the neuron, called 'weight' and 'bias' of the neuron respectively. *f(x)* is a function which transforms the input into output. The functions that are usually employed for this purpose are the sigmoid function or the rectified linear unit (ReLu) function. An ANN is made of up many such layers of neurones. The ANN calibrates it's coefficients 'w' and 'b' in each neuron, as to perform a certain task, e.g. pattern recognition.[6] Convolutional neural networks (CNNs) are a special class of ANNs which take a whole image as input and classifies the image in defined categories. A special operation called 'convolution' is performed in a CNN, which reduced the dimensions by averaging local pixels increase its depth. CNNs have been described in detail by Karpathy et al.[7] Multilayered deep neural networks have successfully been able to classify images in separate categories, from a large and diverse database.[8]

There has been use of neural networks in cytology. Savala et al have used neural networks based on morphometric features, to distinguish follicular adenoma from carcinoma in thyroid FNAC smears. [9] Conventional Pap smears have been successfully analysed with the PAPNET network by Mango et al.[10] However, these studies used morphometric data from the images as inputs to an ANN, and have not used a CNN.

Constructing a CNN requires thousands of training images and computer processing time. However, open source CNNs have been developed which have been trained over a very large dataset of diverse classes of images. Transfer learning is one of the emerging paradigms in machine learning which makes use of these pre-trained CNNs to perform a specific task. In the transfer learning approach, a neural network which has been pre-trained on a data set is used to analyse and predict on a different data set. The method of transfer learning has been described by Karpathy et al.[11] In the present study, we have used the VGG16 [12] neural network, which has been developed by the Visual Geometry Group at University of Oxford. [13] The VGG16 neural network has been pre-trained with the ImageNet dataset, consisting of 32326 different classes of images.[14]

2. Materials and Methods

Fine needle aspiration cytology smears from both benign and malignant lesions of breast were selected from the archives of two tertiary care hospitals of India. All smears were stained with Romanowsky stain (May-Grünwald Giemsa or Leishman Giemsa). The smears where diagnosed cytologically by a group of pathologists. A total of 288 micro photographs from malignant and 127 micro photographs from benign lesions were photographed. Images from one hospital was photographed using Olympus microphotography system on a Labomed LX300 microscope. Images from the other hospital were photographed in a DeWinter Digi510 Microphotography system. All microphotographs were taken at 40x maginification. Images of optimal quality were selected and a python script using openCV was then used to slice the photographs into 8 random slices of 341 x 256 pixels. The cropped images with trainable features specific to either the benign or the malignant categories were selected. A total of 2037 cropped images were thus produced. The images were then split into two data sets: one data set for training the CNN and another for validation. Distribution of the image data is shown in Table 1.

Table 1: Distribution of the processed microphotographs in different categories (N=2037)

	Training	**Validation**	**Total**
Benign	802	233	1035
Malignant	742	260	1002
Total	1544	493	2037

The images were read with the OpenCV library, and resized to a dimension 224 x 224, because the VGG16 neural network operates on this dimension. Weight of a VGG16 model trained on the ImageNet database were imported, using the Keras python library.[15] The model was imported after removing the Fully Connected Layers, i.e. the final step of prediction is skipped and the output of the penultimate layer of the neural network is extracted. We used this output as features vectors extracted from the image. Each of the 1544 training images generated a feature array, which is a 3-dimensional array of size 7 x 7 x 512. The feature array is then flattened into a feature vector of size 7 x 7 x 512 = 25088. This flattened feature vector of 25088 floating point numbers is then fitted with the original labels (0 for 'benign' and 1 for 'malignant) using a logistic regression model. After completion of training, the same algorithm was applied to the validation set of 493 images. Features were extracted by the VGG16 network, and the feature vectors were processed by the LR model to generate predictions, whether the image belongs to class '0' (benign) or '1' (malignant).

3. Results

Diagnostic accuracy achieved with the validation set was 0.89. Table 2 shows a contingency table of the predictions versus the original labels.

Table 2: Contingency table showing original labels and predicted labels of microphotographs of the validation set (N=493)

		Original label (diagnosed by pathologist)		Total
		Benign	**Malignant**	
Predicted by the model	**Benign**	203	25	228
	Malignant	30	235	265
		233	260	493
Sensitivity	90.38%	Specificity	87.12%	
Positive predictive value	88.67%	Negative predictive value	89.03%	

30 false positives (12.8% of benign microphotographs) and 25 false negative cases (9.6% of microphotographs with malignant foci) were reported by the predictor model. A receiver operating characteristic (ROC) curve (Figure 1) showed good performance with an area under curve (AUC) = 0.89.

Figure 2 shows few images from the validation set with their original and predicted labels.

Figure 1: Receiver operating characteristics of the predictor model

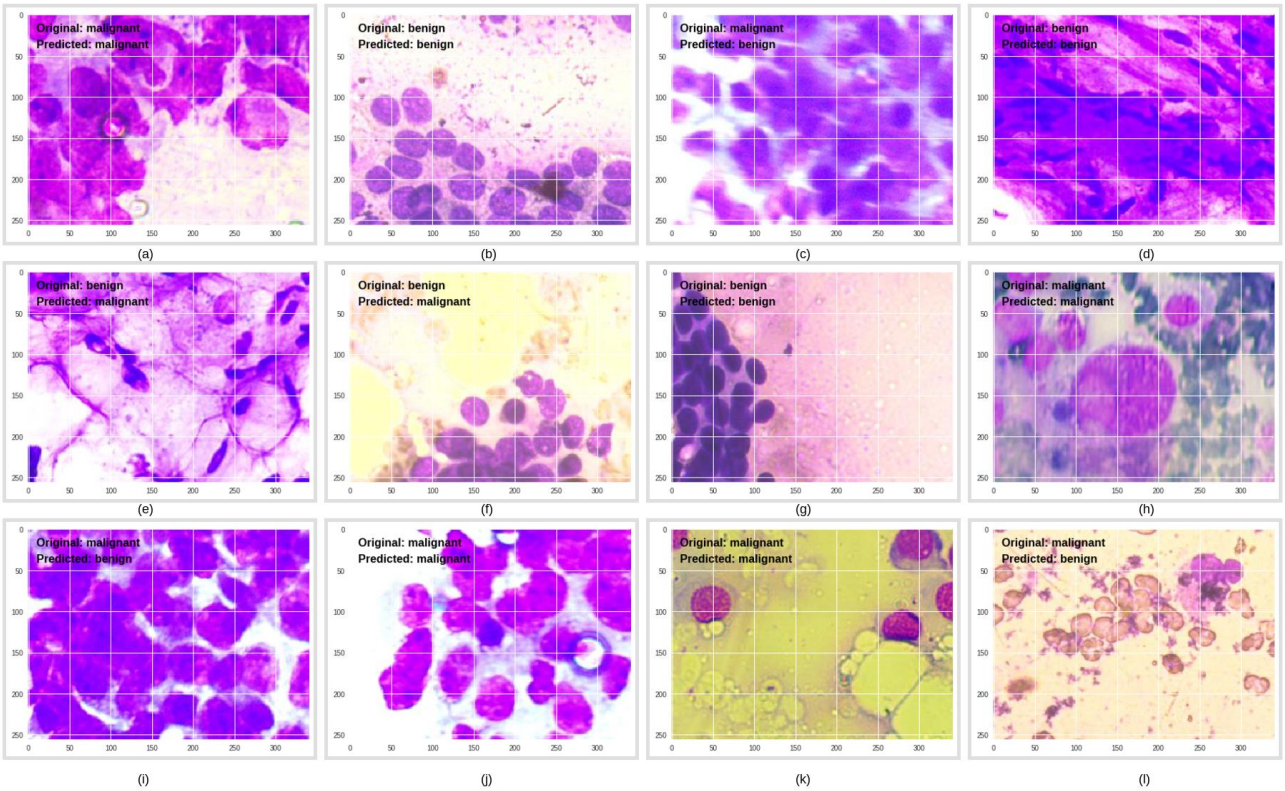

Figure 2: A subset of images in the validation set with original and predicted labels

4. Discussion

In a systematic review by Pouliakis et al in 2016, it was concluded that cytopathology is the ideal ground for application of artificial neural networks.[16] Machine learning tools are now available for screening Papanicolau stained cervical smears. [17] However, such models have not been extended to cytology of other organs. Several studies have used morphometric data from images as inputs to neural networks, which require the intervention of a pathologist or a morphometric device. Previous studies that have employed machine learning in cytological detection of breast cancer have also used morphometric data from images[18], such as radius, texture, perimeter, area, compaction (square perimeter divided by area), concavity, symmetry and fractals [19]. Using morphometric features requires an intermediary, either a human pathologist or a morphometric device. A few other studies have performed automated identification of breast cancer cells on immunohistochemically stained smears; a machine learning tool has been developed by Bolton for predicting immunohistochemical scores from tissue microarrays of breast cancer.[21] Algorithms have also been developed for automated screening of cytokeratin stained sections from sentinel lymph node biopsies.[22] In the present study, we have used whole images of Romanowsky stained cytological smears as input to the machine learning model.

The distinction between benign and malignant breast lesions from microphotographs of FNAC smears is a non-trivial machine learning problem. The criteria for identifying focus of carcinoma cells in FNAC smears from breast lesions include absence of myoepithelial cells, no single bare bipolar nuclei, and loss of cell cohesion moderate to severe nuclear atypia: enlargement, pleomorphism, irregular nuclear membrane and chromatin. In addition fibroblasts and fragments of collagen maybe present depending on presence of stromal desmoplasia. In contrast. FNAC from non-neoplastic breast tissue will include dark staining myoepithelial cells. Variable number of bare nuclei and bipolar cells are found in the background also the epithelial cells will have small round nuclei with bland chromatin.[20] such features are learnt by the trainee pathologist through exposure to several examples both benign and malignant. It is this reinforcement learning over repeated exposure to several training examples that impart learning to the brain. Neural networks operate on a similar principle. Recognition of individual cytological features by geometric methods is not computationally feasible because of the immense variability encountered in any field of biology, including cytology. Neural networks take a holistic approach to the problem and process the image as a whole.

The present study applies a neural network to whole image directly micro photographed from fine needle aspiration cytology smears of breast lesions. The sensitivity, specificity and diagnostic accuracy of the neural network was comparable to conventional fine needle aspiration cytology the network good successful identify features of carcinoma

in the micro photographs from the FNAC smears (Table 2).

Of the 233 images in the 'benign' category, 30 has been reported as false positive (12.8%). However, of the 260 images of the 'malignant' category, only 25 was reported false negative (9.6%). The low false negative rate has resulted in good negative predictive value (89.03%), suggesting potential utility in screening.

Figure 2(a), 2(h), 2(j) and 2(k) shows examples of foci where the model has successfully detected foci of malignancy. In Figure 2(a), 2(h) and 2(j), hyperchromatic, enlarged nuclei with uneven chromatin were recognised by the model. The model was also able to detect the isolated abnormal nuclei in an otherwise hypocellular focus (Figure 2(k)).

Figure 2(c) and 2(i) show a foci which has been reported false negative. In both these cases, the model has failed to distinguish epithelial cells with hyperchromatic nuclei from dark staining myoepithelial cells. In figure 2(j), in spite of obvious features of malignancy in the nuclei, the focus has been reported falsely negative, possibly due to hypocellularity in the focus.

Figure 2(e) and 2(f) represent clearly benign foci, which has been reported as false positives. The darker staining myoepithelial cells and nuclei of adipocytes have been falsely detected to be malignant. This might be attributable to overfitting on the training data, i.e. the model fitting both on the 'signal' (the actual epithelial and myoepithelial cells) as well as 'noise' (adipocytes, debris, background material) in the training data.

The present model depends on extracted features from the original image for its functioning. Each image produces a feature array of shape 7 x 7 x 512. Figure 3 shows systematic slices from this array, i.e. from 0 to 511, at each 60th slice. Each of these slices is a grayscale image, reproduced in figure 3 in artificial colors to highlight the features. All the 512 slices, each of shape 7 x 7, is flattened to produce a vector of 25088 dimensions, which is then fit with logistic regression to original label of the image.

5. Conclusion

We have demonstrated the performance of a predictor module based on transfer learning and logistic regression, to classify microphotographs from fine needle aspiration cytology smears of breast lesions, into two categories benign and malignant. The sensitivity of the model was 90.38%, suggestive of potential utility in screening. With further training with a diverse dataset, the model can evolve into an effective and accurate screening tool for FNAC smears of breast lesions.

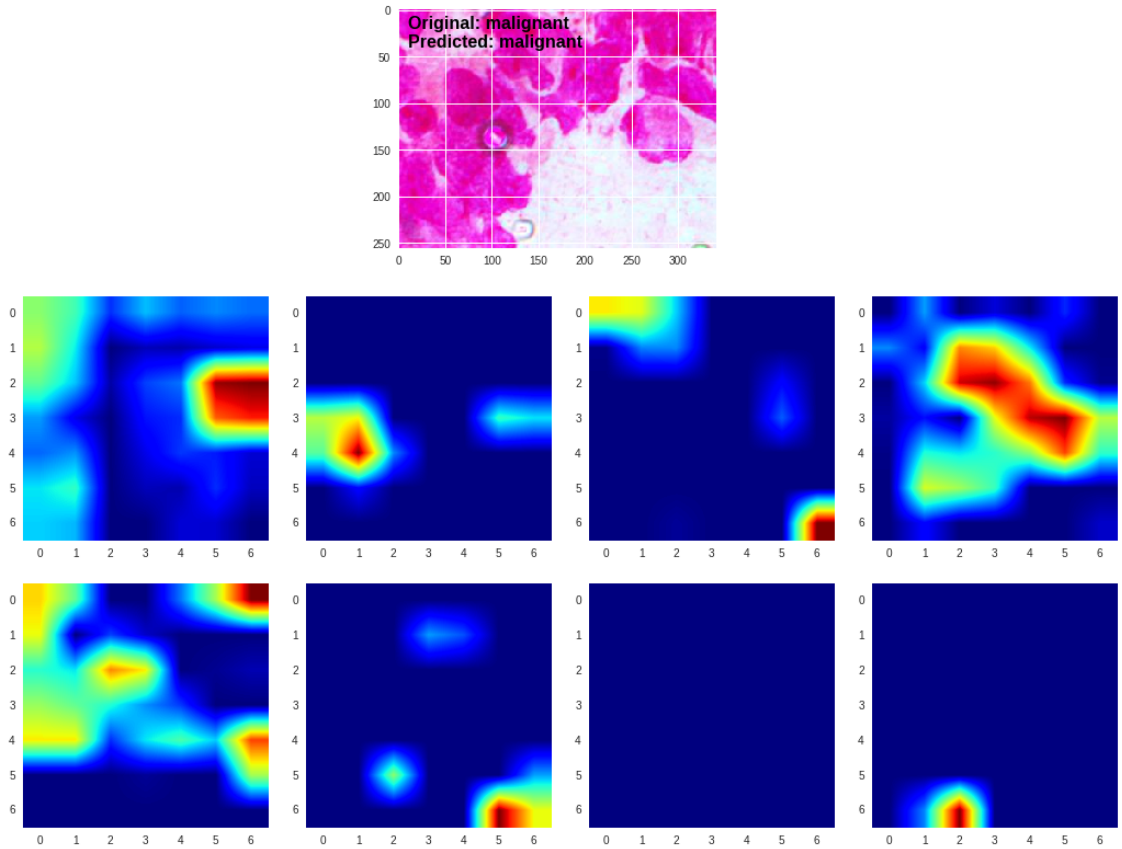

Figure 3: An image from the validation set and subset of its extracted features

5. References

1. Ghoncheh M. Incidence and Mortality and Epidemiology of Breast Cancer in the World. - PubMed - NCBI. Asian Pac J Cancer Prev. 2016; 17(S3):43–6.
2. Breast cancer statistics [Internet]. World Cancer Research Fund. 2018 [cited 2019 Jan 20]. Available from: https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics
3. Moschetta M M, Telegrafo, M. M, Carluccio DA, Jablonska JP, Rella L, Serio G, et al. Comparison between fine needle aspiration cytology (FNAC) and core needle biopsy (CNB) in the diagnosis of breast lesions. Il G Chir. 2014 Sep 1;35(7–8):171–6.
4. Wang M, He X, Chang Y, Sun G, Thabane L. A sensitivity and specificity comparison of fine needle aspiration cytology and core needle biopsy in evaluation of suspicious breast lesions: A systematic review and meta-analysis. Breast Edinb Scotl. 2017 Feb;31:157–66.
5. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. ArXiv151203385 Cs [Internet]. 2015 Dec 10 [cited 2018 Nov 12]; Available from: http://arxiv.org/abs/1512.03385
6. Haykin S. Neural Networks: A Comprehensive Foundation. 2nd ed. New York: Prentice Hall; 1999.
7. CS231n Convolutional Neural Networks for Visual Recognition [Internet]. [cited 2019 Jan 10]. Available from: http://cs231n.github.io/convolutional-networks/
9. Savala R, Dey P, Gupta N. Artificial neural network model to distinguish follicular adenoma from follicular carcinoma on fine needle aspiration of thyroid. Diagn Cytopathol. 2017 Dec 20;46.
10. Mango LJ. Computer-assisted cervical cancer screening using neural networks. Cancer Lett. 1994 Mar 15;77(2):155–62.
11. CS231n Convolutional Neural Networks for Visual Recognition [Internet]. [cited 2019 Feb 13]. Available from: http://cs231n.github.io/transfer-learning/
12. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv14091556 Cs [Internet]. 2014 Sep 4 [cited 2019 Feb 13]; Available from: http://arxiv.org/abs/1409.1556
13. Visual Geometry Group Home Page [Internet]. [cited 2019 Feb 13]. Available from: http://www.robots.ox.ac.uk/~vgg/
14. ImageNet [Internet]. [cited 2019 Feb 13]. Available from: http://www.image-net.org/
15. Home - Keras Documentation [Internet]. [cited 2019 Jan 20]. Available from: https://keras.io/
16. Pouliakis A, Karakitsou E, Margari N, Bountris P, Haritou M, Panayiotides J, et al. Artificial Neural Networks as Decision Support Tools in Cytopathology: Past, Present, and Future. Biomed Eng Comput Biol. 2016 Feb 18;7:1–18.
17. Cervical Cancer Screening - BD [Internet]. [cited 2019 Jan 10]. Available from: https://www.bd.com/en-us/offerings/capabilities/cervical-cancer-screening
18. Isa NAM, Subramaniam E, Mashor MY, Othman NH. Fine Needle Aspiration Cytology Evaluation for Classifying Breast Cancer Using Artificial Neural Network. Am J Appl Sci. 2007 Dec 31;4(12):999–1008.
19. Fiuzy M, Haddadnia J, Mollania N, Hashemian M, Hassanpour K. Introduction of a New Diagnostic Method for Breast Cancer Based on Fine Needle Aspiration (FNA) Test Data and Combining Intelligent Systems. Iran J Cancer Prev. 2012;5(4):169–77.
20. Cangiarella J. Breast. In: Orell & Steretts Fine Needle Aspiration Cytology. 5th ed. Edinburgh: Churchill Livingstone; 2012. p. 162.
21. Bolton KL, Garcia-Closas M, Pfeiffer RM, Duggan MA, Howat WJ, Hewitt SM, et al. Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2010 Apr;19(4):992–9.
22. Holten-Rossing H, Talman M-LM, Jylling AMB, Laenkholm A-V, Kristensson M, Vainer B. Application of automated image analysis reduces the workload of manual screening of sentinel lymph node biopsies in breast cancer. Histopathology. 2017 Dec;71(6):866–73.