

Deep Learning-based Image Analysis Model for Diagnosing Thyroid Carcinoma in Fine Needle Aspiration Cytology (FNAC) Images

Gopinath Balasubramanian¹ and Santhi Ramalingam²

¹Associate Professor, Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore-641049, India

²Assistant Professor, Department of Biochemistry, PSG College of Arts & Science, Coimbatore-641014, India

ABSTRACT

In this work, the diagnostic accuracy of an automated diagnosis system is evaluated using two pre-trained convolutional neural network models, namely AlexNet and VGG16. The diagnosis system is used to identify the cancerous and normal thyroid cells in Fine Needle Aspiration Cytology (FNAC) photographs. The proposed Alexnet and VGG16 models are implemented using deep learning based Transfer Learning (TL) to process multi-stained FNAC images. Initially, the image patches are derived from the cytology images based on the thyroid cell population. These patches are fed to the 8-deep layered AlexNet and 16-deep layered VGG16 as inputs and they are passed through multiple convolution layers, max pooling and dense layers. Through optimal implementation and testing of the models, the AlexNet model achieves a diagnostic accuracy of 92.5% whereas the VGG16 model results a diagnostic accuracy of 96.66% and sensitivity and specificity of 98.75% and 92.5% respectively.

KEY WORDS: BENIGN, CARCINOMA, DEEP LEARNING, MALIGNANT, THYROID.

INTRODUCTION

Thyroid is a gland in human body producing Thyroid Hormone (TH) that regulates metabolic processes in human body for normal growth and development (Mullur et al., 2014). In few cases, Around the thyroid organ, unwanted tissue section is formed which is known as thyroid nodules. Normally, this swelling nature of nodule is harmless as well as treated as benign type of nodules. However, around 8–15% of them are accounted as

malignant and the malignant nodules should be removed by suitable surgical treatment. (Cooper et al., 2009). Hence, accurate discrimination of malignant nodule is an essential thing. High degree of attention must be given for thyroid malignant studies (Kwong et al., 2015). FNAC is a simple cost-effective and widely used method procedure for evaluating the head and neck masses with an accuracy of around 90%. (Wong et al., 2020). In this procedure, the area to be aspirated is cleaned properly and a sufficient small size of biopsy sample is derived using a syringe of 23 gauge needle setup from the thyroid nodule. The smears are prepared by the pathologist using these samples. All the smears are stained with suitable staining protocol in the pathology laboratory and placed under a microscope to study the characteristics of cell structures of thyroid nodule. Based on the experience of the pathologist and nature of appearance of the smear

ARTICLE INFORMATION

*Corresponding Author: gopinath.b.ece@kct.ac.in
Received 11th Oct 2020 Accepted after revision 12th Dec 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRCBA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal



NAAS Journal Score 2020 (4.31) SJIF: 2020 (7.728)
A Society of Science and Nature Publication,
Bhopal India 2020. All rights reserved.
Online Contents Available at: <http://www.bbrc.in/>
Doi: <http://dx.doi.org/10.21786/bbrc/13.11/14>

under microscope, an appropriate diagnostic report is generated. However, this routine manual screening of cytological slides is a tedious task and subjective in nature.

Various automated techniques are currently used by the researchers for diagnosing malignant thyroid nodule. The latest developments in deep learning algorithms are being used to solve the issues in the diagnosis process (Thomas et al., 2020). A combined neural networks and morphometric feature-based model was designed for the discrimination problem and a diagnosis accuracy of 97.8% was achieved (Karakitsos et al., 1999). A k-NN classifier and Bayesian models were used by Würflinger et al., (2004) to classify the cell nuclei in pathology images and a diagnostic accuracy of 86.1% using Bayesian model and 87.5% using k-NN model were achieved. Ahmed et al., (2016) has proposed an intelligent diagnostic system for separating papillary and other thyroid carcinoma levels with neural network classifier model and obtained a diagnostic accuracy of 90.32%.

In the previous work, Gopinath, & Gupta, (2010) designed a computer aided diagnosis tool to differentiate cancerous and non-cancerous thyroid FNAC cytology images using Gabor features and reported 93.33% of diagnostic accuracy. Subsequently, authors (Gopinath, & Shanthi, 2013; Gopinath, & Shanthi, 2015) tested SVM, ENN, k-NN and decision tree classifiers trained with wavelet decomposition based statistical features, for the same problem, from which the single SVM and ENN classifiers resulted an accuracy of 90% to separate cancerous and non-cancerous thyroid nodules. However, an improved accuracy of 96.66% was achieved by combining SVM, ENN and k-NN classifiers using linear combination rule and majority voting rule. Using ultrasound radiology images also, there was an automated screening system developed for detecting the malignancy in Thyroid nodules. For the detection of nodules, initially, a region-based detection network was developed to extract pyramidal features. In the next stage, a classification network was developed to classify the ultrasound thyroid images. The detection and diagnostic accuracy of the deep-learning-based CAD system were observed as 97.5% and 97.1%, respectively (Liu et al., 2019).

The skin lesions were automatically classified using AlexNet model and transfer learning (Hosny et al., 2019). The weights of the architecture were fine-tuned and a softmax layer was used on skin lesions. The proposed method produced a classification accuracy of 97.70%. Chen et al., (2020) examined 345 thyroid sections by applying the deep learning techniques and extracting the patch features and 96.1% of classification accuracy was reported. A two-phase algorithm was developed for automated extraction of red blood cells (RBCs) (Aliyu et al., 2020). Initially, cell regions were extracted and in the second step, AlexNet was used for diagnosing the abnormal conditions in the cells. Around, 9000 images were processed, for the study, using AlexNet transfer learning model. The study resulted a prediction accuracy of 95.92%, sensitivity and specificity of 77% and 98.82%

respectively. A kernel-based classifier and optimization techniques were used with significant feature choice by Shankar et al., (2020) which enhanced the diagnosis process and they achieved a classification accuracy of 97.49%.

The current research work proposes the evaluation of the diagnostic accuracy of an automated diagnosis system using two pre-trained convolutional neural network models, namely AlexNet and VGG16. The FNAC images are processed by implementing the neural network models to diagnose the normal and cancerous thyroid cell regions. The proposed Alexnet and VGG16 models are implemented using deep learning based Transfer Learning (TL) to process multi-stained FNAC images.

MATERIAL AND METHODS

Acquiring Thyroid FNAC images: Fine needle aspiration cytology is a partially invasive procedure for determining cancerous and normal state of the thyroid nodules under study. Thyroid biopsy samples of thyroid nodule are taken by a fine-needle and transferred to the pathology laboratory. Then, the glass slides are prepared and analyzed under a microscope by a pathologist. After the examination of slides, the pathologist generates a diagnosis summary of report. The diagnostic report consists of four classes, namely, malignant, benign, inadequate sample and suspicious state. In general, the FNAC technique produces an accuracy of around 90-95% whereas the falsely identified positive cases and falsely identified negative results are varying in the range of 0-5% (Sinna, & Ezzat, 2012).

From the report of FNAC, the Diagnostic Accuracy (DA), sensitivity and specificity parameters are obtained as given in Eqs. (1-3) respectively. The parameter sensitivity can calculate the count of truly positive (TP) that correctly determines a positive case as positive. Similarly, the specificity calculates the count of truly negative (TN) that correctly identifies a negative case as negative. The DA is a combined measure of true positive and true negative.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (1)$$

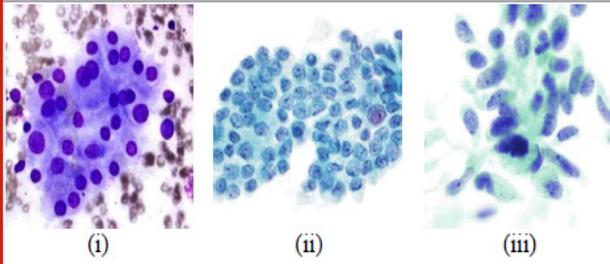
$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2)$$

$$DA = \frac{(TP+TN)}{(FP+FN+TP+TN)} \quad (3)$$

Now, the FNAC cytology images are acquired focusing on the slide's portions of the biopsy sample with a help of a digital camera attached with the microscope. Then, the conventional manual screening method can be automated by developing automated diagnosis system. This system, whenever required, can be utilized as a tool for getting a second opinion to support the decision made by the pathologist. For testing the developed system, the thyroid FNAC cytological images are obtained from image atlas

database released by PapSociety which are evaluated and ratified by a panel of expert committee. The typical normal and cancerous FNAC images are shown in Figure 1. Figure 1(i) shows a benign image. Figures 1(ii) and Figure 1(iii) show papillary image and medullary image respectively.

Figure 1: (i) Benign, (ii) Papillary and (iii) Medullary Carcinoma FNAC images



Pre-trained CNN models: Krizhevsky et al., (2012) developed the AlexNet and it uses ImageNet which was presented by Deng et al., 2009. This research work uses 80 training images in which 40 images are belonging to benign and remaining 40 images are belonging to malignant group. On the other hand, 30 testing images are used in which 10 images are belonging to benign group and remaining 20 images are belonging to malignant group. All the images are manually segmented into 4 fragment patches which are having high concentrated cell population. Now, the total count of patch portions in training group is 320 images. Similarly, due to this manual segmentation, the total count of patches in testing group is 120 images as given in Table 1. These image segments are given as input to AlexNet and VGG16 pre-trained models.

Table 1. FNAC Image Set and its Patches with Training Set and Testing Set

	Total Images	Image Patches	Benign	Image Patches	Malignant	Image Patches
Training Set	80	320	40	160	40	160
Testing Set	30	120	10	40	20	80

The architecture of AlexNet is shown in Figure 2. The first layer of the AlexNet filters the input image. Then, the next convolutional layer receives the input from the filters of previous layer which is connected to the pooling layer. It keeps the significant features and reduces the large number of features into a smaller number of features. In Alex-Net, each layer has more filters. Like most convolutional neural networks, it has a combination of convolution layer and a pooling layer. It uses ReLU activation function which is more biological inspired. The third layer receives its input from the second convolutional layer. It has multiple kernels, and each kernel has a 3x3 size. Also, in fourth convolutional layer, 384 kernels are there. In this sequence, 256 kernels are there in the fifth convolutional layer with a same 3x3

kernel size. The convolutional layer outputs from third, fourth and fifth stages are given to two fully connected layers as inputs (Hosny et al., 2019).

Figure 2: Layout of pre-trained AlexNet Architecture

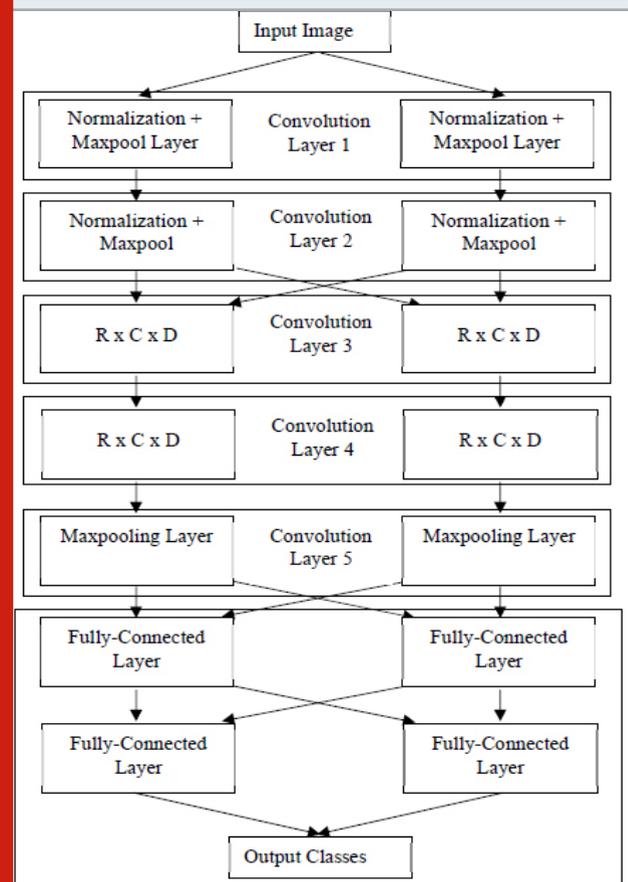
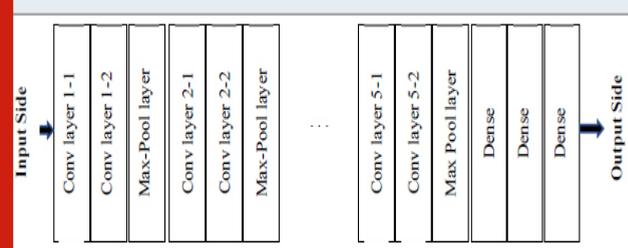


Figure 3 shows the complete layout of the VGG16 architecture. It has five blocks as a combination of two convolution and a max-pooling layers (Zhao et al., 2018; Belaid et al., 2020; Hameed et al., 2020). The first two blocks have two convolution layers and the 3rd, 4th and 5th blocks have three convolution layers. All the blocks are ended with a max-pooling layer. Each layer of VGG16 model uses the activation function 'ReLU'. Finally, the dense layers are being utilised for transferring neuron elements in the networking structure from input side to output side.

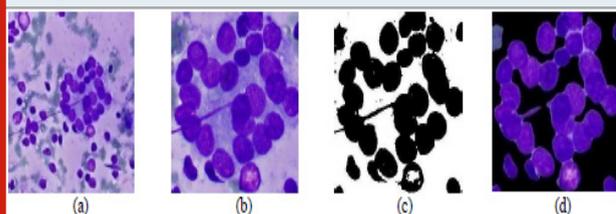
Figure 3: Layout of of pre-trained VGG16 architecture



RESULTS AND DISCUSSION

The proposed automated diagnosis system utilises the pre-trained models AlexNet and VGG16 with transfer learning. The input FNAC images are converted into patch segments. These patch images are pre-processed before the application of pre-trained models using Thresholding operation followed by segmentation of foreground cell regions. The thresholding operation is implemented using Otsu's algorithm on the input patches for the removal of background stain information and mathematical morphology segmentation procedure is used to further remove the small unwanted noise pixels present in the inner and outer sides of the thyroid cell portions.

Figure 4: (a) Input slide image of Thyroid FNAC, (b) Manually cropped patch portion, (c) Output of Thresholding operation and Mathematical morphology, (d) Superimposed image used to perform the classification



In mathematical morphology, the optimised size of disk-shaped structural element has been used to capture the foreground cell regions. The results are presented in Figure 4. The segmented thyroid FNAC image patches are now fed as the inputs to the proposed AlexNet and VGG16 pretrained models. These models are used to perform feature extraction and classification on the given input image patches and the result from these models are the required diagnosis results such as benign or malignant classes. The outcome of the pretrained architectures is evaluated by sensitivity, specificity and diagnostic accuracy.

Table 2. Performance analysis of AlexNet and VGG16

<i>Parameters</i>	<i>AlexNet</i>	<i>VGG16</i>
<i>True Positive</i>	75	79
<i>True Negative</i>	36	37
<i>False Positive</i>	5	1
<i>False Negative</i>	4	3
<i>Diagnostic Accuracy</i>	92.5%	96.66%
<i>Sensitivity</i>	93.75%	98.75%
<i>Specificity</i>	90%	92.5%

The results truly positive cases and truly negative cases are tabulated in Table 2 through which the other performance measures namely, sensitivity, specificity and diagnostic accuracy have been calculated. From Table 2,

it is observed that both the models have performed well to diagnose the malignant thyroid cells in FNAC images. However, the VGG16 model has produced a higher sensitivity of 98.75%. This indicates that only one image patch is misdiagnosed in malignant group whereas five image patches are misdiagnosed by the AlexNet model. However, the VGG16 model has produced a specificity of 92.5% only. By analysing the overall performance of the two pre-trained models, the VGG16 is outperforming well in the diagnosis of thyroid FNAC pathology images with a higher diagnostic accuracy of 96.66%.

CONCLUSION

The performance of two pre-trained convolutional neural network models, namely AlexNet and VGG16 was evaluated to differentiate normal and cancerous thyroid cells in FNAC patch image sets. Both models were tested by implementing Transfer Learning concept in deep-learning based approach. In the initial step, the FNAC images were pre-processed and patch images were cropped based on the thyroid cell population. These patches were fed to the 8-deep layered AlexNet and 16-deep layered VGG16 as inputs and they were passed through multiple convolution layers, max pooling and dense layers. Through optimal implementation and testing of the models, the AlexNet model has achieved a diagnostic accuracy of 92.5% whereas the VGG16 model has reached a higher diagnostic accuracy of 96.66% along with 98.75% of sensitivity and 92.5% of specificity. From the results, it can be evaluated that the pre-trained AlexNet and VGG16 models can play a vital role in effective diagnosis of malignant carcinoma in cyto-pathology images.

REFERENCES

- Ahmed J and Rahman MA (2016) Intelligent diagnostic system for Papillary Thyroid Carcinoma Journal of Applied Environmental and Biological Sciences, Vol 6 No 3 Pages 72-82
- Aliyu HA Razak MAA Sudirman R and Ramli N (2020) A deep learning AlexNet model for classification of red blood cells in sickle cell anemia International Journal of Artificial Intelligence Vol 9 No 2 Pages 221-228
- Belaïd ON and Loudini M (2020) Classification of Brain Tumor by combination of pre-trained VGG16 CNN Journal of Information Technology Management Vol 12 No 2 Pages 13-25
- Chen P Shi X Liang Y Li Y Yang L and Gader PD (2020) Interactive thyroid whole slide image diagnostic system using deep representation Computer Methods and Programs in Biomedicine Vol 195 Pages 105630
- Cooper DS Doherty GM Haugen BR Kloos RT Lee SL Mandel SJ and Sherman SI (2009) Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association (ATA) guidelines

- taskforce on thyroid nodules and differentiated thyroid cancer *Thyroid* Vol 19 No 11 Pages 1167-1214
- Deng J Dong W Socher R Li LJ Li K and Fei-Fei L (2009) ImageNet: A large-scale hierarchical image database in *Proceedings of IEEE Conference Computer Vision and Pattern Recognition* Pages 248-255
- Gopinath B and Gupta BR (2010) Majority voting based classification of thyroid carcinoma *Procedia Computer Science* Vol 2 Pages 265-271
- Gopinath B and Shanthi N (2013) Computer-aided diagnosis system for classifying benign and malignant thyroid nodules in multi-stained FNAB cytological images *Australasian Physical & Engineering Sciences in Medicine* Vol 36 No 2 Pages 219-230
- Gopinath B and Shanthi N (2015) Development of an automated medical diagnosis system for classifying thyroid tumor cells using multiple classifier fusion *Technology in Cancer Research & Treatment* Vol 14 No 5 Pages 653-662
- Hameed Z Zahia S Garcia-Zapirain B Javier Aguirre J and María Vanegas A (2020) Breast cancer histopathology image classification using an ensemble of deep learning models *Sensors* Vol 20 No16 Pages 4373
- Hosny KM Kassem MA and Foad MM (2019) Classification of skin lesions using transfer learning and augmentation with Alex-net *PloS one* Vol 14 No 5 Pages e0217293
- Karakitsos P Cochand-Priollet B Pouliakis A Guillausseau PJ and Ioakim-Liossi A (1999) Learning vector quantizer in the investigation of thyroid lesions *Analytical and Quantitative Cytology and Histology* Vol 21 No 3 Pages 201-208
- Krizhevsky A Sutskever and Hinton G (2012) ImageNet Classification with Deep Convolutional Neural Networks In *Proceedings of Neural Information Processing Systems (NIPS)* Vol 1 Pages 1097-1105
- Kwong N Medici M Angell TE Liu X Marqusee E Cibas ES and Alexander EK (2015) The influence of patient age on thyroid nodule formation, multinodularity, and thyroid cancer risk *The Journal of Clinical Endocrinology & Metabolism* Vol 100 No 12 Pages 4434-4440
- Liu T Guo Q Lian C Ren X Liang S Yu J and Shen D (2019) Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks *Medical Image Analysis* Vol 58 Pages 101555
- Mullur R Liu YY and Brent GA (2014) Thyroid hormone regulation of metabolism *Physiological Reviews* Vol 94 No 2 Pages 355-382
- Shankar K Lakshmanaprabu SK Gupta D Maselena A and De Albuquerque VHC (2020) Optimal feature-based multi-kernel SVM approach for thyroid disease classification *The Journal of Supercomputing* Vol 76 No 2 Pages 1128-1143
- Sinna EA and Ezzat N (2012) Diagnostic accuracy of fine needle aspiration cytology in thyroid lesions *Journal of the Egyptian National Cancer Institute* Vol 24 No 2 Pages 63-70
- Thomas J Ledger GA and Mamillapalli CK (2020) Use of artificial intelligence and machine learning for estimating malignancy risk of thyroid nodules *Current Opinion in Endocrinology, Diabetes and Obesity* Vol 27 No 5 Pages 345-350
- Wong CKH Liu X. and Lang BHH (2020) Cost-effectiveness of fine-needle aspiration cytology (FNAC) and watchful observation for incidental thyroid nodules *Journal of Endocrinological Investigation* Vol 43 Pages 1645-1654
- Würflinger T Stockhausen J Meyer-Ebrecht D and Böcking A (2004) Robust automatic coregistration, segmentation, and classification of cell nuclei in multimodal cytopathological microscopic images. *Computerized Medical Imaging and Graphics* Vol 28 No 1-2 Pages 87-98
- Zhao D Zhu D Lu J Luo Y and Zhang G (2018) Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16 *Symmetry* Vol 10 No 10 Pages 519