VGG Classification Model for Lung Cancer Diagnosis

Nabilah Hanani Hider*, Anis Salwa Mohd Khairuddin, Effariza Hanafi

Department of Electrical Engineering, Faculty of Engineering, Universiti of Malaya, Malaysia

*Corresponding Author

doi: https://doi.org/10.21467/proceedings.141.9

ABSTRACT

Lung cancer is one of the most common cancers worldwide that leads to a small survival rate. It is important to detect the presence of these harmful cells in the human body at early stages to prevent it from worsening. The primary goal of this study is to propose an efficient lung cancer image classification model using a deep learning method. The cancer image classification framework is proposed by using transfer learning with Convolutional Neural Network (CNN) to classify three categories of 5,100 cancer images namely lung adenocarcinoma, lung squamous cell carcinoma and benign lung tissues obtained from the dataset. Several experiments have been performed to improve the VGG19 model performance by varying the optimizers including RMSprop, Adam and SGD. The performance of all experiments conducted were analyzed based on the training and validation curves, classification reports and the confusion metrics.

Keywords: Deep learning, Convolutional Neural Network (CNN), Lung cancer, Optimizer, Learning rate

1 Introduction

Lung cancer is one of the top causes of cancer-related death alongside breast and prostate cancer. It is a poorly understood disease as in most cases, smokers who inhale the tobacco may develop the disease, but non-smokers who are exposed to the smoke can develop it too. The worldwide incident of this diseases is increasing by 2% per year and 80% of the cases are linked to tobacco consumption. It is very harmful as the cancer can spread to other parts of the human body in which this process is called metastasis. The diagnosis of this disease is crucial in order to reduce the mortality rate of the disease. Early detection is important as most of the time, the lung cancer is detected at the time when the treatment is ineffective or has lower rate of success [1].

The existing techniques used for lung cancer detection and classification are based on hand-engineered techniques and their performance in terms of accuracy and other evaluation measures are limited. One of the examples of the existing lung cancer detection technique is by having radiologists examining the CT scan images in the lab. Pathologist review on tissue images is most of the time time-consuming and error prone. The technology has evolved, and classification procedure has been accomplished by using the naïve Bayes algorithm [2]. In a study conducted by Kalaivani et al. [3], the use of this type of classifier has resulted in increase of 8.34% of accuracy rate, 11.76% of sensitivity rate, and 5.26% for specificity. Some researchers focus on the earlier detection of lung cancer in which they focus on using approaches like image processing and artificial neural network. The aim of these research is to detect lung cancer at earlier stage and reduce human error in the manual detection process. As a result, various machine learning and deep learning methods have been introduced for lung cancer classification problems.

2 Research Background

In many CNN-based classification problems, achieving high accuracy while keeping training-validation variance as low as possible is a challenging task. The loss needs to be kept low as well as the loss represents

Proceedings DOI: 10.21467/proceedings.141; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-957605-4-1

the error made during each iteration. Various optimization algorithms are present to help reducing this loss and keep improving at each iteration. However, finding the suitable optimizer for a particular classification problem is a difficult task and needs to be experimented as they showcase different behaviour based on different tasks [4].

Aside from the optimization technique, the learning rate also need to be chosen carefully. It does not affect the performance of the model, but it affects the speed and quality of the training process. As the default learning rate cannot guarantee high quality of model, it has become essential to fine tune it [5].

Thaventhiran & Sekar [6] proposed in their work to implement a novel technique called Target Projection Feature Matching Based Deep Artificial Neural Network (ANN) with LSTM to predict lung cancer. For evaluation, the work uses the Lunar 16 Lung Cancer datasets. The obtained results shows that TPFMDANN-LSTM achieved a better performance with 6% increase in prediction accuracy, 36% reduction for false positive rate, and consume 16% lower time taken for prediction as compared to existing works.

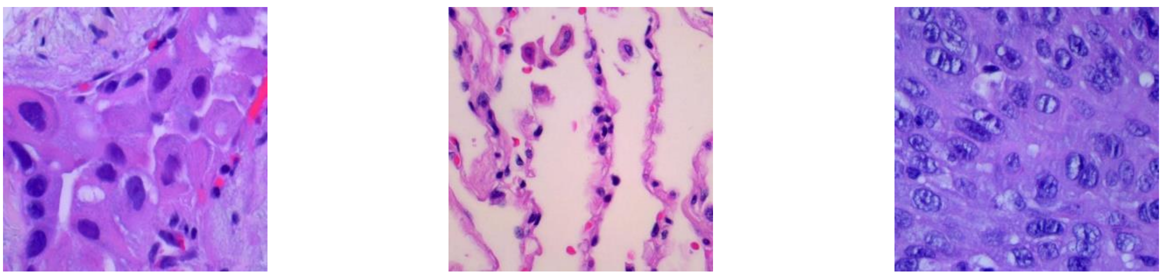
In the work proposed by Oliver et al. [7], the cross-validation classification tree, RF cross-validation classification and random tree are the incorporated machine learning methodologies. The result shows that the classification achieved 94.3% accuracy which is higher than the conventional methodology.

The main goal of a research conducted by Dutta [8] is to propose the use of Random Forest (RF) classifier and evaluate its performance by comparing it with the performance of Support Vector Machine (SVM) and Naïve Bayes (NB) classifiers. It obtained an average accuracy of 93.2%, better precision, recall and f1-score. It also takes shorter time to train compared to the other existing method.

3 Materials and Methods

3.1 Image dataset

The cancer images are obtained from the LC25000 dataset [9] where a total of 750 lung tissue images were obtained from HIPAA and expanded into 15,000 image datasets. The images obtained are grouped together and separated by their respective folders based on the knowledge of the tissue types. All images are in .jpeg format and 768 x 768 pixels in size. The lung tissue types are lung adenocarcinomas (lung_aca), benign lung tissues (lung_n), and lung squamous cell carcinomas (lung_scc). The sample images of the lung tissue types are as shown in Figure 1 (a), (b) and (c) respectively. For the purpose of this study, a total of 5,100 images (1,700 images per class) were chosen and split into training (72%), validation (18%) and testing (10%) datasets.



(a) **(b)** **(c)**

Figure 1: *Image samples consists of (a) lung adenocarcinomas (b) benign lung tissues (c) lung squamous cell carcinomas*

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

3.2 Experimental Setup

As the name suggests, VGG19 is a CNN model that is 19 layers deep. The model is trained with over millions of images from ImageNet datasets and available to be used as classification model [10]. The model was originally trained with image of input size 224×224 RGB channel.

On top of the VGG19 model, two Dense layers of 256 and 128 neurons respectively, with 'relu' activation are added to the model. The last layer added is the output layer with 'softmax' activation function since this is a multi-class classification problem. The epochs chosen for all experiments is 50 epochs due to capacity limitation. The batch size is by the default value which is 32.

3.3 Optimizer

In this paper, three types of optimizers are used, and the results obtained from the result of these three optimizers namely RMSprop, Adam and SGD are compared. RMSprop is a part of the adaptive learning rate method. The main idea of this optimizer is to keep the moving average of the squared gradients for every weight. After that, the gradient is divided by the square root of the mean square. This is the reason why it is named RMSprop (root mean square).

Adaptive Moment Estimation or Adam is more towards the more advanced optimizer type. It can lead the model to converge faster and achieve high accuracy. This optimizer combines the algorithm of RMSprop and Momentum optimizers where it considers the first and second moment.

SGD often deploys and works well in many applications. It has the ability to generalize well, hence making it a good optimization technique. This optimizer updates the parameter using only one randomly selected training instance in each iteration. As compared to the other gradient descent optimization technique, SGD is more popular since it works faster as only one training instance is chosen. This is relatively useful when dealing with a big dataset [11].

In each epoch of the neural networks, it is aimed to minimize the error which can be monitored from the loss function. The loss function is defined in terms of cross entropy. In this paper, the categorical cross entropy function is chosen as the loss function as it deals with multi-class classification with one hot encoded dataset.

3.4 Learning Rate

There are various types of learning rate fine-tuning mechanisms and one of them is the naïve approach where the learning rate is kept constant. In this study, this approach will be implemented. The test will start experimenting at a fixed value of learning rate and continued with other fixed rates. In this study, learning rate of 0.01, 0.001, 0,0001 and 0.00001 are chosen to be analysed.

4 Results and Discussion

The first experiment conducted is to compare the performance of different optimization techniques using VGG19 model with learning rate of 0.001. Line graph in Figure 2 shows the validation accuracy for each optimizer until 50 epochs. Table 1 summarizes the testing accuracy obtained for each optimizer while Table 2 shows the classification report for each optimizer which consists of precision, recall and f1-score.From the results in Figure 2 and Table 1, it is observed that Adam achieves highest accuracy on validation and testing followed by RMSprop and SGD. Adam also showcases fastest convergence speed as compared to the other two. Hence, Adam optimizer is chosen to be further tuned by varying the learning rate.

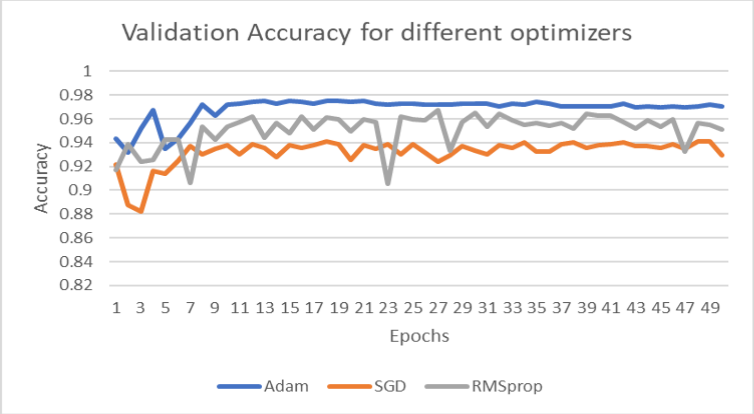


Figure 2: *Validation Accuracy of different optimizers using VGG19 model with learning rate 0.001*

Table 1: *Testing Accuracy for different optimizers using VGG19 model*

Optimizer	Learning Rate	%Prediction Accuracy	Wrong Predictions (out of 510)
RMSprop	0.001	96.7	17
Adam	0.001	98.2	9
SGD	0.001	95.3	24

Table 2: *Classification report for VGG19 model with RMSprop and learning rate of 0.001*

Optimizer	Classes	Precision	Recall	f1-score
RMSprop	Lung_aca	0.93	0.97	0.95
	Lung_n	0.99	0.98	0.99
	Lung_scc	0.98	0.95	0.96
	Weighted Average	0.97	0.97	0.97
Adam	Lung_aca	0.98	0.97	0.97
	Lung_n	0.99	1.00	1.00
	Lung_scc	0.97	0.98	0.98
	Weighted Average	0.98	0.98	0.98
SGD	Lung_aca	0.98	0.89	0.93
	Lung_n	0.99	1.00	0.99
	Lung_scc	0.89	0.97	0.93
	Weighted Average	0.96	0.96	0.95

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

The second experiment analyses the effect of different learning rates used on the performance of the VGG19 classification model with Adam optimizer. The line graph in Figure 3 compares the performance of all learning rates used while Table 3 summarizes the testing accuracy for various learning rates using VGG19 model with Adam optimizer. The result shows that the moderate value of learning rate which is not too high or too low achieves the best fitting curves for validation and training. The highest accuracy achieved is the model with a learning rate of 0.001.

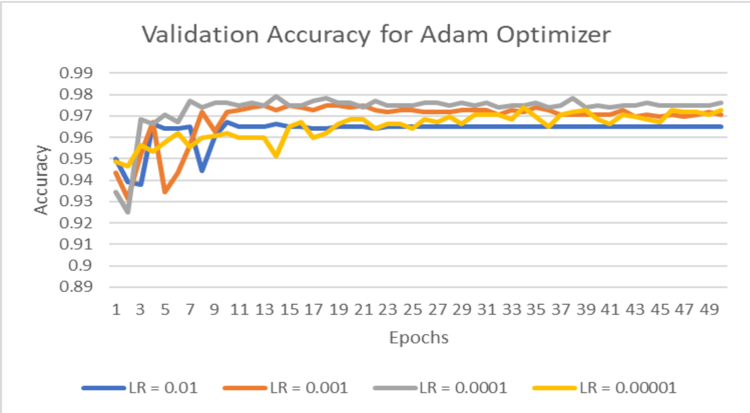


Figure 3: *Validation Accuracy for Adam optimizer with various learning rate*

Table 3: *Testing accuracy for Adam optimizer with various learning rates using VGG19 model*

Optimizer	Learning Rate	% Prediction Accuracy	Wrong Predictions (out of 510)
Adam	0.01	97.45	13
	0.001	98.2	9
	0.0001	97.3	14
	0.00001	97.1	15

5 Conclusions

This paper presents deep learning methods to classify three classes of lung cancer images. The key objectives of this research are to propose the use of pre-trained model namely VGG19 for the lung cancer images as well as to compare the performance of different optimization techniques and various learning rates. At the initial experiment where VGG19 was used with RMSprop optimizer with learning rate of 0.001, the model achieves high accuracy of 96.7%. The model was further tuned by comparing the use of three different optimizers which are RMSprop, Adam and SGD. From the result, Adam was found to achieve better convergence rate and the better accuracy (98.2%) compared to RMSprop (96.7%) and SGD (95.3%) when using learning rate of 0.001 hence Adam is chosen to be further experimented with different learning rates. The accuracy scores obtained for learning rates of 0.01, 0.001, 0.0001, 0.0001 are 97.45%, 98.2%, 97.3% and 97.1% respectively. From the validation curve, it was observed that learning rate of 0.001 and 0.0001 has better convergence than the other two rates. Thus, it is concluded that choosing the middle value that is not too high or low will result in better accuracy and convergence rate. The limitation of this study is mainly on the limited RAM and GPU capacity that leads to less image dataset used for the experiments. In future work, it is recommended to use more image dataset as increasing the training data will yield better results.

6 Declarations

6.1 Acknowledgement

The research funding is provided by Universiti Malaya with Project No. IIRG002C-2021FNW.

6.2 Competing Interests

There is no conflict of interest.

6.3 Publisher's Note

AIJR remains neutral with regard to jurisdiction claims in published maps and institutional affiliations.

References

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L. and Jemal, A. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, 2018, 68(6), pp. 394-424.

[2] Christopher, T., & Banu, J. J. "Study of classification algorithm for lung cancer prediction," *International Journal of Innovative Science, Engineering & Technology*, 2016, 3(2), pp. 42-49.

[3] S. Kalaivani, P. Chatterjee, S. Juyal and R. Gupta, "Lung cancer detection using digital image processing and artificial neural networks," *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, 2017, pp. 100-103, doi: 10.1109/ICECA.2017.8212773.

[4] A. Zohrevand and Z. Imani, "An Empirical Study of the Performance of Different Optimizers in the Deep Neural Networks," *2022 International Conference on Machine Vision and Image Processing (MVIP)*, 2022, pp. 1-5, doi: 10.1109/MVIP53647.2022.9738743.

[5] J. Konar, P. Khandelwal and R. Tripathi, "Comparison of Various Learning Rate Scheduling Techniques on Convolutional Neural Network," *2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS)*, 2020, pp. 1-5, doi: 10.1109/SCEECS48394.2020.94.

[6] Thaventhiran, C., & Sekar, K. R. "Target Projection Feature Matching Based Deep ANN with LSTM for Lung Cancer Prediction," *Intelligent Automation and Soft Computing*, 2022, 31(1), pp. 495-506.

[7] Oliver, A. S., Jayasankar, T., Sekar, K. R., Devi, T. K., Shalini, R., Poojalaxmi, S., & Viswesh, N. G. "Early Detection of Lung Carcinoma Using Machine Learning," *Intelligent Automation and Soft Computing*, 2021, 30(3), pp. 755-770.

[8] Dutta, A. K. "Detecting Lung Cancer Using Machine Learning Techniques," *Intelligent Automation and Soft Computing*, 2021, 31(2), pp. 1007-1023.

[9] Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., & Mastorides, S. M. *Lung and colon cancer histopathological image dataset (LC25000)* ,2019, arXiv preprint arXiv:1912.12142

[10] Karen S, Andrew Z. "Very deep convolutional networks for large-scale image recognition," 2014, arXiv e-prints.

[11] P. Kamsing, P. Torteeka and S. Yooyen, "Deep Convolutional Neural Networks for plane identification on Satellite imagery by exploiting transfer learning with a different optimizer," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 9788-9791, doi: 10.1109/IGARSS.2019.8899206.