Lung Cancer Detection using VGG NET 16 Architecture

Thanzeem Mohamed Sheriff S, Venkat Kumar J, Vigneshwaran S, Aida Jones, and Jose Anand

Corresponding author’s e-mail: thanzeem4@gmail.com

Abstract. Cancer is one of the main reason for loss of human life across the world. All the medical practitioners and researchers are dealing with the demanding situations to fight against cancer. Based on the report in 2019 from American Cancer Society, 96,480 deaths are anticipated due to skin cancers, 142,670 deaths are from lung cancers, 42,260 deaths are from breast cancers, 31,620 deaths are from prostate cancers, and 17,760 deaths are from mind cancers. Initial detection of most cancers has the pinnacle precedence for saving the lives. This paper proposed a lung cancer detection using Deep Learning based on VGG NET architecture. This was one of the famous models submitted to ILSVRC-2014. Visual checkup and manual practices are used on this venture for the various types of cancer diagnoses. This guide interpretation of scientific images that needs massive time intake and is notably susceptible to mistakes. Thus, in this project, we apply deep learning algorithms to identify lung cancer and its presence without the need for several consultations from different doctors. This leads to an earlier prediction of the presence of the disease and allows us to take prior actions immediately to avoid further consequences in an effective and cheap manner avoiding human error rate. In this project lung cancer and its presence is determined. A web application is developed as a hospital application where an input x-ray image is given to detect lung cancer.

Keywords. Lung Cancer, Deep Learning, ILSVRC-2014.

1. Introduction

Lung cancer is also called lung carcinoma. This is a kind of harmful lung tumor considered as unrestrained cell growth in lung tissues. This development spread outside the lung by the method of metastasis into neighboring tissue or additional parts of the body [1]. Most cancers that twitch inside the lung, called primary lung cancers, are carcinomas. The common symptoms are persistent cough, weight problem, tininess of breath, and shoulder, arm, and chest cautions. The mainstream of approximately 85 % of people who cause lung cancer are owing to long-term tobacco smoking. Nearly 10–15 % of belongings happen in people that they not having smoking habit [2]. These belongings are frequently created by a combination of genetic influences and experience to radon gases, asbestos, second-hand smoking, or other kinds of pollution. The Computed Tomography (CT) images from the chest radiography is used to detect lung cancer [3]. The identification is inveterate by biopsy that is characteristically achieved by bronchoscopy or CT-guidance. In 2012, globally lung cancer is affected by 1.8 million people and 1.6 million people have died. This creates is the leading common
description for cancer-related death in men and second shared in women after breast cancer [4]. The greatest shared age for lung cancer analysis is 70 years. In general, nearly 20% of peoples are affected by lung cancer. The survive rate after five years of diseases occurred is very poor, while outcomes on average are worse in the developing world [5]. More about segmentation of image analysis for the detection of non-invasive anemia application is discussed [6] and for the same a ridge regression algorithm is used [7]. Image analysis on Matlab software in analyzed [8] and unsupervised prediction for medical treatment on lung cancer with stereotactic body radiation therapy is mentioned [9]. For the identification of boundaries in medical images and edge vector and mapping method is used [10]. Also for security system image analysis is done on the finger vein [11] and a content based image retrieval method is implemented for texture and shape analysis [12]. Lung cancer detection from CT images are discussed in [13], and CT is used for nodule classification of thorax images using SVM [14]. Image classification based on deep convolution networks is described in [15] and predictive variance for accelerating stochastic gradient detection [16] and deep leaning based algorithm [17] for the detection of images using convolution techniques [18]. Diagnosis of lung cancer using clinical and radiomic features was described [19], and CT based image analysis for predicting lung adenocarcinoma in [20] and a nomogram development of clinical staging using image feature analysis in non-small cell lung cancer is discussed [21].

2. Deep Learning
Deep learning is an algorithm that impersonates the network of neurons during a brain. This is a subset of machine learning which is named as deep learning and this uses deep neural networks. As shown in figure 1 the deep learning technique is built with connected layers. The primary layer is understood because of the input layer. The final layer is the output layer. The layers which is used between input and output layer is the hidden layers. The meaning for deep suggests that the network joins neurons in additional than the two layers.

![Image of Deep Learning Architecture](image.png)

**Figure 1. Deep Learning Architecture**

Every hidden layer consists of number of neurons. The successive neurons are connected with one another. The function of neuron is to process and propagate the received input which is received from the above layer. Signal strength provided to neuron within the successive layer is based on factors such as bias, the load and the activation parameters. Huge volume of input data is consumed by the network for the various process to be performed by the multiple layers. Based on the inputs the network under consideration will learn themselves with the complex functions at each layer. Deep learning approach is an influential technique to obtain better prediction of an actionable output. Deep learning outperforms in pattern identification and knowledge based forecast. Combining both will enhance the performance with extraordinary leads on factors such as innovation, management and production. By the use of deep learning, traditional techniques are better performed. The deep learning approaches are 41 percent more accurate than other image classification machine learning techniques, and the same with 27 percent in face recognition applications and 25 percent better in voice recognition applications.
2.1. Deep Learning Process
A deep neural network delivers progressive accurateness in numerous tasks, from object recognition to speech recognition, and will acquire mechanically, without predefined information expressly coded by the application programmer.

![Figure 2. Learning Process](image)

Figure 2 shows the learning process involved in the deep learning technique. Every layer has an interior layer of data, i.e., the grading of data. In a neural network application with 4 layers will learn additional composite features than that with two levels. Learning happens in two stages. The first stage contains nonlinear transformation inputs and produces a statistical prototypical as output. In the second level the algorithm intents to rise the model with a mathematical process stated as a derivative. Neural network concept repeats the 2 levels hundreds of times or even to thousands of times until it reaches the tolerance level of accuracy. This processes is repeated as iteration until an optimal solution is received.

3. VGG NET 16 Architecture
The trained model using VGG Net 16 architecture is shown in the figure 3. Convolution 1 input layer is a standard image size of 224 x 224 RGB. The input is conceded over all the 16 layers in order to find the best accurate results, where filters are applied with a exact minor reception field of size 3 x 3. In some layout, it also uses 1 x 1 convolution filters, that are seen as straight modification of input stations. Convolution line altered to 1 pixel, the spatial padding of convolution layer input is such that the spatial resolution is stored after the convolution, e.g., Padding is 1-pixel for 3 x 3 convolution layers. The collection is done with max-pooling layers, which follow one of the convolution layers. Max-pooling is done with a 2 x 2 pixel convolution layer.

All the 3 fully connected layers will continue with the convolutional layer with different depths in various structures. The initial two layers have 4096 channels and the third layer connects to ILSVRC classification in 1000 channels making one channel per category. Soft-max layer is the retaining layer and the configuration of completely integrated layers are similar across the entire networks. The entire hidden layers are fitted to non-linear retrieval represented by ReLU. This is also known that there is no network other than one which contain the local response and such practice do not have any improved performance in the ILSVRC database thus leading to memory usage and computational time.

3.1. ResNet Algorithm
The construction of ResNet50 has 4 stages are there. This network will take an input image with width and height, as duplicate 32, and 3 as channel width. Aimed at the description, look at the input size as 224 x 224 x 3. All ResNet drawings make the first conversion and multi-integration using the dimensions of 7 x 7 and 3 x 3 respectively. After the phase 1 the network twitches and consists of 3 remaining blocks covering 3 layers individually. The maximum number of characters used to achieve the convolution function on all 3 layers of block 1 is 64, 64, and 128 respectively. Curved arrows point to the joining of the identity. The combined arrow signifies that the process of the convolution in
Residual Block is done in stride 2, henceforth, the length and width of the input will be cut in half, but the channel width will be doubled. As we move from one section to another, the channel width is folded and the input size is minimized by half.

Figure 3. VGG16 Architecture

In deep networks like ResNet50, ResNet152, etc. and for each remaining F function, 3 layers are arranged in sequence. Convolutions $1 \times 1$, $3 \times 3$, $1 \times 1$ are used in the three levels. Layers of $1 \times 1$ convolution are responsible for dropping and reinstating size. The $3 \times 3$ layer is left as a bottle with a small size of input/output. Finally, the network has a Pooling layer followed by a fully connected layer of 1000 neurons (output of the ImageNet class).

4. Training Process
In this training procedure, we are going to train the model by performing image augmentation first. This image augmentation will help us to train the model effectively, by pre-processing all the images present in the data set and converting each image into multiple images of different angles. Various steps involved in our training are given below which are array pre-processing, rotation, flipping, etc. Then after augmentation train the datasets by extracting the features using a combination of two deep learning algorithms VGG-16 and Resnet algorithm. It undergoes a process called optimization which
will optimize the model and loss minimization which will reduce the noises generated during training. In the last it will be undergoing a process is called model seriation which will be evaluated after generating model using the testing dataset and predict the presence of lung cancer. Then python codes are executed using visual studio software. The image given in figure 4 is the flowchart of the training process in the proposed system. The modules represented in the training process are dataset collection, dataset pre-processing, dataset separation, training and testing the dataset using VGG-16. After the optimization technique finally the validation and evaluation process is carried out to obtain the optimized result.

![Flowchart Representation](4.4.png)

**Figure 4.** Flowchart Representation

4.1. **Importing Necessary Libraries**
The required libraries are imported for training the model. Some libraries like NumPy, OS, Adam, SGD are imported and the process is continued by pre-processing.

4.2. **Performing Image to Array Pre-processing**
The pixels of each image contain a 0:255 aspect ratio and 8bit data. If we give this huge value for training, it will not be trained accordingly. So, to avoid this accuracy problem we are reducing the aspect ratio from 0:255 to 0:1 so that the training will be more efficient [0:1] in the sense we are reducing it to 0:42, 0.25.
4.3. Aspect Aware Pre-processor
During the pre-processing and training of images of the lung in the image should not be affected by any of the operations like shifting, cropping, flipping, rotating, resizing, etc. So, to maintain the original features of an image we are doing Aspect Aware pre-processing.

4.4. Performing Image Augmentation
Here we have converted single image into the multiple images of different categories by Rotating, Flipping, Contrast adjustment, Inverting etc.

4.5. Training and Testing Dataset
The data is divided into two categories: training and testing, with 75% of the data being used for training and 25% for testing.

4.6. Label Binarizer
It converts the name of labels such as Normal, Severe and Mild into Binary format so that the machine can understand it so we can get accurate output, as the machine cannot understand human language it’s advisable to convert human language to Binary.

4.7. Optimization
First an ADAM Optimizer is used and trained up to 10 epochs. Additionally, SGD Optimizer layer is introduced for adjusting the learning write and adding it as a layer to get more Accuracy.

5. Discussion and Evaluation
As the execution get completed the accuracy is plotted for the training using the deep learning algorithm. As discussed [1] the system is not very much accurate and the real-time process is quite complex, and this is overcome by the proposed model with ease implementation and accuracy. The cancer cannot be diagnosed with limited parameters [2], but in the system proposed presence of cancer can be identified at an early stage. There is no satisfactory accuracy in the detected image for cancer diagnosis [13], the system implemented using deep learning algorithm has complete accuracy. The evaluation results are shown in table 1.

|                | Precision | Recall | F1-score | Support |
|----------------|-----------|--------|----------|---------|
| Ignore         | 1.00      | 1.00   | 1.00     | 34      |
| Lung_cancer    | 1.00      | 1.00   | 1.00     | 27      |
| Lung_normal    | 1.00      | 1.00   | 1.00     | 49      |
| Accuracy       |           |        | 1.00     | 110     |
| Macro Average  | 1.00      | 1.00   | 1.00     | 110     |

Figure 5 shows the image for the training loss and the training accuracy validation graph obtained after the training process. In the first training process the accuracy was less and there would be more loss. After repeating the training process, the accuracy was increased and loss was reduced. The trained model graph also automatically stored in the file.
6. Conclusion
In this paper, the recognition of lung cancer is done in early phases. The architecture used is VGG which consist of 16 layers. The input is passed through all the 16 layers in order to find the best accurate results. VGG can also be compared with other architectures or else the other convolutional layers can be increases the accuracy and probability for better results.

References
[1] C Parmar, R T Leijenaar, P Grossmann, E R Velazquez, J Bussink, D Rietveld, M M Rietbergen, B Haibe-Kains, P Lambin and H J Aerts 2015 Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer Scientific Reports 5 11044
[2] M G Vicente, A F H Martinez, Â M S Castrejon, G A J Londoño, J M Udias and P L. Atance 2018 Heterogeneity in [18F] Fluorodeoxyglucose positron emission tomography/computed tomography of non–small cell lung carcinoma and its relationship to metabolic parameters and pathologic staging. Molecular Imaging 13 (9) 7290.2014. 00032
[3] J Duchi, E Hazan and Y Singer 2011 Adaptive Subgradient Methods for Online Learning and Stochastic Optimization Journal of Machine Learning Research 2 2121-2159
[4] Kuruvilla and K Gunavathi 2014 Lung cancer classification using neural networks for CT images Computer Methods and Programs in Biomedicine 113 (1) 202–209
[5] H C Shin, H R Roth, M Gao et al. 2016 Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning IEEE Transactions on Medical Imaging 35 (5) 1285–1298
[6] Sivachandar Kasiviswanathan, Thulasi Bai Vijayan, Lorenzo Simone, Giovanni Dimauro 2020 Semantic Segmentation of Conjunctiva Region for Non-Invasive Anemia Detection Applications. Electronics 2020 9 1309 doi:10.3390/electronics9081309
[7] Sivachandar Kasiviswanathan, Thulasi Bai Vijayan, Sheila John 2020 Ridge regression algorithm based non-invasive anaemia screening using conjunctiva images Journal of Ambient Intelligence and Humanized Computing https://doi.org/10.1007/s12652-020-02618-3.
[8] Subbarayan Sadasivam, S Karthick Ramanathan 2009 Effective watermarking of digital audio and image using Matlab technique International Conference on Machine Vision IEEE

[9] H Li, M Galperin-Aizenberg, D Pryma, C B Simone II and Y Fan 2019 Unsupervised machine learning of radiomic features for predicting treatment response and overall survival of early-stage non-small cell lung cancer patients treated with stereotactic body radiation therapy Radiotherapy and Oncology 129 (2) 218-226

[10] J Anand, and K Sivachandar 2013 An Edge Vector and Edge Map Based Boundary Detection in Medical Images International Journal of Innovative Research in Computer and Communication Engineering 1 (4) 1050-1055

[11] J Anand, T G A Flora, and A S Philip 2013 Finger-Vein based Biometric Security System International Journal of Research in Engineering and Technology 2 (12) 197-200

[12] Sibia E V, Mareena G, and J Anand 2014 Content Based Image Retrieval Technique on Texture and Shape Analysis using Wavelet Feature and Clustering Model International Journal of Enhanced Research in Science Technology & Engineering 3 (8) 224-229

[13] Vaishali C Patil and Shriniwas R Dhotre 2016 Lung Cancer Detection from Images of Computer Tomography Scan International Journal of Advanced Research in Computer and Communication Engineering 5 (7)

[14] H M Orozco and O O V Villegas 2013 Lung nodule classification in CT thorax images using support vector machines 12th Mexican International Conference on Artificial Intelligence 277–283

[15] Rizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks Advances in Neural Information Processing Systems 1097–1105

[16] Johnson R and Zhang T 2013 Accelerating Stochastic Gradient Descent Using Predictive Variance Reduction Proceedings of the 26th International Conference on Neural Information Processing Systems 315-323

[17] Sutskever, J Martens, G Dahl and G Hinton 2013 On the importance of initialization and momentum in deep learning PMLR 1139-1147

[18] C Szegedy, W Liu, Y Jia, P Sermanet, S Reed, D Anguelov et al. 2014 Going deeper with convolutions arxiv preprint arxiv: 1409 4842

[19] H Zhou, D Dong, B Chen, M Fang, Y Cheng, Y Gan, R Zhang, L Zhang, Y Zang and Z Lin 2018 Diagnosis of distant metastasis of lung cancer based on clinical and radiomic features Translational Oncology 11 (1) 31-36

[20] T P Coroller, P Grossmann, Y Hou, E R Velazquez, R T Leijenaar, G Hermann, P Lambin, B Haibe Kains, R H Mak and H JAerts 2015 CT based radiomic signature predicts distant metastasis in lung adenocarcinoma Radiotherapy and Oncology 114 (3) 345-350

[21] M C Desseroit, D Visvikis, F Tixier, M Majdoub, R Perdrisot, R Guillevin, C C Le Rest and M Hatt 2016 Development of a nomogram combining clinical staging with 18-FDG PET/CT image features in non-small cell lung cancer stage I- III European Journal of Nuclear Medicine and Molecular Imaging 43 (8) 1477-1485