International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0342
e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2365-2370

Prediction of Lung Cancer Using VGG19 Training Learning

Dr. V. Sesha Bhargavi[1], *Rida Fatima Khan*[2]
[1,2]*Department of IT, GNITS, Hyderabad-T. S, 500104, India*
Emails: *b.velagaleti@gmail.com*[1], *krida1802@gmail.com*[2]

Abstract

The critical need for accurate prediction of lung cancer, employing artificial intelligence on CT scan images to mitigate the high mortality rate associated with this disease. Deep learning, particularly CNNs, emerges as a powerful tool for achieving superior prediction accuracy compared to traditional machine learning methods. Leveraging a dataset comprising 3000 chest scan images across various types of lung cancer which includes adenocarcinoma, benign and squamous cell carcinoma the effectiveness of multiple machine learning algorithms is evaluated. It confirms CNN as the optimal choice for accurate prediction, with the implementation of VGG-19 further enhancing the assessment of lung cancer severity and precautionary measures. Performance analysis is obtained by using accuracy, precision, recall and loss metrics. For designing this application python software is used and result analysis is performed.

Keywords: *Non-small lung Cancer; VVG -19; Lung Cancer; Deep Learning; Neural Networks.*

1. Introduction

The goal of this work is to predict lung cancer using VGG19 transfer learning. Recent developments in deep learning, especially in transfer learning methods, have demonstrated potential to support medical imaging analysis-based early lung cancer diagnosis and detection. In this study, we explore the application of VGG19 transfer learning, a pre-trained CNN, for the prediction of lung cancer.

The objectives are:

- To explore the effectiveness of VGG19, a pre trained convolutional neural network, in the context of lung cancer prediction.
- To transfer the knowledge learned by VGG19 from ImageNet dataset to the task of lung cancer classification.
- To compare the predictive performance of VGG19 transfer learning with traditional methods and other deep learning architectures.

This dataset comprises 3,000 histopathological images categorized into three classes. The images were created from an initial set of 750 images of lung tissue that adhered to HIPAA regulations and were validated. This initial set included 250 images each of lung squamous cell carcinomas, lung adenocarcinomas, and lung benign tissue. Using the Augmentor package, this set was augmented to produce the final dataset of 3,000 images. Each class in the dataset contains 1,000 images.VGG-19 is a deep CNN with 19 layers, widely used for image classification. It's popular because it employs multiple 3 × 3 filters in each convolutional layer. Trained on the ImageNet database with a million images across 1000 categories, a pre-trained VGG-19 model can classify images into various objects like keyboards, animals, etc. It achieves an impressive accuracy of 95% with a loss of 17%. Lung cancer ranks second in cancer-related deaths, posing a significant threat to the population. Early diagnosis, crucial for recovery, is emphasized by the American Cancer Society.Detecting lung cancer through medical imaging like CT scans is effective, with CT scans being particularly efficient. Automated detection offers improved results over manual checks. CNNs are recognized as effective tools for image-based prediction. Types of lung cancer mentioned in the paper are large cell carcinoma, benign, squamous cell carcinoma with different consequences as well as different characteristics.

1.1. Lung Adenocarcinoma

Lung adenocarcinoma is the most common type of lung cancer, making up approximately 30% of all cases and 40% of non-small cell lung cancer cases.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0342
e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2365-2370

This type of cancer is also found in other organs like the colon, prostate, and breast. Adenocarcinomas in the lungs develop in the glands responsible for aiding breathing and mucus production. Symptoms may include weakness, weight loss, hoarseness, and coughing. [5] Figure 1 shows the Types of Lung Cancer Pathology Images

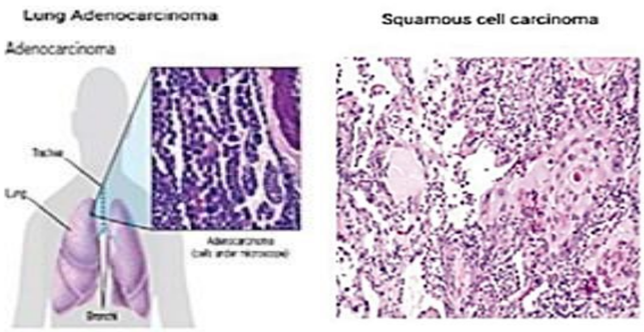

Figure 1 Types of Lung Cancer Pathology Images

1.2. Squamous Cell Carcinoma

Squamous cell lung cancer develops internally within the lung, often in the larger bronchial tubes connecting the trachea to the lung, or in major airway branches. [8] Various medical imaging techniques such as MRI, X-rays, and CT scans are used to detect lung cancer, with CT scans being particularly efficient. Automated detection offers improved results over manual checks. Convolutional Neural Networks are recognized as one of the most effective methods for image-based prediction among existing approaches. [9]

2. Methodology

Prediction of Lung Cancer Using VGG19 Transfer Learning is performed in proposed model. For given dataset, results of VGG-19 are analyzed and different graphs are plotted for results validation.

2.1 Dataset Details

There are 3,000 histological pictures in this collection, divided into three groups. Every image is 768 by 768 pixels and is stored in a jpeg file format. 750 original samples of HIPAA-compliant, vetted sources were used to create the dataset, which was then enhanced to 3,000 pictures of lung tissue using the Augmentor software. There are 1,000 photos in each of the three classifications in the dataset.

2.2 Algorithm/Methodology

2.2.1 VGG-19 Algorithm for Lung Image Classification

VGG-19 is a deep convolutional neural network (CNN) composed of 19 layers, renowned for its effectiveness in image classification tasks. It employs multiple 3 × 3 filters in each convolutional layer, contributing to its ability to extract intricate features from images.[10] Trained on the ImageNet database containing a vast array of images spanning 1000 categories, a pre-trained VGG-19 model can classify images with remarkable accuracy. With an accuracy of 95% and a loss of 17%, VGG-19 outperforms many existing methods, with fewer training samples required, boasting faster training speeds, and higher accuracy. It is a combination layer like convolutional in the quantity of 16 as well as a layer fully connected with the quantity of 3, network architecture of VGG19 is characterized by its utilization of small convolutional filters, enabling it to effectively capture detailed features within images.

2.2 Architecture of VGG-19

2.2.1 Input Image

The VGGNet architecture is designed to accept input images of size 224×224 pixels. Each image with patch of 224×224 is cropped in standard creation of architecture of VGGNet model. This approach ensures consistency in input size across all images processed by the model, facilitating efficient training and classification.

2.2.2 Convolution Layer

To extract horizontal and vertical features more precisely, architecture of VGG model select the compact field for operation as 3x3. Additionally, 1×1 convolution filters serve as a linear transformation of the input, contributing to the network's capacity to extract diverse features. Following each convolutional layer is a Rectified Linear Unit (ReLU), which introduces non-linearity to the network and speeds up training, a significant improvement from AlexNet.[12] The rectified linear unit activation function, or ReLU, improves the network's capacity to recognise intricate patterns by producing zero if the input is negative and the input if it is positive. In order to preserve spatial resolution following convolution and guarantee detailed feature

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0342
e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2365-2370

extraction while shifting over the input matrix, the convolution stride is fixed at 1 pixel.

2.2.3 Hidden Layers

In the VGG network, Rectified Linear Unit (ReLU) activation functions are applied across all hidden layers. Unlike some other architectures, VGG typically does not utilize Local Response Normalization (LRN) due to its tendency to increase both memory usage and training duration. Furthermore, LRN does not typically contribute to enhanced accuracy in VGG networks.

2.2.4 Fully-Connected Layers

Three completely connected layers make up the VGGNet architecture. There are 4096 channels in each of the layers with the first two next, and in the third layer 1000 channels. A channel is allocated to each class in the dataset.

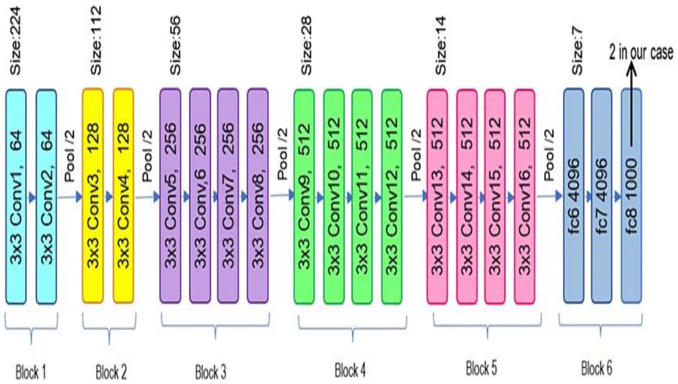

Figure 2 VGG-19 Architecture

The architecture of VGG-19 is composed of five blocks containing 16 convolutional layers. Every block is followed by a Maxpool layer that halves the size of the input picture and twice the number of filters in the next convolutional layer. Block 6's last three dense layers are 4096, 4096, and 1000 in size, respectively. [14] Usually, VGG is trained to classify data into 1000 categories, hence the final dense layer's (fc8) dimension is 1000. However, to accommodate the binary classification job in this study with only two output classes, the dimension of fc8 is set to two.

3. Results and Discussion

The VGG19 model is being used to assess individual images for the ability to forecast the type of lung cancer they contain, as its performance continues to improve. In order to determine the type of lung cancer, the user can choose any image from the dataset for testing. Figure 3 shows the process of the dataset.

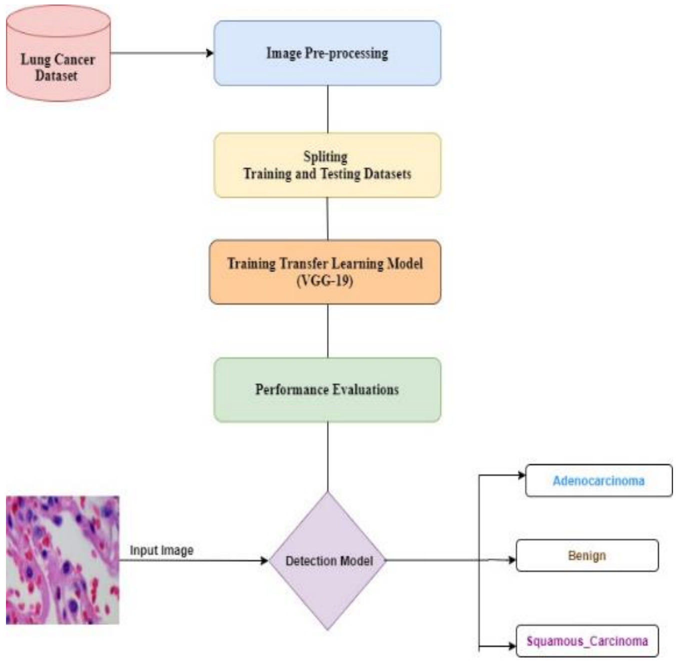

Figure 3 Process of the Dataset

3.1 Output Screenshots

The output screenshots of Figure 4 show the homepage, Figure 5 shows Login page, Figure 6 shows the performance analysis, Figure 7 shows Accuracy& loss graph, Figure 8 shows Precision& recall graph, Figure 9 Shows Physician registration, Figure 10 Physician login, Figure 11 shows Frontend page for Lung Cancer Prediction, Figure 12 shows Prediction Results obtained is "Adenocarcinoma"

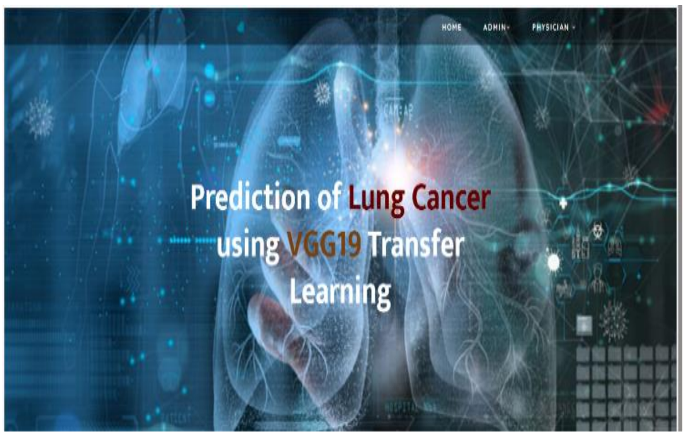

Figure 4 Homepage

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0342

e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2365-2370

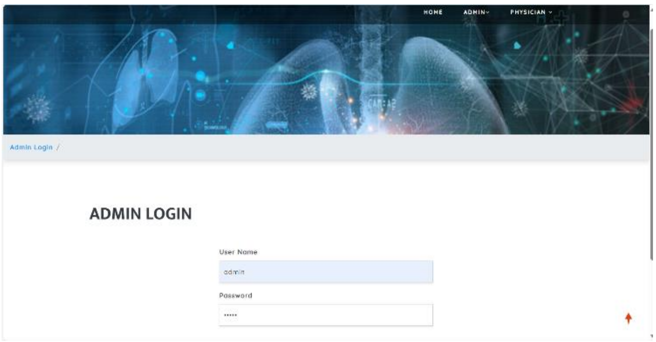

Figure 5 Login Page

Figure 6 Performance Analysis

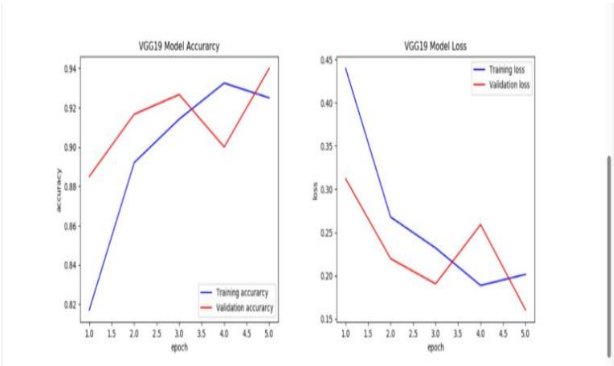

Figure 7 Accuracy & Loss graph

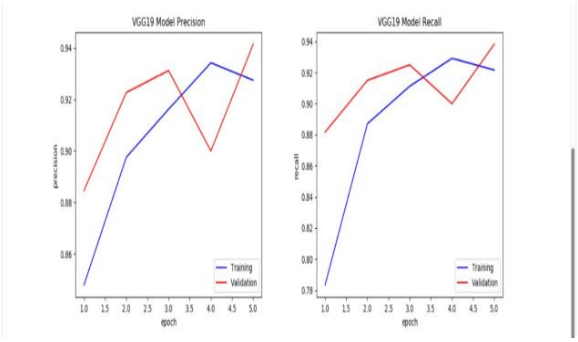

Figure 8 Precision & Recall Graph

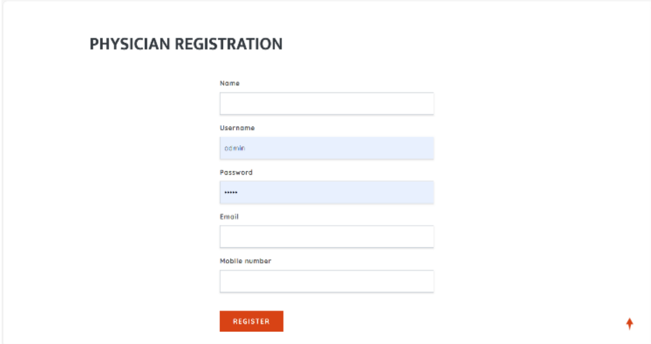

Figure 9 Physician Registration

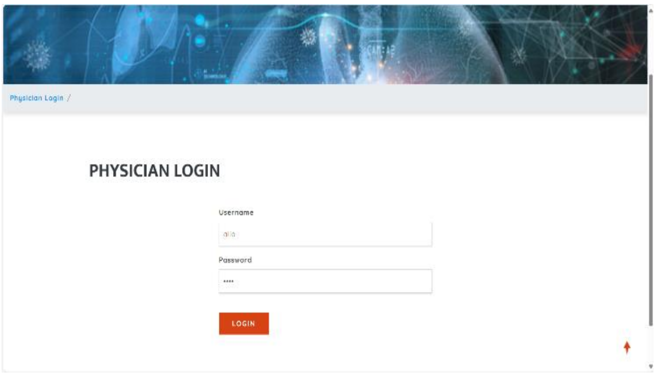

Figure 10 Physician Login

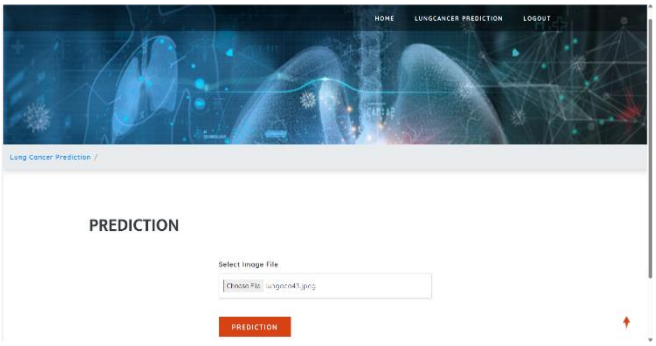

Figure 11 Frontend page for Lung Cancer Prediction

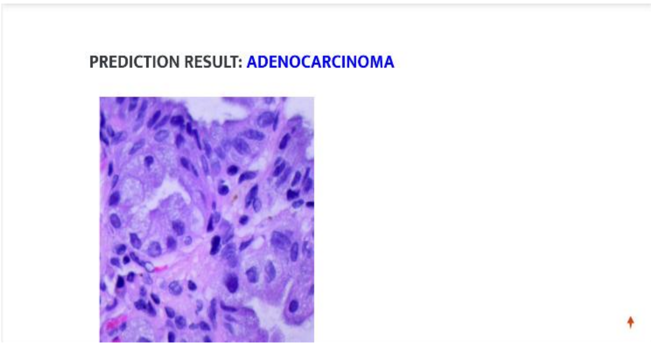

Figure 12 Prediction Results obtained is "Adenocarcinoma"

The proposed model is analyzed using the VGG-19 model for lung cancer prediction using Python software and different relevant libraries.

Conclusion

Squamous cell lung cancer develops internally within the lung, often in the larger bronchial tubes connecting the trachea to the lung, or in major airway branches. [8] Various medical imaging techniques such as MRI, X-rays, and CT scans are used to detect lung cancer, with CT scans being particularly efficient. Automated detection offers improved results over manual checks. Convolutional Neural Networks are recognized as one of the most effective methods for image-based prediction among existing approaches. [9]

References

[1]. M. Norouzi and P. Hardy. (2021). Clinical applications of nanomedicines in lung cancer treatment, Acta Biomaterialia, vol. 121, pp. Doi: 10.1016/j.actbio.2020.12.009
[2]. B.-Y. Wang, J.-Y. Huang, H.-C. Chen, C.-H. Lin, S.-H. Lin, W.-H. Hung, and Y.-F. Cheng. (2020). The adenocarcinoma comparison and between squamous cell carcinoma in lung cancer patients, Journal of cancer research and clinical oncology, vol. 146, no. 1, pp. Doi: 10.1007/s00432-019 03079-8
[3]. D. G. Beer, S. L. Kardia, C.-C. Huang, T. J. Giordano, A. M. Levin, D. E. Misek, L. Lin, G. Chen, T. G. Gharib, D. G. Thomas, et al. (2002). Gene-expression profiles predict survival of patients with lung adeno-carcinoma, Nature medicine, vol. 8, no. 8, pp. 816–824.
[4]. C.-R. Guo, Y. Mao, F. Jiang, C.-X. Juan, G. P. Zhou, and N. Li. (2022). Com- putational detection of a genome instability-derived lncrna signaturefor predicting the clinical outcome of lung adenocarcinoma, Cancer Medicine, vol. 11, no. 3, pp. 864–879. Doi: 10.1002/cam4.4471
[5]. M. A. Gillette, S. Satpathy, S. Cao, S. M. Dhanasekaran, S. V. Vasaikar, K. Krug, F. Petralia, Y. Li, W.-W. Liang, B. Reva, et al. (2020). Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarci- noma, Cell, vol. 182, no. 1, pp. 200–225. Doi: 10.1016/j.cell.2020.06.013
[6]. A. Agaimy, O. Daum, M. Michal, M. W. Schmidt, R. Stoehr, A. Hartmann, and G. Y. Lauwers. (2021). Undifferentiated large cell/rhabdoid carcinoma presenting in the intestines of patients with concurrent or recent non-small cell lung cancer (nsclc):clinicopathologic and molecular analysis of 14 cases indicates an unusual pattern of dedifferentiated metastases, Virchows Archiv, pp. 1–11. Doi: 10.1007/s00428 021-03032-6
[7]. K. I. Tosios, V. Papanikolaou, D. Vlachodimitropoulos, and N. Goutas. (2021) Primary large cell neuroendocrine carcinoma of the parotid gland. report of a rare case, Head and Neck Pathology, pp. 1–8. Doi: 10.1007/s12105-021-01300-0
[8]. P. H. Viale.(2020) The american cancer society's facts & figures: 2020 edition, Journal of the Advanced Practitioner in Oncology, vol. 11,no. 2, p. 135. Doi: 10.6004/jadpro.2020.11.2.1
[9]. S. Li, P. Xu, B. Li, L. Chen, Z. Zhou, H. Hao, Y. Duan, M. Folkert, J. Ma, S. Huang, et al. (2019). Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features," Physics in Medicine & Biology, vol. 64, no. 17, p. 175012. Doi: 10.1088/1361-6560/ab326a
[10]. V. S. Bhargavi, S. V. Raju. (2016). "Enhancing security in MANETS through trust-aware routing," 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2016, pp. 1940-1943, Doi: 10.1109/WiSPNET.2016.7566481.images from biorender, "www.biorender.com," 2022.
[11]. V. S. Bhargavi, M. Seetha and S. Viswanadharaju. (2016) "A trust based secure routing scheme for MANETS," 2016 6th International Conference - Cloud System and Big Data Engineering

(Confluence), Noida, India, pp. 565-570, Doi:10.1109/CONFLUENCE.2016.7508183.

[12]. C. Yao, Y. Qu, B. Jin, L. Guo, C. Li, W. Cui, and L. Feng. (2016). A convolutional neural network model for online medical guidance," IEEEAccess, vol. 4, pp. 4094 4103. Doi: 10.1109/ACCESS.2016.2594839

[13]. .V. S. Bhargavi, M. Seetha and S. Viswanadharaju. (2016). "A hybrid secure routing scheme for MANETS," 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS), Pudukkottai, India, pp. 1-5, Doi: 10.1109/ICETETS.2016.7602991.

[14]. L.-Y. Ye, X.-Y. Miao, W.-S. Cai, and W.-J. Xu. (2022). Medical image diagnosis of prostate tumor based on psp-net+ vgg16 deep learning network," Computer Methods and Programs in Biomedicine, vol. 221, p. 106770. Doi: 10.1016/j.cmpb.2022.106770

[15]. J. Wang, B. Wang, J. Bi, and K. Li. (2013). Prognostic significance of microvascular invasion and microlymphatic permeation in non-smallcell lung cancer," European Journal of Cardio-Thoracic Surgery, vol. 43, no. 6, pp. 1269–1269. Doi: 10.1093/ejcts/ezs632

[16]. S. T. M. Sheriff, J. V. Kumar, S. Vigneshwaran, A. Jones, and J. Anand. (2021). Lung cancer detection using vgg net 16 architecture," in Journal of Physics: Conference Series, vol. 2040, p. 012001, IOP Publishing. Doi: 10.1088/1742-6596/2040/1/012001

[17]. M. Phankokkruad. (2021). Ensemble transfer learning for lung cancer detection," in 2021 4th International Conference on Data Science and Information Technology, pp. 438–442.Doi: 10.1145/3478905.3478995