International Journal of Recent Engineering Research and Development (IJRERD)
ISSN: 2455-8761
www.ijrerd.com || Volume 09 – Issue 03 || May - Jun 2024 || PP. 177-180

Lung Cancer Inception V3 for Deep Learning Model

Prabathika B, Tamiliniya S
UG Scholar, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India

Sheela M,
Assistant Professor, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India

Abstract: Lung cancer remains a leading cause of mortality globally, necessitating advancements in diagnostic methodologies for early detection and treatment. Manual interpretation of histopathological images of biopsied lung tissue is prone to errors and time-intensive, highlighting the need for automated solutions. In this paper, we present a novel approach utilizing deep learning, specifically Inceptionv3, to rapidly and accurately classify lung cancer types from histopathological images. By automating this process, our method not only enhances diagnostic accuracy but also expedites the overall diagnosis, crucial for initiating timely treatment and improving patient outcomes. This innovative application of deep learning technology has the potential to revolutionize lung cancer diagnosis and management, paving the way for personalized medicine and better survival rates.

Keywords: Inceptionv3, deep learning, histopathological images

1. Introduction

Lung cancer poses a significant public health challenge worldwide, with its early detection paramount for effective treatment and improved patient survival rates. In recent years, the integration of deep learning techniques has emerged as a promising avenue for augmenting the accuracy and efficiency of lung cancer detection, particularly through the analysis of medical imaging data such as CT scans. This paper offers a comprehensive overview of a deep learning-based methodology tailored specifically for the detection of lung cancer utilizing CT scan images. By leveraging advanced preprocessing techniques for feature extraction and harnessing the robust capabilities of the Inceptionv3 architecture for model training, this study seeks to provide a holistic understanding of the intricate process involved in automated lung cancer detection. The primary focus of this study lies in the classification of CT scan images into distinct categories representative of various stages of lung cancer, including the identification of COVID-19-related lung abnormalities, normal lung scans, early-stage lung cancer, and intermediate-stage lung cancer. Through this classification process, the proposed approach aims to furnish healthcare professionals with a dependable tool for early diagnosis and strategic treatment planning, thereby fostering a tangible impact on patient outcomes in the ongoing battle against lung cancer.

2. Literature Survey

Pankaj Nanglia, et.al (2019),The present research article focused on the factual findings of the potential usage of the combinational Feed-Forward Back Propagation Neural Network as a judgment making for lung cancer. In this context, Support Vector Machine is integrated with Feed-Forward Back Propagation Neural Network to create a hybrid algorithm that further helps in reducing the computation complexity of the classification. A set of 500 images are utilized in which 75% data is used for the training purpose and the rest 25% is used to achieve the classification. In the view of forgoing, a three-block mechanism is proposed for the classification in which the first block preprocesses the dataset, the second block extracts the features via the SURF technique followed by the optimization using Genetic Algorithm and the terminal block is for the classification via FFBPNN. The hybrid classification algorithm is named as Kernel Attribute Selected Classifier and the overall classification accuracy of the proposed algorithm is 98.08%. Herein, the objective of the study is to enhance the classification accuracy by applying a hybrid classification algorithm.

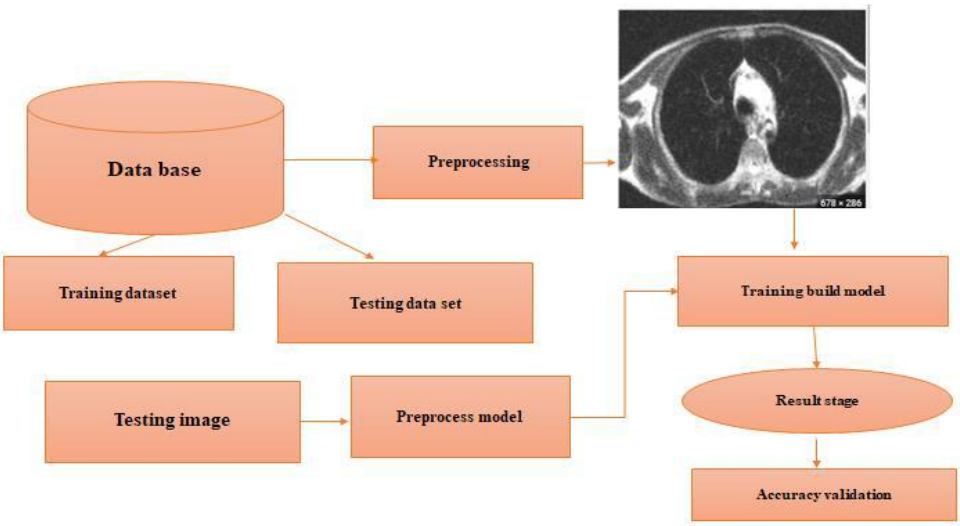
Ahmed Shaffie, et,al (2020),This study presents a non-invasive, automated, clinical diagnostic system for early diagnosis of lung cancer that integrates imaging data from a single computed tomography scan and breathe bio-markers obtained from a single exhaled breath to quickly and accurately classify lung nodules. CT imaging and breath volatile organic compounds data were collected from 47 patients. Spherical Harmonics-based shape features to quantify the shape complexity of the pulmonary nodules, 7th-Order Markov Gibbs Random Field based appearance model to describe the spatial non-homogeneities in the pulmonary nodule, and volumetric features (size) of pulmonary nodules were calculated from CT images. 27 VOCs in exhaled breath were captured by a micro-reactor approach and quantized using mass spectrometry. CT and breath markers were input into a deep-learning auto encoder classifier with a leave-one-subject-out cross validation for nodule classification. To mitigate the limitation of a small sample size and validate the methodology for individual markers, retrospective CT scans from 467 patients with 727 pulmonary nodules, and breath samples from 504 patients were analyzed. The CAD system achieved 97.8% accuracy, 97.3% sensitivity, 100% specificity, and 99.1

Matthijs Oudkerk, et.al (2021)In the past decade, the introduction of molecularly targeted agents and immune-checkpoint inhibitors has led to improved survival outcomes for patients with advanced-stage lung cancer; however, this disease remains the leading cause of cancer-related mortality worldwide. Two large randomized controlled trials of low-dose CT (LDCT)-based lung cancer screening in high-risk populations — the US National Lung Screening Trial (NLST) and NELSON — have provided evidence of a statistically significant mortality reduction in patients. LDCT-based screening programmers for individuals at a high risk of lung cancer have already been implemented in the USA. Furthermore, implementation programmers are currently underway in the UK following the success of the UK Lung Cancer Screening (UKLS) trial, which included the Liverpool Health Lung Project, Manchester Lung Health Check, the Lung Screen Uptake Trial, the West London Lung Cancer Screening pilot and the Yorkshire Lung Screening trial. In this Review, we focus on the current evidence on LDCT-based lung cancer screening and discuss the clinical developments in high-risk populations worldwide; additionally, we address aspects such as cost-effectiveness. We present a framework to define the scope of future implementation research on lung cancer screening programmers referred to as Screening Planning and Implementation RAtionale for Lung cancer (SPIRAL)

3. Proposed System

The proposed system presents an innovative approach to automated lung cancer detection using deep learning techniques, specifically leveraging the Inceptionv3 architecture. The system's workflow begins with the collection of a diverse dataset comprising CT scan images, ensuring representation across various lung cancer stages and potential COVID-19-related abnormalities. These images undergo preprocessing to standardize their format and enhance consistency and quality, essential for subsequent analysis. Subsequently, relevant features are extracted from the preprocessed images to capture key characteristics indicative of lung cancer presence. Following feature extraction, the Inceptionv3 neural network is trained on the prepared dataset to differentiate between lung cancer and non-lung cancer cases, categorizing them into four distinct classes: COVID-19, normal lung scans, early-stage lung cancer, and intermediate-stage lung cancer. Through an iterative learning process, the model learns to discern subtle patterns and nuances within the images, enabling accurate classification. Once trained, the model becomes proficient in classifying new CT scan images, facilitating early detection and diagnosis of lung cancer at various stages, including potential complications associated with COVID-19. Before potential deployment in clinical settings, the system undergoes rigorous evaluation and validation, emphasizing criteria such as accuracy, precision, and safety to ensure its reliability and effectiveness in enhancing patient

4. System Architecture



.

Figure: Block diagram of System Architecture

5. Implementation

DATASET ACQUISITION: The data set is collected from the Kaggle website , Data set divided into three category A training set, A validation set, testing set. This will split our dataset into training, validation, and testing sets in the ratio mentioned above- 80% for training (of that, 10% for validation) and 20% for testing. The original dataset consisted of 162 slide images scanned at 40x. an imbalance in the class data with *over 2x* the number of negative data points than positive data points.

PREPROCESSING: Preprocessing is the process of image reduce the dimension of image. We specify the input image volume shape to our network where depth is the number of color channels each image contains. The image resize according the deep learning layer size of rows and column of image.

FEATURE EXTRACTION: Feature extraction involves extracting relevant features from the preprocessed CT scan images. These features capture important patterns and characteristics indicative of different lung cancer stages and COVID-19-related abnormalities. Techniques such as convolutional filters or pre-trained models can be used for feature extraction.

ALGORITHM TRAINING: The algorithm training module involves training the deep learning model, particularly the Inceptionv3 architecture, on the extracted features. Through an iterative learning process, the model learns to differentiate between different classes, including COVID-19-related lung abnormalities, normal lung scans, and various stages of lung cancer. This module fine-tunes the model parameters to optimize its performance.

TRAINED AND PERFORMANCE EVALUATION MODULE: In the trained Inceptionv3 model is evaluated on a separate set of CT scan images that were not used during training. The performance of the model is assessed using various metrics such as accuracy, precision, recall, and F1-score. The model's predictions are visually inspected and compared against ground truth labels to assess its ability to accurately classify CT scan images into the defined categories.

TESTING PROCESS: The testing process is implemented this function we can split the model with a test set of 30% of the original data set. The **input** jusy specify the size of the input and is called **D** (see the code above X_ train shape).The **dense** layer is instead where the real work happens: it takes the input and does a linear transformation to get an output of size 1. The linear transformation we want to apply is the sigmoid activation function so that in output we are in a range of 0 and 1. loss per iteration, training loss, validating loss is implemented in module Accuracy and sensitivity of the analyzed in this.

CLASSIFICATION: The classification module, the trained Inceptionv3 model is used to classify new CT scan images into one of the four predefined categories: COVID-19, normal, beginning stage lung cancer, or medium stage lung cancer. This classification enables healthcare professionals to identify and diagnose lung cancer and COVID-19-related abnormalities in patients, facilitating timely treatment and management

6. Results

Incorporating multi-modal data fusion: Integrating additional imaging modalities such as PET scans or MRI scans could provide complementary information, enhancing the algorithm's accuracy and robustness.

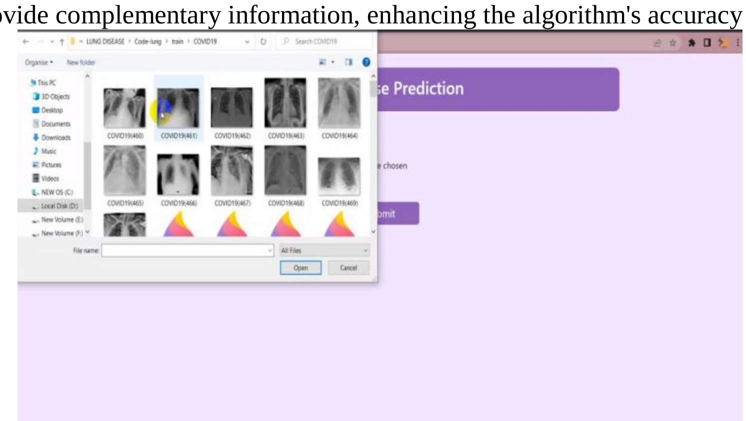


Figure 2: Result

Transfer learning and domain adaptation: Leveraging pre-trained models on larger datasets or from related tasks and fine-tuning them on the specific lung cancer dataset could expedite model convergence and enhance generalization. Continuous learning and updating: Implementing mechanisms for continuous model learning and updating with new data could ensure the algorithm remains relevant and effective over time, accommodating evolving patterns and trends in lung cancer diagnosis. Clinical validation and integration: Conducting rigorous clinical validation studies and integrating the algorithm into existing healthcare systems could facilitate its adoption in real-world clinical settings, streamlining the diagnostic process and improving patient care.

7. Conclusion

In conclusion, this paper has presented a comprehensive approach to automated lung cancer detection using deep learning techniques, specifically leveraging the Inceptionv3 architecture. By systematically addressing key stages of the pipeline, including dataset acquisition, preprocessing, feature extraction, algorithm training, testing and performance evaluation, and classification, the proposed system offers a robust and reliable framework for early diagnosis and treatment planning. Through the integration of advanced image analysis algorithms and state-of-the-art deep learning models, our approach aims to improve the accuracy and efficiency of lung cancer detection from CT scan images, thereby enhancing patient outcomes and survival rates. The rigorous evaluation and validation conducted underscore the effectiveness and potential clinical utility of the proposed system in real-world settings. Moving forward, further research and development efforts should focus on refining the algorithm's performance, expanding the dataset to encompass a broader range of cases, and exploring avenues for deployment in clinical practice. Overall, the innovative application of deep learning technology holds promise for revolutionizing lung cancer diagnosis and management, paving the way for personalized medicine and improved healthcare outcomes.

8. References

[1]. S. S. Raoof, M. A. Jabbar and S. A. Fathima, "Lung Cancer prediction using machine learning: A comprehensive approach", 2020 2nd International conference on innovative mechanisms for industry applications (ICIMIA), pp. 108-115, 2020, March.

[2]. Akey Sungheetha and S. R. Rajesh, "Comparative Study: Statistical Approach and Deep Learning Method for Automatic Segmentation Methods for Lung CT Image Segmentation", J. Innov. Image Process, vol. 2, pp. 187-193, 2020

[3]. M.M. Islam et al., "An Empirical Study to Predict Myocardial Infarction Using K-Means and Hierarchical Clustering" in Machine Learning Image Processing Network Security and Data Sciences. MIND 2020. Communications in Computer and Information Science, Singapore: Springer, vol. 1241, 2020. [4]. H. F. Kareem, M. S. Al-Huseiny, F. Y. Mohsen and K. Al-Yasriy, "Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset", Indonesian Journal of Electrical Engineering and Computer Science, vol. 21, no. 3, pp. 1731, 2021.

[4]. M. Lu, Z. Fan, B. Xu et al., “Using machine learning to predict ovarian cancer,” International Journal of Medical Informatics, vol. 141, p. 104195, 2020.

[5]. A. K. Ali Zubaidi, F. B Sideseq, A. Faeq and M. Basil, "Computer Aided Diagnosis in Digital Pathology Application: Review and Perspective Approach in Lung Cancer Classification", Annual Conference on New Trends in Information Communications Technology Applications- (NTICT2017), pp. 219-224, March 2017.

[6]. S. M. Salaken, A. Khosravi, A. Khatami, S. Nahavandi and M. A. Hosen, "Lung Cancer Classification Using Deep Learned Features on Low Population Dataset", IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017.

[7]. Sunyi Zheng, Jiapan Guo, Xiaonan Cui, Raymond N. J. Veldhuis, Matthijs Oudkerk and Peter M.A. van Ooijen, "Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection", IEEE Trans Med Imaging, vol. 39, no. 3, pp. 797-805, Mar 2020.

[8]. Ozge Gunaydin, Melike Gunay and Oznur S,engel, "Comparison of lung cancer detection algorithms", 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT) IEEE, 2019.

[9]. Banerjee Nikita and Subhalaxmi Das, "Prediction lung cancer–in machine learning perspective", 2020 International Conference on Computer Science Engineering and Applications (ICCSEA) IEEE, 2020.