# BIOMEDICAL IMAGE ANALYSIS USING DEEP LEARNING FOR DISEASE DETECTION

*S.Oyyathevan[1], P.Saravanakumar[2], P.KiranKumar[3] K.uma[4], T.Gopu[5]*
*Associate Professor[1,4,5], Assistant Professor[2], Professor[3]*
*Department of Computer Science & Technology*
*Sasi Institute of Technology and Engineering, Andhrapradesh-534101.*

## ABSTRACT-

Lung - cancer remains one of the leading causes of cancer-related deaths worldwide, highlighting the critical need for accurate and efficient diagnostic methods. This study presents a novel approach for lung tumor detection using the VGG16 convolutional neural network architecture. The VGG16 model is leveraged for feature extraction from lung imaging data, focusing on its deep feature representation capabilities to enhance tumor classification accuracy. Preprocessed CT scan images are input into the VGG16 model, which extracts hierarchical features through its convolutional layers. The extracted features are then classified using machineLearning techniques, such as support vectormachines ( SVM) or logistic regression, to distinguish between malignant and benign tumors. The proposed method demonstrates improved detection performance, emphasizing the importance of deep feature extraction in medical image analysis . The results-indicate that, VGG16-based feature extraction combined with traditional classification methods provides a reliable and efficient solution for lung tumor detection.

**KEYWORDS:** Lung tumor detection, VGG16, feature extraction, CT scan, convolutional neural network, classification, medical imaging.

## 1. INTRODUCTION

Biomedical-image analysis is critical for diagnosing, monitoring, and treating various medical conditions. It provides details into the internal structures and functions of the human body, enabling accurate disease detection and effective treatment planning. Advanced techniques like Machine-Learning & Deep Learning enhance the capabilities, of biomedical imaging , offering improved accuracy and speed in medical diagnostics. Biomedical image analysis plays major role in modern healthcare by facilitating the early detection, diagnosis, and treatment of various diseases, including cancer. The analysis of ` medical-images such as CT-scans, MRI, and X-rays provides valuable insights into the, structure & function of tissues and organs. Traditional image analysis techniques , however, often face challenges in accurately identifying complex patterns and subtle anomalies within these images. [2].Automated systems using Convolutional Neural Networks (CNNs) have proven highly effective in analyzing lung images from modalities like X-rays and CT scans. By leveraging deep learning's ability to extract meaningful `patterns these systems can identify abnormalities such as lung nodules, pleural effusion, or other pathological signs at early stages. It allows better management, ultimately enhancing diagnostic accuracy and enabling improved patient care in clinical settings. [3]. In the context of lung disease detection, CNNs can process high-resolution medical images such as, chest X-rays & CT scans to identify abnormalities. By feeding these images into the CNN, the network processes the data through multiple layers of convolutional filters designed to detect features such as textures, edges, and patterns associated with various lung conditions. These features help the system classify images into categories, identifying early indicators of diseases such as pneumonia, lung cancer, or COVID-19. This automated process can outperform traditional, manual diagnostic methods, offering high precision while reducing subjectivity and time constraints.[4]. Implementing CNNs for lung disease detection involves collecting high resolution chest X-rays and CT scan images, labeling them with relevant disease categories, and preprocessing them to enhance quality & reduce noise. These images are then used to train CNN, which processes data through multiple convolutional layers to identify features like textures and edges indicative of lung abnormalities. Once `trained the model is evaluated for accuracy and robustness , ensuring reliable automated detection of conditions like pneumonia and lung disease. [5]. The use of deep learning in disease detection is not confined to a single application. Approaches such as those utilizing Convolutional Neural Networks (CNNs) can extend to various types of diseases, improving diagnostic accuracy and efficiency. Combining IoT devices with synergic DL models allows for advanced classification techniques, such as detecting and classifying medical anomalies.[6]. The primary goal of this research is to develop a robust system for the early detection and classification of Lung diseases using deep learning, particularly Convolutional Neural Networks (CNNs). By leveraging high-resolution CT scan images, the proposed system aims to accurately identify lung abnormalities and assess severity levels. This involves creating a computational tool capable of automatically detecting and classifying lung tissues to assist medical professionals in diagnosing efficiently. [7]. The motivation behind my work lies in advancing the capabilities of deepLearning for medical imaging, focusing on lung disease detection and diagnosis. Lung diseases, like cancer, pneumonia, and chronic obstructive-pulmonary disease (COPD), pose significant global health challenges. Early detection & accurate diagnosis are crucial for improving, patient outcomes and reducing healthcare burdens. However, traditional `diagnostic methods often rely on manual interpretation, which can be time-consuming and subjective.[8]. The intentions behind my work are rooted in leveraging advancements in deep learning to create efficient and reliable systems for lung disease detection. Lung diseases, such as cancer and pneumonia, are among the leading causes of mortality and require early and accurate diagnosis to improve patient outcomes.[9]. Another key intention behind my work is to leverage deep learning advancements to enhance the precision and efficiency of lung disease diagnosis, addressing the growing demand for reliable and automated medical imaging systems. Lung diseases like cancer and pneumonia require timely and accurate detection, yet traditional diagnostic methods often encounter challenges such as variability in interpretations and delays [10]. The medical field is increasingly relying on AI to enhance diagnostic accuracy and patient care. Deep learning , a subset of AI, is particularly well -suited for medical image analysis , where it can identify , subtle patterns and features that may be missed by human interpretation [11]. In the era of big data, healthcare is increasingly turning to deep~learning for the analysis of complex medical images. By automating the interpretation of imaging data, deeplearningmodels are revolutionizing how clinicians diagnose diseases, improving both speed and accuracy [12]. Deep learning at the fore front of innovations in medical imaging, offering unparalleled accuracy and efficiency in diagnosing diseases. This journal examines how deep-learning techniques are reshaping the landscape of medical image analysis, making healthcare more precise and personalized [13]. As deep learning technology continues to evolve , its applications in medical , image analysis are becoming increasingly sophisticated. From detecting early signs of disease to enhancing the quality of imaging data, deep learning is redefining what is possible in medical diagnostics [14]. The complexity of medical-imaging data often makes it difficult for healthcare professionals to analyze manually. Deep learning is providing a solution by automating the process and enabling faster, more accurate diagnoses, especially in areas such as oncology, cardiology, and neurology [15]. Deep learning has become a transformative force in medical image analysis, allowing healthcare providers to extract - valuable insights from, complex imaging data. By leveraging deep neural networks, clinicians can now make faster, more accurate decisions, that improve patient care [16].

## 2.RELATED WORK

### Machine learning and deep learning approach for medical image analysis : diagnosis to detection

In this journal Meghavi Rana & Megha Bhushan discuss the integration of machine-learning (ML) and deep-learning (DL) techniques for disease detection and diagnosis. It highlights the growth of computer-aided detection using DL and ML, emphasizing the importance , of medical images for accurate diagnosis. The study reviews various ML and DL approaches, including CNN s , GAN s, and addresses challenges such as noise, artifacts, and inter-patient variability. It provides , a detailed analysis, of 40- primary studies and presents experimental results, using MRI datasets. The paper concludes that DL techniques, despite challenges, offer, significant `potential for improving medical-image analysis . [1]

### A review on deep learning in medical image analysis

In their research, S. Suganyadevi, V. Seethalakshmi, and K. Balasamy discuss , how Convolutional Neural-Networks (CNNs) serve as a foundational technique in medical image analysis. Due to their `ability to efficiently process both 2D & 3D data, CNNs are well-suited for tasks like image `classification and feature extraction. The paper examines a variety of deep learning methods and evaluates their usefulness in detecting, categorizing, and analyzing clinical features across multiple areas of medicine, including neuroimaging, retinal diagnostics, and pulmonary imaging. Other domains like breast and bone imaging, gastrointestinal studies, and musculoskeletal systems are also explored. The study outlines `challenges such as the demand , for extensive annotated datasets and high computational power. Furthermore, the authors discuss advanced deep learning architectures—including CNNs, recurrent neural-networks (RNNs) , and autoencoders—and their integration with big data analytics to enhance diagnosis and personalize treatments. The paper underscores how deep learning continues to revolutionize healthcare diagnostics by boosting accuracy and operational efficiency.[2]

### Medical image analysis based on deep learning approach

In this study `Muralikrishna Puttagunta and S. Ravi investigate the role of deep-learning in overcoming limitations of conventional image-processing methods in medicine. The paper highlights the implementation of CNNs, RNNs, and autoencoders in analyzing X-rays, CT and MRI scans, mammograms, and histopathology images. These methods support critical tasks such as disease detection, medical image classification, segmentation, and pattern analysis. The authors provides, an overview of the progression of neural network models and how they're being integrated into diagnostic tools. The review not only examines current applications but also points to challenges and future opportunities, reinforcing the impact of deep learning in streamlining healthcare workflows and improving diagnostic outcomes. [3]

### Deep learning -based image processing in optical microscopy

In this journal, Sindhoora Kaniyala Melanthota, Dharshini Gopal, Shweta Chakrabarti, Anirudh Ameya Kashyap, Raghu Radhakrishnan, and Nirmal Mazumder discuss the transformative role of deep-learning techniques in optical microscopy. The authors `highlight how deep-learning has significantly improved image processing, enabling tasks such as classification, segmentation, and resolution enhancement with higher accuracy and efficiency compared to the traditional methods. Various optical microscopy techniques, including fluorescence, phase-contrast, and non-linear microscopy, are explored in the context of deep learning applications. The review emphasizes advancements in medical and biological research, particularly in automated analysis and mobile-phone-based microscopy for remote diagnostics. It also addresses difficulties such as, computational costs and the need for large datasets, while showcasing state-of-the-art deep-learning models that have enhanced image quality and enabled innovative applications across diverse fields. [4]

### Deep Learning and the Future of Biomedical Image Analysis

In this journal, Sindhoora Kaniyala Melanthota et al. explore, the use of deep-learning in optical microscopy, emphasizing how it has enhanced tasks like image-classification, segmentation, & super-resolution. The review details how various microscopy methods—including fluorescence, phase-contrast, and nonlinear imaging—benefit from DL technologies. The paper highlights its value in medical and biological research and notes applications in remote diagnostics using mobile-based platforms. Despite these advances, it acknowledges limitations such as the high computational cost and the dependency on large annotated datasets for training.[5]

**Internet of Things and Synergic Deep Learning Based Biomedical Tongue Color Image Analysis for Disease Diagnosis and Classification**

In this journal, Romany F. Mansour, `Maha M. Althobaiti, & Amal Adnan Ashour introduce a new automat-system for analyzing tongue color images for medical diagnostics. This system, called ASDL-TCI, integrates IoT technology with synergic deep learning to enhance disease classification. It includes various phases: data collection via IoT devices, preprocessing through median filtering, feature extraction with a `synergic deep-learning network, and final classification using a deep neural-network . The model's parameters are fine-tuned using an `enhanced black widow optimization method. Their results show that ASDL-TCI achieves `outstanding precision, recall, & accuracy, establishing it as a reliable tool for non-invasive diagnosis.[6]

**A Review on Alzheimer's Disease Through Analysis of MRI Images Using Deep Learning Techniques**

In this journal, Battula Srinivasa Rao and Mudiyala Aparna conduct a detailed analysis of deep-learning `applications for MRI image interpretation in Alzheimer's diagnosis. Their paper `highlights the importance of early detection and compares deep learning with conventional methods, noting CNNs' efficiency in processing MRI data. They explore the utility of datasets like ADNI, OASIS, and MICCAI, while addressing segmentation challenges like anatomical complexity and low contrast. The review, also discusses the metrics and improvements made possible by DL technologies, emphasizing their value in improving AD detection. [7]

**Deep-Learning in Medical Ultrasound Image Analysis: A Review**

This journal, Xinke Ge and colleagues review how deep learning enhances ultrasound image analysis. They discuss how CNNs, RNNs, autoencoders, and transfer-learning assist in tasks such as `image enhancement, classification, segmentation, and object detection. The paper emphasizes DL's ability to reduce the subjectivity and limitations found in conventional ultrasound diagnostics, offering improved accuracy and efficiency.They use image enhancement to improve the quality and reliability of ultrasound data. By integrating these technologies, the review showcases how deep learning is transforming ultrasound image analysis, making diagnostics more accurate and efficient.[8]

**Endoscopic Image Analysis for Gastrointestinal Tract Disease Diagnosis Using Nature Inspired Algorithm With Deep Learning Approach**

In this journal, Abdulrahman Alruban and team propose the EIAGTD-NIADL framework, combining ShuffleNet and SLSTM models for GI tract disease diagnosis using endoscopic images. Bilateral filtering is applied for preprocessing, and parameter tuning is performed using a modified spotted hyena optimizer. Tested on the Kvasir dataset, their system shows 98.96% accuracy, demonstrating its experimental value for medical diagnostics. Experimental results, tested on the Kvasir dataset, demonstrate the system's superior performance with an `accuracy of 98.96%, surpassing other contemporary models. The study `highlights the potential of the EIAGTD-NIADL system in advancing medical image analysis, reducing diagnostic errors, and enabling precise disease identification in the GI tract. [9]

**Biomedical Image Analysis for Colon and Lung Cancer Detection Using Tuna Swarm Algorithm With Deep Learning Model**

In this journal, Munya A. Arasi and others propose BICLCD-TSADL, a hybrid deep learning system for lung and colon cancer diagnosis. They use techniques such as Gabor filtering, GhostNet, and Echo State Network, further enhanced by the Tuna

Swarm Algorithm and AFAO for hyperparameter tuning. Their approach yields a remarkable 99.33% accuracy, positioning the model as a highly effective diagnostic solution. Experimental results demonstrate the model's superior performance, achieving a maximum accuracy of 99.33% compared to present methods. The authors highlight the `potential of this approach to revolutionize cancer diagnostics by providing a cost effective, automated, and highly accurate solution for early `detection and treatment planning.[10]

**Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection**

This journal, authored by Iftikhar Naseer and team present a modified U-Net approach for classifying lung cancer, using `CT scan data. This method features enhanced lobe segmentation and precise nodule detection. By targeting nodules with varying shapes and sizes, their model ensures accurate classification and supports early diagnosis, reducing the burden on radiologists. The authors emphasize the capability of this model to transform clinical practices by enabling early diagnosis and timely intervention, which are critical for improving lung cancer survival rates. Furthermore, the automated nature of the model reduces the burden on radiologists, allowing for faster and more consistent diagnoses. The study also discusses future improvements, such as incorporating more diverse datasets and refining the model's ability to handle complex cases involving overlapping or hidden nodules.[11]

**Identification of Lung Tumors in Nude Mice Based on the LIBS With Histogram of Orientation Gradients and Support Vector Machine**

This journal, Qian-Lin Lian and collaborators introduce a classification method that merges LIBS with HOG and SVM for analyzing lung tumors. LIBS captures chemical data from tissues, and HOG extracts texture features. These are classified by SVM, providing a fast, reliable, and non-invasive alternative to traditional methods, particularly beneficial for early-stage cancer detection.[12]

**A Survey of Wound Image Analysis Using Deep Learning: Classification, Detection, and Segmentation**

In this journal, Ruyi Zhang and team provides an overview of deep-learning applications in wound image analysis. Their paper addresses key challenges such as data quality and scarcity while highlighting CNN-based solutions for tasks like segmentation and detection. They emphasize how DL improves the reliability and efficiency of wound diagnosis. The challenges in wound image analysis, such as data scarcity, noise, and color variation, are discussed. The study emphasizes the transformative role of deep-learning in improving the accuracy, efficiency, and accessibility of wound diagnosis and treatment, providing insights into future research and development in this vital area of medical imaging.[13]

**Hybrid Metaheuristics With Deep Learning-Based Fusion Model for Biomedical Image Analysis**

In this journal, Marwa Obayya and colleagues propose a deep-learning and metaheuristic fusion framework, HMDL-MFMBIA, for biomedical image classification. It combines Swin-UNet, Xception, ResNet, GRU, and the HSSA optimizer. Validated on ISIC datasets, this model shows superior performance and demonstrates how combining optimization techniques with deep learning can enhance healthcare diagnostics. With challenges such as improved accuracy and efficiency, the HMDL-MFMBIA technique showcases its effectiveness through experimental validation on datasets like ISIC 2017 and ISIC 2020, demonstrating superior performance compared to existing methods. [14]

**Deep Learning Applications in Medical Image Analysis**

This journal, authored by Justin Ker and co-authors provide a broad overview of deep-learning applications, in the medical imaging, with a focus on CNNs, GANs, and autoencoders. They analyze these tools across multiple imaging modalities and highlight how transfer learning and data augmentation help mitigate data limitations. Their study reflects DL's potential in improving patient outcomes through faster and more accurate diagnostics. This review underscores the potential of deep learning to revolutionize medical diagnostics by improving accuracy, efficiency, and patient outcomes. It underscores how these models have evolves with the advent of big data and GPU advancements, allowing the discovery of hierarchical relationships in medical images without manual feature crafting including CT, MRI, ultrasound, and histology slides. [15]

**Unified deep learning models for enhanced lung cancer prediction with ResNet 50–101 and EfficientNet-B3 using DICOM images**

In this journal Vinod Kumar and collaborators investigate the uses of unified deep-learning models such as ResNet-50, ResNet-101, and EfficientNet-B3 for lung cancer classification using DICOM CT images. Their work emphasizes the success of transfer learning in classifying various cancer subtypes, achieving precision rates up to 100% for squamous cells, and demonstrating the impact of deep-learning in addressing late-stage cancer diagnosis challenges. The study addresses challenges, such as high mortality rates due to late detection and the `difficulty in distinguishing between similar cancer subtypes. Using a dataset of 1,000 DICOM images from the LIDC-IDRI repository, the authors achieved significant results. The Fusion Model demonstrated 100% precision in classifying Squamous Cells, while ResNet-50, EfficientNet-B3, and ResNet-101 achieved 90% precision, with slight variations in performance. [16]

**Table 1:** Comparison among Different Existing Systems

| S. No | Title | Algorithm | Challenges | Data Sets | Results |
|---|---|---|---|---|---|
| 1 | Machine learning and deep learning approach for medical image analysis : diagnosis to detection | ML, DL | Complexity of Medical Imaging Data | MRI, 4D DCE MRI images | CNN got accuracy of 97.6% RF achieved an accuracy of 96.93% |
| 2 | Unified deep learning models for enhanced lung cancer prediction with ResNet - 50–101 and EfficientNet - B3 using DICOM images | Resnet 50, resnet 101, EfficientNet B3, Fusion Model | Overfitting, late detection | LIDC-IDRI | Accuracy of 90% with resnet 50 |
| 3 | A review on deep-learning in medical image analysis | ( CNNs ), (RNN s ) | High Computational Requirements | Lung cancer dataset, Heart-C dataset | CNN achieved the more accuracy of 97.6%, Random Forest (RF) achieved an accuracy of 96.93% |
| 4 | Medical Image Analysis Based on Deep Learning Approach | CNN, RNN, GANs | Need of large labeled datasets | ChestX-ray14,LUNA16 | CNN achieved the highest accuracy of 97.6%, Random Forest ( RF ) achieved an accuracy of 96.93% |
| 5 | Deep learning - based image processing in optical microscopy | U-Net-based CNN, Deep Loc | Data variability and generalizability | Fluorescent microscopy images, Holographic images of living cells | DeepLoc enhanced automated categorization of protein subcellular localization in yeast cell images, achieving 84% precision |

| | | | | | |
|---|---|---|---|---|---|
| 6 | Deep Learning and the Future of Biomedical Image Analysis | Long Short-Term Memory Networks (LSTMs), Encoder-Decoder (ED) architectures | Interpretability of model predictions | EEG, ECG, MEG (Magnetoencephalogy) | CNN models showed an accuracy of up to 97.6%, Random Forest (RF) achieved an accuracy of 96.93%. |
| **7** | Internet of Things and Synergic Deep Learning Based Biomedical Tongue Color Image Analysis for Disease Diagnosis and Classification (2021) | Median Filtering (MF) for image pre-processing, Synergic Deep Learning (SDL) for feature extraction | Data privacy and security concerns | Benchmark tongue images dataset consisting of 936 images under 12 class labels | Precision: 98.4%, Recall: 97.3%, Accuracy: 98.3% |
| 8 | A Review on Alzheimer's Disease Through Analysis of MRI Images Using Deep Learning Techniques (2023) | Deep Neural Networks (DNN), Enhanced Black Widow Optimization (EBWO) | Interpretability of model predictions | OASIS, ADNI | Dice Similarity Coefficient (DSC): Up to 98%, Jaccard Index (JI): 95%, Precision: 98.4%, Recall: 97.3%, Accuracy: 98.3% |
| 9 | Deep Learning in Medical Ultrasound Image Analysis: A Review (2021) | Restricted-Boltzmann Machines (RBMs), Transfer Learning (TL) | Complexity of medical-imaging data | CETUS Challenge dataset, Thyroid dataset, Breast ultrasound image dataset | CNNs showed improved accuracy of 98% in classification tasks. |
| 10 | Endoscopic Image Analysis for Gastrointestinal Tract Disease Diagnosis Using Nature Inspired Algorithm With Deep Learning Approach (2023) | Bilateral Filtering (BF), Improved Spotted Hyena Optimizer (ISHO) | High computational requirements. | Kvasir Dataset | Precision: 95.85%, Recall: 95.85%, Accuracy: 98.96% |

| | | | | | |
|---|---|---|---|---|---|
| 11 | Biomedical Image Analysis for Colon and Lung Cancer Detection Using Tuna Swarm Algorithm With Deep-Learning Model (2023) | Gabor Filtering (GF), Adaptive Firefly Algorithm Optimization (AFAO), `Tuna Swarm Algorithm (TSA) | Training complexity, Need for large annotated datasets | LC25000 Dataset | Accuracy: 99.33%, Precision: 98.31%, Recall: 98.31%, F1 Score: 98.31%, AUC (Area Under Curve): 98.95% |

# 3. PROPOSED METHODOLOGIES

## 3.1 VGG16

This project leverages the power of VGG16 architecture, a well-established deep convolutional neural-network (CNN), to detect and classify lung tumors from CT scan images. VGG16 is particularly effective in image classification tasks due to its consistent use of small 3×3 convolutional filters and deep layered structure , which allows for hierarchical feature extraction . Originally designed for RGB images, the VGG16 `model has been adapted in this project to handle grayscale CT scans by resizing, them to 224 × 224 × 3, ensuring compatibility with the pretrained architecture while maintaining the visual structure of the input.

The proposed system for lung tumor classification involves several sequential stages as outlined in the block diagram (Fig 1). Each stage is important in transforming raw medical imaging data into meaningful predictions that can assist in clinical decision-making. The stages are described below:

**1. Lung Dataset:**

The foundation of this system lies in a carefully curated lung CT image dataset, which contains both healthy and tumor-affected lung scans. The tumor cases are further divided into three distinct categories representing various types of lung cancer. These include Large cell carcinoma, Squamous cell carcinoma, Adenocarcinoma.

**2. Image Preprocessing:**

To ensure effective and uniform input for the neural network, each image undergoes a series of preprocessing steps:

- **Image Resizing:** Adjusts images from their original size to uniform (224 x 224) size for consistent analysis.
- **Image Normalization:** Enhances image quality by scaling `pixel values to a standard, range(0 to 1), improving model performance.

**3. Feature Extraction:**

In the VGG16 architecture, feature extraction from diseased lung images occurs through a deep hierarchical structure of convolutional layers. The initial layers capture low-level features like edges, textures , and boundaries within the lung CT image . As the image passes through deeper layers, the model identifies more complex patterns like abnormal growths, nodules, or density variations associated with tumors. Each convolutional block extracts increasingly abstract and detailed information, allowing the model to differentiate between normal tissue and cancerous regions. By the end of the convolutional stages, VGG16 effectively encodes critical features that represent the unique characteristics of various lung disease types, which are then pass to fully connected layers, for classification.

**4. Classification:**

The extracted features are fed into VGG16 , a DL‘s model for classification. The model categorizes images into four classes:

- Normal (Healthy Lungs)
- Adenocarcinoma
- Large Cell Carcinoma
- Squamous Cell Carcinoma

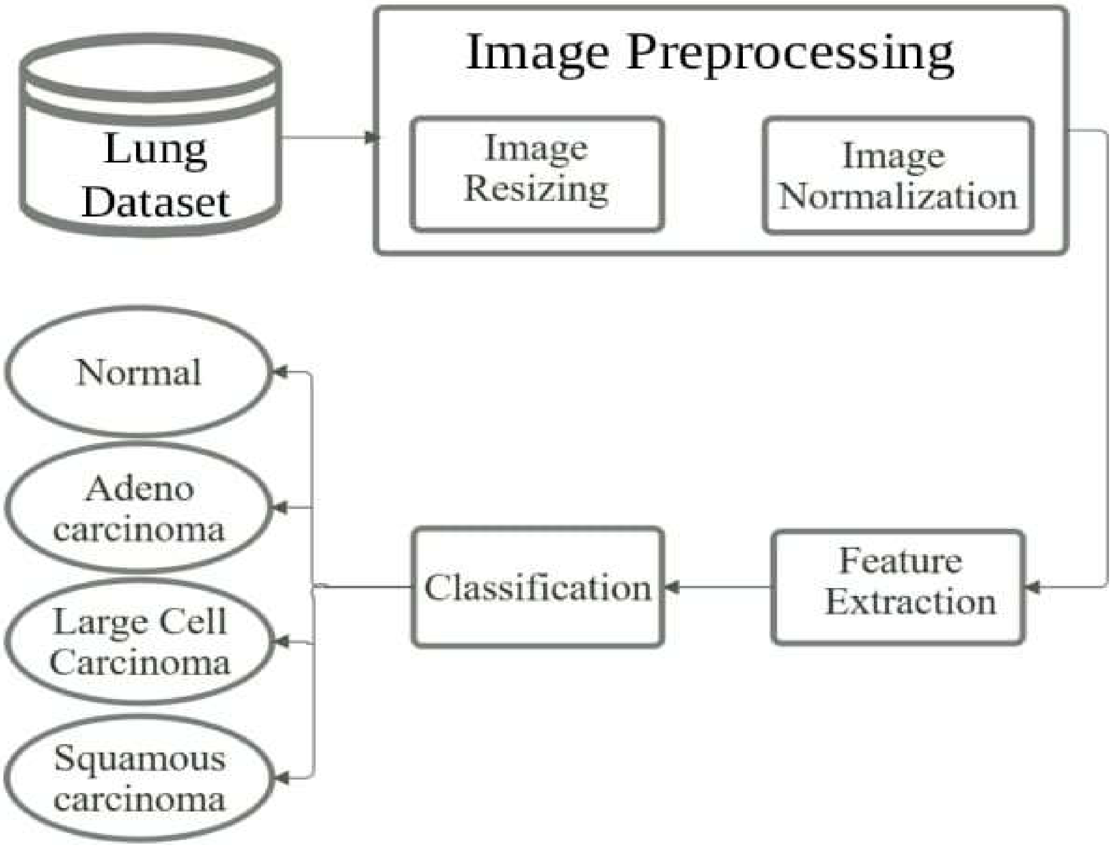



**Fig 1:** Disease Detection Techniques using Biomedical Image Analysis

## 4 DATASETS

The dataset summary in above table 2 includes various medical imaging datasets used for disease detection. OASIS and ADNI provide large sets of MRI brain images for neurological research, while BRATS and the Brain Tumor MRI Dataset focus on brain tumor analysis. Chest-related conditions are targeted by Montgomery (X-ray), JSRT (X-ray), and the Chest CT Scan dataset. Skin cancer diagnosis benefits from dermoscopy images in ISIC 2017 and ISIC 2020. Additionally, GLAS supports colorectal cancer segmentation, and the Covid-19 dataset offers basic CSV data for infection analysis.

**Table 2:** Datasets used in biomedical Image Analysis.

| S.NO | Dataset Name | No of Images | Type |
|---|---|---|---|
| 1. | Oasis | 87k | MRI |
| 2. | Brats (Brain Tumor Segmentation) | 57k | MRI |
| 3. | Montgomery | 139 | X-Ray |
| 4. | Brain Tumor MRI Dataset | 7k | MRI |
| 5. | ADNI | 18k | MRI |
| 6. | JSRT(Chest X-Rays) | 238 | X-Ray |
| 7. | Covid-19 | 100 | CSV |
| 8. | ISIC 2017 | 2000 | Dermoscopy |
| 9. | ISIC 2020 | 33K | Dermoscopy |
| 10. | GLAS | 100 | CSV |
| 11 | Chest CT scan Images Dataset | 1000 | CT scan |

## 5 RESULTS AND DISCUSSION

This section evaluates the performance of various tumor classification models using standard metrics. And also the web interface used to output the predicted disease.

### 4.1 Accuracy:

Accuracy measures how often a model correctly predicts the outcome of a classification task. It represents the proportion of correctly classified cases (both tumor and non-tumor) out of all cases, calculated as:

$$\text{Accuracy} = \frac{(TP + TN)}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})}$$

### 4.2 Precision:

Precision measures how accurately a model predicts positive outcomes. It represents the proportion of correctly predicted tumor cases out of all cases predicted as tumors, calculated as:

$$\text{Precision} = \frac{TP}{(\mathrm{TP} + \mathrm{FP})}$$

### 4.3 Recall:

Recall is that it measures how well a model identifies positive outcomes. It represents the proportion of correctly predicted tumor cases out of all actual tumor cases, calculated as:

$$\text{Recall} = \frac{TP}{(\mathrm{TP} + \mathrm{FN})}$$

### 4.4 F1-Score:

F1 - Score combines precision and recall to provide a single measure of a model's performance. It represents the harmonic mean of precision and recall, calculated as:

$$\text{F1-Score} = \mathbf{2} \times \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

### 4.5 Mean dice or Dice Coefficient:

It is used to evaluate the similarity between two sets, commonly used in image segmentation tasks. It represents the overlap between predicted and actual tumor regions, calculated as:

$$\text{Mean Dice} = \frac{(2 * \mathrm{TP})}{(2 * \mathrm{TP} + \mathrm{FP} + \mathrm{FN})}$$

### 4.6 Specificity:

Specificity measures the ability of a model to correctly identify negative outcomes. It represents the proportion of correctly predicted non-tumor cases out of all actual non-tumor cases, calculated as:

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

Through keen observation of the below images we know that we are getting a val accuracy of 98% and val loss of 2%.

### 4.7 Accuracy and Loss Plots

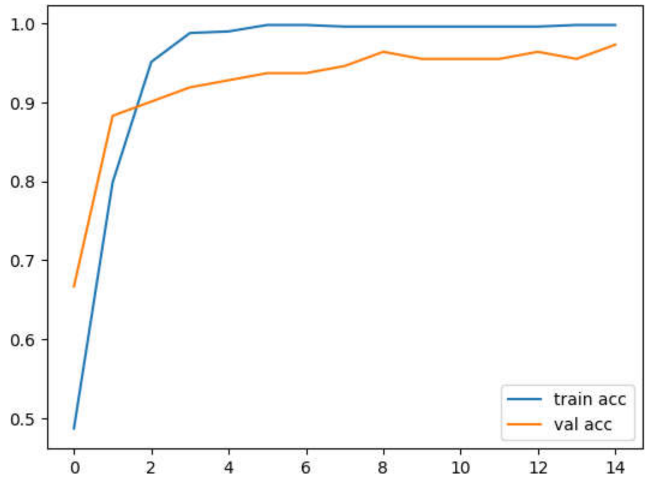
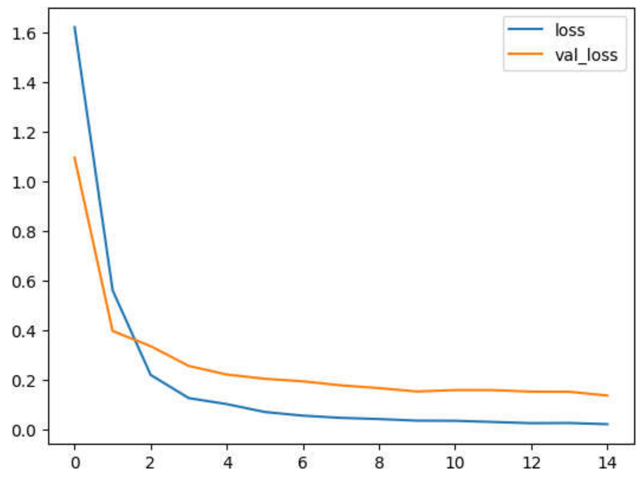



**Fig 2:** Train and val accuracy

**Train accuracy:** It measures how well the model performs on training data, while val accuracy shows performance on unseen validation data. **Train loss:** It indicates the model’s error on training data, and val loss measures error on validation data. A large gap between train and val accuracy suggests overfitting. Ideally, both accuracy should be high, and both losses should decrease steadily.

- Train and val accuracy should be close.
- Train and val loss should both decrease steadily.
- If val loss increases while train loss decreases, the model is overfitting.




**(b):Fig 3:** Train and val loss

## 4.8 Confusion Matrix

The confusion matrix provides a detailed evaluation of the model's ability to classify lung disease types, including adeno, large cell, normal, and squamous. The majority of predictions are correct, with 32 adeno cases, 21 large cell cases, 21 normal cases, and 33 squamous cases accurately classified. Specifically, 2 adeno cases were misclassified as large cell, while 1 large cell case was incorrectly predicted as adenocarcinoma.

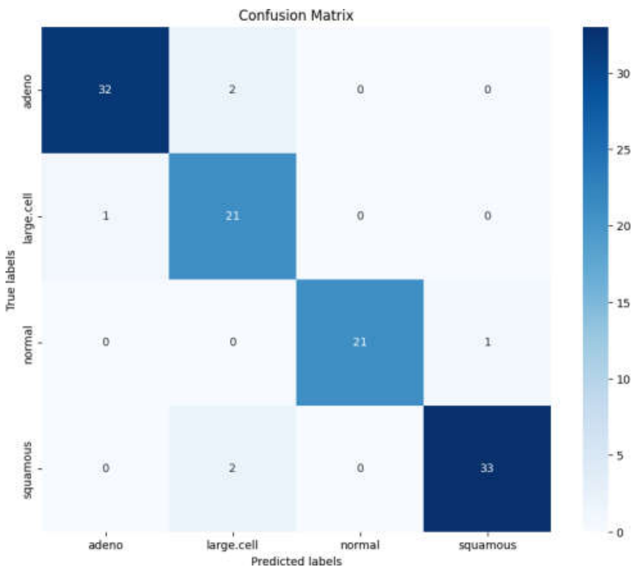



**Fig 4:** Confusion Matrix

## 4.9 Streamlit Input and Output Interface

The Streamlit interface for this lung tumor detection project provides a clean and intuitive user experience. As shown in the Figure 5, at the top, it displays the project title with a lung themed image. Users are prompted to upload a CT scan image (in JPG, JPEG, or PNG format) using the file uploader component. Once the image is uploaded, it is displayed on the screen, giving immediate visual feedback that the upload was successful.

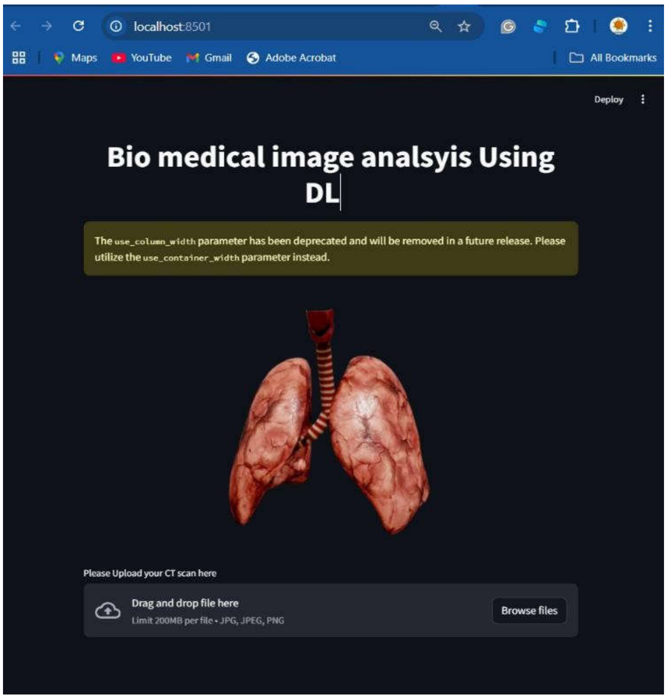



**Fig 5:** Streamlit Input Interface

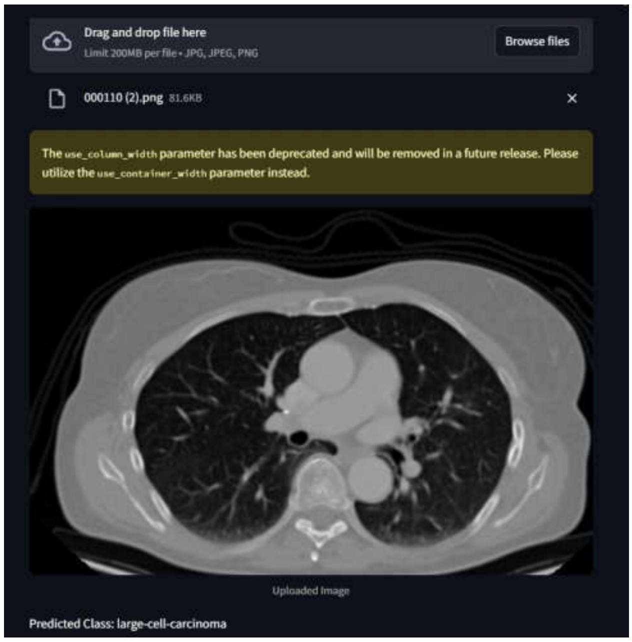



**Fig 6:** Streamlit Output Interface

From Figure 6 we can know that once the image is shown, the system automatically resizes it to 224x224 pixels, normalizes it, and feeds it into the VGG16 model for prediction. The model processes the input and outputs the probabilities for each class (Adenocarcinoma, Large Cell Carcinoma, Normal, and Squamous Cell Carcinoma). This streamlined workflow allows users medical professionals or researchers to upload a scan and receive an efficient output.

**Table 3 :** Performance of VGG16 vs. Existing Models

| Model | Accuracy(%) | Precision(%) | Sensitivity(%) | Specificity(%) |
|---|---|---|---|---|
| Proposed VGG16 Model | 97 | 94 | 94 | 94 |
| Res-Net 50 | 95 | 90 | - | - |
| Res-Net 101 | 94 | 89 | - | - |

From the Table 3 we can observe that the proposed VGG16 model demonstrates superior performance with an accuracy of 97%, precision of 94%, F1 score of 94% and recall at 94%. This indicates that VGG16 effectively balances sensitivity and specificity, making it highly reliable for lung disease classification. The high precision suggests that the model minimizes false positives, while the recall value confirms its ability to correctly identify actual disease cases. This performance can be attributed to VGG16's deep convolutional layers, which efficiently extracts the `hierarchical spatial features from medical images. The structured architecture of VGG16 ensures effective feature learning. ResNet50 achieves an accuracy of 94% with a precision of 90%, while ResNet101 attains a slightly lower accuracy of 94% and precision of 89%. The F1 score and recall values were not explicitly mentioned in the base paper, which lacks a detailed comparison to balanced classification. Although ResNet models leverage deep RL to overcome vanishing gradient issues, their lower accuracy in this context suggests that increasing network depth, does not necessarily translate to better classification performance. The marginal difference between ResNet50 and ResNet101 highlights that adding more layers may introduce complexity without significant performance gains. While ResNet known for their robust feature extraction, VGG16 appears to be more effective in this specific lung disease detection task due to its structured and efficient feature learning process.

**Table 4 :** Performance of different types of diseases

| Type of Disease | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Adeno | 0.9487 | 0.9487 | 0.9487 | 39 |
| Larg cell | 0.9167 | 0.9565 | 0.9362 | 23 |
| Normal | 0.9310 | 0.9310 | 0.9310 | 29 |
| Squamous | 0.9667 | 0.9355 | 0.9508 | 31 |

The performance metrics in the above Table 4 highlight VGG16's strong ability to classify four lung tumor types accurately, with Adenocarcinoma, Normal, and Squamous Cell Carcinoma all showing precision and F1-scores above 0.93. Large Cell Carcinoma also shows high recall at 0.95, indicating effective identification of true positive cases. These consistent results demonstrates that, the VGG16 model is robust, well-generalized, and capable of learning important patterns in lung CT scan images. In contrast, ResNet50 and ResNet101 performed slightly lower, likely due to their deeper and more complex architectures which tend to overfit on smaller medical datasets without sufficient regularization. Moreover, as these models are pretrained on RGB images from natural scenes, they may not fully adapt to the grayscale CT scan domain without extensive fine-tuning. VGG16's sequential structure and smaller filters make it more effective in extracting subtle features from lung tissue, contributing to its higher accuracy and reliability in tumor detection. Overall, the results affirm that VGG16 offers a more optimal balance of performance and simplicity for medical image classification tasks like lung cancer detection.

## 5 CONCLUSION

Based `on the state-of-the-art methods, development of a Streamlit-based web application using the VGG16 model for lung cancer detection, represents a `significant advancement in medical diagnostics. The user-friendly steamlit interface allows healthcare professionals and researchers to easily upload lung CT scan images for analysis, facilitating early and accurate detection of different types of `lung cancer, including squamous cell carcinoma, large cell carcinoma, and adenocarcinoma. Beyond classification, the application provides details about the identified condition, including symptoms, possible causes, and recommended treatment strategies. By offering actionable insights, the system empowers medical professionals to make informed decisions, potentially improving patient outcomes and enabling timely interventions. The uses of the VGG16 model enhances the accuracy and reliability of the classification process, ensuring consistent and precise results. This `technology not only supports early diagnosis but also serves as an educational tool, increasing awareness about lung cancer types and their management among medical practitioners and researchers.

## 6.REFERENCES

[1] “Meghavi Rana and Megha Bhushan," Machine Learning and Deep `Learning Approach for ~Medical Image Analysis Diagnosis to Detection, Multimedia Tools and Applications, vol. 82, pp. 26731–26769, 2023.

[2] “S. Suganyadevi, V. Seethalakshmi, and K. Balasamy”, 'A review on deep-learning in `medical image analysis,' International Journal of Multimedia Information Retrieval, vol. 11, no. 1, pp. 19–38, 2022.

[3] “Muralikrishna Puttagunta and S. Ravi”, “MedicalImage Analysis Based on Deep Learning Approach ”, Multimedia Tools and Applications, vol. 80, pp. 24365–24398, 2021.

[4] “Sindhoora Kaniyala Melanthota, Dharshini Gopal, Shweta Chakrabarti, Anirudh Ameya Kashyap, Raghu Radhakrishnan, and Nirmal Mazumder “, “Deep Learning- Based Image Processing in Optical Microscopy”, Biophysical Reviews, vol. 14, pp. 463–481, 2022.

[5] “Monika Jyotiyana and Nishtha Kesswani,” “Deep-Learning and the Future of Biomedical Image Analysis”, DeepLearning Techniques for `Biomedical and Health Informatics, Studies in Big Data, vol. 68, pp. 329–342, 2020.

[6] “Romany F. Mansour, Maham Althobaiti, and Amal Adnan Ashour”, “Internet of Things and Synergic Deep-Learning Based Biomedical Tongue Color Image Analysis for _Disease Diagnosis and Classification”, IEEE Access, vol. 9, pp. 94769-94777, 2021.

[7] "Battula Srinivasa Rao and Mudiyala Aparna, 'A Review on Alzheimer's Disease Through Analysis of MRI ImagesUsing Deep_Learning Techniques,' IEEE Access, vol. 11, pp. 71542-71556, 2023.

[8] Xinke Ge, He Ma, Shouliang Qi, Yu Wang, and Yudong Yao, "Deep Learning in Medical Ultrasound Image Analysis: A Review," IEEE Access, vol. 9, pp. 54310-54325, 2021.

[9] Abdulrahman Alruban, Abdullah Alharbi, Eatedal Alabdulkreem, Majdy M. Eltahir, Imène Issaoui, and Ahmed Sayed, "Endoscopic Image Analysis for Gastrointestinal Tract Disease Diagnosis Using Nature-Inspired Algorithm With Deep Learning Approach," IEEE Access, vol. 11, pp. 130022-130029, 2023.

[10] Munya A. Arasi, Marwa Obayya, Abdullah Mohamed, Nuha Alruwais, Raed Alsini, and Ishfaq Yaseen, "Biomedical Image-Analysis for Colon &

Lung Cancer Detection Using Tuna Swarm Algorithm With Deep Learning Model," IEEE Access, vol. 11, pp. 94705-94712, 2023.

[11] Iftikhar Naseer, Arfan Jaffar, Sheeraz Akram, Tehreem Masood, and Muhammad Rashid, "Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection," IEEE Access, vol. 11, pp. 60279-60287, 2023.

[12] Qian-Lin Lian, Xiang-You Li, Bing Lu, Chen-Wei Zhu, Jiang-Tao Li, and Jian-Jun Chen, "Identification of Lung Tumors in Nude Mice Depending on the LIBS With Histogram of Orientation Gradients and Support-Vector Machine," IEEE Access, vol. 11, pp. 141915-141921, 2023.

[13] RuyiZhang , Ding cheng Tian , Dechao Xu, Wei Qian , and Yudong Yao, " A Survey of Wound Image Analysis Using Deep Learning: Classification, Detection, and Segmentation," IEEE, Access, vol. 10, pp. 79502-79509, 2022.

[14] Marwa Obayya, Saud S. Alotaibi, Muhammad Kashif Saeed, Nuha Alruwais, Mohammed Assiri, and Ahmed S. Salama, "Hybrid Metaheuristics With Deep. Learning Based Fusion-Model for Biomedical Image Analysis," IEEE Access, vol. 11, pp. 117149-117155, 2023.

[15] Justin Ker, Lipo-Wang, Jai Rao, and Tchoyoson Lim, "Deep Learning `Applications in Medical-Image Analysis," IEEE Access, vol. 6, pp. 9375-9440, 2018.

[16] Vinod Kumar, ChanderPrabha, Preeti Sharma , Nitin Mittal, S . Askar et.al, “Unified deep-learning models for , enhanced lung cancer prediction with ResNet-50–101 and EfficientNet-B3 using DICOM images, vol. 24, no. 1, pp. 1–21, 2024.