ENHANCING EARLY LUNG CANCER DETECTION THROUGH TRANSFER LEARNING: A COMPARATIVE ANALYSIS OF XGBOOST AND DENSENET201 MODELS ON THE IQ-OTH/NCCD DATASETS

Shubham Bhattarai
Computer Science and Engineering
JNTUH University College of Engineering, Science and Technology.
Hyderabad, Telangana.
India.

Dr. M. Chandra Mohan
Computer Science and Engineering (Professor)
JNTUH University College of Engineering, Science and Technology.
Hyderabad, Telangana.
India.

Abstract—Lung cancer remains a leading cause of cancer-related deaths globally, with early detection being critical for improving survival rates. The existing systems for lung cancer detection primarily rely on imaging techniques such as X-rays and computed tomography (CT) scans, which are manually interpreted by radiologists. While these methods have proven effective, they are limited by their dependency on human expertise, which is subject to fatigue, variability, and error. Detecting early-stage lung cancer is particularly challenging due to the subtlety of its early signs, such as small nodules that are difficult to discern. Additionally, manual interpretation is labor-intensive, time-consuming, and often inconsistent, especially in high-volume clinical settings. Even with the advent of computer-aided detection (CAD) systems, the accuracy remains suboptimal due to high false-positive rates and a reliance on radiologists for final decisions. This project addresses the limitations of the existing system by employing a hybrid artificial intelligence-based approach, combining the strengths of DenseNet201 and XGBoost. DenseNet201 is leveraged for feature extraction from CT scan images using transfer learning, while XGBoost performs robust classification based on these features. This methodology significantly improves diagnostic accuracy, efficiency, and scalability. By automating nodule detection and classification, the proposed system reduces the risk of false positives and negatives, streamlining the diagnostic process. The integration of advanced AI models provides a scalable and consistent solution, laying the groundwork for more reliable, accessible, and early lung cancer detection in real-world clinical environments

Keywords: Deep DenseNet201, XGBoost, Computed Tomography (CT), Computer Aided Detection (CAD)

INTRODUCTION

OVERVIEW

The project titled aims to develop an efficient and accurate system for diagnosing lung cancer from CT scan images. Lung cancer remains one of the leading causes of cancer-related deaths worldwide, and early detection plays a pivotal role in improving survival rates. However, traditional diagnostic methods, such as biopsies and radiological analysis, are time-consuming, subjective, and prone to human error. This project addresses these limitations by leveraging advancements in deep learning and machine learning techniques. The system utilizes a hybrid approach combining "**Densenet201**", a deep convolutional neural network (CNN) known for its high efficiency in feature extraction, and "**XGBoost**", a gradient-boosting algorithm renowned for its performance in structured data classification. The project pipeline involves preprocessing the CT scan images, extracting meaningful features using Densenet201, and classifying the data with XGBoost. This combination ensures a robust and accurate classification of lung cancer, even with limited labeled datasets, by leveraging transfer learning and gradient-boosting methodologies. The primary goal of this project is to create a system capable of automating the lung cancer detection process, reducing diagnostic errors, and enabling faster decision-making. The system is evaluated using performance metrics such as accuracy, precision, recall, and F1-score, ensuring its reliability for real-world applications. Additionally, it addresses scalability and deployment considerations, making it a valuable tool for integration into clinical workflows.

PROBLEM STATEMENT

The problem addressed by this project centers around the challenges in accurately and efficiently diagnosing lung cancer from CT scan images. Firstly, traditional diagnostic methods, such as radiological analysis and biopsies, are time-intensive, prone to human error, and often result in delayed detection, which is critical for patient survival. Secondly, existing automated systems for lung cancer detection face limitations in terms of accuracy, scalability, and their ability to generalize across diverse datasets due to the complex nature of medical imaging data. Lastly, the lack of a unified approach that combines the strengths of deep learning for feature extraction and machine learning for structured data classification hinders the development of robust diagnostic tools. This project aims to address these challenges by proposing a hybrid system utilizing Densenet201 for feature extraction and XGBoost for classification, ensuring improved accuracy, efficiency, and reliability in lung cancer detection.

PROJECT DESCRIPTION

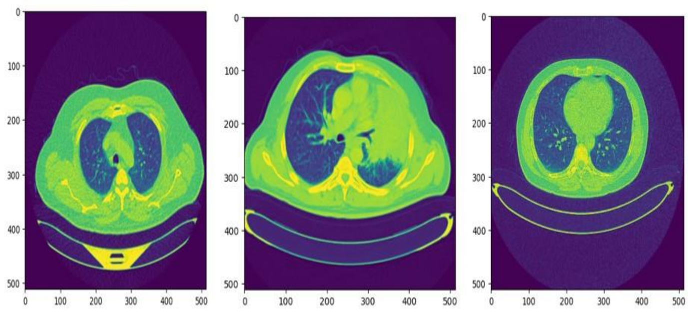
The early detection of lung cancer remains one of the most critical challenges in the medical field, with timely diagnosis often making the difference between life and death. In recent years, artificial intelligence (AI) and deep learning have emerged as transformative tools in medical diagnostics, offering new avenues for enhancing the accuracy and efficiency of disease detection. This study delves into an innovative approach to lung cancer detection, leveraging the power of transfer learning models—specifically, the XGBoost and DenseNet201 architectures applied to the IQ-OTH/NCCD dataset, a comprehensive collection of medical images. Transfer learning, which involves fine-tuning pre-trained models on specific datasets, has proven to be a highly effective strategy in various domains, including medical imaging. By utilizing the pre-existing knowledge embedded within the XGBoost and DenseNet201 models, this research aimed to adapt these deep learning architectures for the specialized task of lung nodule classification. The IQ-OTH/NCCD dataset provided a robust foundation for this endeavor, allowing the models to be trained and validated on a diverse array of medical images, representative of real-world clinical scenarios. The study employed rigorous methodologies to evaluate the performance of the fine-tuned models, focusing on critical metrics such as sensitivity, specificity, and accuracy. These metrics are vital in assessing the models' ability to correctly identify lung nodules, distinguishing between benign and malignant cases with high precision. The results from this research were promising, showcasing the models' capability to effectively contribute to early lung cancer detection, potentially leading to better patient outcomes. Beyond the technical aspects, this study also placed a strong emphasis on ethical considerations, particularly regarding patient privacy and data security. As AI and machine learning increasingly permeate the healthcare sector, safeguarding sensitive medical data becomes paramount. This research adhered to stringent ethical standards, ensuring that all data handling processes were conducted with the utmost care and in compliance with relevant regulations. In conclusion, this study represents a significant step forward in the integration of AI into medical diagnostics, specifically in the context of lung cancer detection. The use of transfer learning models like XGBoost and DenseNet201, combined with the rich IQ-OTH/NCCD dataset, demonstrates the potential to enhance screening processes and improve early diagnosis. By addressing both the technical challenges and ethical considerations, this research not only contributes to the scientific community but also offers practical implications for the future of healthcare, where AI-driven tools may become integral to saving lives through more accurate and efficient disease detection.

BACKGROUND

Lung cancer remains one of the most lethal forms of cancer worldwide, primarily due to its late detection and the aggressive nature of the disease. Early detection significantly increases the chances of successful treatment and survival, yet it poses a substantial challenge, given that the symptoms often appear at advanced stages. Traditionally, radiologists rely on imaging techniques such as X-rays and CT scans to detect lung nodules, which may be indicative of cancer. However, the manual interpretation of these images is subject to human error, variability in expertise, and can be time-consuming. These challenges highlight the urgent need for more reliable, efficient, and accurate diagnostic tools. In recent years, the advent of artificial intelligence (AI) and machine learning has revolutionized various industries, including healthcare. These technologies, particularly deep learning, have shown tremendous promise in automating and enhancing the diagnostic process. Deep learning models, especially convolutional neural networks (CNNs), have been effectively employed in medical image analysis, outperforming traditional methods in some cases. Transfer learning, a technique where a model developed for one task is reused as the starting point for another related task, has gained popularity in this domain. It allows for the application of pre-trained models on large datasets to new, specific tasks with limited data, reducing the computational resources and time required for model development. In the context of lung cancer detection, transfer learning offers a valuable approach. By fine-tuning pre-existing, well-trained models on datasets specific to lung nodule classification, researchers can create highly specialized models capable of identifying cancerous nodules with high accuracy. The IQ-OTH/NCCD dataset, which contains a large collection of medical images, provides an ideal resource for training such models. This dataset includes various types of lung nodules, representing a wide range of cases encountered in clinical practice. The application of transfer learning to lung cancer detection has the potential to significantly enhance early diagnosis, leading to better patient outcomes. However, as with any technology in the medical field, the use of AI and machine learning comes with ethical considerations. Ensuring patient privacy, securing sensitive medical data, and maintaining transparency in AI decision-making processes are critical to the responsible deployment of these technologies. This background sets the stage for understanding the significance of the study, which focuses on applying advanced transfer learning models, specifically XGBoost and DenseNet201, to the IQ-OTH/NCCD dataset. The research aims to fine-tune these models for lung nodule classification, evaluate their performance using key metrics, and address the ethical considerations associated with AI in healthcare. The findings from this study contribute to the growing body of knowledge in AI-driven medical diagnostics, offering a promising tool for improving the early detection and treatment of lung cancer.

PROPOSED FRAMEWORK

The proposed system introduces an advanced AI-driven approach for the detection of lung cancer, leveraging transfer learning techniques and state-of-the-art machine learning models to enhance the accuracy, efficiency, and consistency of lung nodule classification. This system addresses the limitations of the existing manual and CAD-based methods by automating the detection process and providing a robust, scalable solution that can be integrated into clinical practice. **DenseNet201 Model**: The proposed system employs DenseNet201, a deep convolutional neural network (CNN) architecture known for its densely connected layers, which allow for more efficient feature reuse and gradient flow. DenseNet201 has been pre-trained on large-scale image datasets (such as ImageNet), enabling it to recognize complex patterns and features in images. **XGBoost Model**: The proposed system integrates the DenseNet201 model with XGBoost, a powerful gradient boosting algorithm. After the DenseNet201 model processes the input images, it extracts high-level features representing the lung nodules. These features are then fed into the XGBoost classifier, which uses them to make the final classification decision.



Benign Malignant Normal

SYSTEM ARCHITECTURE

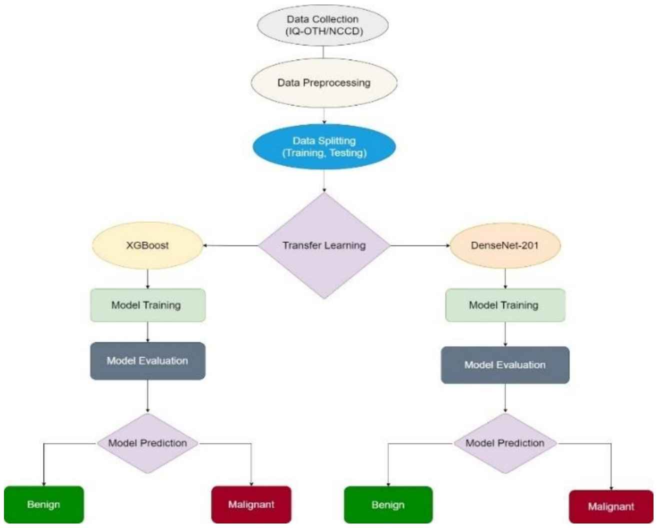


Fig.1. System Architecture

Data Collection: The project begins with the collection of a dataset, which is a crucial step in building an accurate and reliable model for lung cancer detection. The dataset primarily comprises CT scan images of lungs, sourced from well-known repositories such as IQ-OTH/NCCD or similar databases. These datasets contain labeled images that classify each scan as either benign (non-cancerous) or malignant (cancerous). Alongside the image data, metadata such as patient details and diagnostic labels are included. This rich dataset serves as the foundation for the project, providing the raw information required to train and evaluate the models.

Data Pre-processing: Once the dataset is collected, it undergoes preprocessing to prepare it for further analysis and model training. Preprocessing ensures that the data is clean, uniform, and ready for computational tasks. Image data is normalized, which involves scaling the pixel values to a range that the model can efficiently interpret. Additionally, the CT scan images are resized to match the input dimensions required by the Densenet201 model. If the dataset is small, data augmentation techniques such as rotation, flipping, and cropping may be applied to increase its size and diversity. Labels are encoded into numerical formats to make them suitable for training machine learning models. Preprocessing is a critical step to enhance model performance and ensure reliable outcomes.

Data Splitting: After preprocessing, the dataset is split into two subsets: training and testing data. The training set is used to train the models, while the testing set is reserved for evaluating their performance. This separation ensures that the models are tested on unseen data, which provides an accurate measure of their ability to generalize to new cases. Typically, the data is split in a ratio such as 80:20 or 70:30, depending on the dataset size and project requirements. Proper data splitting is essential to avoid overfitting and ensure the reliability of the final models.

Transfer Learning with Densenet201: Densenet201, a pre-trained convolutional neural network, is utilized as the backbone for feature extraction in this project. Transfer learning leverages the pre-trained knowledge of Densenet201, which has already been trained on large datasets like ImageNet. By fine-tuning the model on the lung cancer dataset, it adapts to the specific task of detecting cancer from CT scans. Densenet201 extracts high-level features, such as edges, textures, and patterns, which are critical for distinguishing between benign and malignant cases. This step reduces the computational complexity and training time while improving the model accuracy.

Model Training: Densenet201: Once the feature extraction process is complete, the Densenet201 model is trained using the preprocessed training dataset. During this phase, the model learns to identify patterns and structures in the CT scan images that indicate whether a case is benign or malignant. Training involves optimizing the model's parameters using techniques like backpropagation and gradient descent. Hyperparameters such as learning rate, batch size, and the number of epochs is fine-tuned to achieve the best possible performance. The trained Densenet201 model serves as the first layer of classification in the hybrid system.

Feature Extraction for XGBoost: After the Densenet201 model is trained, intermediate features are extracted from its penultimate layer. These features represent a condensed form of the information learned by the deep learning model and are highly descriptive. These extracted features are then used as input for the XGBoost model. This step is crucial because it bridges the deep learning and machine learning components of the project, combining their strengths to achieve better performance.

Model Training: XGBoost: With the extracted features from Densenet201, the XGBoost classifier is trained to further classify the data into benign or malignant categories. XGBoost is a gradient-boosting algorithm that excels at structured data classification tasks. It refines the classification process by focusing on cases where the Densenet201 model might have uncertainties. Training the XGBoost model involves splitting the feature data into training and validation sets, optimizing its hyperparameters, and ensuring that it effectively learns patterns in the data. This additional layer of classification enhances the overall accuracy and reliability of the system.

Model Evaluation: Both the Densenet201 and XGBoost models are evaluated on the testing dataset to determine their effectiveness in detecting lung cancer. Evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrices are computed to analyze the models' performance. This step ensures that the models generalize well to unseen data and provides insights into any areas for improvement. The evaluation process confirms that the hybrid system is robust and capable of delivering reliable predictions in real-world scenarios.

Model Prediction: After evaluation, the trained models are used to make predictions on new CT scan images. For each input, the Densenet201 model processes the image and extracts features, which are passed to the XGBoost classifier for final classification. The models predict whether the scan represents a benign or malignant case. This dual-layered prediction process ensures accuracy and confidence in the results, with predictions being consistent and aligned with clinical requirements.

Final Output: The final output of the system provides a clear classification for each CT scan image. Benign cases are labeled and highlighted in green, indicating no presence of cancer, while malignant cases are labeled in red, indicating the presence of cancerous growths. This output can be integrated into clinical systems to assist medical professionals in making timely and accurate diagnoses. The system’s fast and reliable predictions can significantly improve early detection rates and patient outcomes.

METHODOLOGY

DATASET PREPARATION

The dataset preparation phase is crucial for the success of the project, as it lays the foundation for building accurate models. The dataset consists of CT scan images of lungs, sourced from publicly available datasets such as IQ-OTH/NCCD or proprietary databases. These images are annotated with labels indicating whether the cases are benign (non-cancerous) or malignant (cancerous).

Steps Involved in Dataset Preparation:

Data Collection: CT scan images are collected from reliable sources. The dataset includes a wide variety of images to ensure that the system can generalize across different patient profiles and medical conditions.

Metadata Collection: Along with the images, metadata such as patient demographics and diagnostic labels are gathered to aid in further analysis.

Labeling: Each image is labeled as either benign or malignant based on clinical annotations. This labeling is critical for supervised learning, as it allows the model to learn patterns associated with each class.

Data Cleaning: The dataset is cleaned by removing duplicate, irrelevant, or low-quality images (e.g., blurred or incorrectly labeled images). This step ensures the integrity of the dataset and prevents noise from affecting the model's performance.

Data Augmentation: To enhance the dataset's diversity and prevent overfitting, augmentation techniques are applied. These include rotation, flipping, scaling, and contrast adjustments to simulate variations in image acquisition conditions.

DATASET PREPROCESSING

Dataset preprocessing ensures that the CT scan images are in a format compatible with deep learning and machine learning models. Proper preprocessing enhances the efficiency and accuracy of the training process.

Steps Involved in Dataset Preprocessing:

Image Resizing: All images are resized to 224x224 pixels, which is the input size required by the Densenet201 model. Resizing ensures uniformity and reduces computational costs.

Normalization: Pixel values are normalized to a range of 0 to 1 by dividing each pixel by 255. This scaling helps the model converge faster during training by standardizing the input data.

Color Channel Handling: The CT scan images, typically grayscale, are converted into a format suitable for deep learning models, which often expect three-channel (RGB) inputs.

Splitting the Dataset: The dataset is divided into training, validation, and testing sets. A typical split might allocate 70% of the data for training, 15% for validation, and 15% for testing. This step ensures the models are trained and evaluated on distinct subsets of data.

Data Shuffling: Before splitting, the dataset is shuffled to ensure randomness, which prevents any biases in the training, validation, or testing sets.

DEEP AND MACHINE LEARNING AGORITHMS

DenseNet201: (Deep Learning Model)

DenseNet201 is a convolutional neural network (CNN) that uses dense connections to improve feature propagation and reduce the vanishing gradient problem. It consists of densely connected layers, where each layer receives input from all preceding layers, enhancing feature reuse and efficiency.

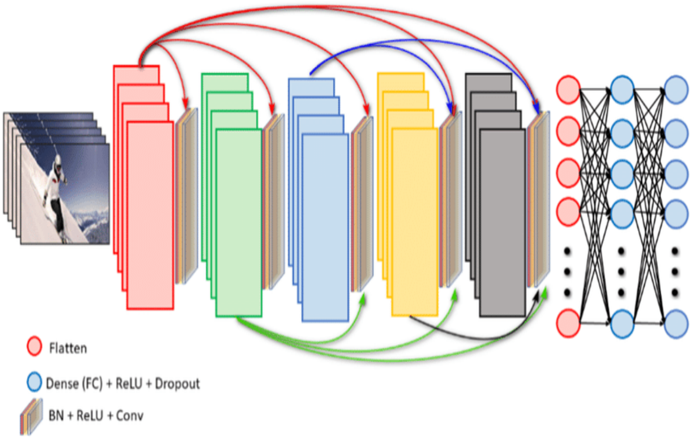


Fig.2. Architecture of DenseNet201

Working of DenseNet201:
Input Layer: CT scan images, resized to 224x224 pixels, are passed into the model.

Convolutional Layers: The image is processed through a series of convolutional layers, which detect low-level features like edges and textures.

Dense Blocks: Dense connections ensure that each layer has direct access to the gradients and features of all preceding layers, leading to improved learning efficiency.

Global Pooling: A global average pooling layer reduces the spatial dimensions of the feature map, making it computationally efficient. Output Layer: The output is either a probability distribution over the classes (benign or malignant) or extracted features for further classification.

Key Strengths of DenseNet201:

- Efficient feature propagation through dense connections.
- Reduced number of parameters compared to traditional CNNs.
- High accuracy in detecting patterns in medical images.

XGBOOST: (Machine Learning Algorithm)

XGBoost is a gradient-boosting algorithm that excels in structured data classification. It uses decision tree ensembles to predict the class labels based on the features extracted by Densenet201.

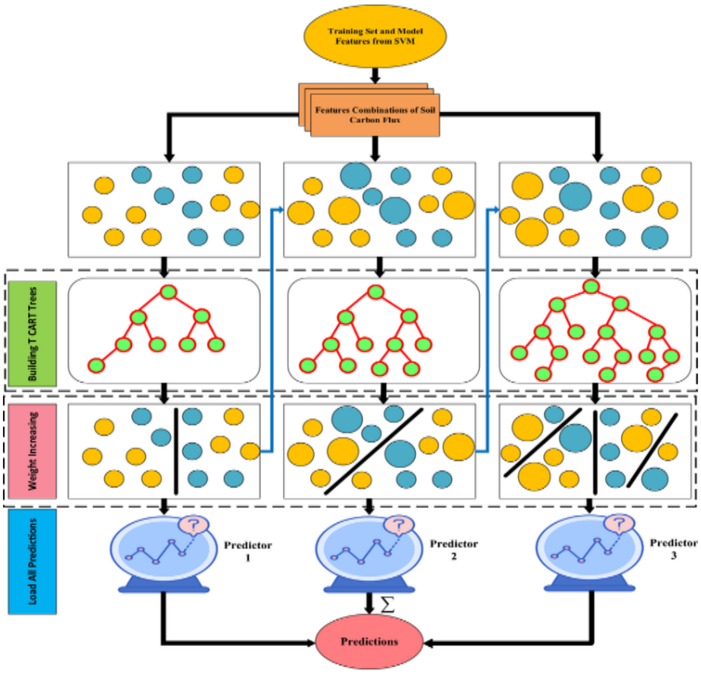


Fig.3. Architecture of XGBoost

Working of XGBoost:

Input Features: Features extracted from the penultimate layer of Densenet201 are fed into the XGBoost classifier.

Training: The classifier learns patterns in the data by building an ensemble of decision trees, each focusing on reducing the error of the previous tree.

Prediction: The trained XGBoost model predicts the likelihood of each image belonging to the benign or malignant class.

Key Strengths of XGBoost:

- Handles Structured data efficiently.
- Optimized for speed and accuracy
- Robust to overfitting due to regularization techniques.

MODEL EVALUATION

After training, both DenseNet201 and XGBoost are evaluated to ensure their performance meets the project objectives. Evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrices are computed. The hybrid system is expected to provide reliable predictions for lung cancer detection, with each component contributing uniquely to the final classification. This methodology ensures a robust pipeline for data processing, feature extraction, model training, and evaluation, ultimately leading to an efficient lung cancer detection system.

OUTCOME AND RESULTS

```
# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
print("Classification Report:\n", classification_report(y_test, y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
```

```
Accuracy: 0.9672727272727273
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.76      0.86        21
           1       0.97      1.00      0.98       150
           2       0.96      0.96      0.96       104

    accuracy                           0.97       275
   macro avg       0.98      0.91      0.94       275
weighted avg       0.97      0.97      0.97       275

Confusion Matrix:
 [[ 16   1   4]
 [  0 150   0]
 [  0   4 100]]
```

Fig.4. Code And output Screenshot of Evaluating XGBoost Model

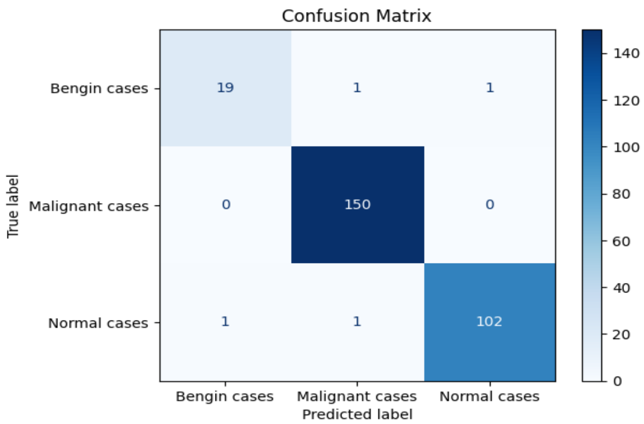


Fig.5. Output Screenshot for Confusion Matrix of XGBoost

```
9/9 [==============================] - 18s 890ms/step
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        29
           1       0.99      0.92      0.96       145
           2       0.71      0.99      0.83       101

    accuracy                           0.85       275
   macro avg       0.57      0.64      0.60       275
weighted avg       0.79      0.85      0.81       275

[[  0   0  29]
 [  0 134  11]
 [  0   1 100]]
```

Fig.6. Accuracy of DenseNet201

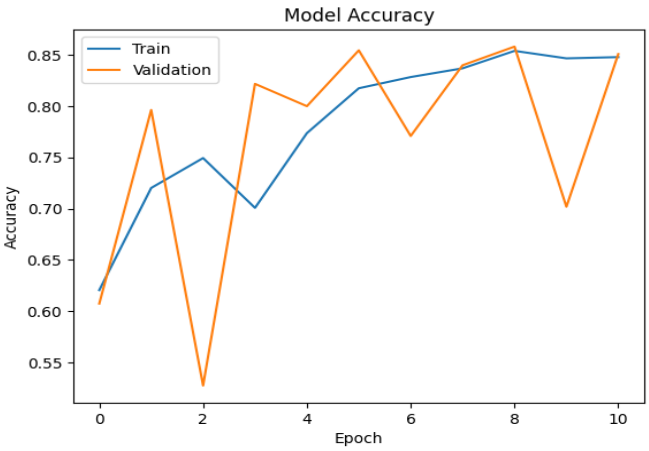


Fig.7. Output Accuracy graph of Densenet201

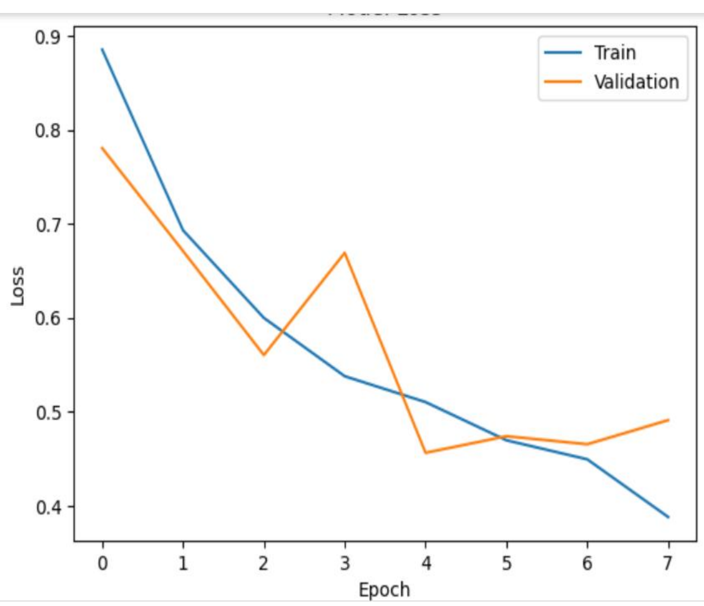


Fig. 8. Output for Loss graph of DenseNet201

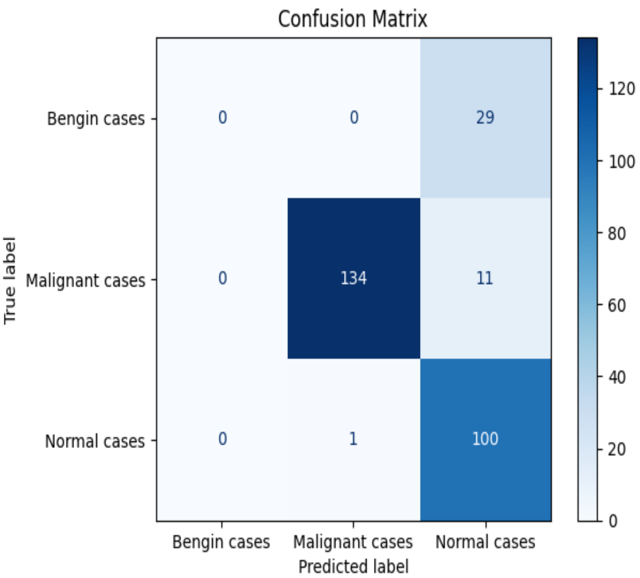


Fig.9. Output Screenshot of Confusion Matrix for DensenNet201

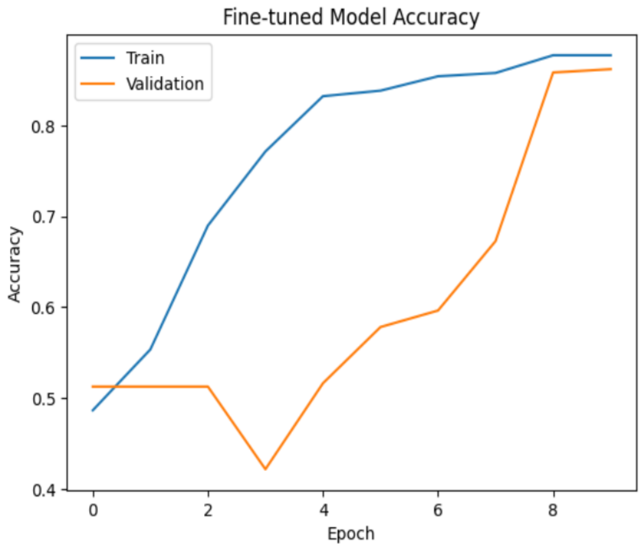


Fig. 9. Output Graph of accuracy after fine tuning of DensNet201

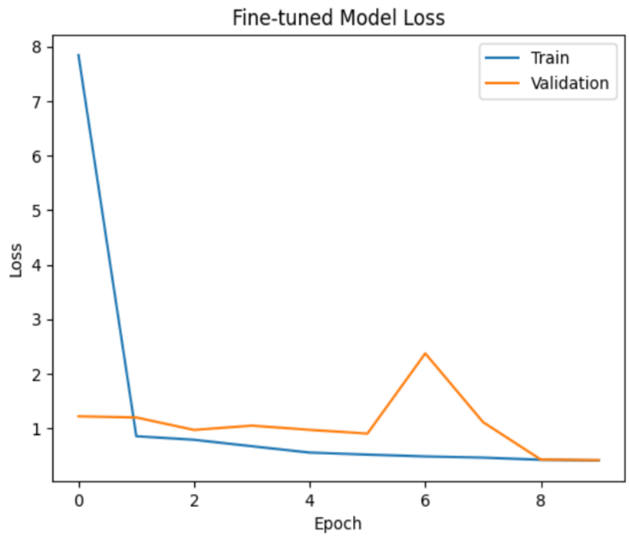


Fig. 10. Output Graph of loss after fine tuning of DensNet201

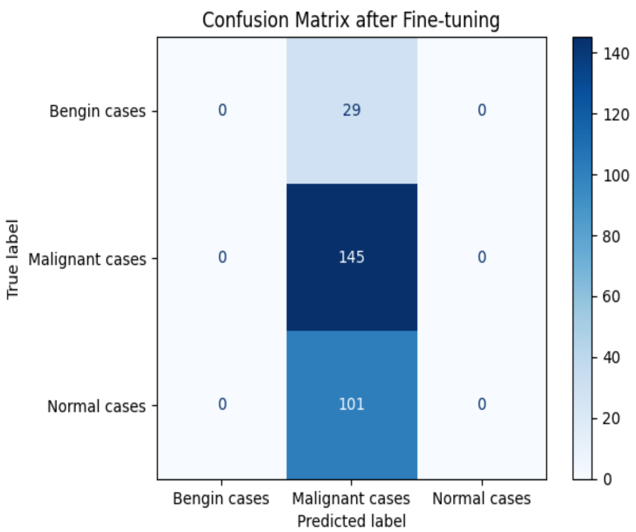


Fig. 11. Output Screenshot of Confusion Matrix Confusion Matrix for DenseNet201 after Fine-Tuning

CONCLUSION

In conclusion, this project successfully demonstrates the potential of transfer learning models, specifically XGBoost and DenseNet201, in the early detection of lung cancer. By leveraging the IQ-OTH/NCCD dataset, the fine-tuned models were able to achieve promising results in lung nodule classification, showcasing strong performance across key metrics such as sensitivity, specificity, and accuracy. The rigorous attention to ethical considerations, particularly in the areas of patient privacy and data security, underscores the feasibility of integrating such AI-driven approaches in clinical settings. This research not only contributes to the growing body of knowledge in medical diagnostics but also highlights the potential of artificial intelligence to enhance the efficiency and precision of lung cancer screening, offering a valuable tool in the fight against this critical health challenge.

The project demonstrates the effectiveness of transfer learning models like DenseNet201 and XGBoost for early lung cancer detection but identifies several areas for future enhancement to improve performance, scalability, and clinical applicability. One promising direction is the integration of multi-modal data, such as combining imaging data (e.g., CT scans, MRIs) with non-imaging data like genetic profiles, medical histories, and laboratory results. This approach can provide a more comprehensive understanding of a patient's health, enhancing diagnostic accuracy and enabling personalized lung cancer detection and treatment.

Improving model interpretability is another critical area for future research, as deep learning models are often perceived as "black boxes." Explainable AI (XAI) techniques, such as Grad-CAM, can be employed to visualize which features or regions in an image contribute most to predictions, making the model's decision-making process more transparent. Providing clinicians with clear justifications for AI-driven results would increase trust and adoption in clinical workflows. Enhanced interpretability ensures that healthcare professionals can verify predictions before making critical decisions.

Further optimization of model architectures could lead to significant improvements in efficiency, speed, and scalability. Lightweight architectures like EfficientNet or simplified DenseNet variants could reduce computational costs, making AI solutions accessible in resource-constrained environments. Techniques like model pruning, quantization, and knowledge distillation can also help compress model size without sacrificing performance, enabling deployment in real-time clinical settings.

Collaborative diagnostic systems, where AI works alongside human radiologists, represent a practical application for enhancing diagnostic accuracy. AI can act as a second opinion, flagging suspicious nodules for further review, thereby reducing false negatives. Real-time feedback systems could also support radiologists during live screenings, helping to improve efficiency in high-stress or high-volume environments.

Validating these models in real-world clinical settings is essential for ensuring generalizability and reliability. Large-scale clinical trials across diverse populations and healthcare systems can help confirm the models' effectiveness in varied contexts. Additionally, post-deployment monitoring can ensure continued accuracy and adaptability as new data is introduced, maintaining the model's relevance over time.

Finally, future work could focus on customizing AI models to detect specific subtypes of lung cancer, such as small cell and non-small cell lung cancer. Subtype-specific or adaptive models could provide more precise and clinically relevant information, evolving with advancements in medical understanding of the disease. These improvements would ensure that AI tools remain at the forefront of lung cancer detection and treatment.

REFERENCE

[1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning." 521(7553), 436-444. **https://www.nature.com/articles/nature14539**, **https://sci-hub.st/10.1038/nature14539**

[2] Pan, S. J., & Yang, Q. (2010). "A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering," 22(10), 1345-1359. **https://ieeexplore.ieee.org/document/5288526**

[3] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). "Convolutional neural networks for medical image analysis: Full training or fine tuning?" IEEE Transactions on Medical Imaging, 35(5), 1299-1312. **https://ieeexplore.ieee.org/document/7436640**

[4] Chen, T., & Guestrin, C. (2016). "XGBoost: A scalable tree boosting system." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. **https://dl.acm.org/doi/10.1145/2939672.2939785**

[5] Hussein, S., Gillies, R., Cao, K., Song, Q., & Bagci, U. (2017). "TumorNet: Lung nodule characterization using multi-view convolutional neural networks." IEEE International Symposium on Biomedical Imaging (ISBI), 1007-1010. **https://ieeexplore.ieee.org/document/7950670**

[6] Shen, W., Zhou, M., Yang, F., Yang, C., & Tian, J. (2015). "Multi-scale convolutional neural networks for lung nodule classification." Information Processing in Medical Imaging, 588-599. **https://link.springer.com/chapter/10.1007/978-3-319-19992-4_46**

[7] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). "The ethics of algorithms: Mapping the debate." Big Data & Society, 3(2), 2053951716679679. **https://journals.sagepub.com/doi/10.1177/2053951716679679**

[8] Char, D. S., Shah, N. H., & Magnus, D. (2018). "Implementing machine learning in health care Addressing ethical challenges." The New England Journal of Medicine, 378(11), 981-983. **https://www.nejm.org/doi/full/10.1056/NEJMp1714229**