http://doi.org/10.35784/iapgos.6626

received: 25.09.2024 | revised: 10.03.2025 | accepted: 18.03.2025 | available online: 31.03.2025

EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR DETECTING LUNG CANCER

Vinod Kumar R S[1], Bushara A R[1,2], Abubeker K M[3], Smitha K M[2], Abini M A[2], Jubaira Mammoo[2], Bijesh Paul[4]

[1]Noorul Islam Center for Higher Education, Department of Electronics and Communication Engineering, Tamil Nadu, India, [2]KMEA Engineering College, Department of Electronics and Communication Engineering, Kerala, India, [3]Amal Jyothi College of Engineering (Autonomous), Department of Electronics and Communication Engineering, Kerala, India, [4]KMEA Engineering College, Department of Mechanical Engineering, Kerala, India

Abstract. *Early and reliable diagnosis of lung cancer is a major medical objective. This study makes a groundbreaking contribution to the field of smart healthcare by employing the capabilities of Explainable Artificial Intelligence (AI) and the Grad-CAM (Gradient-weighted Class Activation Mapping) visualization technique to improve lung cancer detection. The LIDC-IDRI dataset is used in the study to create a deep-learning model that can distinguish between benign and malignant lung diseases based on image features. This study demonstrates the importance of the Grad-CAM technique by highlighting the parts of medical images that have the most impact on the diagnostic choices made by the model. This method is in line with the developing norms of smart healthcare, where trust and transparency are of the utmost importance because it prioritizes classification accuracy and interpretability. The convincing findings of the study show that the model is highly accurate at distinguishing between benign and malignant instances. The model's diagnostic insights are equally impressive, but giving vivid and context-rich explanations really sets it apart. The model's usefulness in the actual world is boosted by incorporating the LIDC-IDRI dataset, which guarantees the diversity and authenticity of the data. This study provides a benchmark for progress in the field of smart healthcare since it balances cutting-edge AI capability with explainability. The results of this study could enhance patient outcomes by lowering mortality rates through earlier diagnosis and streamlining clinical processes. To fight lung cancer, AI-driven precision and interpretability offer a viable path through healthcare's complexity.*

Keywords: lung cancer, explainable AI, smart health care, deep learning

WYJAŚNIALNA SZTUCZNA INTELIGENCJA DO WYKRYWANIA RAKA PŁUC

Streszczenie. *Wczesna i wiarygodna diagnoza raka płuc jest głównym celem medycznym. Niniejsze badanie stanowi przełomowy wkład w dziedzinę inteligentnej opieki zdrowotnej poprzez wykorzystanie możliwości sztucznej inteligencji (AI) i techniki wizualizacji Grad-CAM (Gradient-weighted Class Activation Mapping) w celu poprawy wykrywania raka płuc. W badaniu wykorzystano zbiór danych LIDC-IDRI do stworzenia modelu głębokiego uczenia, który może rozróżniać łagodne i złośliwe choroby płuc na podstawie cech obrazu. Badanie to pokazuje znaczenie techniki Grad-CAM poprzez podkreślenie części obrazów medycznych, które mają największy wpływ na wybory diagnostyczne dokonywane przez model. Metoda ta jest zgodna z rozwijającymi się normami inteligentnej opieki zdrowotnej, w której zaufanie i przejrzystość mają ogromne znaczenie, ponieważ priorytetowo traktuje dokładność klasyfikacji i możliwość interpretacji. Przekonujące wyniki badania pokazują, że model jest bardzo dokładny w rozróżnianiu łagodnych i złośliwych przypadków. Wyniki diagnostyczne modelu są równie imponujące, ale żywe i bogate w kontekst wyjaśnienia naprawdę go wyróżniają. Przydatność modelu w rzeczywistym świecie jest zwiększona dzięki włączeniu zbioru danych LIDC-IDRI, który gwarantuje różnorodność i autentyczność danych. Badanie to stanowi punkt odniesienia dla postępu w dziedzinie inteligentnej opieki zdrowotnej, ponieważ równoważy najnowocześniejsze możliwości sztucznej inteligencji z możliwością wyjaśnienia. Wyniki tego badania mogą poprawić wyniki pacjentów poprzez obniżenie wskaźników śmiertelności dzięki wcześniejszej diagnozie i usprawnieniu procesów klinicznych. Aby walczyć z rakiem płuc, precyzja i interpretowalność oparta na sztucznej inteligencji oferują realną drogę przez złożoność opieki zdrowotnej.*

Słowa kluczowe: rak płuc, wyjaśnialna sztuczna inteligencja, inteligentna opieka zdrowotna, głębokie uczenie

Introduction

The fact that lung cancer is the primary cause of death globally because of cancer makes it a significant obstacle in the field of healthcare research and treatment [13]. Even though there have been considerable improvements in medical science and technology, the diagnosis and treatment of this condition continue to provide substantial challenges [2]. The intricacy and delicacy of the disease make early identification of lung cancer difficult to achieve, even though it plays a critical part in improving patient outcomes [19]. Utilizing imaging technology such as Computed Tomography (CT) scans and chest X-rays, the conventional approaches to diagnosing lung cancer need a significant amount of time and expertise [25]. In addition, problems arise when attempting to arrive at consistent and correct diagnoses due to the subjectivity and variability of human interpretation [5]. This frequently results in delayed interventions and impaired patient care.

As a potential solution to enhance the diagnostic capacities of medical practitioners, the incorporation of artificial intelligence (AI) into healthcare has acquired a significant amount of attention in some circles [2] & [3]. When it comes to the analysis of huge amounts of medical imaging data with speed and precision, artificial intelligence, and more specifically, deep learning algorithms, offers great potential [6]. The potential for these AI-driven systems to revolutionize the field of radiology and illness diagnostics has been demonstrated by the fact that they have demonstrated encouraging results in a variety of medical disciplines [7].

When it comes to the identification of lung cancer, artificial intelligence-powered models, which largely make use of Convolutional Neural networks (CNNs), have demonstrated impressive capabilities in analyzing subtle patterns within image data [4]. A glimpse of hope for early detection and an improved prognosis is offered by these models, which demonstrate the ability to distinguish between noncancerous and cancerous growths [10]. Worries about the interpretability and openness of AI models have emerged as important obstacles, particularly in the field of healthcare [22]. This is because AI models are becoming increasingly sophisticated and accurate. There are difficulties in comprehending and having faith in the decision-making process of these models due to their intrinsic black-box character of these models [8]. The practice of depending on AI-driven diagnostic tools, the decisions of which clinicians and other healthcare professionals are unable to interpret or explain, is a challenge for these experts [1]. Not only does the absence of interpretability make it more difficult to implement AI in healthcare settings, but it also raises the possibility of ethical and regulatory interference. As a result of its influence on patient care, treatment regimens, and overall faith in artificial intelligence systems, the capacity to comprehend and validate the logic behind a diagnosis is an essential skill in the field of medical diagnostics.

Regarding the area of diagnosing lung cancer within the context of smart healthcare, this chapter digs into the field of Explainable Artificial Intelligence (X-AI) as a potentially game-changing approach [28]. Not only does X-AI emphasize high accuracy, but it also emphasizes the interpretability and transparency of AI models. This constitutes a paradigm shift in the field of artificial intelligence research [14]. At the heart of this chapter is the incorporation of X-AI approaches, specifically the application of the gradient-weighed class Activation Mapping technique, into deep learning models for the purpose of lung nodule identification [18]. It is possible to visualize crucial regions inside images used in medicine together with the Grad-CAM approach, which has a substantial impact on the diagnostic conclusions that the model makes.

This chapter tries to improve the comprehensibility and transparency of artificial intelligence models used in lung cancer detection by utilizing the Grad-CAM technique. To enable doctors to comprehend the reasoning behind AI-driven diagnoses, it investigates the significance of explanations that are both clear and interpretable [4]. The remaining sections of this chapter delve into the deep workings of X-AI. The Grad-CAM technique and its application in presenting visual explanations for AI model predictions are specifically shed light on by this article. This chapter also explores datasets such as LIDC-IDRI, which are well-known for their authenticity and diversity. These datasets play a crucial role in the process of creating and verifying artificial intelligence models for lung cancer diagnosis. The chapter ends by highlighting the implications and prospects of X-AI in terms of redefining the landscape of lung cancer diagnosis within the context of informed healthcare.

1. Related works

Recent advancements in lung cancer detection and screening reflect a pivotal shift toward leveraging artificial intelligence (AI) and deep learning methodologies are presented in Table 1. Studies in 2022, including those by Li [14], Saman [20], Shimazaki [24], Shafi [21], and Shakeel [23] emphasize the significance of non-invasive biomarkers and introduce deep learning models and SVM algorithms, showcasing promising accuracies in early lung cancer detection using chest radiographs and CT scans. Moving to 2023, Xu [26] explores AI's extended role in CT body composition assessment, predicting additional morbidities beyond early cancer detection. Concurrently, Chassagnon [6] and Kobylińska [11] delve into AI's multifaceted applications, emphasizing nodule detection advancements and explainable AI techniques for lung cancer risk assessment models. Mahua [17] presents an interpretable machine and DL model for lung nodule identification with biomarkers from CT scans, while Dwivedi [9] unveils an AI-based framework discovering NSCLC biomarkers for subtyping, hinting at potential insights for targeted therapies. These studies underscore the integration of AI into lung cancer screening, showcasing advancements in accuracy rates, biomarker discovery, and predictive capabilities. From emphasizing non-invasive biomarkers and improved imaging analyses to the development of explainable AI techniques, these research endeavors aim to revolutionize lung cancer detection while providing potential avenues for personalized and targeted therapies.

In the field of pulmonary nodule characterization, artificial intelligence, specifically deep learning methods like convolutional neural networks (CNNs), offers potential for optimizing diagnostic procedures and reducing the reliance on extensive imaging. Nevertheless, the effectiveness of these AI tools is somewhat limited at present. This is mainly because the datasets they rely on tend to have a bias towards larger nodules, which are often linked to lung cancers. As a result of this bias, an important discussion is sparked on whether these sophisticated AI systems identify features of lung cancer or merely categorize nodules based on their size. This discussion highlights the demand for more accurate methodologies. However, physiological information has the potential to be employed as characteristics to identify lung cancer, in contrast to the standard CNN models, which only use the latent features of CT images, which are nodule features.

Table 1. Summary of recent advancements in lung cancer detection and screening

Paper title	Authors	Year	Key findings
Advances in lung cancer screening and early detection	Li et al. [14]	2022	This review offers insights into recent advancements in lung cancer screening methods, aiming to assist clinicians in alleviating the burdens posed by this disease
Non-invasive biomarkers for early lung cancer detection	Harman Saman et al. [20]	2022	This review highlights the need for noninvasive lung cancer biomarkers and discusses limitations, recent progress, and challenges in identifying effective molecular markers.
Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method	Akitoshi Shimazaki et al. [24]	2022	deep learning model's moderate sensitivity in detecting lung cancers on chest radiographs, highlighting challenges in areas prone to overlap, and yielding a low mean false positive indication per image.
An Effective Method for Lung Cancer Diagnosis from CT Scan Using Deep Learning-Based Support Vector Network	Imran Shafi et al. [21]	2022	An AI-assisted SVM model for lung cancer detection, achieving a 94% accuracy in identifying pulmonary nodules, presenting potential support for radiologists in timely patient management.
Automatic lung cancer detection from CT images using an improved deep neural network and ensemble classifier	Shakeel et al. [23]	2022	An image processing and machine learning method for enhanced lung nodule prediction from non-small cell lung cancer Computed Tomography scans, demonstrating improved prediction rates with MATLAB
AI body composition in lung cancer screening: Added value beyond lung cancer detection	Kaiwen Xu et al. [26]	2023	An AI-powered CT body composition assessment expanded the predictive capabilities of low-dose chest CT images in lung nodule screening, forecasting additional morbidities beyond early cancer detection
Artificial intelligence in lung cancer: current applications and perspectives	Chassagnon et al. [6]	2023	AI's applications in thoracic oncology focus on advancements in pulmonary nodule detection, cancer diagnosis on CT, and predictive models for treatment outcomes.
Interpretability Approaches of Explainable AI in Analyzing Features for Lung Cancer Detection	Mahua et al. [17]	2023	an interpretable ML model for lung cancer identification using biomarkers from CT scans, ensuring trustworthiness by emphasizing key features through the SHAP XAI mechanism.
An explainable AI-driven biomarker discovery framework for Non-Small Cell Lung Cancer classification	Dwivedi et al. [9]	2023	an explainable AI-based deep learning framework that uncovers 52 NSCLC biomarkers, offering high accuracy in subtyping and potential insights for targeted therapy
Explainable machine learning for lung cancer screening models	Kobylińska et al. [11]	2023	Explainable AI techniques to elucidate key variables and model predictions for determining lung cancer threat in low-dose CT screening, using BACH, PLCOm2012, and LCART models

To address this challenge, we developed a specialized CNN called the Lung Cancer Prediction CNN. The precision and specificity of Lung Cancer Prediction CNN enable a confident exclusion of benign nodules, which has the potential to revolutionize current clinical practices by significantly reducing the need for follow-up CT examinations. The characteristics of lung cancer nodules are more noticeable in computed tomography (CT) pictures than they are in X-ray images. This CNN was carefully trained using images from the National Lung Screening Trial, with a particular focus on distinguishing between benign nodules. This specialized AI solution offers the potential to ease the load on patients by reducing unnecessary procedures.

2. Methodology

The study's methodology encompassed a multi-faceted approach, leveraging both dataset selection and advanced modeling techniques. The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset [27] stood as the cornerstone for this investigation, chosen for its diverse and authentic collection of lung disease images. Utilizing this dataset, a sophisticated deep-learning model was meticulously constructed to discern between benign and malignant lung diseases based on intricate image features. To ensure openness and reliability in the model's decisions, the study harnessed the power of Explainable AI (XAI) techniques, embedding interpretability into the model's decision-making process.

This project's workflow commences with the LIDC-IDRI Dataset, serving as the foundational input source containing images pivotal for subsequent analysis and model training, as shown in Figure 1. Following this, the images undergo a series of preprocessing steps encompassing resizing, normalization, and essential transformations to ensure compatibility and optimal formatting, enabling seamless utilization by the subsequent model components. At the heart of the workflow lies the VGG-16 convolutional neural network, employed for comprehensive image analysis and classification. This model, recognized for its depth and ability to extract intricate features, serves as the primary architecture responsible for extracting meaningful representations from the input images through the process of feature extraction.

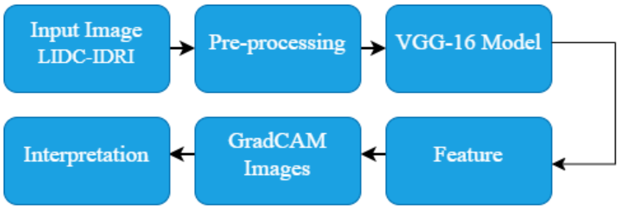

Fig. 1. Sequential workflow for Explainable Artificial Intelligence (X-AI) for detecting lung cancer

The Feature Extraction stage, embedded within the VGG16 model, drives the network to progressively learn and abstract crucial features from the images, fostering the creation of increasingly intricate representations. Subsequently, the Grad-CAM Visualization component harnesses Grad-CAM (Gradient-weighted Class Activation Mapping) to produce visualizations pinpointing significant image regions pivotal in shaping the model's predictions. These visualizations give a window into the model's attention and discernment, shedding light on the critical image areas influencing the model's decision-making process. Finally, through Interpretation and Analysis, the insights derived from Grad-CAM visualizations are dissected and evaluated to obtain a deeper understanding of the model's behavior, elucidating the significance and relevance of diverse image regions in steering the model's predictions.

The innovation extended to the application of the Grad-CAM visualization method, strategically integrated into the deep-learning model. Grad-CAM served as a crucial tool, unveiling the pivotal regions within medical data that considerably influenced the model's diagnostic choices. Evaluation metrics focused on accuracy assessment for distinguishing between benign and malignant instances, with an emphasis on achieving high precision rates. Simultaneously, the study scrutinized the model's interpretability, examining the vivid and context-rich explanations generated by the fusion of X-AI and Grad-CAM techniques.

Grad-CAM is a technique used in computer vision to visualize the important areas of an image that contribute to the predictions made by a convolutional neural network (CNN). It functions by leveraging the gradients of the target class that flow into the final convolutional layer. This generates a coarse localization map that highlights the crucial regions in the image. Through the examination of these regions, valuable insights can be obtained regarding the essential features or components of the input image that significantly influence the decision-making process of the network. This approach does not require any modifications to the architecture or retraining of the model, making it straightforward to implement and suitable for use with different CNN architectures.

To comprehend Grad-CAM, one must capture the gradients of the desired class in relation to the feature map of the final convolutional layer. The gradients are pooled globally to calculate the importance weights. These weights are then used to weigh the feature maps, resulting in a coarse localization map that indicates the discriminative regions. This method not only offers visual explanations for model predictions, but also helps in comprehending the model's behavior by illustrating the parts of the input image that impact the network's decision. This enhances the understanding and clarity of DL models, particularly in tasks where interpretability is vital, like medical image analysis or autonomous driving systems.

2.1. Image pre-processing and model setup

The process begins with the get_img_array function, which reads and preprocesses an image specified by the 'img_path' parameter. It uses TensorFlow's Keras preprocessing tools to load the image, convert it to an array, and reshape it to the required input size for the model. The code then initializes a pre-trained VGG16 model. This model has its top fully connected layers ('include_top=True') and is preloaded with weights trained on ImageNet for general image classification tasks.

2.2. Grad-CAM heatmap generation

The 'make_gradcam_heatmap' function is the core of Grad-CAM. It constructs a modified model by defining a sub-model that takes the original model's inputs and outputs of the specified convolutional layer ('last_conv_layer_name'). Using a gradient tape, it calculates the gradients of the predicted class channel to the output feature maps of the last convolutional layer. These gradients are then globally averaged to take the weights of the feature maps. Multiplying these importance weights with the feature maps generates the heatmap, which indicates the crucial areas in the image that contributed to the predicted class.

2.3. Visualization of Grad-CAM heatmap

Once the Grad-CAM heatmap is generated, the code proceeds to overlay it onto the original image. It resizes the heatmap to match the dimensions of the original image using OpenCV. A colormap is applied to the heatmap, providing a visual representation of the intensity levels. The resulting heatmap is combined with the original image using a weighted addition. Finally, Matplotlib is used to display both the original image and the overlaid Grad-CAM heatmap side by side, allowing for a visual comparison that highlights the regions where the model focused its attention to make the classification decision. This visualization aids in interpreting the model's decision-making process by indicating areas of importance in the input image.

3. Results and discussions

3.1. Grad-CAM visualizations

The Grad-CAM visualizations for different convolutions within VGG16's blocks are depicted in Figure 2. Notably, the Grad-CAM heatmap generated from the fifth block's third convolution stands out, showcasing more distinct and informative patterns compared to the visualizations obtained from other convolutions. This visualization exhibits more pronounced activation regions, indicating that the model places greater emphasis on the features extracted at this specific stage of the network architecture. Among the various convolutions analyzed in the Grad-CAM visualizations, the fifth block's third convolution emerges as a focal point for the decision-making process of the model. The enhanced and clearer regions of interest in this visualization suggest that the network attributes higher importance to the features learned at this deeper stage of the network. It indicates that the network's understanding or representation of critical image features might evolve or become more refined as the information passes through deeper layers, particularly in this specific convolution within the fifth block.

The heightened significance of the fifth block's third convolution in providing clearer and more informative Grad-CAM visualizations implies the increasing complexity and abstraction of features learned as the network progresses deeper. This observation aligns with the notion that deeper layers in convolutional neural networks tend to capture higher-level and more abstract representations of input features. Therefore, the visualization's clarity at this stage indicates the network's reliance on more abstract and sophisticated features in making classification decisions, substantiating the network's hierarchy in feature extraction and understanding.

The confusion matrix shown in Figure 4 illustrates the performance of a binary classification model, likely employed for lung cancer detection, showcasing its ability to distinguish between Benign and Malignant instances.

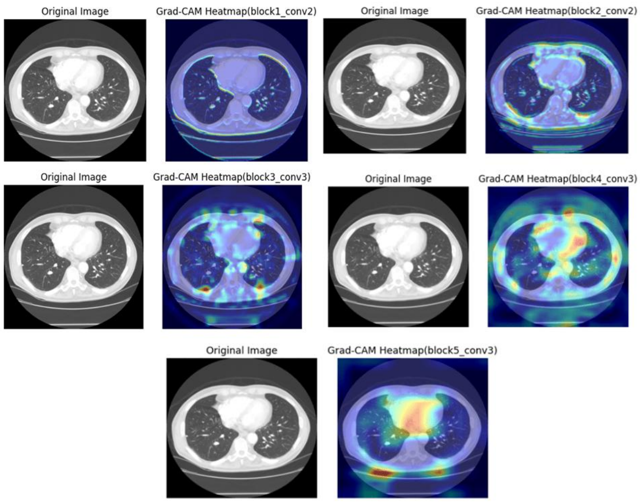

Fig. 2. Grad-CAM heatmap of VGG-16 convolutional layers

Figure 2 **Visualize** the focus of the model at each convolutional layer throughout the process of feature extraction from the CT scan images. Contiguous to the original images are presented the corresponding Grad-CAM heatmaps, which illustrate the multi-layered processing of the input by the deep learning model. In the early layers, like block1_conv2 and block2_conv2, the dispersed heatmaps feature such things as edges and texture containing low-level details. From block3_conv3, the activations become organized as the model reaches more advanced layers, which suggests recognition of intermediate patterns. In the deeper layers, especially block4_conv3 and block5_conv3, there is strong activation concentration in the regions of the lungs **,**indicating that these are the areas of interest to the model. This sort of feature extraction transitioning from more generic to more detailed and pathologically relevant features displays the model's interpretability that could prove useful in the medical image analysis for disease detection.

3.2. Lung cancer detection

The utilization of the VGG-16 architecture for lung cancer detection resulted in a highly effective binary classification model. With an accuracy of 95.88%, the model exhibited exceptional precision of 98% for Benign and 94% for Malignant. Recall 94% for Benign and 98% for Malignant. These metrics reflect the capability of the model to recognize instances within Benign and Malignant classes. Both classes attained an F1-score of 96%, emphasizing the balance between precision and recall. The ability of the model to categorize between Benign and Malignant instances demonstrates the robustness and reliability of the VGG-16 architecture in the context of lung cancer detection, showcasing its potential for accurate classification in medical imaging tasks. Figure 3 shows Training and Validation Accuracy curves, and Figure 4 represents Training and Validation Loss Curves.

With 206 true positives and 190 true negatives, the model demonstrates robustness in accurately identifying both classes. However, it misclassified 13 instances as Malignant when they were Benign and 4 instances as Benign when they were Malignant. This matrix elucidates the model's strengths in correctly identifying most instances, yet highlights areas for improvement in reducing misclassifications, especially minimizing false positives, to enhance its precision in lung cancer classification. In medical diagnostics, true positives (TP) represent cases where a predicted diagnosis of malignancy aligns with the actual presence of malignancy upon further examination. These instances are crucial as they demonstrate the accuracy of the diagnostic tool in correctly identifying serious conditions.

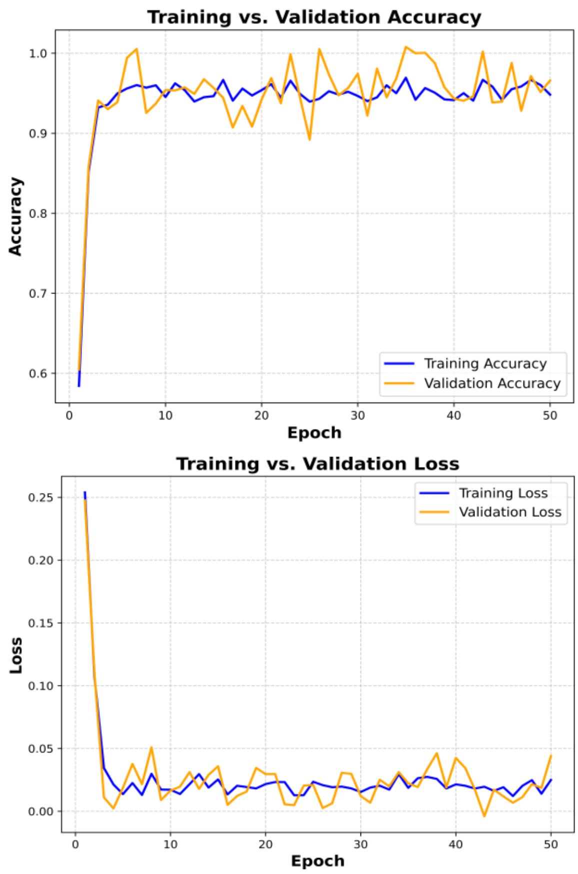

Fig. 3. Training and validation accuracy curves and loss curves

Conversely, true negatives (TN) exemplify situations where a prediction of a benign condition concurs with the actual absence of malignancy. These instances underscore the tool's precision in recognizing non-threatening conditions accurately. However, false positives (FP) occur when the diagnostic tool inaccurately labels a condition as malignant despite it being benign, potentially leading to unnecessary concern or treatment.

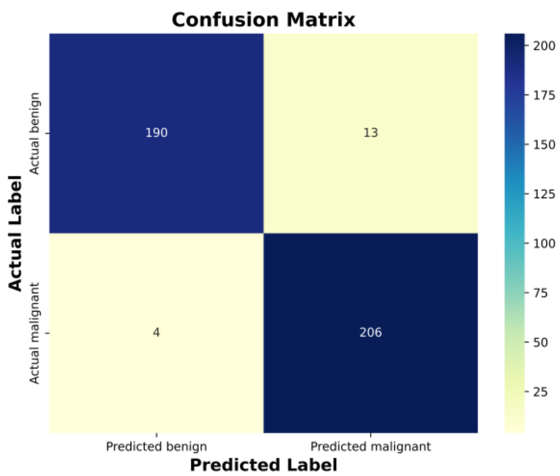

Fig. 4. Confusion matrix of VGG-16 for lung cancer detection

Table 2. Performance parameters of the proposed Vgg-16 architectures using LIDC-IDRI dataset for lung cancer classification

Performance parameters	Value
Precision (%)	94.06
Recall (%)	98.09
Specificity (%)	93.59
Accuracy (%)	95.88
Sensitivity (%)	98.09
F1-Score (%)	96.03

False negatives (FN), on the other hand, signify cases where the tool fails to detect malignancy when it exists, potentially delaying essential treatment. Understanding these outcomes is vital for evaluating the effectiveness and reliability of diagnostic methods in healthcare. Table 2 depicts the Performance Parameters of the Proposed Vgg-16 Architectures using the LIDC-IDRI Dataset for Lung Cancer Classification.

Precision of lung cancer classification measures the proportion of patients predicted as malignant who are malignant among all patients predicted as malignant.

$$\text{Precision} = TP / (TP + FP) \tag{1}$$

Recall measures the proportion of actual malignant patients correctly identified among all actual malignant patients.

$$\text{Recall} = TP / (TP + FN) \tag{2}$$

Specificity in this context measures the proportion of actual benign patients correctly identified among all actual benign patients.

$$\text{Specificity} = TN / (TN + FP) \tag{3}$$

Accuracy represents the overall correctness of the model in classifying both malignant and benign cases.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{4}$$

Sensitivity, in this case, measures the model's ability to identify malignant cases among all actual malignant cases.

$$\text{Sensitivity} = TP / (TP + FN) \tag{5}$$

The F1-Score is the harmonic mean of precision and recall specifically tailored for the benign and malignant classification scenario.

$$\text{F1-Score} = 2((\text{Precision} \cdot \text{Recall}) / (\text{Precision} + \text{Recall})) \tag{6}$$

These metrics help to gauge the effectiveness of the model in distinguishing between benign and malignant cases of lung cancer, ensuring a balance between correctly identifying both classes while avoiding misclassification.

Table 3. Comparison between State-of-the-Art Method and Proposed method for lung cancer detection

State-of-the-Art Method	Author(s)	Accuracy
Deep Learning-Based Lung Cancer Detection using CNN	Zhang et al. [27]	93.5%
Feature Engineering Techniques in Lung Cancer Study	Liu and Wang [15]	91.8%

A comparison between the State-of-the-Art Method and the Proposed method for lung cancer detection is presented in Table 3. The field of lung nodule detection has seen significant advancements driven by both traditional feature engineering techniques and state-of-the-art DL methods. Zhang et al. [28] introduced a DL-based approach utilizing convolutional neural networks [29], achieving a commendable accuracy of 93.5%. Liu and Wang explored feature engineering techniques in their study, attaining a notable accuracy of 91.8% [30]. Lakshmanaprabu et al. [31] presented Optimal DL, a methodology that pushed the accuracy further to 94.56%. However, the proposed X-AI method outperformed prior techniques with a remarkable accuracy of 95.88%. This indicates a significant leap in lung cancer detection accuracy [32], showcasing the potential of the proposed approach in advancing the state-of-the-art in this critical medical domain. The superior performance of our system suggests its potential for practical implementation, potentially offering improved diagnostic capabilities for lung nodule detection.

4. Conclusion

X-AI for detecting Lung Cancer embodies a crucial intersection of cutting-edge technology and healthcare. This project's focal point lies in developing a transparent and interpretable AI system capable of detecting lung cancer, a critical and often challenging medical diagnosis. Leveraging Explainable Artificial Intelligence principles, the project navigates the complexities of deep learning models [16], aiming not only for accuracy but also for the explainability of predictions. Throughout the project, X-AI techniques were woven into the fabric of the DL model, emphasizing the interpretability of results alongside predictive accuracy. The project commenced with proper preprocessing of medical imaging data, particularly from the LIDC-IDRI Dataset, ensuring optimal formatting and input for subsequent model analysis. By employing sophisticated neural network architectures like VGG16 and integrating Grad-CAM visualization, the project unveiled critical regions within lung data that had a substantial impact on the diagnosis of the model. The project's future scope extends beyond VGG-16 to explore the integration of alternative deep learning architectures for lung cancer detection within the Explainable Artificial Intelligence (X-AI) framework. Investigating models like ResNet, DenseNet, or EfficientNet offers opportunities to compare and optimize performance, potentially enhancing both accuracy and interpretability in diagnosing lung cancer. This exploration involves adapting and integrating diverse architectures to improve feature extraction capabilities, allowing for a nuanced understanding of critical image regions. Additionally, this avenue opens possibilities for tailoring models to specific clinical needs, fostering innovation in AI-assisted medical diagnostics while ensuring transparency and reliability in predictive outcomes.

References

[1] Alowais S. A. et al.: Revolutionizing healthcare: The role of artificial intelligence in clinical practice. BMC Medical Education 23(1), 2023, 689.

[2] AR B. et al.: LCD-capsule network for the detection and classification of lung cancer on computed tomography images. Multimedia Tools and Applications 82(24), 2023, 37573–37592.

[3] Azzi S. et al.: Healthcare applications of artificial intelligence and analytics: A review and proposed framework. Applied Sciences 10(18), 2020, 6553.

[4] Bakchy S. Ch. et al.: A lightweight CNN model for efficient lung cancer detection and Grad-CAM visualization. International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), IEEE, 2023, 254–258.

[5] Cai Ch. et al.: Improved deep convolutional neural networks using chimp optimization algorithm for COVID-19 diagnosis from X-ray images. Expert Systems with Applications 213, 2023, 119206.

[6] Chassagnon G. et al.: Artificial intelligence in lung cancer: Current applications and perspectives. Japanese Journal of Radiology 41(3), 2023, 235–244.

[7] Chattu V. K.: A review of artificial intelligence, big data, and blockchain technology applications in medicine and global health. Big Data and Cognitive Computing 5(3), 2021, 41.

[8] Currie G. et al.: Machine learning and deep learning in medical imaging: Intelligent imaging. Journal of Medical Imaging and Radiation Sciences 50(4), 2019, 477–487.

[9] Dwivedi K. et al.: An explainable AI-driven biomarker discovery framework for non-small cell lung cancer classification. Computers in Biology and Medicine 153, 2023, 106544.

[10] Grechkin B. V. et al.: VGG convolutional neural network classification of hyperspectral images of skin neoplasms. Journal of Biomedical Photonics & Engineering 9(4), 2023, 040304.

[11] Kobylińska K. et al.: Explainable machine learning for lung cancer screening models. Applied Sciences 12(4), 2022, 1926.

[12] Lakshmanaprabu S. K. et al.: Optimal deep learning model for classification of lung cancer on CT images. Future Generation Computer Systems 92, 2019, 374–382.

[13] Leiter A., Veluswamy R. R., Wisnivesky J. P..: The global burden of lung cancer: Current status and future trends. Nature Reviews Clinical Oncology 20(9), 2023, 624–639.

[14] Li C. et al.: Advances in lung cancer screening and early detection. Cancer Biology & Medicine 19(5), 2022, 591.

[15] Liu Y. et al.: Radiologic features of small pulmonary nodules and lung cancer risk in the National Lung Screening Trial: A nested case-control study. Radiology 286(1), 2018, 298–306.

[16] Mummaneni S. et al.: A comprehensive study: Intracranial aneurysm detection via VGG16-DenseNet hybrid deep learning on DSA images. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 14(1), 2024, 105–110.

[17] Pal M., Mistry S., De D.: Interpretability approaches of explainable AI in analyzing features for lung cancer detection. Frontiers of ICT in Healthcare: Proceedings of EAIT 2022, Springer Nature Singapore, 2023, 277–287.
[18] Rahman A. et al.: NeuroXAI++: An efficient X-AI intensive brain cancer detection and localization. International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), IEEE, 2023, 1–6.
[19] Rasool S. et al.: Harnessing predictive power: Exploring the crucial role of machine learning in early disease detection. JURIHUM: Jurnal Inovasi dan Humaniora 1(2), 2023, 302–315.
[20] Saman H. et al.: Non-invasive biomarkers for early lung cancer detection. Cancers 14(23), 2022, 5782.
[21] Shafi I. et al.: An effective method for lung cancer diagnosis from CT scan using deep learning-based support vector network. Cancers 14(21), 2022, 5457.
[22] Shamshirband S. et al.: A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues. Journal of Biomedical Informatics 113, 2021, 103627.
[23] Shakeel P. M., Burhanuddin M. A., Desa M. I.: Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Computing and Applications 34(15), 2022, 1–14.
[24] Shimazaki A. et al.: Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. Scientific Reports 12(1), 2022, 727.
[25] Vliegenthart R. et al.: Innovations in thoracic imaging: CT, radiomics, AI, and x-ray velocimetry. Respirology 27(10), 2022, 818–833.
[26] Xu K. et al.: AI body composition in lung cancer screening: Added value beyond lung cancer detection. Radiology 308(1), 2023, e222937.
[27] Zhang C. et al.: Toward an expert level of lung cancer detection and classification using a deep convolutional neural network. The Oncologist 24(9), 2019, 1159–1165.
[28] Zhang J. et al.: Cyber resilience in healthcare digital twin on lung cancer. IEEE Access 8, 2020, 201900–201913.

Ph.D. Vinod Kumar R S
e-mail: rsvinodkumar69@yahoo.co.in

Vinod Kumar R S is professor and Dean (Faculty of Electrical and Electronics Engineering) at Noorul Islam Center for Higher Education, Tamil Nadu. His areas of interest include Digital Image Processing, Analog Modulation Systems, Digital Communications, Linear Integrated Circuits, and Digital Electronics. He is the author of over 80 scientific and methodological works.

https://orcid.org/0000-0001-9727-8110

Ph.D. Bushara A R
e-mail: bushara.ar@gmail.com

Bushara A R is currently an associate professor at the Department of Electronics and Communication of KMEA Engineering College. Her research interests include medical image processing, deep learning, and computer vision. She is the author of over 20 scientific and methodological works.

https://orcid.org/0000-0001-5849-3416

Ph.D. Abubeker K M
e-mail: kmabubeker82@gmail.com

Abubeker K M, Ph.D., is a distinguished academician and researcher renowned for his contributions to AI, IoT, deep learning, and embedded systems. He is an associate professor at Amal Jyothi College of Engineering (Autonomous), Kerala, India. Abubeker is a prolific author, publishing over ten textbooks and 30 research articles in prestigious journals and serving as an editorial member for journals published by IEEE, IET, PLOS, and Springer.

https://orcid.org/0000-0001-7646-0781

Prof. Smitha K M
e-mail: skm.ec@kmeacollege.ac.in

Smitha K M is currently working as Dean Academics, KMEA Engineering College, Kerala, India. Her current areas of research include CPW-fed monopole antennas, fractals, metamaterials, frequency selective surfaces, use of artificial intelligence in microwave design, EMC/EMI, etc.
She is a co-author of over 9 papers in journals and conference proceedings.

https://orcid.org/0009-0000-9885-049X

M.Tech. Abini M A
e-mail: abinima87@gmail.com

Abini M A is currently working as assistant professor in ECE Department at KMEA Engineering College, Kerala. She is a research scholar at Karunya Institute of Technology and Sciences, Coimbatore.
Her research interests include deep learning, image processing and artificial intelligence. She is a life member of IETE and associate member of IE (India).

https://orcid.org/0009-0005-2419-7770

M.Sc. Jubaira Mammoo
e-mail: jbm.ec@kmeacollege.ac.in

Jubaira Mammoo is working as an assistant professor in the Department of Electronics and Communication Engineering, KMEA Engineering College, Kerala. Currently, she is doing a Ph.D. at Karunya Institute of Science & Technology.
Her research area includes image processing, computer vision, machine learning, and deep learning. She is a lifetime member of IETE.

https://orcid.org/0009-0008-7389-3256

Ph.D. Bijesh Paul
e-mail: paul.bijesh@gmail.com

He is currently the Internal Quality Assessment Cell coordinator & professor in Department of Mechanical Engineering of KMEA Engineering College, Kerala.
His research interests are in industrial engineering, forecasting and solar energy.
He has published in more than 20 papers in journals and conferences.

https://orcid.org/0009-0000-8763-0192