Original Article

Check for updates

Available online at
www.heca-analitika.com/ijcr

Indonesian Journal of Case Reports

Vol. 2, No. 1, 2024

Explainable Artificial Intelligence in Medical Imaging: A Case Study on Enhancing Lung Cancer Detection through CT Images

Teuku Rizky Noviandy [1], Aga Maulana [2], Teuku Zulfikar [3], Asep Rusyana [4], Seyi Samson Enitan [5] and Rinaldi Idroes [6,*]

[1] Interdisciplinary Innovation Research Unit, Graha Primera Saintifika, Aceh Besar 23771, Indonesia; trizkynoviandy@gmail.com (T.R.N.)
[2] Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; agamaulana@usk.ac.id (A.M.)
[3] Department of Pulmonology and Respiratory Medicines, Faculty of Syiah Kuala, Universitas Syiah Kuala/Zainoel Abidin Hospital, Banda Aceh, Indonesia; teukuzulfikar@ymail.com (T.Z.)
[4] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; asep.rusyana@usk.ac.id (A.R.)
[5] Department of Medical Laboratory Science, Babcock University, Ilishan-Remo, Nigeria; enitans@babcock.edu.ng (S.S.E.)
[6] School of Mathematics and Applied Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; rinaldi.idroes@usk.ac.id (R.I.)

***** Correspondence: rinaldi.idroes@usk.ac.id

Article History

Received 25 February 2024
Revised 18 April 2024
Accepted 26 April 2024
Available Online 4 May 2024

Keywords:
Deep learning
ResNet50
Grad-CAM
Computer-aided detection
XAI

Abstract

This study tackles the pressing challenge of lung cancer detection, the foremost cause of cancer-related mortality worldwide, hindered by late detection and diagnostic limitations. Aiming to improve early detection rates and diagnostic reliability, we propose an approach integrating Deep Convolutional Neural Networks (DCNN) with Explainable Artificial Intelligence (XAI) techniques, specifically focusing on the Residual Network (ResNet) architecture and Gradient-weighted Class Activation Mapping (Grad-CAM). Utilizing a dataset of 1,000 CT scans, categorized into normal, non-cancerous, and three types of lung cancer images, we adapted the ResNet50 model through transfer learning and fine-tuning for enhanced specificity in lung cancer subtype detection. Our methodology demonstrated the modified ResNet50 model's effectiveness, significantly outperforming the original architecture in accuracy (91.11%), precision (91.66%), sensitivity (91.11%), specificity (96.63%), and F1-score (91.10%). The inclusion of Grad-CAM provided insightful visual explanations for the model's predictions, fostering transparency and trust in computer-assisted diagnostics. The study highlights the potential of combining DCNN with XAI to advance lung cancer detection, suggesting future research should expand dataset diversity and explore multimodal data integration for broader applicability and improved diagnostic capabilities.

1. Introduction

Lung cancer remains the leading cause of cancer-related deaths worldwide [1], with its high mortality rate largely attributed to late-stage detection [2] and the intricacies involved in treatment [3]. The disease originates from unregulated cell growth in the lung tissues [4], which, if not identified and managed at an early stage, may metastasize to other body parts, drastically reducing survival prospects [5]. The prevalence of lung cancer varies across regions, with higher rates observed in countries where tobacco smoking is widespread [6]. Additionally, exposure to environmental pollutants such as secondhand smoke, radon, asbestos, and air pollution

also significantly contribute to the development of lung cancer [7–9].

Lung cancer is broadly categorized into two main types based on the appearance of lung cancer cells under a microscope: Non-Small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC) [10]. NSCLC is the most common type of lung cancer, accounting for about 80% to 85% of all cases [11]. NSCLC is further divided into three main subtypes based on the type of cells found in the tumor: Adenocarcinoma, Squamous Cell Carcinoma, and Large Cell Carcinoma [12]. Adenocarcinoma originates in the cells that secrete substances such as mucus and is the most common form of lung cancer among non-smokers [13]. Squamous Cell Carcinoma arises from the flat cells lining the inside of the lungs and is closely associated with a smoking history [14]. Large Cell Carcinoma, known for its large and abnormal-looking cells, can appear in any part of the lung and tends to grow and spread rapidly, making it more difficult to treat [15].

Conventionally, lung cancer detection has relied on methods such as chest X-rays and computed tomography (CT) scans, followed by biopsy for confirmation [16, 17]. While these methods are essential in identifying abnormalities in the lung, they have notable limitations. For instance, chest X-rays can miss small tumors or fail to distinguish between tumors and other abnormalities, leading to false negatives or positives [18, 19]. CT scans provide more detailed images but also present challenges in accurately differentiating between benign and malignant nodules, requiring invasive procedures like biopsies for definitive diagnoses [20].

In recent years, the incorporation of artificial intelligence (AI) into diagnostic assistance has represented a major advancement [21–23], mainly through the use of a method known as deep learning [24–26]. Deep learning is a method where computers are trained to recognize patterns by analyzing large amounts of data. Specifically, one type of deep learning, called Deep Convolutional Neural Networks (DCNN), has proven very effective [27, 28]. These networks can analyze complex images, such as CT scans, to find signs of lung cancer [29]. Unlike traditional methods, which might require more manual effort and be less accurate, DCNN can more precisely identify different types of lung cancer and distinguish between harmful and harmless nodules with little human oversight. This technology promises to make the detection of lung cancer earlier, less invasive, and more reliable, potentially leading to better patient outcomes.

However, a limitation exists the need for explainable AI [30]. Deep learning systems, including DCNN, operate in a way that is often described as a 'black box,' meaning their decision-making processes are not easily understood by humans [31]. This lack of transparency can be a barrier, particularly in sensitive areas such as healthcare [32, 33]. Clinicians and patients must be able to comprehend how AI models, like those used in lung cancer detection, arrive at their conclusions. This understanding is important not just for acceptance but also for integrating AI tools into clinical decision-making processes. The necessity for this understanding brings XAI into focus.

XAI seeks to bridge the gap between AI decision-making and human interpretability by developing methods and models that clarify, in human-understandable terms, how AI systems make decisions. This includes techniques that can provide visual explanations, decompose the decision-making process into understandable components, and offer insights into the data features most important for predictions. For example, in lung cancer detection, an explainable AI system could highlight which aspects of a lung image were most indicative of malignancy, thereby providing clinicians with understandable evidence to support the AI's diagnosis.

In this study, we aim to enhance traditional lung cancer detection methods by integrating the Residual Network (ResNet) model with Gradient-weighted Class Activation Mapping (Grad-CAM) to improve the precision and efficiency of screenings. Our objectives are to advance detection accuracy, enable the interpretability of AI decisions through visual heat maps, and build trust among healthcare professionals. This approach is designed to make AI's decision-making processes transparent, supporting better clinical decision-making and facilitating wider clinical adoption. The innovative combination of ResNet's robust image processing with Grad-CAM's explanatory power addresses the technical and ethical challenges in early lung cancer detection. By improving both the technology's capability and accountability, our methodology sets a new standard for the responsible and effective use of AI in medical diagnostics

2. Materials and Methods

2.1. Dataset

The dataset used in this study was obtained from Kaggle, a well-known platform recognized for its diverse collection of datasets spanning various domains, such as healthcare and medical imaging [34]. This specific dataset consists of 1,000 CT scan images, classified into four distinct categories. These categories include normal scans, non-cancerous abnormalities, as well as three primary types of chest cancer: squamous

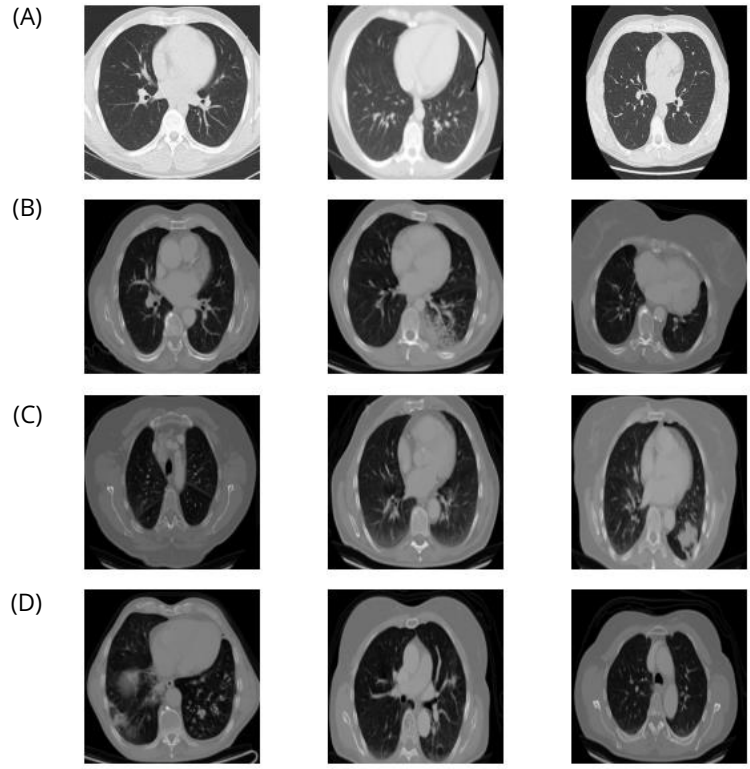

Figure 1. CT Scan images in the dataset: (A) normal; (B) squamous cell carcinoma; (C) adenocarcinoma; (D) large cell carcinoma.

Table 1. Distribution of CT scan images across subsets and classes.

Subset	Class				Total
	Normal	Squamous Cell Carcinoma	Adenocarcinoma	Large Cell Carcinoma	
Training	148	155	195	115	613
Validation	13	15	23	21	72
Testing	54	90	120	51	315
Total					1000

cell carcinoma, adenocarcinoma, and large cell carcinoma. An overview of the CT scan images in this dataset is illustrated in Figure 1.

This dataset is divided into three subsets: the training set, validation set, and testing set, containing 613, 72, and 315 CT scan images, respectively. The training set, consisting of the majority of the data, is utilized to train the DCNN model, allowing it to learn patterns and features present in the images. The validation set, with its smaller size, serves as a means to fine-tune the model's hyperparameters and assess its performance during training, aiding in preventing overfitting, which occurs when a model learns the training data too well, including its noise and outliers, leading to poor performance on new, unseen data. Finally, the testing set, comprising 315 images, provides an independent evaluation of the model's performance after training, helping to gauge its ability to generalize to unseen data and make accurate predictions in real-world scenarios. The distribution of CT scan images across each subset is presented in Table 1.

2.2. ResNet50

Residual Network (ResNet) is a deep learning model designed to solve the vanishing gradient problem encountered in training very deep neural networks [35]. This issue impedes learning by significantly diminishing the gradients. ResNet introduces innovative "skip connections" that permit inputs to skip certain layers, thereby maintaining the flow of gradients and enabling the effective training of much deeper networks. ResNet50, a variant of ResNet, embodies this principle, offering enhanced capabilities for image recognition. Its ability to maintain training efficiency and effectiveness in deep network configurations makes it particularly well-suited for medical imaging, where precision and reliability are critical.

Table 2. Parameter used to train the model.

Parameter	Value
Batch Size	32
Optimizer	Adam
Learning Rate	1e-5
Decay Rate	1e-6
Loss Function	Categorical Cross Entropy
Epoch	100
Early Stopping Patience	20

To adapt ResNet50 for lung cancer detection, we employed a technique known as transfer learning [36]. This involved starting with a ResNet50 model pre-trained on a comprehensive dataset of images called ImageNet, which was then fine-tuned using our specific collection of CT scans. The advantage of this method is that the model has already learned to identify a wide array of general image features, which can be leveraged to detect anomalies within lung images. We further tailored the ResNet50 model to our needs by incorporating three custom layers atop the pre-trained model. These additional layers were designed to hone in on the characteristics most pertinent for differentiating various types of lung cancer and healthy lung tissue. Through this fine-tuning, the model is able to utilize its inherent image recognition prowess for the specialized task of detecting lung cancer subtypes in CT scans.

In our methodology, we specifically modified the ResNet50 architecture to suit our lung cancer detection objectives by setting the layers of the ResNet50 model as non-trainable if they are not part of the 'conv5' block. This approach ensures that only the most advanced features are fine-tuned for our task, preserving the pre-learned weights in the initial layers for general image recognition while focusing the training effort on the deeper layers more relevant to identifying lung cancer. Additionally, we augmented the model with a custom top layer sequence: a Dropout layer set at 0.6 to prevent overfitting, followed by a Flatten layer to transform the output into a 1D array. This is then passed through a Batch Normalization layer to normalize the activations, another Dropout layer at 0.6 for further overfitting mitigation, and finally, a Dense layer with a softmax activation function designed to output predictions across four classes. This tailored architecture aims to enhance the model's ability to discern subtle features indicative of different lung cancer stages from CT scans.

The parameters used to train the model are shown in Table 2. For the training process, a batch size of 32 was chosen to balance the computational efficiency and model accuracy, allowing the model to learn from a sufficiently large subset of data while managing memory resources effectively. The Adam optimizer was selected for its adaptive learning rate capabilities. Unlike traditional fixed learning rate optimizers, Adam adjusts the learning rate for each parameter dynamically based on estimates of the first and second moments of the gradients. This feature enhances the convergence speed towards the optimal solution, leading to faster training times and improved performance on complex datasets [37]. A learning rate of 1e-5 was set to ensure gradual adjustments to the model's weights, preventing overshooting of the global minimum, with a decay rate of 1e-6 applied to decrease the learning rate over time, further stabilizing the training process. The loss function used was Categorical Cross Entropy, suitable for multi-class classification problems, measuring the performance of the model's output relative to the true labels. The model was trained for 100 epochs to allow ample opportunity for learning, with an early stopping mechanism set with a patience of 20 epochs to prevent overfitting by halting training if the validation loss does not improve, ensuring the model's generalizability and efficiency.

2.3. Data Preprocessing

We employed a meticulous approach to prepare the CT scan images for analysis with our deep learning model. Initially, each image underwent resizing to a uniform dimension of 450x450 pixels. This standardization is important for ensuring that the input data is consistent in size, which aids the model in efficiently learning and recognizing patterns across all images [38].

Following the resizing, we converted each image's color scheme from RGB (Red, Green, Blue) to BGR (Blue, Green, Red). This conversion aligns with the color channel ordering convention used in many deep learning frameworks and models pre-trained on the ImageNet dataset, facilitating compatibility and leveraging pre-existing model architectures and weights more effectively [39].

Finally, we conducted a zero-centering process on each color channel of the images, relative to the mean values of the ImageNet dataset. Zero-centering involves adjusting the values of each pixel in such a way that the mean of the pixel intensities across each color channel approximates zero. It's important to note that this adjustment was made without scaling the pixel values. This step is integral to normalizing the input data, reducing the variance among the images, and helping the model to focus on the essential features for classification.

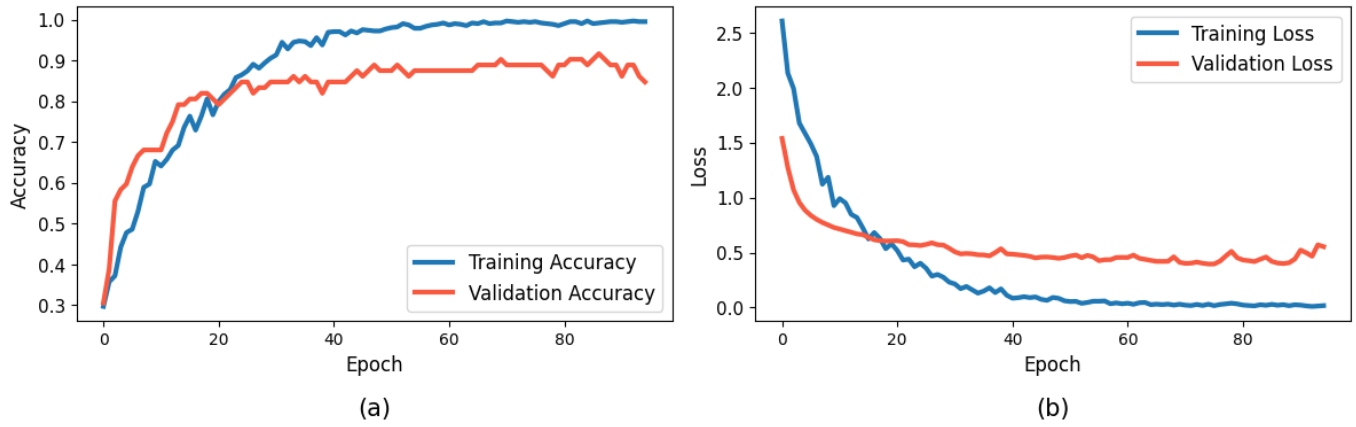

Figure 2. Visualization of the modified ResNet50V2 training and validation: (a) accuracy; (b) loss.

2.4. Evaluation Metrics

In assessing the performance of our lung cancer detection model, we employed a range of evaluation metrics including accuracy, precision, sensitivity, specificity, and F1-score [40]. Accuracy represents the overall correctness of predictions, while precision measures the proportion of true positive predictions among all positive predictions made by the model. Sensitivity, also known as recall, quantifies the model's ability to correctly identify positive instances among all actual positive instances. On the other hand, specificity evaluates the model's ability to correctly identify negative instances among all actual negative instances. The F1-score combines precision and recall into a single metric, providing a balanced measure of a model's performance.

We adopted a weighted average approach for each metric due to the presence of four distinct classes within our dataset. This approach is crucial for providing a more accurate representation of the model's performance across these classes, especially given their varying prevalence. The weighted average method calculates the metrics of accuracy, precision, sensitivity, specificity, and F1-score for each class by considering the proportion of each class's total sample. This ensures that the model's performance in detecting each type of lung cancer and normal tissue is weighted according to the class's representation in the dataset. Consequently, this method offers a nuanced evaluation, reflecting the model's effectiveness in accurately identifying each class, compensating for any imbalance among the classes.

2.5. Explainable Artificial Intelligence (XAI)

To improve the transparency and interpretability of our model's decision-making, we implemented XAI techniques using Grad-CAM [41, 42]. Grad-CAM is a powerful tool within XAI, as it visually indicates the regions within input images that significantly impact the model's classification decisions. This method illuminates the specific features and patterns the model leverages for making its predictions. Through the visualization of these heatmaps, we are provided with a deeper understanding of the underlying rationale for the model's decisions [43]. This insight is crucial for validating the model's predictions and enhancing trust in its capabilities by offering a clear view into how and why certain decisions are made.

3. Results and Discussion

The results obtained from the training of the modified ResNet model are illustrated in Figure 2. As shown in Figure 2a, the training accuracy (blue line) starts at a lower value and exhibits a steady increase as the number of epochs progresses. It reaches a plateau around epoch 50, indicating that the model becomes more consistent in correctly identifying the different categories of lung cancer as it learns from the training data. The validation accuracy (red line) also increases and follows a similar trend as the training accuracy, which suggests that the model generalizes well to new, unseen data. However, there is a noticeable gap between the training and validation accuracy, which could imply a slight overfitting of the model to the training data. Nonetheless, the validation accuracy does not decrease, and its plateau indicates the model's robustness.

Turning to Figure 2b, the training loss (blue line) and validation loss (red line) both decrease over time, with a sharp drop in the initial epochs followed by a more gradual decline. This decline reflects the model's improving performance in correctly classifying the CT scan images as it learns. The training loss continues to decrease slightly throughout the epochs, but the validation loss reaches its lowest point at epoch 75. This

Table 3. Performance comparison of ResNet50 and modified ResNet50 models in predicting lung cancer.

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
ResNet50	89.21	89.59	89.21	94.87	88.97
Modified ResNet50	91.11	91.66	91.11	96.63	91.10

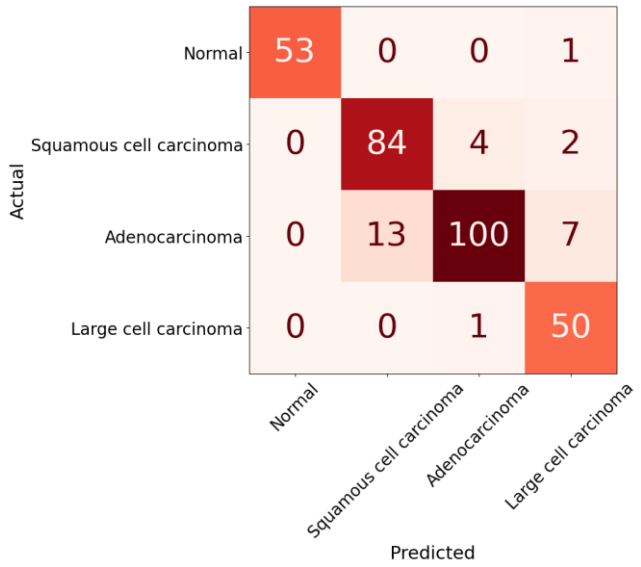

Figure 3. Confusion matrix of the testing set prediction from modified ResNet50 model.

lowest point represents the best generalization performance of the model on the validation set.

Post epoch 75, the validation loss shows minor fluctuations but does not achieve any significant improvement. The early stopping mechanism is designed to halt training when there is no improvement in validation loss for a certain number of consecutive epochs, which in this study is set to 20. As there was no improvement in the validation loss from epoch 75 to epoch 95, the early stopping criterion was met, and the training was terminated to prevent overfitting and unnecessary computations.

The performance of our lung cancer detection models, as outlined in Table 3, showcases the effectiveness of employing deep learning architectures for medical image analysis. The table compares the performance metrics of the original ResNet50 model and our modified ResNet50 across five key metrics: accuracy, precision, sensitivity, specificity, and F1-score.

The modified ResNet50 model demonstrates superior performance across all evaluated metrics compared to the original ResNet50. Specifically, the modified ResNet50 achieved an accuracy of 91.11%, precision of 91.66%, sensitivity of 91.66%, specificity of 96.63%, and an F1-score of 91.10%. These improvements underline the impact of the modifications made to the original ResNet50 architecture, which were aimed at enhancing the model's ability to discern and classify lung cancer from CT scans with greater accuracy.

The increase in accuracy and sensitivity suggests that the modified ResNet50 is more adept at correctly identifying cases of lung cancer, which is critical for early detection and timely treatment. The specificity improvement indicates a reduction in false positives, which is vital for minimizing unnecessary anxiety and invasive follow-up procedures for patients. Furthermore, the enhanced F1-score reflects a balanced improvement in precision and recall, highlighting the model's overall reliability in lung cancer detection.

The confusion matrix for the modified ResNet50 model provides an insightful breakdown of the model's predictive performance across the different classes of lung cancer (Figure 3). A confusion matrix is a table often used in classification to visualize the performance of an algorithm. Each row of the matrix represents the instances in an actual class, while each column represents the instances in a predicted class. The diagonal cells represent correct predictions, while the off-diagonal cells indicate errors.

The matrix shows that the model performs exceptionally well in distinguishing between different types of lung cancer and normal scans. For instance, it correctly identified 53 out of 54 normal cases, with only one misclassification. In the case of squamous cell carcinoma, the model identified 84 out of 90 cases correctly, with 4 cases misclassified as adenocarcinoma and two as large cell carcinoma. For adenocarcinoma, the model showed remarkable precision, correctly classifying 100 out of 120 cases, with the remaining misclassified as squamous cell carcinoma or large cell carcinoma. Lastly, for large cell carcinoma, the model accurately diagnosed 50 out of 51 cases, with only one misclassification.

The high numbers on the diagonal indicate that the model has a strong ability to correctly predict the class of each CT scan, which is crucial for reliable lung cancer screening. The few off-diagonal numbers represent misclassifications, which are relatively low, suggesting that the model's precision and recall are high for each class. This is supported by the performance metrics reported, which show high sensitivity and specificity,

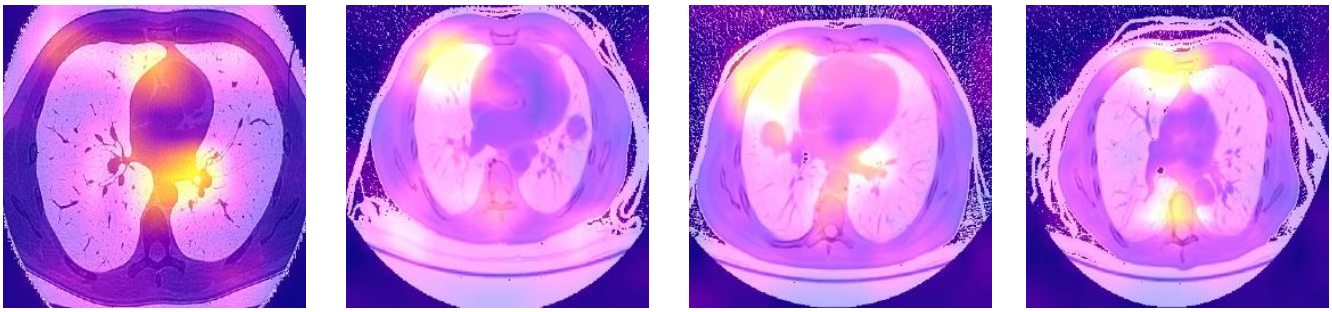

Figure 4. Confusion matrix of the testing set prediction from modified ResNet50V2 model.

minimizing the chances of false negatives and false positives.

The prediction results visualized with Grad-CAM heatmaps are shown in Figure 4. The provided image offers a clear insight into the decision-making process of the modified ResNet50 model. Each heatmap provides a visual explanation for why the model made a particular prediction, highlighting the regions in the CT images that most influenced the model's decision. The heatmaps show that the model is focusing on appropriate areas of the CT scans to distinguish between normal tissue and various types of cancer. For example, in the adenocarcinoma prediction, the heatmap is concentrated around the abnormal growth, aligning with the expected patterns of this cancer type.

The accompanying confidence scores reflect the model's certainty in its predictions, with high confidence in its correct predictions for normal tissue (100%) and large cell carcinoma (98.02%). The confidence scores for squamous cell carcinoma and adenocarcinoma are slightly lower at 91.67% and 91.37%, respectively, which might correlate with the more challenging nature of distinguishing these conditions or the variability in their appearance on CT scans.

The outcomes of this research possess substantial implications for the field of medical imaging and oncology. The high accuracy, precision, sensitivity, specificity, and F1-scores achieved by the modified ResNet50 model suggest that deep learning can significantly enhance lung cancer detection, thereby potentially improving patient outcomes. Early and accurate detection is paramount in cancer treatment, and AI-assisted diagnostics could lead to earlier interventions, more targeted therapies, and, consequently, better survival rates. Furthermore, the explainability aspect introduced by Grad-CAM provides clinicians with valuable insights into the AI decision-making process, fostering a greater level of trust in AI tools.

For the integration of AI models like the modified ResNet50 into clinical settings, several practical considerations must be addressed. Clinicians require tools that not only deliver high performance but also fit seamlessly into the existing workflow, ensuring that they complement, rather than complicate, the diagnostic process. The visual explanations provided by Grad-CAM can serve as a communication medium, helping clinicians to understand and verify the AI's recommendations, which is crucial for acceptance and ethical accountability in medical practice.

Despite the promising results, this study has limitations. The dataset, while diverse, is limited in size and sourced from a single platform, which could affect the model's generalizability to broader populations [44]. Moreover, the complexity of AI models can lead to challenges in clinical interpretation, and the reliance on visual explanations does not fully elucidate the intricate patterns learned by the deep learning model. There is also the risk of the model encountering novel presentations of lung cancer not represented in the training dataset, which could lead to misclassifications.

Future research should prioritize expanding the dataset to encompass a more diverse demographic, including multi-institutional data that captures a broader spectrum of lung cancer presentations. This expansion is crucial for enhancing the model's robustness and extending its applicability across different populations. Additionally, there is potential in exploring the integration of multimodal data, such as combining CT images with patient medical histories and genetic information, to enhance the model's diagnostic capabilities. However, integrating multimodal data presents challenges, such as data heterogeneity, alignment issues, and fusion complexities, which may complicate the training process

and impact the model's performance. Future research efforts could concentrate on developing advanced algorithms for effective data integration and exploring techniques like deep learning architectures tailored to handle multimodal information. These approaches hold promise in overcoming current limitations and further enhancing the model's diagnostic accuracy.

4. Conclusions

This study has successfully demonstrated the potential of a modified ResNet50 model with explainable AI techniques for the detection of lung cancer in CT images. By achieving high accuracy, precision, sensitivity, specificity, and F1-scores, the modified ResNet50 model shows promise in significantly enhancing the early detection of lung cancer, which is crucial for improving patient prognosis and survival rates. The implementation of Grad-CAM provides valuable visual explanations of the model's decision-making process, addressing the critical need for transparency in AI applications in healthcare.

Author Contributions: Conceptualization, T.R.N., A.M., and R.I.; methodology, T.R.N., A.M. and A.R.; software, T.R.N. and A.M.; validation, T.Z., A.R., S.S.E. and R.I.; formal analysis, T.R.N.; investigation, T.R.N. and A.M.; resources, T.Z. and R.I.; data curation, T.Z., S.S.E. and R.I.; writing—original draft preparation, T.R.N., A.M., and A.R.; writing—review and editing, T.Z., S.S.E., and R.I.; visualization, T.R.N.; supervision, T.Z. and R.I.; project administration, R.I.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study does not receive external funding.

Ethical Clearance: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study was obtained from Kaggle (https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images, accessed 27 November 2023) and was made available by Mohamed Hany. We acknowledge the contributions of the data providers and the Kaggle platform for making this dataset publicly accessible

Conflicts of Interest: All the authors declare no conflicts of interest.

References

1. Barta, J. A., Powell, C. A., and Wisnivesky, J. P. (2019). Global Epidemiology of Lung Cancer, *Annals of Global Health*, Vol. 85, No. 1. doi:10.5334/aogh.2419.
2. Schabath, M. B., and Cote, M. L. (2019). Cancer Progress and Priorities: Lung Cancer, *Cancer Epidemiology, Biomarkers & Prevention*, Vol. 28, No. 10, 1563–1579. doi:10.1158/1055-9965.EPI-19-0221.
3. Leiter, A., Veluswamy, R. R., and Wisnivesky, J. P. (2023). The Global Burden of Lung Cancer: Current Status and Future Trends, *Nature Reviews Clinical Oncology*, Vol. 20, No. 9, 624–639. doi:10.1038/s41571-023-00798-3.
4. Lundin, A., and Driscoll, B. (2013). Lung Cancer Stem Cells: Progress and Prospects, *Cancer Letters*, Vol. 338, No. 1, 89–93. doi:10.1016/j.canlet.2012.08.014.
5. Heuvers, M. E., Hegmans, J. P., Stricker, B. H., and Aerts, J. G. (2012). Improving Lung Cancer Survival; Time to Move On, *BMC Pulmonary Medicine*, Vol. 12, No. 1, 77. doi:10.1186/1471-2466-12-77.
6. Chaitanya Thandra, K., Barsouk, A., Saginala, K., Sukumar Aluru, J., and Barsouk, A. (2021). Epidemiology of Lung Cancer, *Współczesna Onkologia*, Vol. 25, No. 1, 45–52. doi:10.5114/wo.2021.103829.
7. Cani, M., Turco, F., Butticè, S., Vogl, U. M., Buttigliero, C., Novello, S., and Capelletto, E. (2023). How Does Environmental and Occupational Exposure Contribute to Carcinogenesis in Genitourinary and Lung Cancers?, *Cancers*, Vol. 15, No. 10, 2836. doi:10.3390/cancers15102836.
8. Xue, Y., Wang, L., Zhang, Y., Zhao, Y., and Liu, Y. (2022). Air Pollution: A Culprit of Lung Cancer, *Journal of Hazardous Materials*, Vol. 434, 128937. doi:10.1016/j.jhazmat.2022.128937.
9. S Cheng, E., Weber, M., Steinberg, J., and Qin Yu, X. (2021). Lung Cancer Risk in Never-Smokers: An Overview of Environmental and Genetic Factors, *Chinese Journal of Cancer Research*, Vol. 33, No. 5, 548–562. doi:10.21147/j.issn.1000-9604.2021.05.02.
10. Araujo, L. H., Horn, L., Merritt, R. E., Shilo, K., Xu-Welliver, M., and Carbone, D. P. (2020). Cancer of the Lung, *Abeloff's Clinical Oncology*, Elsevier, 1108-1158.e16. doi:10.1016/B978-0-323-47674-4.00069-4.
11. Padinharayil, H., Varghese, J., John, M. C., Rajanikant, G. K., Wilson, C. M., Al-Yozbaki, M., Renu, K., Dewanjee, S., Sanyal, R., Dey, A., Mukherjee, A. G., Wanjari, U. R., Gopalakrishnan, A. V., and George, A. (2023). Non-Small Cell Lung Carcinoma (Nsclc): Implications on Molecular Pathology and Advances in Early Diagnostics and Therapeutics, *Genes & Diseases*, Vol. 10, No. 3, 960–989. doi:10.1016/j.gendis.2022.07.023.
12. Qu, Y., Cheng, B., Shao, N., Jia, Y., Song, Q., Tan, B., and Wang, J. (2020). Prognostic Value of Immune-Related Genes in the Tumor Microenvironment of Lung Adenocarcinoma and Lung Squamous Cell Carcinoma, *Aging*, Vol. 12, No. 6, 4757–4777. doi:10.18632/aging.102871.
13. Corrales, L., Rosell, R., Cardona, A. F., Martín, C., Zatarain-Barrón, Z. L., and Arrieta, O. (2020). Lung Cancer in Never Smokers: The Role of Different Risk Factors Other Than Tobacco Smoking, *Critical Reviews in Oncology/Hematology*, Vol. 148, 102895. doi:10.1016/j.critrevonc.2020.102895.
14. Wang, B.-Y., Huang, J.-Y., Chen, H.-C., Lin, C.-H., Lin, S.-H., Hung, W.-H., and Cheng, Y.-F. (2020). The Comparison between Adenocarcinoma and Squamous Cell Carcinoma in Lung Cancer Patients, *Journal of Cancer Research and Clinical Oncology*, Vol. 146, No. 1, 43–52. doi:10.1007/s00432-019-03079-8.
15. Travis, W. D. (2020). Lung Cancer Pathology, *Clinics in Chest Medicine*, Vol. 41, No. 1, 67–85. doi:10.1016/j.ccm.2019.11.001.
16. Demirci, N. Y. (2023). Diagnostic Workup for Lung Cancer, C. Cingi; A. Yorgancıoğlu; N. Bayar Muluk; A. A. Cruz (Eds.), , Springer International Publishing, Cham, 1–16. doi:10.1007/978-3-031-22483-6_62-1.
17. Hyldgaard, C., Trolle, C., Harders, S. M. W., Engberg, H., Rasmussen, T. R., and Møller, H. (2022). Increased Use of Diagnostic Ct Imaging Increases the Detection of Stage IA Lung Cancer: Pathways and Patient Characteristics, *BMC Cancer*, Vol. 22, No. 1, 464. doi:10.1186/s12885-022-09585-2.
18. Ciello, A. del, Franchi, P., Contegiacomo, A., Cicchetti, G., Bonomo, L., and Larici, A. R. (2017). Missed Lung Cancer: When, Where, and Why?, *Diagnostic and Interventional Radiology*, Vol. 23, No. 2, 118–126. doi:10.5152/dir.2016.16187.

19. Bradley, S. H., Abraham, S., Callister, M. E., Grice, A., Hamilton, W. T., Lopez, R. R., Shinkins, B., and Neal, R. D. (2019). Sensitivity of Chest X-Ray for Detecting Lung Cancer in People Presenting with Symptoms: A Systematic Review, *British Journal of General Practice*, Vol. 69, No. 689, e827–e835. doi:10.3399/bjgp19X706853.

20. Loverdos, K., Fotiadis, A., Kontogianni, C., Iliopoulou, M., and Gaga, M. (2019). Lung Nodules: A Comprehensive Review on Current Approach and Management, *Annals of Thoracic Medicine*, Vol. 14, No. 4, 226. doi:10.4103/atm.ATM_110_19.

21. Noviandy, T. R., Nainggolan, S. I., Raihan, R., Firmansyah, I., and Idroes, R. (2023). Maternal Health Risk Detection Using Light Gradient Boosting Machine Approach, *Infolitika Journal of Data Science*, Vol. 1, No. 2, 48–55. doi:10.60084/ijds.v1i2.123.

22. Maulana, A., Faisal, F. R., Noviandy, T. R., Rizkia, T., Idroes, G. M., Tallei, T. E., El-Shazly, M., and Idroes, R. (2023). Machine Learning Approach for Diabetes Detection Using Fine-Tuned XGBoost Algorithm, *Infolitika Journal of Data Science*, Vol. 1, No. 1, 1–7. doi:10.60084/ijds.v1i1.72.

23. Suhendra, R., Suryadi, S., Husdayanti, N., Maulana, A., and Rizky, T. (2023). Evaluation of Gradient Boosted Classifier in Atopic Dermatitis Severity Score Classification, *Heca Journal of Applied Sciences*, Vol. 1, No. 2, 54–61. doi:10.60084/hjas.v1i2.85.

24. Tran, K. A., Kondrashova, O., Bradley, A., Williams, E. D., Pearson, J. V., and Waddell, N. (2021). Deep Learning in Cancer Diagnosis, Prognosis and Treatment Selection, *Genome Medicine*, Vol. 13, No. 1, 152. doi:10.1186/s13073-021-00968-x.

25. Bakator, M., and Radosav, D. (2018). Deep Learning and Medical Diagnosis: A Review of Literature, *Multimodal Technologies and Interaction*, Vol. 2, No. 3, 47. doi:10.3390/mti2030047.

26. Liu, X., Wang, H., Li, Z., and Qin, L. (2021). Deep Learning in Ecg Diagnosis: A Review, *Knowledge-Based Systems*, Vol. 227, 107187. doi:10.1016/j.knosys.2021.107187.

27. Maulana, A., Noviandy, T. R., Suhendra, R., Earlia, N., Bulqiah, M., Idroes, G. M., Niode, N. J., Sofyan, H., Subianto, M., and Idroes, R. (2023). Evaluation of Atopic Dermatitis Severity Using Artificial Intelligence, *Narra J*, Vol. 3, No. 3, e511. doi:10.52225/narra.v3i3.511.

28. Talukder, M. A., Islam, M. M., Uddin, M. A., Akhter, A., Pramanik, M. A. J., Aryal, S., Almoyad, M. A. A., Hasan, K. F., and Moni, M. A. (2023). An Efficient Deep Learning Model to Categorize Brain Tumor Using Reconstruction and Fine-Tuning. doi:10.48550/arXiv.2305.12844.

29. Cellina, M., Cacioppa, L. M., Cè, M., Chiarpenello, V., Costa, M., Vincenzo, Z., Pais, D., Bausano, M. V., Rossini, N., Bruno, A., and Floridi, C. (2023). Artificial Intelligence in Lung Cancer Screening: The Future Is Now, *Cancers*, Vol. 15, No. 17, 4344. doi:10.3390/cancers15174344.

30. Amann, J., Blasimme, A., Vayena, E., Frey, D., and Madai, V. I. (2020). Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective, *BMC Medical Informatics and Decision Making*, Vol. 20, No. 1, 310. doi:10.1186/s12911-020-01332-6.

31. Noviandy, T. R., Maulana, A., Idroes, G. M., Suhendra, R., Adam, M., Rusyana, A., and Sofyan, H. (2023). Deep Learning-Based Bitcoin Price Forecasting Using Neural Prophet, *Ekonomikalia Journal of Economics*, Vol. 1, No. 1, 19–25. doi:10.60084/eje.v1i1.51.

32. Holzinger, A., Biemann, C., Pattichis, C. S., and Kell, D. B. (2017). What Do We Need to Build Explainable AI Systems for the Medical Domain?, *ArXiv Preprint ArXiv:1712.09923*.

33. Ali, S., Akhlaq, F., Imran, A. S., Kastrati, Z., Daudpota, S. M., and Moosa, M. (2023). The Enlightening Role of Explainable Artificial Intelligence in Medical & Healthcare Domains: A Systematic Literature Review, *Computers in Biology and Medicine*, Vol. 166, 107555. doi:10.1016/j.compbiomed.2023.107555.

34. Hany, M. (2020). Chest CT-Scan Images Dataset, from https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images/data, accessed 27-11-2023.

35. He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition, *Computer Vision and Pattern Recognition*.

36. Idroes, G. M., Maulana, A., Suhendra, R., Lala, A., Karma, T., Kusumo, F., Hewindati, Y. T., and Noviandy, T. R. (2023). TeutongNet: A Fine-Tuned Deep Learning Model for Improved Forest Fire Detection, *Leuser Journal of Environmental Studies*, Vol. 1, No. 1, 1–8. doi:10.60084/ljes.v1i1.42.

37. Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization, *ArXiv Preprint ArXiv:1412.6980*.

38. Vasuki, P., Kanimozhi, J., and Devi, M. B. (2017). A Survey on Image Preprocessing Techniques for Diverse Fields of Medical Imagery, *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, IEEE, 1–6. doi:10.1109/ICEICE.2017.8192443.

39. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A Large-Scale Hierarchical Image Database, *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Ieee, 248–255.

40. Idroes, G. M., Noviandy, T. R., Maulana, A., Zahriah, Z., Suhendrayatna, S., Suhartono, E., Khairan, K., Kusumo, F., Helwani, Z., and Abd Rahman, S. (2023). Urban Air Quality Classification Using Machine Learning Approach to Enhance Environmental Monitoring, *Leuser Journal of Environmental Studies*, Vol. 1, No. 2, 62–68. doi:10.60084/ljes.v1i2.99.

41. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, 618–626. doi:10.1109/ICCV.2017.74.

42. Noviandy, T. R., Maulana, A., Khowarizmi, F., and Muchtar, K. (2023). Effect of CLAHE-based Enhancement on Bean Leaf Disease Classification through Explainable AI, *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE)*, IEEE, 515–516. doi:10.1109/GCCE59613.2023.10315394.

43. Samek, W., Wiegand, T., and Müller, K.-R. (2017). Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models, *ArXiv Preprint ArXiv:1708.08296*.

44. Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., Folio, L. R., Summers, R. M., Rubin, D. L., and Lungren, M. P. (2020). Preparing Medical Imaging Data for Machine Learning, *Radiology*, Vol. 295, No. 1, 4–15. doi:10.1148/radiol.2020192224.