International Journal on Engineering Technology (InJET)

www.kec.edu.np/journal
Volume 1, issue no. 1, Nov 2023, Page 223-231
Received Date: July 8, 2023
Accepted Date: Nov 2, 2023
ISSN: 3021-940X (print)

Unlocking the Potential of CT scans: An Explanation-Driven Deep Learning Model for Predicting Lung Cancer

Yuba Raj Oli[1], Loknath Regmi[2], Prakash Poudel[3], Satish Kumar Karna[3], Mohan Bhandari[4*]

[1]*Phoenix College of Management, Kathmandu, Nepal, yubaraj.oli@pcmgmt.edu.np*
[2]*IOE, Pulchowk Campus, Nepal, lnregmi046@pcampus.edu.np*
[3]*Nepal College of Information Technology, Nepal, visitppdl@gmail.com; satishkarna76@gmail.com*
[4]*Samriddhi College, Bhaktapur, Nepal, mail2mohanbhandari@gmail.com*

Abstract

This research study aims to evaluate the effectiveness of transfer learning with the ResNet50 model for classifying CT scan images of lungs as having cancer or not. Additionally, it explores the interpretability methods of LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (SHapley Additive exPlanations) to provide explanations for the predictions made by the ResNet50 model on CT scan images. The research objectives include developing a deep learning model based on the ResNet50 architecture, evaluating its performance using various metrics, and explaining the predictions using LIME and SHAP techniques. The dataset consists of a collection of CT scan images of lungs, with labels indicating the presence or absence of cancer. Through k-fold cross-validation, the model achieves high accuracy and low loss, demonstrating its effectiveness in classifying lung cancer. The interpretability methods of LIME and SHAP shed light on the crucial features and regions in the CT scan images that contribute to the model's predictions, enhancing the understanding of the model's decision-making process. The results highlight the potential of transfer learning and interpretability techniques in improving the accuracy and explainability of lung cancer detection models. Future directions may involve applying the developed model to larger datasets, classifying different stages of cancer, and identifying the specific regions within the lungs where cancer cells are detected.

Keywords: Transfer learning, ResNet50, CT scan images, Lung cancer, Classification, Explainable AI, LIME, SHAP, Deep learning, Predictive modeling

1. Introduction

Lung cancer is a significant global health concern, accounting for many cancer-related deaths worldwide. Early and accurate detection of lung cancer plays a crucial role in improving patient outcomes and survival rates. Recent advancements in deep learning algorithms, coupled with the availability of large-scale medical imaging datasets, have shown promising results in automating the diagnosis of lung cancer using CT scan images. However, the inherent complexity of deep learning models and the lack of interpretability hinder their widespread adoption in clinical practice.

This paper addresses the challenges of classifying lung cancer in CT scan images by leveraging the power of deep learning techniques while enhancing interpretability. The research aims to investigate the effectiveness of transfer learning with the ResNet50 model, a state-of-the-art architecture, in accurately classifying CT scan images as either malignant or benign. Additionally, the research employs two popular interpretability methods, LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (SHapley Additive exPlanations), to explain the predictions of the deep learning model and gain insights into the features that are most relevant for determining the presence of lung cancer.

**Corresponding Author*

The objectives of the study encompass the development of a deep learning model based on the ResNet50 architecture, evaluating its performance using various metrics, and providing interpretability analysis to enhance trust and transparency in the model's predictions. We hypothesize that transfer learning, combined with LIME and SHAP techniques, can significantly improve both the accuracy and interpretability of lung cancer classification in CT scan images.

By addressing the challenges of accuracy and interpretability in lung cancer classification, this research contributes to the field of medical image analysis and has the potential to improve diagnostic accuracy, patient outcomes, and overall healthcare decision-making in lung cancer management.

The contribution of this study is what follows:

1. Evaluate the accuracy of the ResNet50 model in classifying CT scan images of lungs as malignant or benign, aiming to enhance diagnostic precision and promote early lung cancer detection.

2. Enhance model interpretability by employing LIME and SHAP techniques to explain the deep learning model's predictions, fostering trust and facilitating its adoption in clinical practice.

2. Related Works

In the field of lung cancer detection, various traditional, machine learning (ML), and deep learning (DL) approaches have been explored. Traditional approaches, relying on handcrafted features and classical machine learning algorithms, provided foundational insights but had limited performance compared to more recent techniques. ML approaches, such as the supervised CNN method proposed by Kumar et al. (2020), demonstrated remarkable accuracy in identifying relevant regions in PET-CT images, surpassing traditional methods. Kumar, A. et. al. achieved an exceptional accuracy of 99.29% in foreground detection.

DL approaches, like the DL-based CAD system introduced by Hanan et al. (2018), showcased the power of automatically learning intricate patterns, achieving high accuracy and sensitivity in lung cancer detection. Their model achieved a classification accuracy rate of 99.06%, a sensitivity of 100%, and a specificity of 99.2% (Hanan et al., 2018). Suren Makaju et al. (2018) employed a prototype DL model for lung cancer detection, resulting in an accuracy of 92%. Although specific details regarding the algorithm and dataset were not provided, the achieved accuracy signifies the potential of DL approaches in improving the accuracy and aiding in the early detection of lung cancer.

Allison and A.M.R.W. (2017) utilized DL models with manifold preprocessing methods, achieving an accuracy of 97.5% with a false positive rate below 10%. By employing these techniques, the study showcased the potential of DL in accurately identifying and diagnosing lung cancer cases.

Aggarwal, Furquan, and Kalra introduced a model designed to facilitate the differentiation between nodules and the regular anatomical structures found in the lungs. This approach involves the extraction of geometric, statistical, and grayscale attributes. The classifier employed in this method is Linear Discriminant Analysis (LDA), combined with optimal thresholding for segmentation purposes. The system demonstrates a commendable 84% accuracy, along with a high sensitivity rate of 97.14%. However, the specificity, which stands at 53.33%, remains notably lower. Despite its capacity to detect cancer nodules, the model's overall accuracy remains unsatisfactory.

Jin, Zhang, and Jin employed a convolutional neural network as the classifier within their CAD (Computer-Aided Detection) system for the identification of lung cancer. The system achieved notable performance metrics with an accuracy of 84.6%, sensitivity of 82.5%, and specificity of 86.7%. One notable strength of this model lies in its utilization of circular filters during the Region of Interest (ROI) extraction phase, effectively mitigating the training and recognition expenses. However, it is worth noting that despite the reduction in implementation costs, the model's accuracy remains suboptimal.

3. Methodology

The framework used in this study aims to develop an accurate and efficient system for lung cancer detection using deep learning techniques. It consists of multiple stages, starting with data preprocessing, where the images from the dataset are resized and pixel values are normalized. The preprocessed data is then used to train a deep convolutional neural network (CNN) model, designed to learn intricate patterns and features indicative of lung cancer. Transfer learning with the ResNet50 architecture is employed to leverage pre-trained weights and improve the model's performance. The trained model is evaluated using various metrics such as accuracy, precision, recall, and F1-score. To enhance interpretability, the LIME and SHAP methods are applied to explain the model's predictions and identify essential features in the CT scan images. The framework holds the potential to provide accurate and reliable lung cancer detection, contributing to early diagnosis and improved patient outcomes.

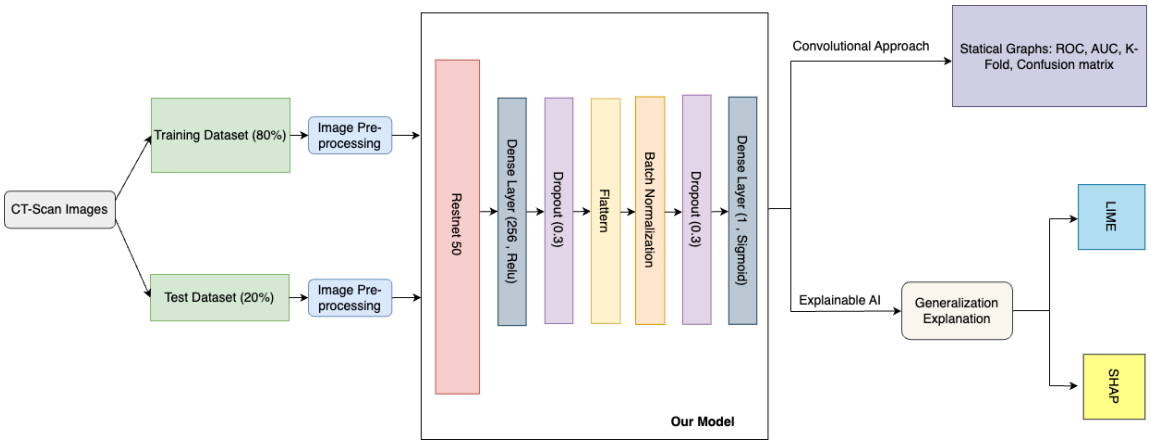

Figure 1. Framework used

The following algorithm (Gaur, L., Bhandari, M., Razdan, T., Mallik, S., & Zhao, Z. 2022) is used to calculate the Shapley values in this study.

Input: Number of iterations M, instance of interest x, feature index j, data matrix X, and (SVM or LR Model)

for Every Iteration 1 up to M **do**

Draw random instance z from the data matrix X

Choose a random permutation of the feature values

Order instance x: ( ) (Equation 1)

Order instance z: ( ) (Equation 2)

Construct two new instances:

with(( )) (Equation 3)

Without(( )) (Equation 4)

Compute Marginal Distribution:

$$\phi_j^m = f(x_{+j}) - f(x_{-j}) \quad \text{(Equation 5)}$$

end for

Compute Shapely values:

$$\phi_j(x) = \frac{1}{M}\sum_{m=1}^{M}\phi_j^m \qquad \text{(Equation 6)}$$

3.1 Dataset

In this study, the "Lung Cancer Dataset" was created by merging data from two sources: "The IQ-OTHNCCD lung cancer dataset" (Alyasriy, H., & AL-Huseiny, M. 2021) and "Chest Cancer Detection Transfer Learning" (Hany, M. 2020). With a total of 2020 images, including 1444 cancerous and 576 non-cancerous samples, the merged dataset provides a more extensive and more diverse collection for robust analysis and model training. By combining datasets from different sources, the "Lung Cancer Dataset" captures variations in imaging techniques, patient demographics, and disease presentations, enhancing the models' ability to discriminate between cancerous and non-cancerous cases. This comprehensive dataset in Google Drive facilitates accurate and reliable lung cancer analysis and detection.

3.1.1. Data Pre-processing

The images in the dataset were resized to a standardized dimension of 150 × 150 × 3 to ensure consistent learning and efficient processing. Subsequently, the pixel values of each image were scaled down to a range of 0 to 1 for optimal calculations. The dataset was then split into training and test sets, with an 80:20 ratio, where 80% of the data was used for training the model and 20% for independent testing. To improve convergence and prevent the CNN model from learning the training order, all images, regardless of their original size, were resized to uniform dimensions. Gaussian noise, with a mean of 0 and a standard deviation of 100.5, was introduced as an augmentation technique to enhance the deep learning process. Figure 1 provides an example of a tumor instance from the dataset, showcasing one of the tumor categories for analysis.

4. Result and Analysis

4.1. Accuracy and Loss

After the classification process, the evaluation of CNN models is conducted by assessing their accuracy and the number of incorrect predictions. The results obtained from the conventional CNN approach are illustrated in Figure 2, showcasing the corresponding curves. These metrics provide valuable insights into the effectiveness and reliability of the CNN models in accurately classifying the given data.

A. B.

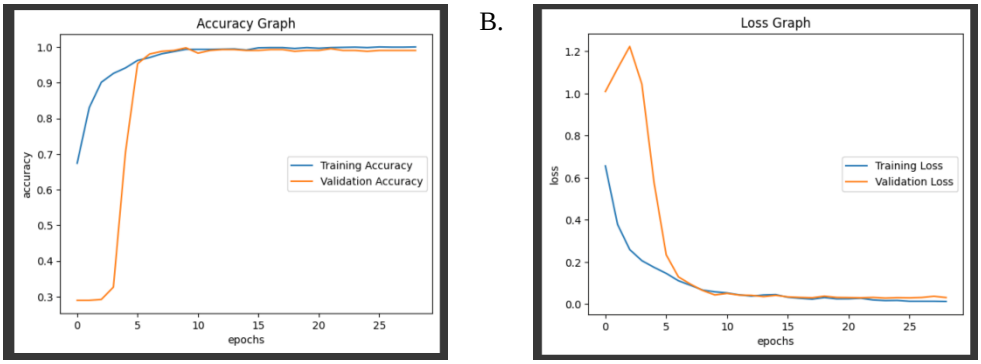

Figure 2. Training and Validation results of the model. Figure A shows the training and validation accuracy of 99.64% and 99.25%, respectively. In figure B it shows the training and validation loss of 0.1149 and 0.153 respectively.

4.2. Confusion Matrix

The confusion matrix, shown in Figure 3, provides a comprehensive overview of the classification results achieved by the CNN model. It demonstrates the model's effectiveness in accurately identifying cancerous instances (class 0), with 286 out of 290 cases correctly classified. Additionally, the model performs well in distinguishing non-cancerous instances (class 1), with 114 out of 118 cases accurately classified. Overall, the

high accuracy of the model's predictions indicates its potential for reliable cancer detection. However, it is crucial to further investigate the misclassification of 4 non-cancerous instances as cancerous, as this could have implications in real-world scenarios. The reason behind this misclassification is small set of data and will go further low in number if we train our model with large dataset.

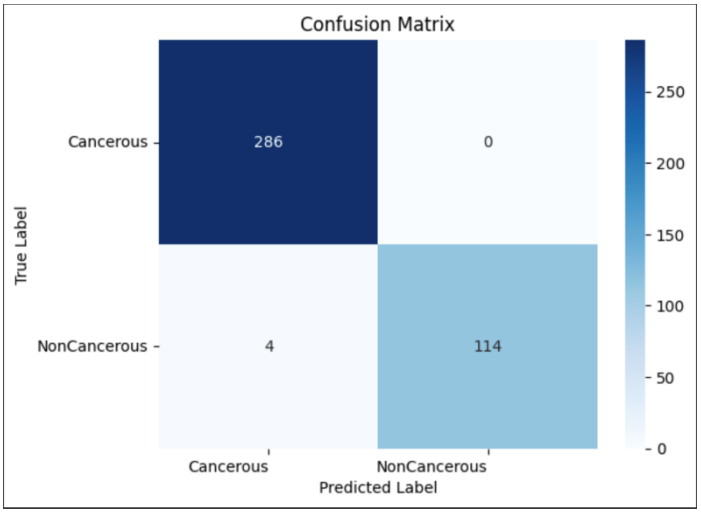

Figure 3: Effectiveness of model with confusion matrix.

4.3. K- Fold Analysis

The k-fold analysis produced exceptional results, with a perfect accuracy score of 1.0000 in each fold except for fold 8, which had a slightly lower accuracy of 0.9876. This is due to the insufficient dataset while randomly picking the training and test data. The average accuracy across all folds was an impressive 0.9987, indicating the robustness and reliability of the CNN model in accurately predicting lung cancer. Additionally, the average loss value of 0.0071 demonstrates the model's ability to minimize errors and make precise predictions. These findings suggest that the proposed CNN model performs excellently in classifying lung cancer cases. The consistently achieved high accuracy across multiple folds reinforces the model's effectiveness and strengthens its potential for practical applications in lung cancer diagnosis.

Fold	Training Accuracy	Training Loss	Test Accuracy	Test Loss
1	0.9988	0.0142	1.0000	0.0040
2	0.9972	0.0168	1.0000	0.0023
3	0.9987	0.0145	1.0000	0.0038
4	0.9987	0.0114	1.0000	0.0029
5	0.9978	0.0124	1.0000	0.0094
6	0.9934	0.0280	1.0000	0.0046
7	0.9999	0.0092	1.0000	0.0011

8	0.9942	0.0287	0.9876	0.0361
9	0.9961	0.0199	1.0000	0.0058
10	0.9990	0.0093	1.0000	0.0012
Average	**0.9974**	**0.0164**	**0.9987**	**0.0071**

Table 1: K-fold analysis of the proposed model

4.4. SHAP

The CNN model employed in this study exhibits complex mathematical behavior, which can pose challenges in interpreting its predictions directly. To address this issue, SHAP (SHapley Additive exPlanations) was utilized to elucidate the impact of individual input features on the model's output. The resulting SHAP values were visualized in Figures 4 and 5, where red pixels indicate positive values that contribute to higher class probabilities, and blue pixels represent negative values that decrease class probabilities. These visual representations provide valuable insights into how specific features influence the model's predictions. By leveraging SHAP, this study enhances the interpretability of the CNN model, enabling a better understanding of its decision-making process and feature importance.

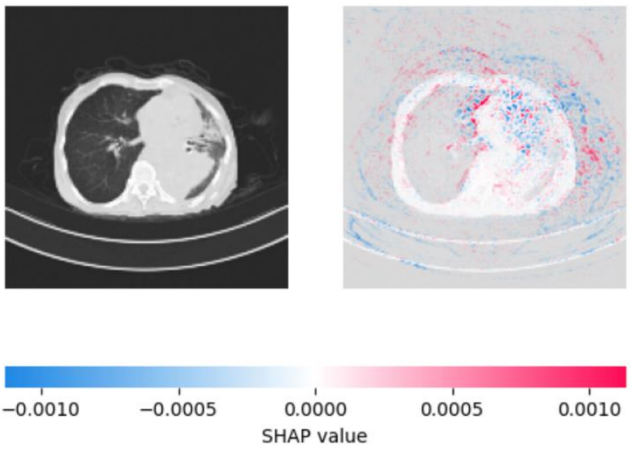

Figure 4: Based on the high SHAP value displayed in the explanation image (second image), we can say the CT image is diagnosed as Cancerous.

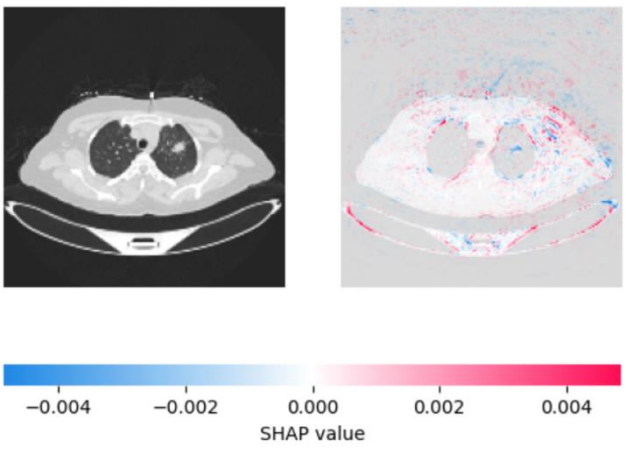

Figure 5: Based on the high SHAP value displayed in explanation image (second image), we can say the CT image is diagnosed as non-cancerous.

4.5. LIME

This study employed LIME to provide localized explanations for the model's predictions. LIME accomplishes this by perturbing the test image and generating a set of random perturbations in a matrix format consisting of ones and zeros. Each row in the matrix represents a perturbation, while the columns correspond to superpixels. A superpixel is labeled as "ON" if assigned a value of 1, indicating its relevance, or "OFF" if assigned a value of 0, indicating its lack of influence. The length of the vector displayed in the visualizations reflects the total number of superpixels in the image, which in our case is 150*150 perturbations. In Figure 6, the third image depicts the final perturbed image for the cancerous test image, while in Figure 7, the third image displays the perturbed image for the test image with standard conditions. These visualizations highlight specific regions within the image that play a crucial role in the classification process, emphasizing the portions that have a significant impact on determining the presence or absence of cancer. The localized interpretability provided by LIME enhances our understanding of the model's decision-making process, fostering transparency and trust in the model's predictions (Vedaldi and Soatto, 2008).

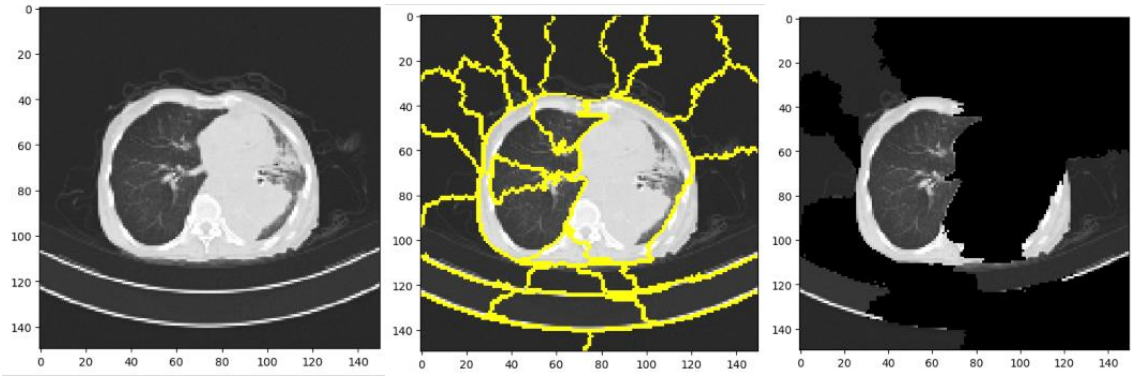

Figure 6: Interpretations generated by LIME for cancerous tissue. The leftmost is sample of cancerous CT images from the test image. The middle one represents super pixels generated from quick-shift segmentation to create perturbations. The rightmost represents the final perturbed image showing the presence of cancerous tissue.

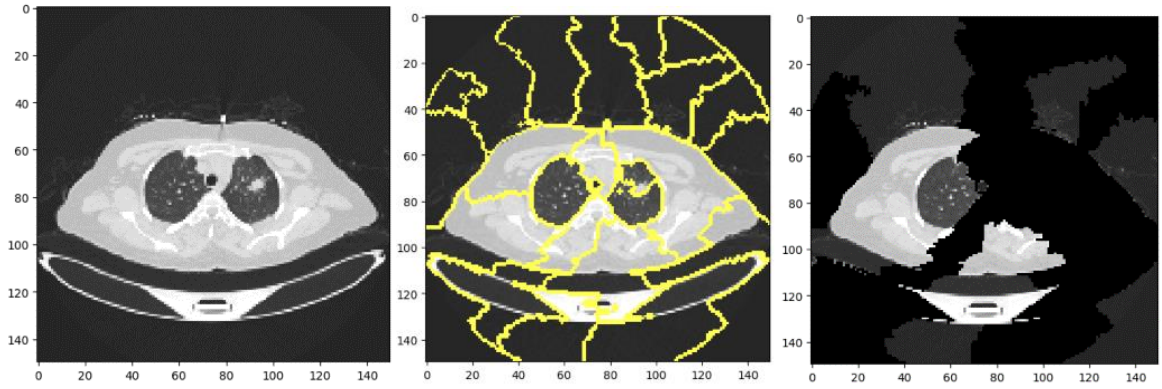

Figure 7: Interpretations generated by LIME for a normal image. Leftmost is a sample of the normal image from the test image. The middle one represents super pixels generated from a sample of the normal image from test image quick-shift segmentation to create perturbations. The last one, rightmost represents the final perturbed image for the normal image.

5. Conclusion and Further Direction

In conclusion, this study successfully developed a deep learning model for the classification of lung cancer with high accuracy on the provided dataset. However, future research should focus on scaling the model to larger datasets to ensure its performance and generalizability. Additionally, extending the model to classify different stages of lung cancer would enhance its clinical utility and assist in personalized treatment planning. Moreover, there is potential for further research in localizing cancerous regions within the lungs to improve targeted interventions. By addressing these aspects, we can advance the field of lung cancer detection and contribute to improving patient outcomes through early diagnosis and personalized treatment strategies. The interpretability methods of LIME and SHAP also played a crucial role in enhancing the transparency of the model's predictions and providing valuable insights into the critical features for classification. Overall, this study demonstrates the potential of deep learning models and opens avenues for further advancements in the field of lung cancer diagnosis.

References

Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). *Lung Cancer Detection using CT Scan Images. Procedia Computer Science,* 125, 107-114. https://doi.org/10.1016/j.procs.2017.12.016.

Amer, H.M., et al., *A Computer Aided Early Detection model of Pulmonary Nodules in CT Scan Image*, 7th International Conference on Software and Information Engineering -ICSIE '18. 2018, ACM: Cairo, Egypt. p. 81-86

Kumar, A., Fulham, M., Feng, D., & Kim, J. (2020). *Co-Learning Feature Fusion Maps From PET-CT Images of Lung Cancer. IEEE Transactions on Medical Imaging*, 39(1), 204-217. https://doi.org/10.1109/TMI.2019.2923601

Elnakib A, Amer H., et. al., *Computer aided detection system for early cancerous pulmonary nodules by optimizing deep learning features,* ACM International Conference Proceeding Series, 2019. P.75-79

Zhou, A.M.R.W. *Deep Learning for Categorizations of Lungs Cancer CT Image, IEEE/ACM International conference on Connected Health: Applications,* Systems and Engineering Technologies (CHASE). 2017. Philadelphia, PA, USA: IEEE.

Pawar, Vikul J., *Lung Cancer Detection System Using Image Processing and Machine Learning Techniques*, International Journal of Advanced Trends in Computer Science and Engineering, 2020. 9: p.5956-5963

Firdaus, Qurina, Sigit, et. al., *Lung cancer detection based on ct-scan images with detection features using gray level co-occurrence matrix (glcm) and support vector machine (svm) methods*, IES 2020 - International Electronics Symposium: The Role of Autonomous and Intelligent Systems for Human Life and Comfort, 2020. P:643-648

Rohit Y. Bhalerao, et. al., *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS).*, Institute of Electrical and Electronics Engineers. Madras Section, ISBN: 9781538695333

Bray F, et al. *Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin. 2021; 71:209-249.

Revel MP, et al. *Can computed tomography features predict respectability of lung cancer?* Diagn Interv Imaging. 2015; 96:1017-1025.

Saini S. *Challenges in imaging lung cancer.* J Nat Sci Biol Med. 2013; 4:8-14.

Ardila D, et al. *End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography*. Nat Med. 2019; 25:954-961.

Jacobs C, et al. *Deep learning in lung cancer detection.* Ann Transl Med. 2019; 7:590.

Vedaldi, A. and Soatto, S. (2008). *Quick shift and kernelmethods for mode seeking. Proceedings of European Conference on Computer Vision*, pages 705–718.

Chen IY, et al. *Explainable artificial intelligence for breast cancer detection from mammography*. Sci Rep. 2020; 10:16430.

Ribeiro MT, et al. *"Why should I trust you?": Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM; 2016. p. 1135

Gaur, L., Bhandari, M., Razdan, T., Mallik, S., & Zhao, Z. (2022). *Explanation-Driven Deep Learning Model for Prediction of Brain Tumor Status Using MRI Image Data. Frontiers in Genetics*, 13. Retrieved from https://www.frontiersin.org/articles/10.3389/fgene.2022.822666. doi: 10.3389/fgene.2022.822666. ISSN 1664-8021

Alyasriy, Hamdalla; AL-Huseiny, Muayed (2021), *"The IQ-OTHNCCD lung cancer dataset"*, Mendeley Data, V2, doi: 10.17632/bhmdr45bh2.2

Hany, M. (2020). *Chest CT-Scan images Dataset (Version 1)* [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images

Jin, X., Zhang, Y., & Jin, Q. (2016) "Pulmonary Nodule Detection Based on CT Images Using Convolution Neural Network." 2016 9Th International Symposium On Computational Intelligence And Design (ISCID). DOI: 10.1109/ISCID.2016.1053.

Aggarwal, T., Furqan, A., & Kalra, K. (2015) "Feature extraction and LDA based classification of lung nodules in chest CT scan images." 2015 International Conf