International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Visual Explainability-Driven DL framework for Lung Nodule Classification

I.Tejaswini, T.Thanmai, M.Apphia, M.Mohit, and A.Sagar

Anil Neerukonda Institute Of Technology and Sciences

Abstract. The large number of images and the subtle characteristics of pulmonary nodules make it challenging to detect the lung cancer from CT scans. Complex and small nodules may provide a less accurate diagnosis, and manual examination is time-consuming and result to variation among observers. Previous studies have shown excellent accuracy but often fall short in terms of organization and understandability. This paper introduces an automated and explainable deep learning framework for lung nodule classification. CNN is used to classify CT scans. We use consistent preprocessing methods to make sure that model is strong and delivers reliable results. Experiments conducted on public lung CT datasets show an accuracy of about 95%.

Keywords: Deep Learning · CNN · Machine Learning · Nodule Classification

1 Introduction

Lung cancer continues to be significant health issue worldwide, leading to a significant number of deaths related to cancer annually. Detecting pulmonary nodules early is crucial for improving survival rates; however, analysing lung CT scans is challenging due to the vast amount of imaging data and the subtle characteristics of nodules. These nodules capable to vary in size, shape, texture, and location. This variation makes manual diagnosis which subject to human error and time-consuming. As a result, there is an increasing need for automated and dependable computer-aided diagnosis (CAD) systems that can help radiologists accurately and consistently detect and classify lung nodules.

Convolutional neural networks [3] (CNNs) have significantly improved the medical image processing in our model by impulsively classifying lung nodules from the computed tomography [2] (CT) scans using contemporary deep learning architectures to keep the focus on the anatomically relevant body region, this work proposes a holistic deep learning pipeline for the automated segmentation of the lungs the standardization of image preprocessing.

Through the usage of saliency maps and other visual explanation tools, it is possible to highlight the features that guide the CNN in making predictions. The results can be assessed by clinicians, and diagnostic reports may be created through the upload of the images and the usage of a web-based GUI. The usage of public lung CT image datasets has recorded an accuracy of about 95%.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

In the past few years, deep learning has recorded outstanding advancements in the area of computer vision, automatically identifying the most important features in images. This is contrast with the traditional approaches that relied on the manual specification of the features. CNNs have the capability of learning spatial and texture features in images. However, the CNN is a "black box," and there is a need to increase the importance of the development of explainable deep learning models with ability to achieve the high accuracy and visual insights.

2 Literature Survey

Detection of lung nodules is one of the prominent issues, especially with the increase in lung cancer cases, emphasising the importance of early detection using CT scans. Initially, conventional machine learning models such as Support Vector Machine[6] (SVM), Random Forest[10], and Naive Bayes[4] were employed, which required hand-crafted feature extraction models such as texture, shape, and intensity features.

Liu et al. [1] delivered a real-time deep learning- based detector called as STBi-YOLO, which is specifically designed to detect the lung nodules using CT scans. This method is based on the YOLO framework, where stochastic pooling is incorporated, enabling a feature pyramid to effectively utilize the multi-scale information. Evaluation is done using the dataset containing annotated CT scan images, where performance is measured using detection real-time inference speed and accuracy. This method is initiate to have improved performance compared to other YOLO variants, where focus is only on detecting nodules, not classifying them.

Nguyen et al. [2] proposed a framework using Faster R-CNN, where Adaptive Anchor Boxes were incorporated to effectively detect nodules of various sizes. Residual connections were also included to reduce false detections, where datasets were used to show improved sensitivity using various lesion sizes. However, this framework is limited to detecting nodules, not classifying them.

Saha et al. [3] created Lung-AttNet, an attention-based convolutional neural network. In order to identify discriminative areas in CT scans, their approach uses lightweight attention modules. Additionally, predictions are interpreted using visualisation methods like Grad-CAM and LIME. When compared to conventional CNN designs, the experimental results show better classification performance. However, real-time deployment is difficult due to the lack of localisation and the higher processing burden.

Shakir et al. [4] introduced a radiomics-driven Bayesian inversion approach for malignancy and stage prediction. The system extracts from quantitative imaging biomarkers, including shape irregularity, entropy, and the surface features, which are often used for probabilistic inference. Results indicate reliable stage estimation under the controlled conditions. Nevertheless, this approach is highly depends on accurate segmentation and can degrade when boundaries are unclear.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Javed R. [5] conducted a systematic mapping study of deep learning for lung cancer classification, structuring their analysis into the three phases: Study Preparation, Conduct of Study, and Analysis and Results. The study was preceded by a comprehensive Systematic Literature Review (SLR) to classify methodologies and identify deep learning application trends for a lung cancer detection. It incorporated with state-of-the-art practices, focusing on newer advancements in CNN Architectures.

3 System Architecture

Lung CT images from the IQ-OTHNCCD dataset are prepared for the CNN input by resizing, normalizing, and formatting them according to the suggested system. Random oversampling balances the class distribution before splitting the data into validation, testing sets and training.A multi-layer convolutional neural network uses convolution, pooling, and dropout operations to extract spatial features. After flattening, the learned features move to fully connected layers for the classification[11]. The model's performance is actually evaluated during training with the Adam optimizer and categorical cross entropy loss, which allows seamless communication between backend inference modules and frontend interfaces for the visualization and reporting.

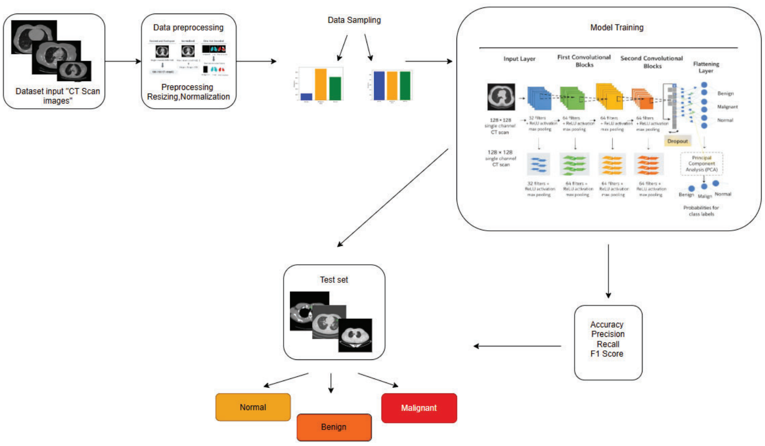

Fig. 1. Workflow of proposed methodology

4 Dataset

The IQ-OTHNCCD Lung Cancer Dataset is used in this study to classify lung cancer using images obtained through a CT scan. This dataset includes grayscale images of CT scans classified into three types: benign, malignant, and normal,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

as classified by experts. This dataset has variations in intensity, contrast, and structure, which are representative of real-world scenarios. The IQ- OTHNCCD Lung Cancer Dataset is used in this study to classify lung cancer using images obtained through a CT scan. This dataset includes grayscale images of CT scans classified into three types: benign, malignant, and normal, as classified by experts. This dataset has variations in intensity, contrast, and structure, which are representative of real-world scenarios.

5 Data Preprocessing

All images were modified to a standard resolution of 128 x 128 pixels and then reshaped into a single-channel format for use with a convolutional neural network. The pixel values were also normalized to a range of [0,1] by dividing by 255. The class labels were converted into parts like one-hot encoded vectors for the multiclassification.

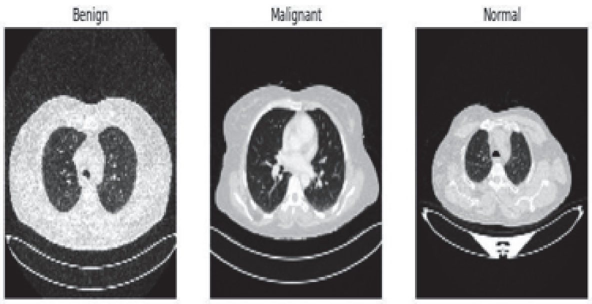

Fig. 2. Dataset images

6 Data Sampling and Dataset Splitting

To tackle the problem of imbalanced data in the benign, malignant, and normal classes, the Random Oversampling (ROS) technique was employed on the flattened data to attain a balanced data distribution. The data is then reshaped to the original image form and split into validation, training and test sets. The data is initially split into 70% for testing and the remaining 30% for training, with 20% of the 30% being used for validation

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

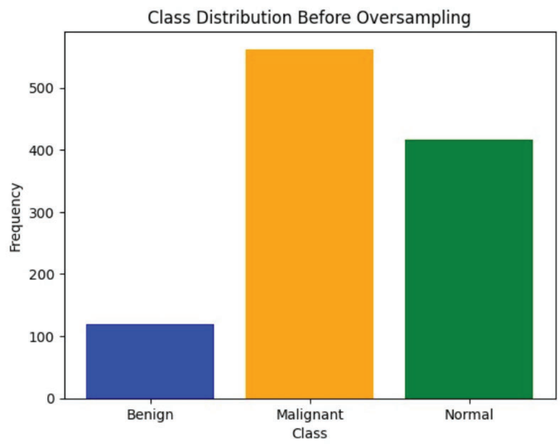

Fig. 3. Original class distribution

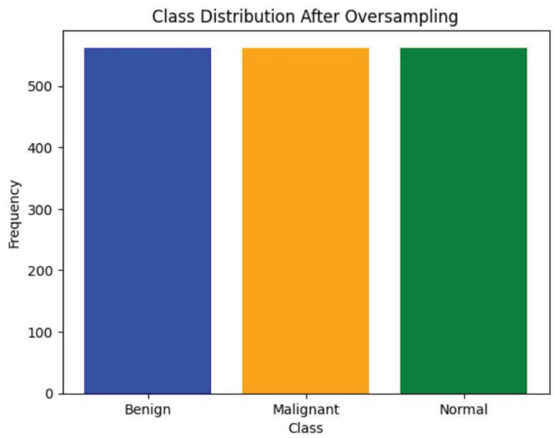

Fig. 4. Class distribution after Random oversampling (ROS)

7 Model Architecture

Convolutional Neural Network: A convolutional neural network is used to classify lung CT scan images into benign, malignant, and normal classes. The CNN is trained to effectively learn spatial and structural features from grayscale CT scan images of size $128 \times 128 \times 1$. The CNN architecture includes pooling, dropout, convolutional and fully connected layers.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

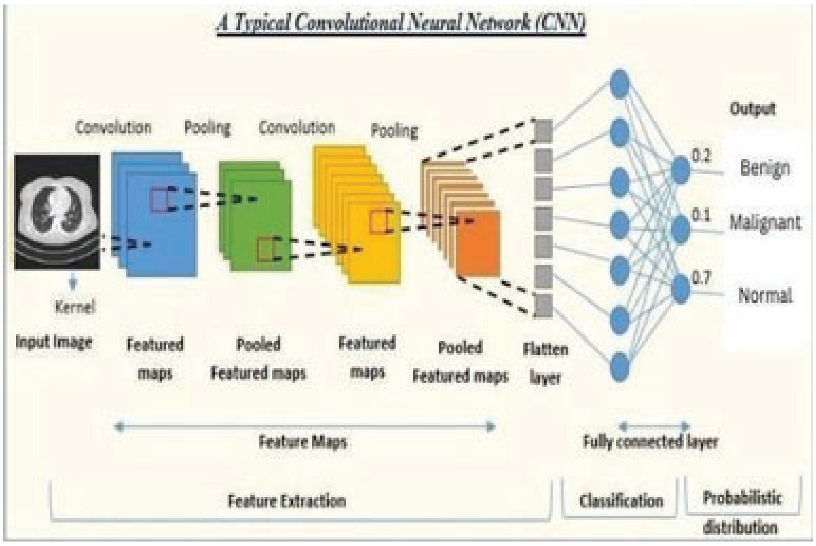

Fig. 5. Proposed CNN Architecture

Input Layer: The input layer takes pre- processed grayscale CT scan images of size 128 × 128 × 1. This is a standard size, ensuring consistency throughout all images.

First Convolutional Block: The first convolutional layer employs 32 filters of the size 3 × 3 with ReLU activation function. The goal of this layer is focused on learning low-level characteristics such as contours, edges and basic texture features present in lung CT scan images. A max pooling layer is used to reduce spatial dimensions of size 2 × 2 while preserving essential features. In order to avoid overfitting, a dropout layer is included with a dropout rate of 0.01 by randomly dropping a few neurons while training. Second Convolutional Block: The second convolutional layer uses 64 filters of size 3 x 3 with ReLU activation function[1]. This layer is used to learn more abstract features by utilising the features learned in the previous block. A 2 x 2 max pooling layer is used to reduce the feature map size. A dropout layer is then used with a dropout rate of 0.01. Third Convolutional Block: The third convolutional block uses the same parameters as the previous block: 64 filters with ReLU function. This block is used to learn complex features that will help the network distinguish between different lung tissue types. The max pooling and dropout layers are again used to avoid overfitting.

Flattening Layer: The feature maps learned in the convolutional blocks need to be crushed into a single dimensional vector. This is done in the flattening layer. Fully Connected Layer: A dense layer which contains 128 neurons and ReLU anchor activation function is useful to learn the features from the flattened data. This is the final

feature extraction block. A dropout layer of 0.01 dropout rate is used to avoid overfitting.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Output Layer: This layer uses three neurons, one for each of the three classes. Softmax activation function is used in this layer so the network can assign a confidence value to each prediction.

Training Configuration and Batch Processing: The CNN model is trained with the help of Adam optimiser, which adaptively adjusts the step size to ensure stable and efficient convergence. Tier cross entropy is used as the loss function, as the task include multi-class classification. Training is arranged for 10 epochs with a batch size of 32, meaning that 32 images are processed simultaneously in each iteration. Batch-wise training improves computational efficiency, stabilises gradient updates, and allows effective utilisation of available memory. Model performance is monitored using validation data during training.

8 Principal Component Analysis

In this context, PCA is used as a dimensionality reduction technique to extract important features using CNN feature representations. This reduces the dimensions while preserving the most important features in the feature space. This makes the feature space computationally efficient and reduces the redundancy. PCA[9] is a technique for dimensionality reduction that encompasses a series of steps such as feature extraction, covariance matrix calculation, eigenvalue decomposition, transformation and selection of principal components.

The first move in PCA is feature extraction using the CNN model. Here, the CNN feature

extractor is used to generate high-dimensional feature representations for the input CT scan images. The CNN feature extractor is composed of all the layers in the CNN architecture except for the last two fully connected layers[3]. This is because the spatial information is preserved in the last two layers during CNN training. Once the features are extracted for the training and test datasets, they are used as input to the PCA algorithm. PCA normalises the feature vectors. This makes achieved by subtracting the mean value of each feature for every data point in the feature space. This is to maintain the feature space at a neutral position in the coordinate system, i.e., at the origin. This is because any existing bias in the feature space due to pixel intensity may be reduced.

After mean aligning, the covariance matrix is calculated. The covariance matrix maintains connection between the various features, showing the relationship as they vary with respect to one another. High covariance between the features implies redundancy, while low covariance implies that the features are independent of one another. PCA[9] identifies the direction in which the data shows maximum variance using the covariance matrix.

To find the main components, the covariance matrix is decomposed using the eigenvalue decomposition method, which gives the eigenvalues and the corresponding eigenvectors. The eigenvectors are the directions in the feature space in which the variance of the data is maximum, while the eigenvalues are the magnitudes of the corresponding eigenvectors. The eigenvectors with the max-

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

imum eigenvalues contain the maximum amount of descriptive data about the variance in the data.

Principal component selection is the selection of the eigenvectors with the maximum eigenvalues, i.e., PCA selects the first two components (n-components = 2) from the CNN features, which maps the data onto a space with two dimensions.

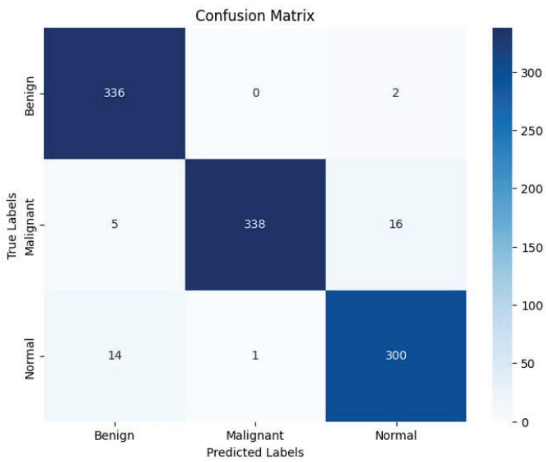

Fig. 6. Confusion matrix

Table 1. Model Comparison Table

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Naive Bayes	79.17	80.51	79.17	78.90
Decision Tree	81.37	81.36	81.37	81.34
KNN	83.15	83.46	83.15	83.24
SVM	90.01	90.17	90.01	90.06
LDA	90.86	91.26	90.86	90.94
CNN (Proposed)	95.09	95.19	95.09	95.09

9 Results

When we consider the performance of the suggested CNN classifier for the lung nodule images, we can see that the accuracy, precision, recall, and F1-score are the key performance indicators. These results are demonstrated in the Evaluation Metrics Bar Graph. The accuracy of the model is around 95%, which is a clear

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

indicator of good performance for the classification of images. This also shows that the CNN is successfully extracting useful spatial features from the images. Similarly, a precision of about 95% is an indicator of fewer false positive instances and better positive predictions. The recall is about 95%, which is a clear indicator that most of the malignant cases are being identified. This is a great performance, especially when considering the early diagnosis of lung cancer. The F1-score is also estimated at about 95%, which is a clear indicator of a good balance of precision and recall. This is true for all classes, and when the results are above 90%, it is a clear indicator of the robustness of the model.

10 Conclusion

This paper presented a CNN-based system for classifying lung nodules using the IQ-OTHNCCD Lung CT dataset, dividing them into Benign, Malignant, and Normal classes. After precise preprocessing, using a Random Over Sampling method, and a well-configured CNN, the system achieves remarkable performance: 95.09% accuracy, 95.19% precision, 95.09% recall, and 95.09% F1 score. Its high recall rate is evidence of its effectiveness in a medical context, where early identification of malignant nodules is critical, while its strong F1 score is evidence of fair classification performance. Thus, this system is a highly effective tool to assist radiologists in lung cancer diagnosis, showing considerable potential for future upgrades.

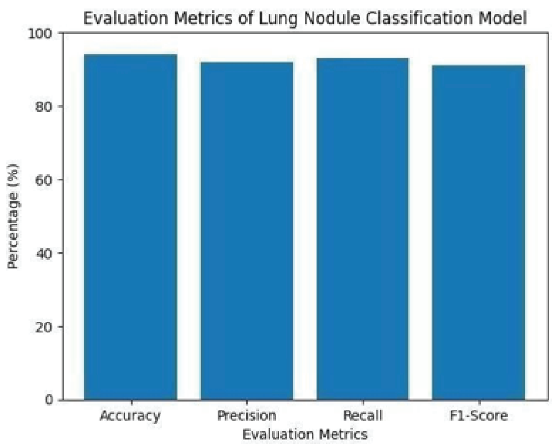

Fig. 7. Evaluation Metrics

References

1. Liu, K.: Stbi-yolo: A real-time object detection method for lung nodule recognition. IEEE Access **10**, 75385–75394 (2022)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

2. Nguyen, C.C., et al.: Pulmonary nodule detection based on faster R-CNN with adaptive anchor box. IEEE Access **9**, 154740–154751 (2021)
3. Saha, C., et al.: Lung-attnet: An attention mechanism-based CNN architecture for lung cancer detection with federated learning. IEEE Access (2025)
4. Shakir, H., et al.: Radiomics-based Bayesian inversion method for prediction of cancer and pathological stage. IEEE Journal of Translational Engineering in Health and Medicine **9**, 1–8 (2021)
5. Javed, R., et al.: Deep learning for lung cancer detection: a review. Artificial Intelligence Review **57**(8), 197 (2024)
6. Li, Y., et al.: Machine learning for lung cancer diagnosis, treatment, and prognosis. Genomics, Proteomics & Bioinformatics **20**(5), 850–866 (2022)
7. Saleh, A.Y., et al.: Lung cancer medical images classification using hybrid CNN-SVM (2021)
8. Swain, S., et al.: Innovative approach for early detection and risk assessment of lung cancer using machine learning techniques. In: International Conference on Intelligent Computing and Sustainable Innovations in Technology (IC-SIT), IEEE (2024)
9. Bhargav, S.P., et al.: Impact of PCA on lung cancer dataset classification: A comparative analysis of machine learning models. In: International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), IEEE (2024)
10. Nair, S.S., Devi, V.N.M., Bhasi, S.: Enhanced lung cancer detection: Integrating improved random walker segmentation with artificial neural network and random forest classifier. Heliyon **10**(7) (2024)
11. Nimmagadda, S.M., et al.: Lung cancer prediction and classification using machine learning algorithms. In: International Conference on Expert Clouds and Applications (ICOECA), IEEE (2024)
12. Nageswaran, S., et al.: Lung cancer classification and prediction using machine learning and image processing. BioMed Research International (2022)
13. Das, S., Majumder, S.: Lung cancer detection using deep learning network: A comparative analysis. In: Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE (2020)
14. Slack, D., et al.: Explaining machine learning models with interactive natural language conversations using TalkToModel. Nature Machine Intelligence **5**(8), 873–883 (2023)
15. Shravya, V., et al.: Encouraging discriminative attention through contrastive explainability learning for lung cancer diagnosis. IEEE Access (2025)