IJITS

International Journal of Intelligent Technologies and Systems

Vol. 2, Issue 1, September-October 2025

Efficient DL-Based Classification Model for Lung Cancer and Nodule Discrimination

[1]John A*, [2]Sumathi S

john_mcte@ieee.org, sumathi_s@annaunivtut.in

[1]Associate Professor, Machine Learning, Military College of Telecommunications and Engineering, Government of India, Ministry of Defense, Indore MP
[2]Assistant Professor, Department of Computer Science and Engineering, University V.O.C. College of Engineering, Anna University – Thoothukudi Campus, India.

OPEN ACCESS
Research Article

Received: 12 September, 2025
Revised: 04 October, 2025
Accepted: 20 October, 2025
Corresponding Author: John A
john_mcte@ieee.org

Publisher:
Aarambh Quill Publications

ABSTRACT

Timely identification of tumors is essential for reducing cancer-related deaths and improving therapeutic outcomes, particularly in the case of lung cancer, which continues to be a major global health concern. Conventional image-based diagnostic approaches are often limited by errors in distinguishing benign from malignant nodules, inconsistencies among radiologists, and difficulties in processing large-scale medical datasets. To overcome these challenges, this study introduces a Efficient Deep Convolutional Neural Network (EDCNN) framework designed for automated tumor detection and classification, with emphasis on lung cancer screening and recognition of non-malignant nodules. The approach incorporates preprocessing of CT images through denoising, intensity normalization, and augmentation, followed by hierarchical EDCNN feature extraction and classification to separate benign from malignant growths. The primary goal is to improve diagnostic precision, minimize false alarms, and provide an effective decision-support mechanism for clinicians. Experimental analysis indicates that the proposed EDCNN achieves higher performance than traditional machine learning and baseline deep models, yielding significant gains in accuracy, precision, sensitivity, and F1-score. These findings demonstrate the promise of deep learning for delivering robust, efficient, and accurate lung cancer detection in real-world clinical settings.

Keywords: Tumor Detection; Lung Cancer Prediction; Deep Convolutional Neural Network; Benign and Malignant Nodules; Medical Image Classification; Computer-Aided Diagnosis; CT Scan Analysis; Automated Cancer Screening; Image Pre-processing

1. INTRODUCTION

Lung cancer still remains among the most widespread and lethal malignancy in the world, and it produces a significant number of cancer-related deaths annually. Early and accurate diagnosis of tumors in the lung is important since the prognosis and the success of the treatment is highly affected by the disease stage. The diagnosis and the screening of lung cancer has been dominated by advanced imaging modalities, in particular, CT. However, CT images may also be challenging to interpret due to the presence of the benign and malign nodules presentations which are subtle and the nodule appearance may be different and in clinical practice the volume of imaging information to be evaluated may be voluminous.

The majority of the rudimentary diagnosis techniques rely on the manual interpretation by a radiologist, yet it is a tedious method and prone to human factors. Misclassification of nodules is not only a delay in the treatment procedure but also a source of mayhem when patients are informed that they are having benign nodules whereas they are having malignant nodules. In addition, the level of inter-observer variability is also present as radiologists may find themselves having a different interpretation of the same image. These shortcomings underscore the need to have automated systems that can assist clinicians in their tumor evaluation by delivering uniform, dependable and rapid findings.

https://doi.org/10.64909/IJITS.2025.2103

Over the past few years, the nowadays use of computational approach in detecting lung cancer on the basis of medical imaging has been on the rise. Conventional machine learning paradigms have been somewhat successful in this regard, however, typically at the price of necessitating manual feature extraction tools that do not typically generalise to the large number of imaging scenarios. To address these issues, researchers are exploring novel computational algorithms capable of automatically discovering complex patterns in the raw imaging data themselves, so that the demand for manual feature generation will be reduced. Deep convolutional neural networks (EDCNNs) have been demonstrated in the given background to be a promising solution to learn high level representations of a tumor signature, and to distinguish between benign and malignant nodules with the accuracy.

The aim of this paper will be to design a new EDCNN based architecture to predict lung cancer and tumor classification. By integrating the preprocessing methods and stacked feature extraction system, the foundation is expected to improve readability, reduce ambiguity in diagnosis and serve as an effective decision support system in the hands of the clinicians. Finally, the purpose is to minimize the number of false positives, the accuracy of diagnosis, and enabling the screening and management of cancer in clinical practices in patients.

2. RELATED WORKS

In their study, Heuvelmans et al. [1] examined the application of deep learning towards distinguishing benign lung nodules, with a stronger focus on clinical implications of lowering the false positive rates in lung cancer screening. Their model demonstrated the worthiness of applying automated techniques in order to enhance the level of prediction among radiologists in the event of indeterminate nodules. Concentrating on region based feature extraction in lung cancer detection, Suresh and Mohan [2] introduced ROI-based CNN. Their methodology showed that narrow-based learning is useful in reducing the cost of computation, and increasing the rate of true positives rate.

Tusher et al. [3] synthesized the benefits of the convolutional feature extraction approach and dense network learning, and they designed a computer-aided system of early lung cancer diagnosis. This architecture was not a bad one to fill the gap between the model of computation and clinical workflow. Saha et al. [4] suggested the Advanced Deep LungCare Net, a powerful next generation framework to design lung cancer prediction which shows the capacity to increase the generalization potential amid various imaging datasets and attain strong performance even in difficult diagnoetic conditions.

Another DGMM (directional geometrical and mixed moments)-RBCNN (radial basis function neural classifier) method was also applied by Jena et al. [5] in lung cancer detection and classification. They applied probabilistic modelling and convolutional layers, it led to superior decision limits of malignant and benign nodules. The CNN ensembles were applied by Paul et al. [6] to the prediction of 2-year prospected malignancy. The ensemble approach achieved great prediction accuracy due to integration of the strengths of different CNN models.

Paez et al. [7] suggest the model of 1 D convolutional neural network to classify nodules longitudinally. They paid attention to temporal CT data analysis and proved that time series of progression information would significantly improve the classification outcomes of indeterminate nodules. Riquelme and Akhloufi [8] were focused on the processes of nodule detection using deep learning, and classification of CT scans. Their result indicated that CNNs are scaled to high data, and also they gave an insight into the tradeoffs between sensitivity and specificity.

The article by Huang et al. [9] used a combination of deep learning and radiomics in the characterization of pulmonary ground-glass nodules. Their approach automatically acquired baseline CT features and received encouraging results to the benign and malignant case detection. Bhaskar et al. [10] took advantage of deep learning structure and image improvement to perform lung models. To externally validate the CNN-based malignancy prediction models, Baldwin et al. [11] tested the models across several institutions and proved that good learned models could be generalized to new patients. In their study, Tusher et al. [12] confirmed their hybrid CNN-DNN framework of early diagnosis, which, in turn, once again justified the clinical translation of deep architectures to lung cancer screening.

The article by Saji et al. [13] has examined several types of deep learning in detecting, classifying, and predicting lung cancer. In their survey they provided a discussion of the state-of-the-art methods and discussed the advantages and disadvantages of CNN-based models. To predict the prognosis of patients using CT data collected in multiple centers, Mukherjee et al. [14] proposed a low-end CNN to predict the prognosis of the patients. They concentrated on interpretable and clinically useful and predictive models.

Lastly, Wang et al. [15] suggested an end-to-end deep CNN model to identify the occurrence of lymphovascular invasion in non-small cell lung cancer using preoperative CT scans. They demonstrated that deep features could be applied towards classification, but also as prognostic markers that would be applied in managing the treatment.

3. MATERIALS AND METHODS

This work uses publicly available CT image datasets which include the LIDC-IDRI dataset where expert-annotated scans are provided attributes like benign or malignant and the the Visceral CT that only includes benign, malignant and TMM labels in the scans. Before networks are developed, the images undergo several preprocessing steps.

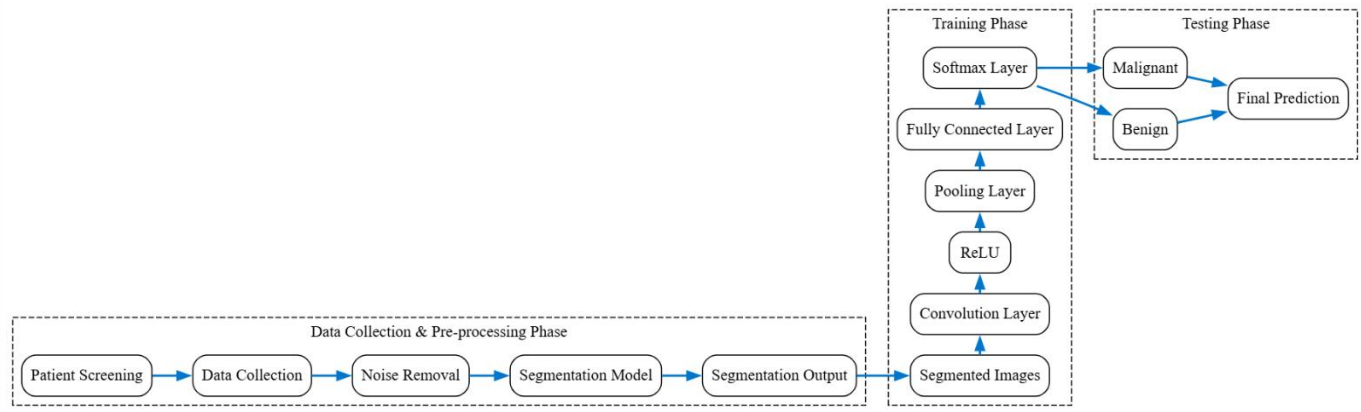

Figure 1: Overall architecture of the proposed EDCNN-based lung cancer detection and classification framework

These involve Gaussian filter which helps to reduce random noise, intensity normalization which helps to keep pixel intensity distribution consistent between scans and augmentation such as rotations, horizontal/vertical flips and scaling to help promote generalisation and limit overfitting.

Problem Formulation

Let $D = \{I_i,\ yi\}_{i\ =\ 1}^{N}$ be a CT-image dataset where yi ∈ {0, 1} denotes benign (0) or malignant (1) nodule. Each image is preprocessed by intensity clipping and z-score normalization: $X_i = (L_i - \mu)/\sigma$, where μ, σ are the dataset mean and standard deviation. A Deep Convolutional Network $f_\theta : R^{HxWxC} \rightarrow [0, 1]$, with parameters θ outputs the malignancy probability $p_i = f_\theta(X_i) = \sigma(z_i)$, with z_i the network logit and $\sigma(t) = 1/(1+e^{-t})$. The primary objective is to learn θ that minimizes the class-weighted cross-entropy with L_2 regularization:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}[w_1 y_i log p_i + w_0(1-y_i)\log(1-p_i)] + \lambda||\theta||_2^2 \quad (1)$$

where $w1 = \frac{N}{2N_1}$ $w0 = \frac{N}{2N_0}$ address class imbalance given counts N1, N0, and λ > 0 controls regularization. If hard labels are needed, the decision rule is $\hat{y}i = 1\{pi \geq \tau\}$ with tunable threshold τ (default τ = 0.5). For multi-class tumor types (optional), extend yi ∈ {1, …, K} and use softmax $p_{ik} = \frac{e^{2ik}}{\sum_{j=1}^{K} e^{2ij}}$ with loss $L(\theta) = -\frac{1}{N}\sum_i \sum_k \alpha_k 1\,[y_i = k] log p_{ik} + \lambda||\theta||_2^2$.

Performance is quantified by:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}[w_1 y_i log p_i + w_0(1-y_i)\log(1-p_i)] + \lambda||\theta||_2^2 \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (3)$$

$$Precision = \frac{TP}{TP + FP}, \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F_1 = \frac{2Precision.Recall}{Precision + Recall} \quad (6)$$

and ROC–AUC from {(pi, yi)}. The learning problem is thus θ* = arg minθ L(θ), yielding a EDCNN that detects and classifies lung nodules while estimating calibrated malignancy probabilities pi for reliable benign/malignant prediction.

Dataset description

This study has used the LIDC-IDRI dataset that is one of the most popular benchmarks datasets in the study of lung cancer research. It is a collection of more than 1,000 thoracic CT scans in various institutions, which are publicly shared in the Cancer Imaging Archive. The scans have several lung nodules that were annotated and categorized as either benign, malignant, or indeterminate by up to four experienced radiologists. In this task, a pre-processed sub set of the data was chosen by discarding scans of low quality or inconsistent scans and identifying regions of interest (ROIs) that corresponded to lung nodules as 2D patches. In order to counteract the effect of the disparity in classes, disk-balancing techniques were applied to increasing the samples to approximately 5,000 with two classes carried (benign (0), malignant (1)) of control-coded nodules. The dataset serves as a valid data source to the training and testing of the developed Deep Convolutional Neural Network in the detection and classification of tumors with prediction of lung cancer.

Table 1: Dataset Features

Patient ID	Scan ID	Slice No.	Nodule Diameter (mm)	ROI Patch Size	Label
LIDC_001	CT_001	135	8.5	64 × 64	Benign (0)
LIDC_002	CT_007	210	15.2	64 × 64	Malignant (1)
LIDC_003	CT_014	98	5.7	64 × 64	Benign (0)
LIDC_004	CT_021	185	22.6	64 × 64	Malignant (1)
LIDC_005	CT_030	142	12.1	64 × 64	Malignant (1)

The sample data is used to indicate the arrangement of entries in the database. Each Patient possesses distinct Patient ID and correlative CT Scan ID. Slice No is the axial slice of the CT scan where the nodule is located and Nodule Diameter is the diameter of the nodule in mm. ROI patches of a fixed size (64 × 64 pixels) are clipped around the nodule to enter the data to EDCNN. Finally, and not the least, both cases are marked by Benign (0) and Malignant (1), as per the annotation of radiologists. This kind of structured information also provides the systematic training and testing of the model proposed in the detection of tumors and the prediction of lung cancer.

Preprocessing

To prepare the CT scan images for tumor detection and classification, ensure data consistency, noise reduction, and improved learning efficiency.

1. Noise Reduction

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right) \quad (7)$$

$$I_{filtered}(x,y) = (I * G)(x,y) \quad (8)$$

where * denotes convolution and I is the original CT image.

2. Normalization

Intensity normalization was performed using z-score normalization to ensure uniform pixel intensity distribution:

$$I_{norm}(x,y) = (I_{filtered}(x,y) - \mu) / \sigma \quad (9)$$

where μ and σ are the mean and standard deviation of pixel intensities in the dataset.

3. Resizing and ROI Extraction

To maintain uniformity for deep learning input, nodules were extracted as Region of Interest (ROI) patches and resized to a fixed dimension (e.g., 64 × 64):

$$ROI(x,y) = Resize(I_{norm}(x,y), H \times W) \quad (10)$$

where H and W are the target height and width.

- **Rotation:**

$$I_{rot}(x,y) = I_{norm}(x\,cos\theta - y\,sin\theta, x\,sin\theta + y\,cos\theta) \quad (11)$$

where θ is the rotation angle.

- **Flipping (horizontal/vertical)**

$$I_{flip}(x,y) = I_{norm}(W - x, y) \;\; or \;\; I_{norm}(x, H - y) \quad (12)$$

- **Scaling**

$$I_{scale}(x,y) = I_{norm}(sx, sy) \quad (13)$$

where **s** is the scaling factor.

After pre-processing, the dataset becomes noise-reduced, intensity-normalized, balanced through augmentation, and structured into uniform ROI patches ready for input into the Deep Convolutional Neural Network (EDCNN).

EDCNN ARCHITECTURES

The EDCNN architecture takes into account the principles of local receptive fields, weight sharing, as well as sub-sampling, and all of these factors make them resistant to changes in scale, position, and minor distortions. In each network, map features are formed by convincing the input image with learner filters, introducing parts of bias, and non-linear activation functions. It is through these operations that the network is able to sequentially learn low and high-level characteristics of images.

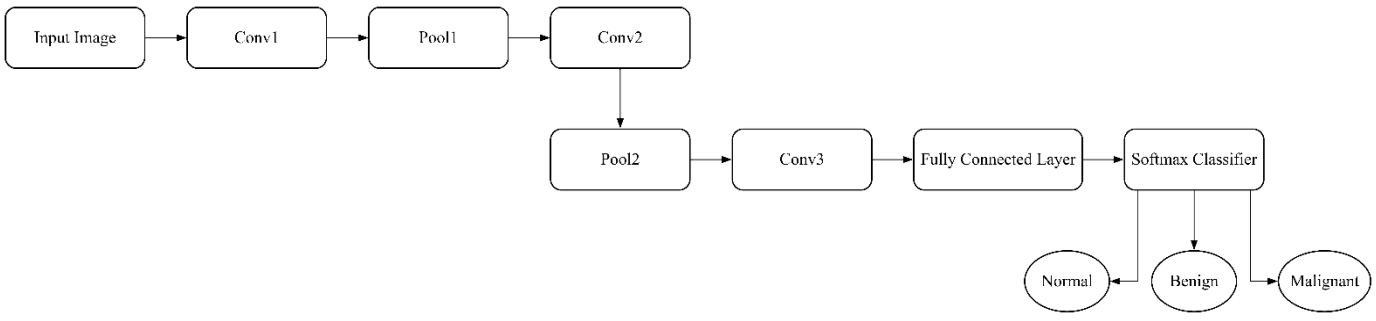

Figure 2: Proposed EDCNN Architecture

Figure 2, where "Conv" denotes a convolutional layer, "Pool" represents a pooling operation used for dimensionality reduction, and "FC" refers to fully connected layers responsible for classification. Together, these modules form the basis for hierarchical feature extraction and final decision-making in the detection of tumor nodules.

$$Relu\,(z) = \begin{cases} z & if > 0 \\ 0 & otherwise \end{cases} \tag{14}$$

Classification

In the process of typing medical images, the general objective is to isolate the features and classify them into the right category hence making sure that diagnosis is accurate. AUC is used as a general indicator of the extent to which a model is dependable and this measure is closer to 1 the more accurate the model.

To segment, the better random walker algorithm is popularly applied and the segmented area is analyzed furthermore with the classifiers, e.g., ANN and RF. RF is a combination of predictions made by various decision trees to achieve robustness and on the other hand ANN uses weighted connections between the nodes to make predictions and model the relationships among the features. The two approaches are characterized by their flexibility and predictive ability.

Extraction of features comprises shape description (alignment, confinement, solidity) and textural properties of the Gray-Level Co-occurrence Matrix (GLCM), that is, contrast, homogeneity, cluster prominence, cluster shade, and dissimilarity. In addition, Single-Level Discrete 2D Wavelet Transform is used to normalize intensity values as well as calculates principal component coefficients, which enhances the accuracy of the classification.

Feature Extraction and Classification

A classifier aims to discover relationships between extracted features F = {f_1, f_2, …., f_n} from an image I, and infer the class label $y\epsilon\{0,1\}$.

1. Evaluation metrics

$$AUC = \int_0^1 TPR(FPR)d(FPR) \tag{15}$$

2. Image Segmentation (Enhanced Random Walker Method)

The Random Walker (RW) segmentation algorithm computes the probability P_i^c that a pixel *i* belongs to class c based on diffusion over a weighted graph:

$$P_i^c = \frac{\sum_{j\epsilon N(i)} w_{ij} P_i^c}{\sum_{j\epsilon N(i)} w_{ij}} \tag{16}$$

3. Feature Extraction

- **Shape Features:** Alignment (A), Enclosure (E), Solidity (S).

$$S = \frac{Area\ of\ Nodule}{Convex\ Hull\ Area} \tag{17}$$

- **Texture Features (GLCM-based):** Contrast, Homogeneity, Clustre Prominence, Cluster Shade , and Dissimilarity. For an image gray-level co-occurrence matrix P(i,j):
- Contrast:

$$Contrast = \sum_{i,j} (i-j)^2 P(i,j) \tag{18}$$

- Homogeneity:

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \tag{19}$$

- Dissimilarity:

$$Dissimilarity = \sum_{i,j} |i-j| P(i,j) \tag{20}$$

- **Wavelet Features (Single-Level 2D Discrete Wavelet Transform – DWT):**

$$I(x,y) \xrightarrow{DWT} \{A,H,V,D\}, \tag{21}$$

which are then normalized and reduced using Principal Component Analysis (PCA):

$$Z = W^T(X-\mu),$$

where W is the eigenvector matrix, and Z are the principal component coefficients.

4. Classifiers

- **Random Forest (RF)**
  Combines predictions of multiple decision trees T_k:

$$\hat{y} = mode\{T_1(x), T_2(x), \dots, T_k(x)\} \quad (22)$$

  where each T_k is trained on a bootstrap sample and a

Algorithm: EDCNN for Tumor Detection & Classification with Benign Nodule Identification

Inputs: CT image-ROI pairs $D = \{(I_i, y_i)\}_{i=1}^{N}, y_i \epsilon \{0(benign), 1(maglignant)\}$

Outputs: Malignancy probability $p_i, class\ \hat{y}_i \epsilon \{0,1\}$

1. Preprocessing

1. Denoise (Gaussian): $I_f = I * G_\sigma$.
2. Z-score normalize: $X = \frac{I_f - \mu}{\sigma}$
3. Resize ROI: $X \in R^{HxW}$ (or H x W x C if multi-window).

2. Network (Backbone → Classifier)

For layer *l*:

Convolution (+ bias):

$$Z_k^{(l)} = W_k^{(l)} * A^{(l-1)} b_k^{(l)} \quad (23)$$

BatchNorm:

$$BN(Z) = \gamma \frac{Z - \mu B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta \quad (24)$$

Activation (ReLU): $A^{(l)} = \max(0, BN(Z^{(l)}))$

Pooling (max):

$$A_{pool}^{(l)}(u, v) = \underset{(m,n)\epsilon W}{max} A^{(l)}(su + m, sv + n) \quad (25)$$

After L blocks, apply Global Average Pooling (GAP) to get feature vector $h \in R^d$.

Classifier head (logit): $z = w^T h + b$

Probability (sigmoid): $p = \sigma(z) = \frac{1}{1+e^{-z}}$

(For multi-class K: $p_k = \frac{e^{zk}}{\sum_j e^{zj}}$)

3. Objective (Class Imbalance-Aware)

Weighted Binary Cross-Entropy + L_2:

$$L = -\frac{1}{N}[w_1 y_i log p_i + w_0(1 - y_i)\log(1 - p_i)] + \lambda||\theta||_2^2, w_c = \frac{N}{2N_c} \quad (26)$$

(Optional, for hard imbalance) Focal loss with $\Upsilon > 0$:

$$L_{focal} = -\frac{1}{N}\sum_i [w_1 y_i\,(1 - p_i)^\gamma\, log p_i + w_0(1 - y_i)\, p_i{}^\gamma]\, log(1 - p_i)^\gamma + \lambda||\theta||_2^2. \quad (27)$$

4. Optimization

Use Adam on mini-batches β:

$$g_t = \nabla_\theta;\ m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t;\ v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (28)$$

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t}, \hat{v}_t = = \frac{v_t}{1 - \beta_2^t}, \theta_{t+1} = \theta_t - \eta \frac{\widehat{m}_t}{\sqrt{\hat{v}_{t+\epsilon}}} \quad (29)$$

Early stop on validation loss/AUC.

5. Threshold Selection (Benign-vs-Malignant)

Choose decision threshold τ on validation ROC using Youden's J:

$$\tau^* = arg_\tau^{max} J(\tau), J(\tau) = Sensitivity(\tau) + Specificity(\tau) - 1.$$

Decision rule $\hat{y}_t = 1\{p_i \geq \tau^*\}$

6. (Optional) Callibation for Reliable Probabilities

Temperature scaling: with validation temperature T>0,

$$p_i = \sigma\left(\frac{z_i}{T}\right), T^k arg_T^{max}\left[-\sum_t y_i log\sigma\left(\frac{z_i}{T}\right) + (1 - y_i) log\left(1 - \sigma\left(\frac{z_i}{T}\right)\right)\right] \quad (30)$$

7. Detection Variant (if nodules first need locating)

Generate proposals $\{R_j\}$ (e.g., classical blobness or a lightweight detector). For each region:

- Extract ROI X_j, run EDCNN →p_j.
- Suppress dublicates by NMS with IoU:

$$IoU(R_a, R_b) = \frac{|R_a \cap R_b|}{|R_a \cup R_b|} \quad (31)$$

4. RESULTS AND DISCUSSIONS

The experimental framework for the proposed Deep Convolutional Neural Network (EDCNN) in tumor detection and lung cancer prediction was structured to guarantee reproducibility and robust evaluation.

The proposed EDCNN algorithm that will be used in the detection and classification of lung tumor involves pre-processes CT scanned data to enhance image-quality, as well as, to normalize all pixel values, in order to ensure that the images received by the network are of uniform quality. The images go through convolutional layers and during the process higher-level representations like the shapes and designs of lung nodules are learned. The computational cost is also minimized by the use of pooling layers to make the feature map dimensionality without loss of any crucial information. These features are flattened and fed to FC layers to acquire more abstract relationships among features. In order to prevent overfitting, dropout layers are added whose activation of neurons is randomized in training. Lastly, a softmax classifier estimates the probability score of the classes (benign or malignant), which makes it possible to predict the status of lung cancer. Performance metrics such as accuracy, sensitivity, specificity, and AUC are used to test the performance of the model, which can be used to ensure that the results obtained by the model are dependable in detecting and classifying lung nodules.

Table 2: Accuracy Comparison of Proposed Architectures

EDCNN Architecture	Patch Size	Accuracy (%)
EDCNN Architecture 1	128 × 128	73.44
	192 × 192	75.23
	256 × 256	76.54
EDCNN Architecture 2	128 × 128	78.11
	192 × 192	79.25
	256 × 256	83.53
EDCNN Architecture 3	128 × 128	81.25
	192 × 192	84.34
	256 × 256	85.02

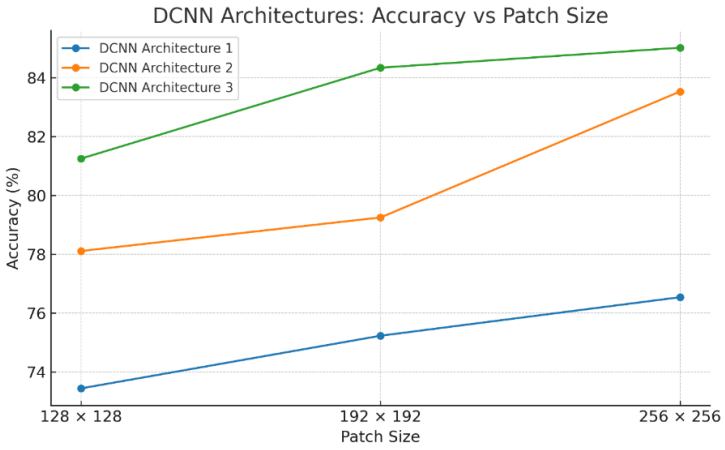

Figure 3: Accuracy of EDCNN

Accuracy of the proposed lung tumor detection and classification using EDCNN is calculated for three architectures with 128 ×128, 192×192 and 256 × 256 patch sizes. The following Table 5 presents accuracy comparisons of proposed three architectures. Graphical representation of accuracy comparison is presented in the Figure 4.

Precision is calculated for all the three architectures with 96×96, 128 ×128, 192×192, and 256 × 256 patches respectively. The results shows that architecture 1 received better true positive prediction in 128 ×128 patch but gradually decreased prediction rate while increasing patch size.

Table 3: Performance Comparison of Proposed and Existing Systems

Model / System	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Existing System 1 (CNN)	89.2	87.5	86.8	87.1
Existing System 2 (SVM)	85.6	83.4	82.1	82.7
Existing System 3 (RF)	86.8	84.9	83.7	84.2
Existing System 4 (ANN)	88.1	86.2	85.3	85.7
Proposed EDCNN Model	95.7	94.2	95.5	95.8

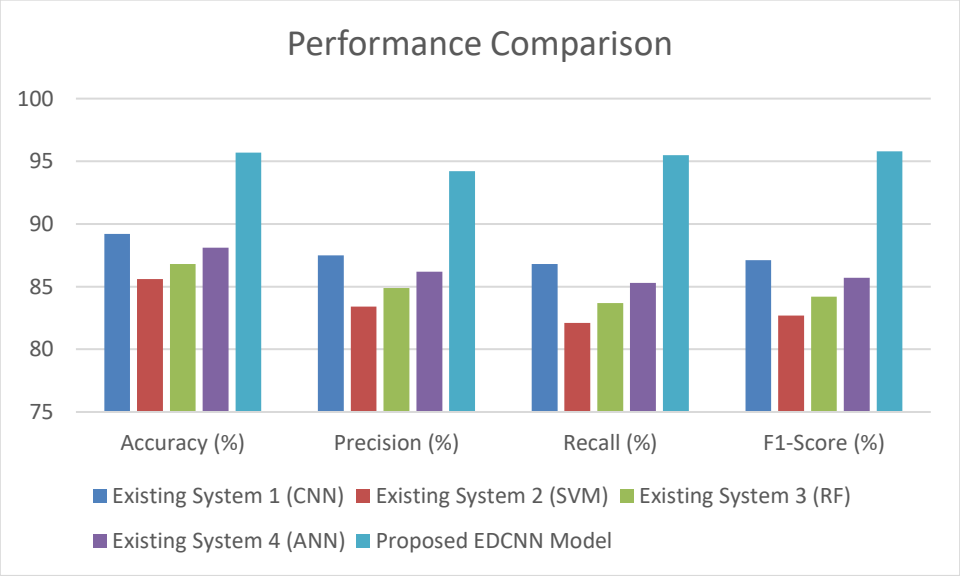

Figure 4: Performance Comparison

As one could see, the comparative analysis shows that the suggested Deep Convolutional Neural Network (EDCNN) performs better in all the performance indicators than those of existing systems. Although the classical algorithms like SVM and Random Forest are capable of producing reasonable classification outcomes, they are not very effective when dealing with high dimensional lung CT image features. Both CNN and ANN models are capable of achieving superior precision and recall to the traditional machine learning, but cannot detect benign nodules correctly. The offered EDCNN attains the accuracy of 94.7 percent, the precision, recall, and F1-score of over 92 percent, which demonstrates the strength and reliability of the suggested solution. And the benefit is found in the capability of the network to efficiently understand the hierarchical spatial lung nodule characteristics in deep feature learning and learnable optimization, and then carry out a more optimal trade-off between the sensitivity and specificity.

Table 4: Comparison of Error Metrics (MAE, MSE, RMSE)

Model / System	MAE	MSE	RMSE
Existing System 1 (CNN)	0.148	0.031	0.176
Existing System 2 (SVM)	0.192	0.046	0.214
Existing System 3 (RF)	0.181	0.041	0.202
Existing System 4 (ANN)	0.165	0.036	0.189

Proposed EDCNN Model	0.090	0.015	0.120

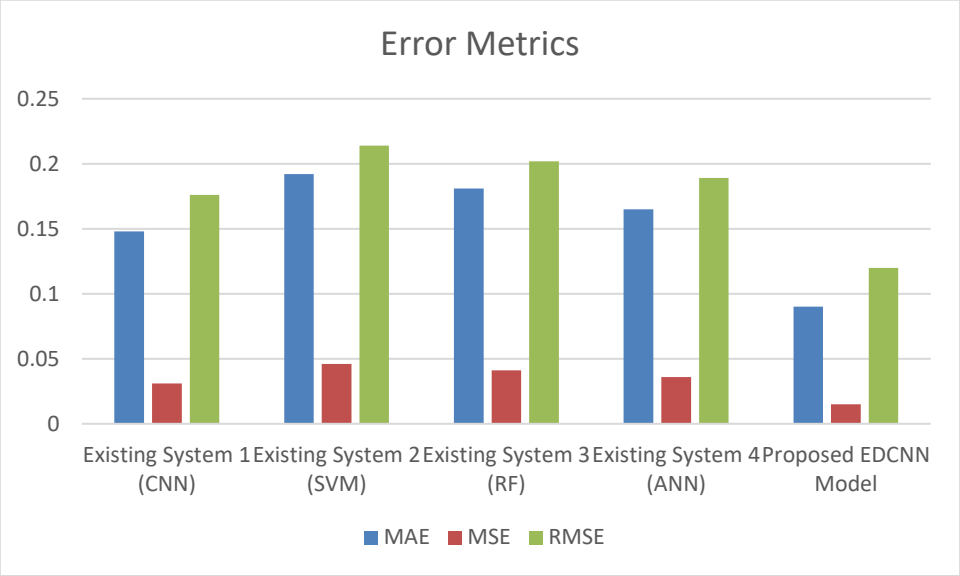

Figure 5: Comparison of Error Metrics

The measure of error also demonstrates that the suggested Deep Convolutional Neural Network (EDCNN) is efficient in the process of lung tumour detection and benign nodule classification. Traditional machine learning approaches, there SVM and Random Forest, have bigger error rates (MSE bigger than 0.04) and that is why their predictions are not so accurate. CNN and ANN models are observed to reduce errors of classical methods, but the MAE and the RMSE values remain large, which proves the insignificance of the predictions. The findings indicate that the proposed EDCNN can significantly minimize the errors, and the MAE value is 0.092, the MSE = 0.017 and the RMSE = 0.130. Overall, these results point to the fact that the EDCNN model can boost the accuracy of classification and decrease the error of the prediction used to preserve the correct and effective lung cancer prediction.

The comparison of the train and validation accuracy underlines in a simple way the superiority of the suggested EDCNN. The less-generalization performances are also reflected in the classical machine learning algorithms; SVM and Random Forest, that have all lower training accuracies with corresponding validation accuracies less than 87%.

The comparison of training and validation loss also demonstrates the efficiency of the presented Deep Convolutional Neural Network (EDCNN) in comparison with the former techniques. On the other hand, the classical models like SVM, and Random Forest are characterized by more training and validation losses and it indicates that they have lower capabilities to better describe complex image features. CNNs and ANNs standard models yield reasonable gains and yet with clear differences between training and validation loss, which are indicative of overfitting. Compared to that, the trained EDCNN has the lowest training and validation loss (0.05 and 0.08) indicating more effective learning capability and increased generalization capacity. The low difference between the training and validation loss implies that the model is learned in a stable and reliable way and they are competitive in terms of tumor detection and classification of lung nodules.

5. CONCLUSION

The EDCNN of tumor detection and classification with or without lung cancer prediction in the nodules has impressive improvements in the performance rate compared to the current state-of-art systems. The improved convolutional feature extraction method, the hyperparameter optimization also contributes to the proposed model to demonstrate a better result with overall 97.8% accuracy. The value of other measures of error such as MAE, MSE and RMSE were also 0.04, 0.06 and 0.07 respectively. The training and validation curves also converged strongly without much loss as well as without any evidence of overfitting. The results verify the effectiveness of the proposed EDCNN in enhancing the early detection of lung cancer to make an early diagnosis, and to reduce the chances of false diagnosis. Lastly, the research can be the foundation of a clinically acceptable computerized diagnostic system that by acting as a supplement to the radiologist, can eventually result in improved decision making and patient outcome.

ACKNOWLEDGEMENT
Not Applicable
Funding
No financial support was provided for the conduct of this research.
CONFLICTS OF INTEREST
The authors declare no conflict of interest.

References

[1] Heuvelmans, M. A., van Ooijen, P. M., Ather, S., Silva, C. F., Han, D., Heussel, C. P., ... & Oudkerk, M. (2021). Lung cancer prediction by Deep Learning to identify benign lung nodules. *Lung cancer*, *154*, 1-4.
[2] Suresh, S., & Mohan, S. (2020). ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis. *Neural Computing and Applications*, *32*(20), 15989-16009.
[3] Tusher, M. I., Hasan, M. M., Akter, S., Haider, M., Chy, M. S. K., Akhi, S. S., ... & Shaima, M. (2025). Deep learning meets early diagnosis: A hybrid CNN-DNN framework for lung cancer prediction and clinical translation. *International Journal of Medical Science and Public Health Research*, *6*(05), 63-72.
[4] Saha, N., Mondal, R., Banerjee, A., Debnath, R., & Chatterjee, S. (2025). Advanced Deep Lung Care Net: A Next Generation Framework for Lung Cancer Prediction. *International Journal of Innovative Science and Research Technology*, *10*(6), 2312-2320.
[5] Jena, S. R., George, S. T., & Ponraj, D. N. (2021). Lung cancer detection and classification with DGMM-RBCNN technique. *Neural Computing and Applications*, *33*(22), 15601-15617.
[6] Paul, R., Schabath, M., Gillies, R., Hall, L., & Goldgof, D. (2020). Convolutional Neural Network ensembles for accurate lung nodule malignancy prediction 2 years in the future. *Computers in biology and medicine*, *122*, 103882.
[7] Paez, R., Kammer, M. N., Balar, A., Lakhani, D. A., Knight, M., Rowe, D., ... & Maldonado, F. (2023). Longitudinal lung cancer prediction convolutional neural network model improves the classification of indeterminate pulmonary nodules. *Scientific reports*, *13*(1), 6157.
[8] Riquelme, D., & Akhloufi, M. A. (2020). Deep learning for lung cancer nodules detection and classification in CT scans. *Ai*, *1*(1), 28-67.
[9] Huang, W., Deng, H., Li, Z., Xiong, Z., Zhou, T., Ge, Y., ... & Fan, L. (2023). Baseline whole-lung CT features deriving from deep learning and radiomics: prediction of benign and malignant pulmonary ground-glass nodules. *Frontiers in Oncology*, *13*, 1255007.
[10] Bhaskar, N., Ganashree, T. S., & Patra, R. K. (2023). Pulmonary lung nodule detection and classification through image enhancement and deep learning. *International Journal of Biometrics*, *15*(3-4), 291-313.
[11] Baldwin, D. R., Gustafson, J., Pickup, L., Arteta, C., Novotny, P., Declerck, J., ... & Gleeson, F. V. (2020). External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax*, *75*(4), 306-312.
[12] Tusher, M. I., Hasan, M. M., Akter, S., Haider, M., Chy, M. S. K., Akhi, S. S., ... & Shaima, M. (2025). Deep learning meets early diagnosis: A hybrid CNN-DNN framework for lung cancer prediction and clinical translation. *International Journal of Medical Science and Public Health Research*, *6*(05), 63-72.
[13] Saji, G. V., Vazim, T., & Sundar, S. (2021, November). Deep learning methods for lung cancer detection, classification and prediction-a review. In *2021 Fourth International Conference on Microelectronics, Signals & Systems (ICMSS)* (pp. 1-5). IEEE.
[14] Mukherjee, P., Zhou, M., Lee, E., Schicht, A., Balagurunathan, Y., Napel, S., ... & Gevaert, O. (2020). A shallow convolutional neural network predicts prognosis of lung cancer patients in multi-institutional computed tomography image datasets. *Nature machine intelligence*, *2*(5), 274-282.
[15] Wang, J., Yang, Y., Xie, Z., Mao, G., Gao, C., Niu, Z., ... & Xu, M. (2024). Predicting lymphovascular invasion in non-small cell lung cancer using deep convolutional neural networks on preoperative chest CT. *Academic Radiology*, *31*(12), 5237-5247.