International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 05 | May 2025 www.irjet.net p-ISSN: 2395-0072

LUNG NODULE CLASSIFICATION AND DETECTION ON COMPUTED TOMOGRAPHY IMAGES USING CNN

P.S. Mayura Veena[1], L. Hymavathi[2], A. Tarun Teja[2], N. Jahnavi[2], K. Sateesh[2]

[1]Assistant Professor, Department of ECE, Anil Neerukonda Institute of Technology and Sciences, Andhra Pradesh, India
[2]UG student, Department of ECE, Anil Neerukonda Institute of Technology and Sciences, Andhra Pradesh, India

---

Abstract*—Lung cancer remains one of the deadliest cancers in the world, mostly because it is challenging to detect earlier. This project proposes an approach for detecting lung cancer in a hybrid manner involving both classification and object detection methods. The classification model utilises the Xception model to differentiate the cancer cells and has an accuracy of 94.4%. A YOLOv5 object detection framework is proposed to detect lung nodules in CT images. With training on the dataset, this model achieved 99% accuracy, showing its performance in detecting and bounding lung nodules with high accuracy.*

Keywords-convolutional neural networks, Xception model, yolov5, lung CT scans, lung cancer

1.INTRODUCTION

Lung cancer remains one of the most prevalent and fatal forms of cancer across the world, taking up a significant share of total cancer centers. Its early diagnosis plays a very vital role in improving patient survival rates; nevertheless, detection of lung nodules in their initial stages is particularly challenging due to their minimal size, contrast and intricate look in CT images. The classical diagnostic procedure on the basis of radiologists evaluating large volumes of thoracic CT scans manually is time-consuming, labour-intensive and susceptible to errors in diagnosis owing to fatigue and personal subjective interpretation. [1]

Recent developments in AI have revolutionized the field of medical imaging by allowing for fully automated, highly accurate diagnostic tools. Of the many deep learning methods, Convolutional Neural Networks (CNNs) have proven to be at the state-of-the-art for image classification, segmentation and object detection tasks [2]. Building on these developments, this research introduces a dual-model framework to tackle two important facets of lung cancer diagnosis: the diagnosis of cancerous conditions and the localization of pulmonary nodules in CT scans.

The first step of the proposed system is the implementation of a deep learning classification model using the Xception architecture. As an extension of the Inception architecture, François Chollet created the deep convolutional neural network called "Xception," or "Extreme Inception." It is based on the assumption that cross-channel correlations and spatial correlations in convolutional neural networks are completely separable. To this purpose, Xception substitutes standard convolutional layers with depthwise separable convolutions, which decompose a typical convolution into a depthwise spatial convolution followed by a pointwise (1×1) convolution. This structure drastically lowers the number of parameters and computational overhead while preserving or even improving the model's representational capabilities. This system trains the Xception model to classify slices of lung CT scans into four groups: normal, adenocarcinoma, large cell carcinoma, and squamous cell carcinoma. Xception is uniquely suited to identify the fine-grained textural and structural patterns characteristic of various lung diseases due to its architectural depth and parameter efficiency. Furthermore, the model is refined on a carefully selected dataset using transfer learning techniques, which enables the system to have high diagnostic accuracy with improved generalization.[3]

Parallel to this, the system incorporates an object detection model in real time with YOLOv5 (You Only Look Once v5), a highly successful architecture for fast and accurate object detection in a variety of applications. YOLOv5 operates on a full image in a single pass and generates bounding boxes and class predictions for objects that are detected. The suggested approach utilizes YOLOv5 for detection of lung nodules in CT scans and provides spatial context, together with visual interpretability to the classification outcome [4]. The model gets trained on the NCCD dataset available in Kaggle platform, with annotated CT scans appropriate for both detection and classification tasks.

2.METHODOLOGY

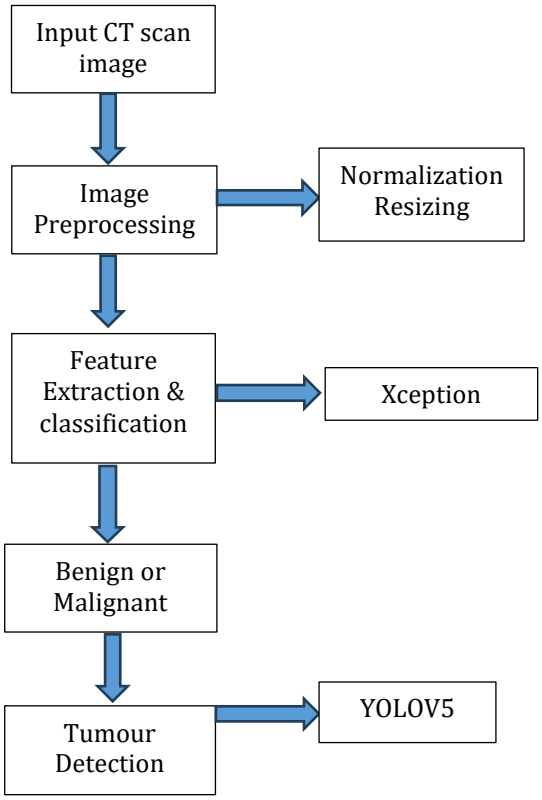

Fig 1: Workflow of the project

2.1. Data Collection

The dataset is collected from NCCD (NATIONAL CENTER FOR CANCER DISEASES) in Kaggle site and it consists of lung CT scan images.

The dataset is divided into three parts: training, validation and test sets. The training portion is used to train the model, the validation set is used for hyperparameter tuning, and the test set is used to evaluate final performance.

2.2. Data Preprocessing

To make the model stronger, the images go through multiple preprocessing operations.

Initially, resizing of all the images takes place so that they all have the same dimensions.

Next, the normalization of pixel intensities takes place to fall between 0 and 1, which really helps the model learn more effectively.

2.3. Feature Extraction & Classification

The Xception is a deep CNN architecture designed to give the high performance in classification tasks.

In this project the xception architecture is used for both feature extraction and classification. As the input image passes through the successive layers of the xception model, it identifies and learn features form low-level details like edges and textures to high level patterns such as structure of lung tumours. These learned features are utilized to classify the image into different cancerous and non-cancerous cells. [5][6]

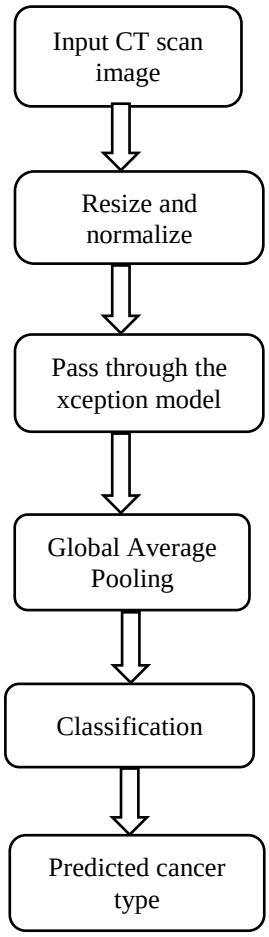

Fig 2: Flowchart for Xception model

Entry flow:

To process the input image the xception model begins with a few standard convolutions.

Then the xception model uses depth wise convolutions with max pooling to reduce the spatial dimensions.

In this the model captures the basic features like edges and simple textures and input image size gets reduced.[7]

Middle flow:

The middle flow contains number of identical modules.

each module has three depthwise convolution layers with residual connections middle flow allows learning of

complex patterns without increasing the computational cost. [7]

Exit flow:

In this further processing of features takes place using separable convolutions. It is followed by global average pooling and fully connected layers for final classification. [7]

2.3.1 Xception Model Performance

The Xception model achieves an accuracy of 94.4% across four distant classes.

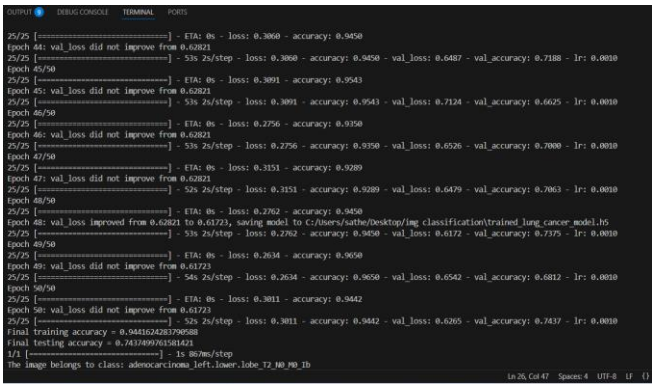

Fig 3: Output showing the accuracy of classification using 50 epochs

2.3.2 Confusion matrix for classification model

In this project xception model is for classification of lung nodules.

In order to know whether the model is giving accurate results or not confusion matrix is used as the performance tool. the confusion matrix is in the square table format and it is the visual representation of how well the model is performing by comparing actual labels with the labels predicted by the model.[8]

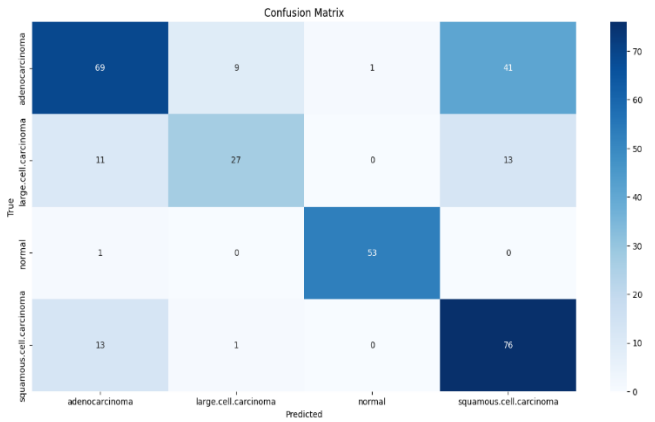

Fig 4: Confusion matrix

The matrix is divided into four key components.

True positive - In this case the Xception model accurately predicts a positive class.

True Negative -Here the model correctly predicts a negative class.

False Positive -here the model incorrectly predicts a positive class.

False Negative -In this case the model incorrectly predicts a negative class.

Also, there are certain evaluation metrices to evaluate the performance of the model.[8][9]

Accuracy: It is a measure of how often a prediction model makes the correct prediction, the formula is

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: It is the measure of how many of the positives predicted by the model are actually correct, the formula is

$$Precision = \frac{TP}{TP+FP}$$

Recall: It measures how many of the actual positives were correctly identified by the model.

$$Recall = \frac{TP}{TP+FN}$$

F1-score: It is the measure of harmonic mean of Precision and recall

$$F1\ score = 2 * \frac{Precision*Recall}{Precisiom+Recall}$$

2.3.3 Classification output images:

Fig 5: Predicted as normal healthy non-cancerous tissue

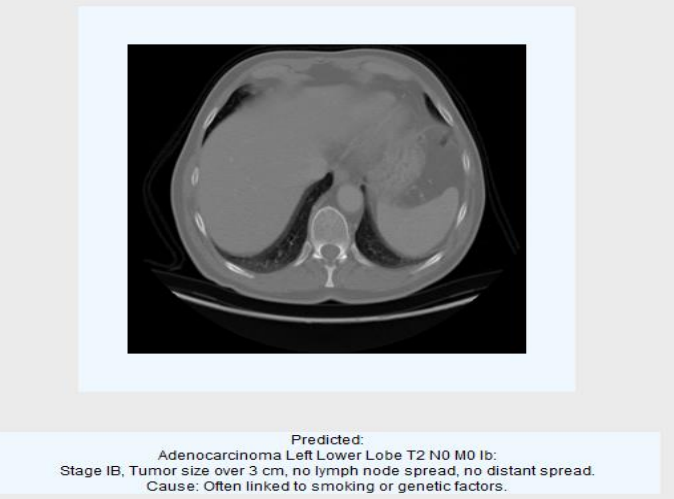

Fig 6: Predicted as cancer cell- adenocarcinoma

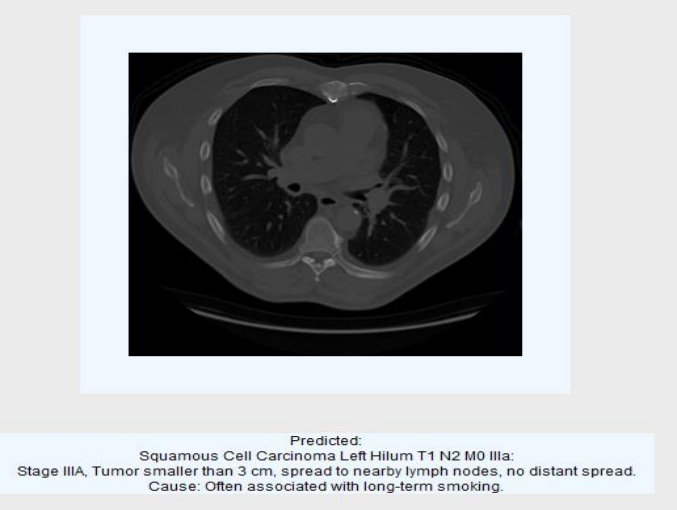

Fig 7: Predicted as cancer cell-Squamous Cell Carcinoma

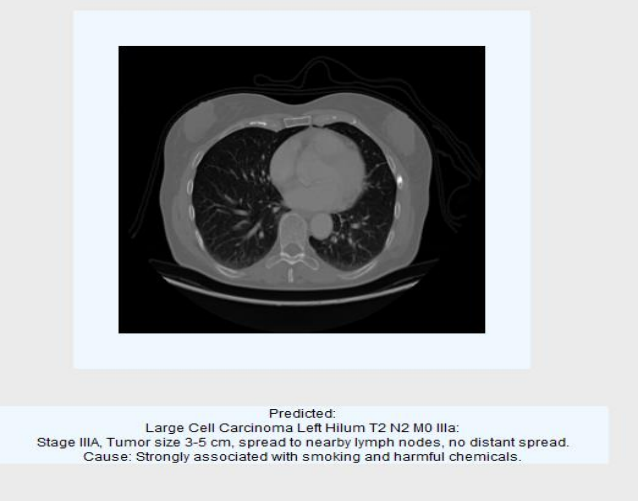

Fig 8: Predicted as cancer cell-Large Cell Carcinoma

Evaluation metrices	Precision	Recall	F1 score	Support
Adeno Carcinoma	0.69	0.63	0.66	120
Large cell Carcinoma	0.76	0.37	0.50	51
Squamous cell Carcinoma	0.60	0.84	0.70	90
Normal	0.98	0.96	0.97	54

Table 1: Classification Report

2.4 Lung Nodule detection using yolov5 architecture

The evolution of object detection models has experienced some incredible advances in computer-aided diagnostic (CAD) systems for medical imaging in the past few years. Among these, the YOLO (You Only Look Once) series of models has become a prominent model class for real-time object detection. The fifth iteration in the YOLO series, YOLOv5, has shown unprecedented advances in detection performance, computation speed, and deployment efficiency. These characteristics make it particularly well-suited for advanced medical tasks such as the localization and classification of lung nodules in computed tomography (CT) images. [10]

Yolo v5 is a one stage object detector so that it detects and classifies the objects like tumours in one pass, this means it predicts bounding boxes and class probabilities simultaneously.so here there is a minimal loss in accuracy. Previous models like faster R-CNN first identify the object locations(proposals) and then classify these proposals. Thus, Yolov5 minimizes the processing delay and optimizes the performance.

Yolov5 architecture has three main parts: the backbone, the neck and the head.[11]

Backbone:

The function of this backbone is to extract features from the input image. It begins with a focus module, which slices the input into patches and concatenates them so that spatial information is retained and size also get reduced.

This is followed by CSP (Cross Stage Partial) bottle neck splits feature maps into two parts and later merges them. It allows reduction of the computational cost.

The backbone also uses C3 modules, which are variant of CSP blocks and applies SiLU (Sigmoid-weighted Linear Unit) activation for better non-linear transformations.[11]

Neck:

The function of this neck part is to aggregate and combine features at different scales by using a combination of Feature

Pyramid Network (FPN) and Path Aggregation Network (PAN). FPN and PAN ensures information flows efficiently in both top-down and bottom-up directions.

Head:

The head predicts the bounding boxes, objectness scores and class probabilities from the features.

The final output consists of bounding boxes and confidence scores for detected lung nodules.[11][12]

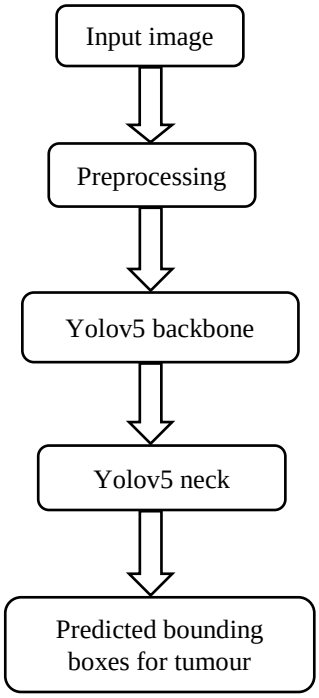

Fig 9: Flowchart for yolov5 architecture

2.4.1 Training and validation performance:

Training and validation dataset are meticulously prepared to have a variety of image slices that reflect real clinical situations, such as nodules of different sizes, shapes, density, and anatomical locations. The images are preprocessed to standardize pixel intensities, regulate contrast, and rescale to a standard input resolution compatible with the YOLOv5 architecture.

The information is separated into training and validation sets, wherein the training subset is applied to optimize parameters of the model and the validation set is used to assess generalization and avoid overfitting.

The YOLOv5 model is trained in a supervised learning fashion, where detection and classification error are minimized for all labeled image slices.

The model takes input images and ground-truth labels as input and learns to output bounding boxes and class probabilities. The configuration of the architecture involves choosing a suitable YOLOv5 based on hardware capability and performance compromises. The model is initialized from pretrained weights to borrow domain knowledge and speed up convergence.[13]

Here the following Evaluation metrics were used to analyze the performance of the detection method.

Evaluation metrices curves:

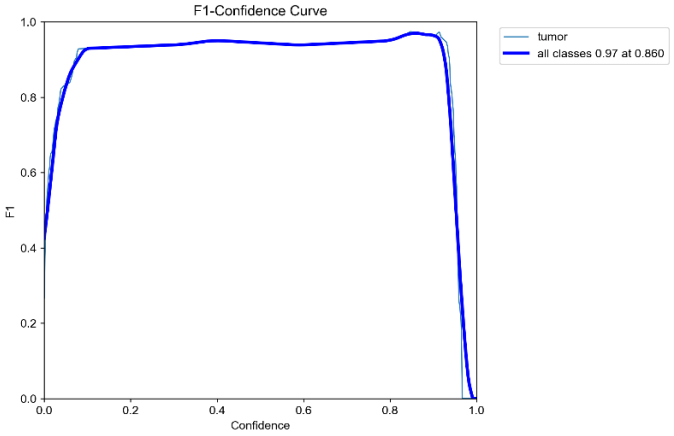

Fig 10: F1 curve

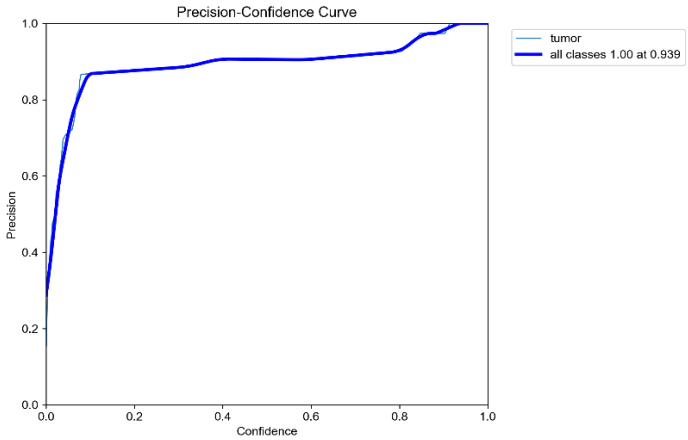

Fig 11: Precision curve

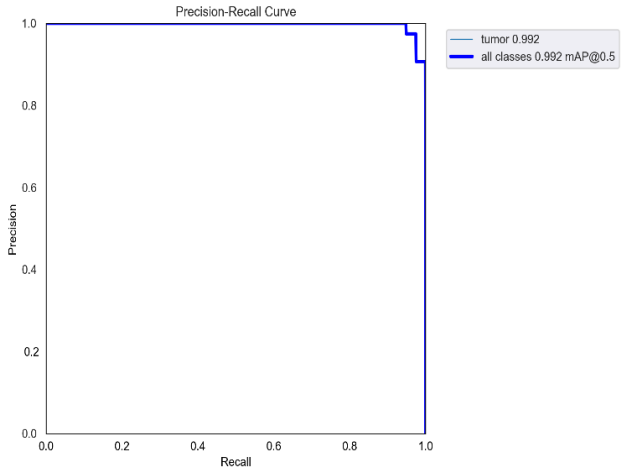

Fig 12: Precision-Recall curve

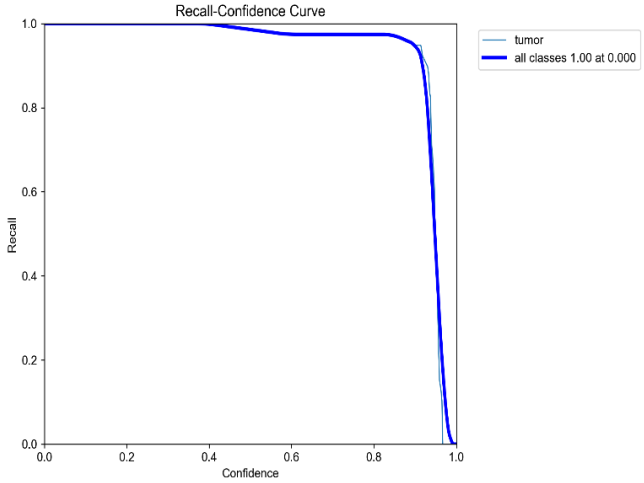

Fig 13: Recall Curve

3.RESULTS

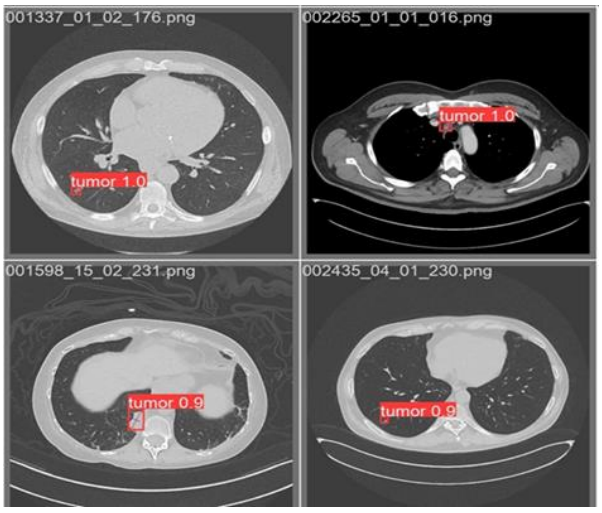

Fig 14: Tumor detection output images

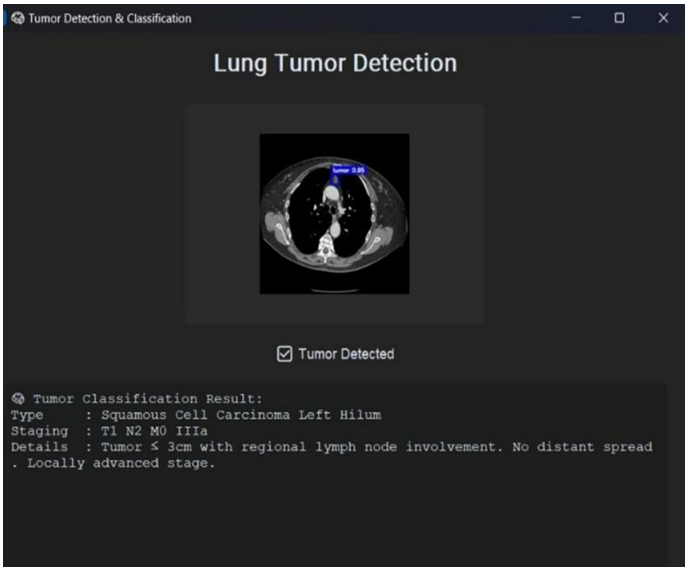

Fig 15: Output image of classification and detection

4.CONCLUSION & FUTURE SCOPE

The Xception model which was used for the classification achieved 94.4% accuracy in distinguishing between various types of lung cancer cells. The yolov5 architecture which was used for detection achieved 99% accuracy for lung nodule localization which also minimizes the false positives. These two models make finding and correctly diagnosing lung cancer more reliable and accurate.

Following the good results obtained applying Xception in classification and YOLOv5 in detection, future research can investigate various improvements. A direction is constructing an end-to-end hybrid network that combines classification and detection as a single process, minimizing the processing time and maximizing system resources.

Another area to consider is semi-supervised and self-supervised learning, which would minimize the dependence on large labeled datasets, a typical shortcoming in medical imaging. Adding attention mechanisms to existing architectures will also enhance the network's ability to pay attention to diagnostically important areas, more enhancing accuracy as well as interpretability.

Besides, the system would be scalable to support real-time diagnostics in clinical settings with optimized deployment methods such as model quantization and edge inference for low-resource environments.

REFERENCES

[1] T. Mahmmod, N. Ayesha, M. Mujahid and A. Rehman, "Customized Deep Learning Framework with Advanced Sampling Techniques for Lung Cancer Detection using CT Scans," *2024 Seventh International Women in Data Science*

Conference at Prince Sultan University (WiDS PSU), Riyadh, Saudi Arabia, 2024, pp. 110-115, doi: 10.1109/WiDS-PSU61003.2024.00035.

[2] Geetha, G., et al. "Image Captioning Using Deep Convolutional Neural Networks (CNNs)." Journal of Physics: Conference Series, vol. 1712, no. 1, Dec. 2020, p. 012015. DOI.org (Crossref), https://doi.org/10.1088/1742-6596/1712/1/012015.

[3] T.T. Al-Shouka and K. M. A. Alheeti, "A Transfer Learning for Intelligent Prediction of Lung Cancer Detection," *2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT)*, Al-Muthana, Iraq, 2023, pp. 54-59, doi: 10.1109/AICCIT57614.2023.10217967.

[4] Kumar, N. Komal, et al. "Detection and Recognition of Objects in Image Caption Generator System: A Deep Learning Approach." 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), IEEE, 2019, pp. 107–09. DOI.org (Crossref), https://doi.org/10.1109/ICACCS.2019.8728516 .

[5] M. A. Javad Zadeh Barzaki *et al*., "Using Deep Learning for Classification of Lung Cancer on CT Images in Ardabil Province: Classification of Lung Cancer using Xception," *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, Islamic Republic of, 2023, pp. 375-382, doi: 10.1109/ICCKE60553.2023.10326262.

[6] P. M. Bruntha et al., "Lung Nodule Classification using Shallow CNNs and Deep Transfer Learning CNNs," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1474-1478, doi: 10.1109/ICACCS51430.2021.9441702.

[7] N. Valarmathi, S. B, S. V, S. F. J and S. M, "Pneumonia Disease Classification Using Xception CNN Model," *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, Namakkal, India, 2024, pp. 1-6, doi: 10.1109/AIMLA59606.2024.10531453.

[8] Aishwarya Maroju et al. "Image Caption Generating Deep Learning Model". 2021 International Journal of Engineering Research & Technology (IJERT). ISSN: 2278-0181, vol 10, issue 09, sep 2021.

[9] R. V. M. d. Nóbrega, S. A. Peixoto, S. P. P. da Silva and P. P. R. Filho, "Lung Nodule Classification via Deep Transfer Learning in CT Lung Images," *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, Karlstad, Sweden, 2018, pp. 244-249, doi: 10.1109/CBMS.2018.00050.

[10] A. M. Harale, V. K. Bairagi, E. Boonchieng and M. R. Bachute, "Nodules Detection in Lungs CT Images Using Improved YOLOV5 and Classification of Types of Nodules by CNN-SVM," in *IEEE Access*, vol. 12, pp. 140456-140471, 2024, doi: 10.1109/ACCESS.2024.3466292

[11] X. Dong, N. Xu, L. Zhang and Z. Jiang, "An improved YOLOv5 network for lung nodule detection," *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS)*, Changchun, China, 2021, pp. 733-736, doi: 10.1109/EIECS53707.2021.9588065.

[12] M. A. Muqeet, H. K. Gill, M. A. N. Hasnain, A. Fatima, A. Mohammed and M. K. Saifullah, "Detection of Lung Cancer using Deep Learning based Convolution Neural Network Models," *2023 IEEE Fifth International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, Bengaluru, India, 2023, pp. 1-5, doi: 10.1109/ICAECC59324.2023.10560318.

[13] Vishal, S. Mehta and A. Rathour, "Optimizing Hybrid CNN-SVM Architectures for Enhanced Detection of Subsolid Lung Nodules: A Comparative Study with Traditional Methods," *2024 5th IEEE Global Conference for Advancement in Technology (GCAT)*, Bangalore, India, 2024, pp. 1-5, doi: 10.1109/GCAT62922.2024.10923969.