Lung Nodule Detection and Segmentation with Cancer Stage Prognosis using Hybrid Deep Learning Model

by

Kazi Sati (192291)
Pinky Akter (192294)
Md. Zuleyenine Ibne Noman (192319)

A Research Project Report submitted to the
Institute of Information Technology
in partial fulfillment of the requirements for the degree of
Bachelor of Science (Hons.) in
Information and Communication Technology

Supervisor: Dr. Mohammad Abu Yousuf, Professor

Institute of Information Technology
Jahangirnagar University
Savar, Dhaka-1342
November, 2023

DECLARATION

We hereby declare that this thesis is based on the results found by ourselves. Materials of work found by other researcher are mentioned by reference. This thesis, neither in whole nor in part, has been previously submitted for any degree.

Kazi Sati	Pinky Akter	Md. Zuleyenine Ibne Noman
Roll: 192291	Roll: 192294	Roll: 192319

CERTIFICATE

This is to certify that the thesis proposal report entitled **"Lung Nodule Detection and Segmentation with Cancer Stage Prognosis using Hybrid Deep Learning Model"** submitted by **Kazi Sati, Pinky Akter, Md. Zuleyenine Ibne Noman**, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science (Hons.) in Information and Communication Technology on November, 2023.

Prof. Dr. Mohammad Abu Yousuf
Supervisor

Accepted and approved in partial fulfilment of the requirement for the degree Bachelor of Science (Hons.) in Information and Communication Technology.

BOARD OF EXAMINERS

Prof. Dr. Fahima Tabassum
Chairman

Prof. Dr. Mohammad Shahidul Islam
Member

Prof. Dr. M. Mesbahuddin Sarker
Member

Prof. Dr. Md. Hasanul Kabir
External member

ACKNOWLEDGEMENTS

We feel pleased to have the opportunity of expressing out heartfelt thanks and gratitude to those who rendered their cooperation in making this report.

This thesis is performed under the supervision of Dr Mohammad Abu Yousuf, Professor, Institute of Information Technology (IIT), Jahangirnagar University, Savar, Dhaka. During the work, he has supplied us with several books, journals, and materials related to the present investigation. Without his help, kind support, and the generous period he has given, we could not perform the project work successfully in due time probably. First and foremost, we wish to acknowledge our profound and sincere gratitude to him for his guidance, valuable suggestions, encouragement, and cordial cooperation.

We express out utmost gratitude to Prof. Dr. M Shamim Kaiser, Director of IIT, Jahangirnagar University, Savar, Dhaka, for his support to complete the work within the time frame. Moreover, we would also like to thank the other faculty members of IIT who have helped me directly or indirectly by providing their valuable support in completing this work.

We express our gratitude to all other sources from where we have found helpful. We are indebted to those who have helped us directly or indirectly in completing this work.

Last but not least, we would like to thank all the staff of IIT, Jahangirnagar University, and our friends and family who have helped us by giving their encouragement and cooperation throughout the work.

ABSTRACT

Lung cancer being one of the deadliest causes one-third of the deaths because of cancer. There is no prior syndrome of it. As a result, it is caught in the last stage. Early detection of lung cancer can increase survival probability of lung cancer. In recent years, computer-aided diagnosis (CAD) has shown remarkable performance in predicting lung cancer from low-dose computed tomography (LDCT) scans. Though there are lung nodules, not all of them put a risk of cancer. Depending on its morphology, the risk it can develop can be different. Therefore, identifying cancerous nodules and predicting the risk of lung cancer based on each malignant nodule is important. This study aims to implement a robust model for detecting and segmenting malignant lung nodules for accurate prediction of lung cancer stage based on the aggregate risk that all the malignant nodules contribute. In the proposed system, 3D convolutional neural network (3D-CNN), U-Net, hierarchical recurrent neural network (RNN) separately detect lung nodules. From their ensembled result, twelve morphological features are extracted using five well-known transfer learning models, i.e MobileNet, VGG16, VGG19, ResNet50, DenseNet201. The features are fed to six machine learning models. The top performing machine learning models are then selected for ensemble learning on detecting the malignancy of a nodule. Based on the malignancy of each nodule, an aggregated decision about the lung cancer stage is achieved. Later, the model performance is evaluated. The LIDC-IDRI dataset is proposed to be used for this model.

Keywords: Computer-Aided Diagnosis (CAD), Low-Dose Computed Tomography (LDCT), Malignant, Convolutional Neural Network (CNN), Transfer Learning, Machine Learning, Recurrent Neural Network (RNN), Ensemble Learning (EL).

LIST OF ABBREVIATIONS

AUC	Area Under the ROC Curve
CAD	Computer-aided Diagnosis
CNN	Convolutional Neural Network
CT	Computed Tomography
CV	Cross-Validation
DL	Deep Learning
EL	Ensemble Learning
FC	Fully Connected
GAAHE	Genetic Algorithm based Adaptive Histogram Equilization
LDCT	Low-Dose Computed Tomography
LGB	Light Gradient Boosting
LR	Logistic Regression
MAE	Mean Average Error
ML	Machine Learning
MLP	Multi-layer Perceptron
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TL	Transfer Learning
VGG	Visual Geometric Group
XGBoost	Extreme Gradient Boosting

LIST OF NOTATIONS

I_{rgb}	Image constructed from gray-scale image
$I_{(\text{gray, R})}$	Red channel equivalent gray value
$I_{(\text{gray, G})}$	Green channel equivalent gray value
$I_{(\text{gray, B})}$	Blue channel equivalent gray value

LIST OF FIGURES

Figure

TABLE OF CONTENTS

CHAPTER I

Introduction

1.1 Overview

This chapter comes up with an illustration of the background, motivation, limitations and objectives of this work. At the chapter's end, the contribution of the research work and outline of the thesis are described.

1.2 Background Story

Lung cancer is the one of those cancers that leads to death. Deaths caused by two-third of all cancers are because of lung cancer [3]. It is the deadliest cancer which is one the the most common type of cancer[4]. Survival rate of patient of lung cancer is 18% [5].But chances of survival can greately increased if lung cancer can be diagnosed at early stage[6**?**]. Since 1970s, CXR and CT images have been used for lung nodule detection for lung cancer diagnosis [7]. More advanced spiral low-dose computed tomography (LDCT) can capture the entire chest in a single breath-hold [8], thus less noise is received. An experiment done by National Ling Screening Trail (NLST) has demonstrated that 20% mortality was reduced by using LDCT images for diagnosis compared to CXR in 2000s [9]. Hence, LDCT shows a greater capability useful for proper diagnosis. In this study, LDCT images is decided to be used. A lung nodule is a small, rounded or oval-shaped cluster of cells within the lung.. There are two types of lung nodule, benign (non-cancerous) and malignant (cancerous). These nodules are quite similar in their appearance with a variance in texture, shape and size[10]. With respect to size, nodules bigger than 3mm is known as lung nodule, mainly benign. But the nodules with bigger size are more likely to become malignant nodules [11]. Usually, these nodules are detected by a radiologist. To check the whole 3D lung CT-scan, the detection takes almost 10 minutes even for an experienced radiologist. Then oncologists reviews the detected nodules to decide whether the nodule is malignant or not. Based on found result risk of cancer is determined of a patient in general. Due to small variation of morphology between the benign and malignant nodules, different radiologists and doctors may provide different diagnosis decisions.[12] Both radiologists and oncologists must have expertise which requires years of experiences. Thus, both time and cost is high for a patient.

1.3 Motivation of the Study

Computer-aided diagnosis (CAD) comes in help of oncologist and radiologist with deep learning algorithm in decision making process for lung cancer[13, 14]. Clinicians can provide more accurate diagnosis decision but with reduced work-load at the same time using this[15**?**]. Computer vision can leverage a substantial amount of data, rendering it well-suited for computer-aided diagnosis (CAD) for helping in the task of diagnosing lung cancer by detecting malignant lung nodules. Deep learning can learn patterns, thus can detect nodules or decide malignancy of these nodules. Among deep learning models, convolutional neural network (CNN) takes huge attention for its promising results and Effectiveness in minimizing reliance on field-specific knowledge. While CNN learns from spatial structure, recurrent neural network (RNN) learns for sequential structure. Thus, both of them can detect nodules that another probably cannot. The sequence of CT scan images in important for 3D CNN, while 2D CNN uses Using a solitary image slice as input is deemed to be computationally more economical. For the proposed system the 2D CNN is fed synthesized images which are images with each channel representing a slice before and after from the middle slice. All of these three can contribute to detecting more nodules. The morphological features are then extracted using transfer learning model, later machine learning is applied to detect the malignancy of nodules based on the features. This model aims to classify lung cancer stage considering the malignancy of each detect malignant nodule to successfully diagnose as real-life oncologists. In general deep learning models are treated to be black-box because of the complex structure for interpretation[16, 17]. Without having proper interpretation oncologist as well as radiologist cannot verify the validity of the result achieved form these models[18]. Hence, another aim of the proposed model is to provide medically interpretable information helpful regarding diagnosis decision making.

1.4 Limitations of Existing Works

The limitations of the existing models are stated below :

1. To our current understanding, no paper suggests evaluation of cancer risk on the aggregated result of each malignant nodule.
2. Most of the paper only focused on lung nodule detection.
3. To our current understanding, no work provides medically interpretable and required information about each detected nodule based on which the risk of cancer must be determined.

1.5 Problem Statement

It is proposed to create a machine-learning based hybrid deep learning model using ensemble learning and assess it for lung nodule identification and classification

with lung cancer stage determination. When given CT images, the model will automatically assess the malignancy of nodules and based on each detected malignant nodule the model will determine the stage of lung cancer while also providing the information of each detected nodule.

1.6 Objective

The main objectives are stated below:

1. To build a robust hybrid model which detects malignant nodules from low-dose computed tomography and can successfully determine cancer stage based on the aggregation result of the malignancy of each malignant nodule for proper evaluation of cancer risk.

2. To not only detect nodule but also classify them as malignant or benign and further, take each of them in consideration for final classification of lung cancer stage.

3. To provide morphological information as well as information required to justify the level of malignancy of all detected nodules, so that the determined cancer stage can be evaluated and the process will not seem to be just a black box.

1.7 Research Outline

The remainder of the report follows this structure: In **Chapter II**, a comprehensive review of related work is presented through a literature study. **Chapter III** introduces the proposed system framework and finally a conclusion is drawn in **Chapter IV**.

1.8 Conclusion

This chapter includes full background research, the motivation for undertaking this work, limitations of existing works, our work objectives, and ultimately our contribution. The book outline is displayed at the end.

CHAPTER II

Literature Review

2.1 Overview

This chapter illustrates an overview of the background, motivations, limitations, and objectives of this research.

2.2 Deep Learning Approaches

2.2.1 2D CNN approach

A study shows that A novel CAD system using image processing techniques and CNN approach is proposed to detect and classify lung cancer in CT scan images, achieving promising performance on sensitivity, specificity, and accuracy[19].

A study at [20] proposed synthetic image which is mainly sequenced CT images on red, green and blue channels. Changes in subsequent slices are reflected as a different color pattern on the image. Thus, one image can represent 3 subsequent slices of CT scan, thus providing more information. The changes are captured in the y-axis, namely time domain. Future works may concentrate on capturing changes in x-axis and z-axis. Being 2D input, these images reduce parameters needed for learning networks. This work implemented U-net as the basis for implementing deep learning networks, which gives 10% more accurate results than state-of-art. Semi-sequential means the image can be made based on CT scan that are 1 or 2 slices before and after the middle slice. Thus, skipping some sequence of the entire CT scan is termed as semi-sequential. This paper proposes synthetic images which are more efficient than conventional gray-scale CT scan images.

Another study at [21]introduces an end-to-end approach for automatically segmenting lung nodules in CT images, addressing challenges like nodule variability. Their model achieved high accuracy on a large dataset and outperformed existing methods. It also generated detailed intra-nodular heterogeneity images, aiding radiological diagnosis. This approach eliminates the need for human intervention and additional preprocessing steps.Faster R-CNN is used, in which a region proposal network (RPN) is modeled for efficient foreground candidate selection. It relies on a

backbone (like ResNet) for feature extraction and defines anchors of various sizes and scales. Positive region proposals (RPs) with high overlap are used for ROI classification in Fast R-CNN. While Smooth L1 loss is traditionally used for bounding box regression, recent studies suggest GIoU loss can be more effective for detection tasks in such networks.A deep learning method using a size-related damper block achieves 86% accuracy in predicting nodal metastasis from the primary tumor in lung cancer using gemstone spectral imaging (GSI) dual-energy[22].

Another study of [23] proposes a model that is an advanced hybrid deep-convolutional neural network model designed to enhance lung cancer diagnosis. It combines CT scan and medical IoT data to achieve a high accuracy of 96.81% in Categorizing lung cancer into five distinct classes.. It also successfully categorizes sub-stages with 91.6% accuracy. LungNet shows great promise for automated lung cancer diagnosis and sub-classification, outperforming similar CNN-based systems.A study of [24] indicates that a new method for classifying lung nodules using CT scans. The basis of their method is a deep learning model that grabs multiple cross-sections of the nodule as input and produces a classification decision. The model is able to achieve state-of-the-art performance on two public datasets, and it is also lightweight and portable to mobile devices. This makes it a promising tool for assisting practitioners in the diagnosis of lung cancer.Another research [25] identifies that presents a novel approach for improving lung cancer detection using contrast stretching and feature fusion. The method involves enhancing image contrast, extracting multiple texture and geometric features, and applying a feature selection process. The approach achieves a high accuracy of 99.4% on a public dataset, outperforming existing methods. This innovation aims to address the challenges of early diagnosis of lung cancer and improve survival rates.

2.2.2 3D CNN approach

A newly developed interpretable deep hierarchical semantic convolutional neural network, abbreviated as HSCNN, is proposed for lung cancer prediction, achieving better results than 3D CNN alone by providing low-level semantic features that explain model predictions in an expert-interpretable manner[26].A recent study at [27] that used a GAN model to generate 3D CT images of lung nodules based on their size. The study evaluated the quality of the generated images by comparing them with true images and testing their usefulness for nodule size classification. The study found that the generated images were realistic enough to fool radiologists and to train a classification model with similar accuracy as the true images.

A study introduces a computer-aided diagnosis (CAD) system for lung cancer, addressing its high global mortality rate. A 3-D lung segmentation tool, combining active contour model and local image bias field, efficiently identifies tumor regions in CT images with intensity variations. It achieves 97% accuracy in tumor extraction and quantifies centroid displacement. The images are subsequently categorized

through an Improved CNN Classifier, showcasing the CAD system's remarkable precision in automating lung tumor diagnosis..[28]

2.3 Machine Learning Approach

Lung cancer stands as a prevalent and fatal form of cancer globally. Detecting it in its early stages holds the potential to enhance patient survival rates and overall quality of life. Yet, diagnosing lung cancer is an intricate process that necessitates expert proficiency and experience. Machine learning (ML), a field within artificial intelligence, possesses the capability to glean insights from data and provide predictions and informed decisions.

2.3.1 KNN approach

There is another study that shows that focuses on early lung cancer diagnosis using a k-Nearest-Neighbors technique with genetic algorithm-based feature selection. The algorithm aims to enhance accuracy and classifier efficiency. The approach achieved 100% accuracy when tested on a lung cancer database, showing its potential for efficient lung cancer staging diagnosis[29].

2.4 Hybrid Approach

A recent research of [30] that proposes an end-to-end system consisting of 2 modules, cancer risk evaluation module and nodule detection. The first module separately trains 3D CNN to capture nodules using spatial structural features and RNN model to capture nodules using sequential structural features, then combine their results. Thus increasing the number of accurate nodule detection. Increased false positives are reduced in the classification stage. 9 3D-CNN models are trained to extract and evaluate 9 morphological features of each nodule.Following this, the XGBoost regression model is employed to establish a mapping from morphological features to ground truth labels. The evaluation of results is conducted using the logloss function. The feature evaluator provides intermediate information that is medically interpretable, including aspects such as the grade of lobulation, spiculation, sphericity, and more. The first module trained by LUNA-16 whereas the second one trained with LIDC.This work only included one nodule with highest malignancy value to assess the cancer probability whereas oncologists in real life consider all nodules present in a CT-scan to evaluate cancer risk.process. In the process of detecting nodules, there is a possibility of multiple counting of the same nodule, especially for very large nodules. To avoid this, an algorithm needs to be developed that can accurately identify and track each nodule throughout the process. This algorithm should be able to differentiate between different nodules and ensure that each nodule is counted only once.

Another study of [31] illustrates that For lung and colon cancer detection a hybrid ensemble feature extraction model is used which is evaluated on histopathological lung

and colon dataset. Deep feature extraction technique is accomplished by selecting one transfer learning model (TL) among five transfer learning models namely, VGG16, VGG19, MobileNet, DenseNet169, and DenseNet20. For each TL model, after the features are extracted, these are fit into six well-known machine learning (ML) models. Among these six ML models, top performing three models are selected for ensemble learning (EL) in the next phase with k-fold cross validation. Between soft-voting and hard-voting, soft-voting outperformed for each evaluation. Later,performance metrics such as MAE, accuracy, recall, f1-score, MSE, precision,RMSE, confusion matrix, ROC curve and AUC score are used for the evaluation of the selected model of the experiment. Ensembling high performing model increases the performance significantly.

2.5 Conclusion

In this chapter, we have undertaken an extensive examination of prior research efforts conducted by a multitude of scholars, employing a wide array of methodologies and approaches.

CHAPTER III

System Model

3.1 Overview

The system workflow is discussed in section 3.2 and proposed system architecture and the mechanisms used for implementing the proposed work is discussed in section 3.3. Pre-processing steps are discussed in 3.4. Feature extraction techniques and machine learning classifiers are discussed in 3.5and 3.6 respectively.

3.2 Basic Work-flow Diagram

The workflow of the proposed model is shown below, The project flow chart is established in Figure 3.1. This indicates that the suggested lung cancer prognosis system consists of four modules, i.e. nodule detection, nodule feature extraction, nodule classification based on extracted features, cancer stage determination based on malignancy of each detected malignant nodule. After the deep learning model selects nodules, their ensembled result will be used for feature extraction. Feature extraction will be accomplished using five TL models. For each model the performance of the system will be evaluated. In the end, the best performing TL will be selected as feature extractor for the system. The extracted features will be used to classify nodules as malignant or not using ML models. Among these models, the top performing models will be selected for ensemble learning to detect malignancy of nodules. Based on the malignancy of all the malignant nodules, the system will classify lung cancer into six stages as follows.

1. **Stage 0:** This stage refers to healthy lung with no nodules.

2. **Stage I:** It refers to cancer that is only present in the top layers of cells lining the air passages 2. in the lungs and has not invaded deeper tissues. It's considered a very early, non-invasive stage.

3. **Stage II:** At this stage, cancer is localized and confined to the lung. It may involve a small tumor in the lung (Stage IIA) or a larger tumor or one that has spread to nearby structures (Stage IIB).

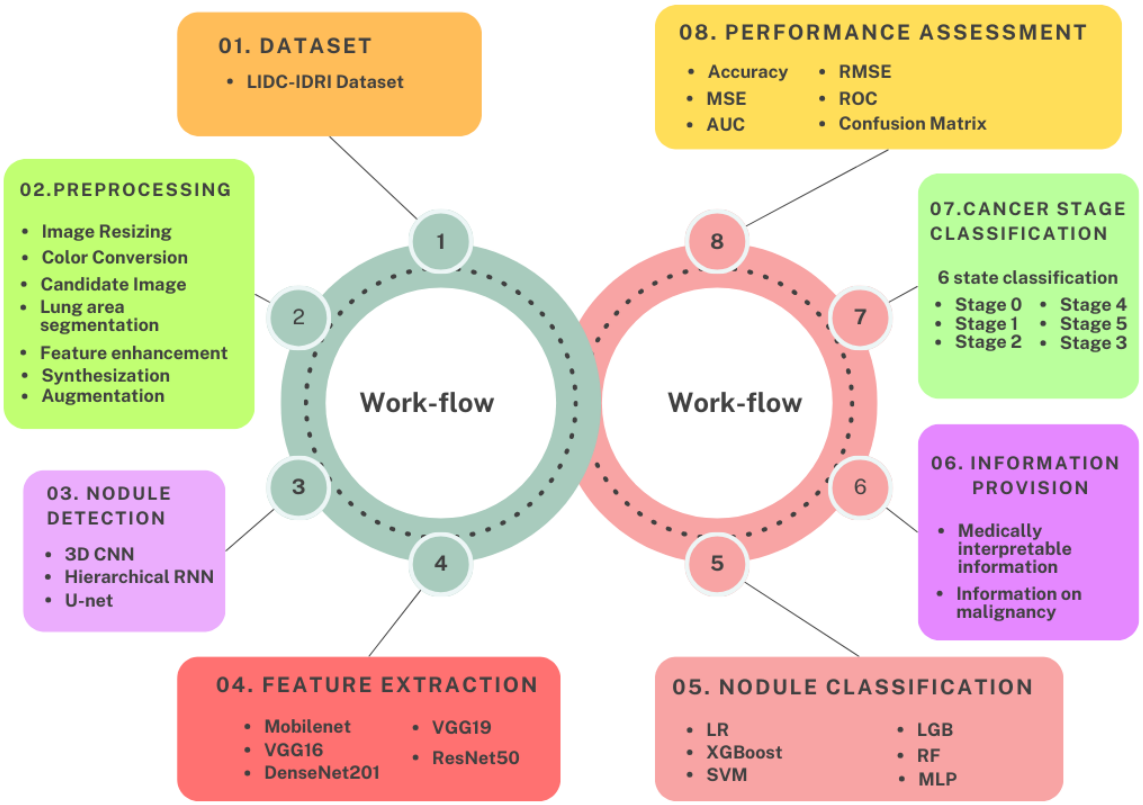

Figure 3.1: Basic work-flow of the proposed system.

4. **Stage III:** At this point, the cancer has extended to neighboring lymph nodes (Stage IIIA) or may have invaded nearby structures like the chest wall or diaphragm (Stage IIIB).

5. **Stage IV:** This stage indicates locally advanced lung cancer, where cancer has spread to lymph nodes farther away from the primary tumor within the chest (Stage IVA) or may involve nearby organs like the heart, great vessels, or the esophagus (Stage IVB).

6. **Stage V:** Cancer is distinguished by the existence of remote metastases, signifying that the malignancy has disseminated to additional bodily organs, such as the cerebral part, skeletal system, hepatic region, or other remote areas within the physique. This state is deemed as progressive or metastatic pulmonary cancer.

Along with this the system will also provide information about malignancy for each detected nodule.

3.3 The Proposed System Architecture

The suggested system for predicting lung cancer prognosis comprises four modules, namely, nodule detection, nodule feature extraction of the detected nodules, nodule classification from extracted features with TL models, and lastly cancer stage is determined using malignancy of each detected malignant nodule.

In the nodule detection step, there are three nodule detectors, i.e. U-Net3.2, 3D CNN3.3 and a Hierarchical RNN proposed in by Wang et al. [1] depicted in 3.4 for their unique qualifications of object detection. Then each of the three deep learning models are fed the images for nodule detection. Here, images are further processed for converting them into synthetic CT images before being fed to U-Net. The 3D CNN nodule detector is an object detection algorithm. It slides over small window through 3D space of the CT scan to detect nodules. Architecture of this model is a VGG network 3D extension proposed by Julian et al. [32].

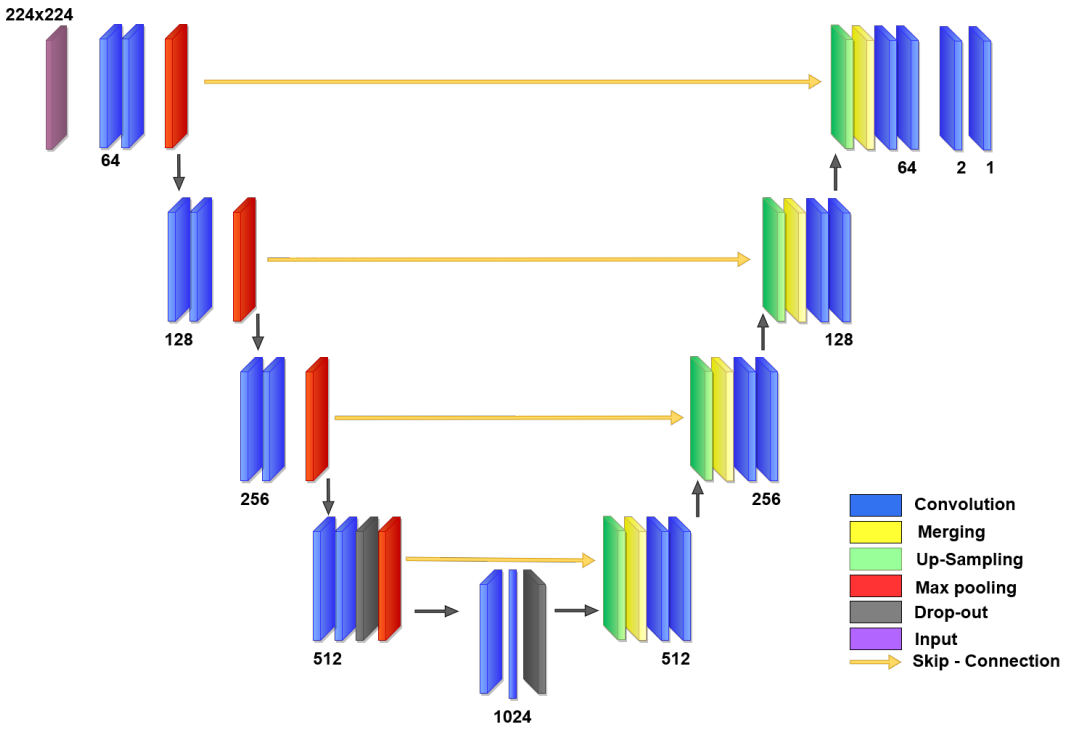

Figure 3.2: Proposed U-Net architecture for nodule detection.

The system architecture is established in Figure 3.5. This indicates that the input CT images will be preprocessed by resizing, restoring orientation, and removing noise. Then for better performance the features of the images are enhanced in the feature enhancement step. Then the images are augmented for having a large amount of data to train the model. Nodule detection is performed using DL models.

Then feature extraction is performed using five transfer learning models,i.e. VGG16, VGG19, ResNet50, MobileNet, DenseNet201 of the detected nodules from these models. For each transfer learning (TL) model, the extracted features are fed to six machine learning algorithm, i.e. logistic regression (LR), extreme gradient boosting

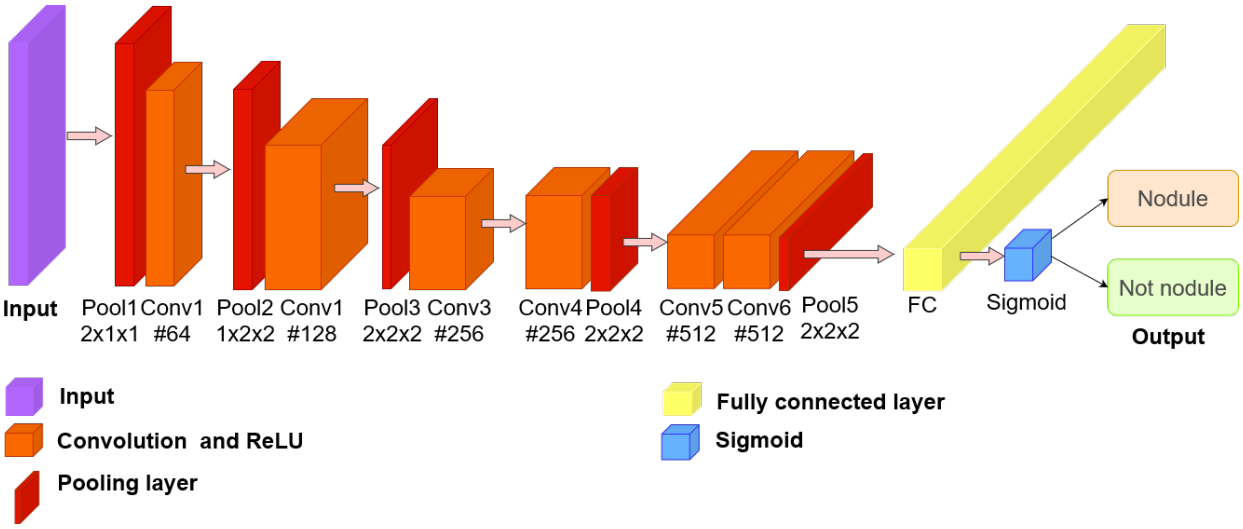

Figure 3.3: Proposed 3D CNN architecture for nodule detection.

(XGBoost), support vector machine (SVM), light gradient boosting (LGB), random forest (RF), multi-layer perceptron (MLP). Among these machine learning the top performing machine learning are selected for ensemble learning. After the nodules are determined as benign or malignant through ensemble learning step, based on the malignancy of each nodule cancer stage is determined. All of the five TL models are evaluated. The best performing TL model is selected for the proposed model. Then an evaluation is performed using recall, accuracy, precision, confusion matrix, f1-score, MAE, ROC curve, confusion matrix, MSE and RMSE to evaluate our model and also to show the comparison with existing models.

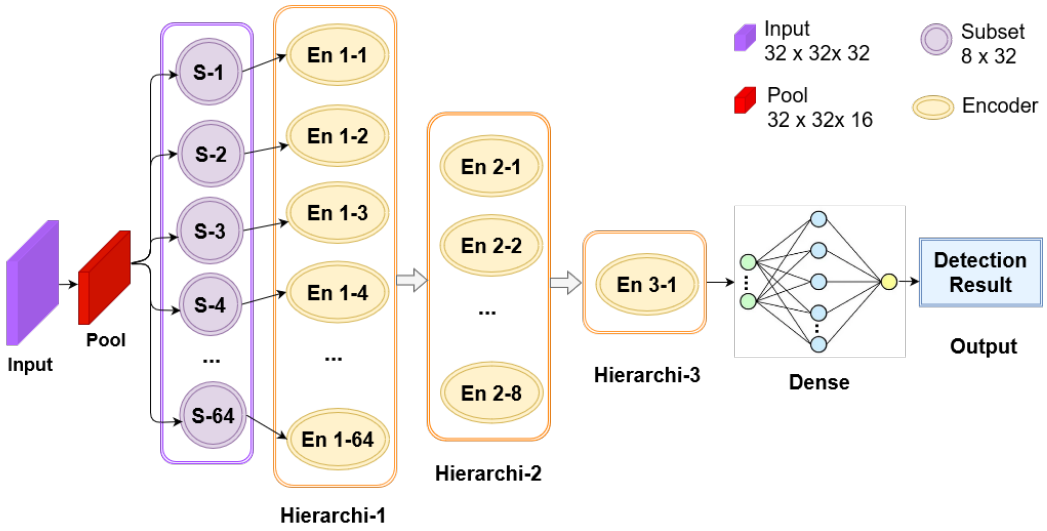

Figure 3.4: Hierarchical RNN architecture for nodule detection proposed in [1].

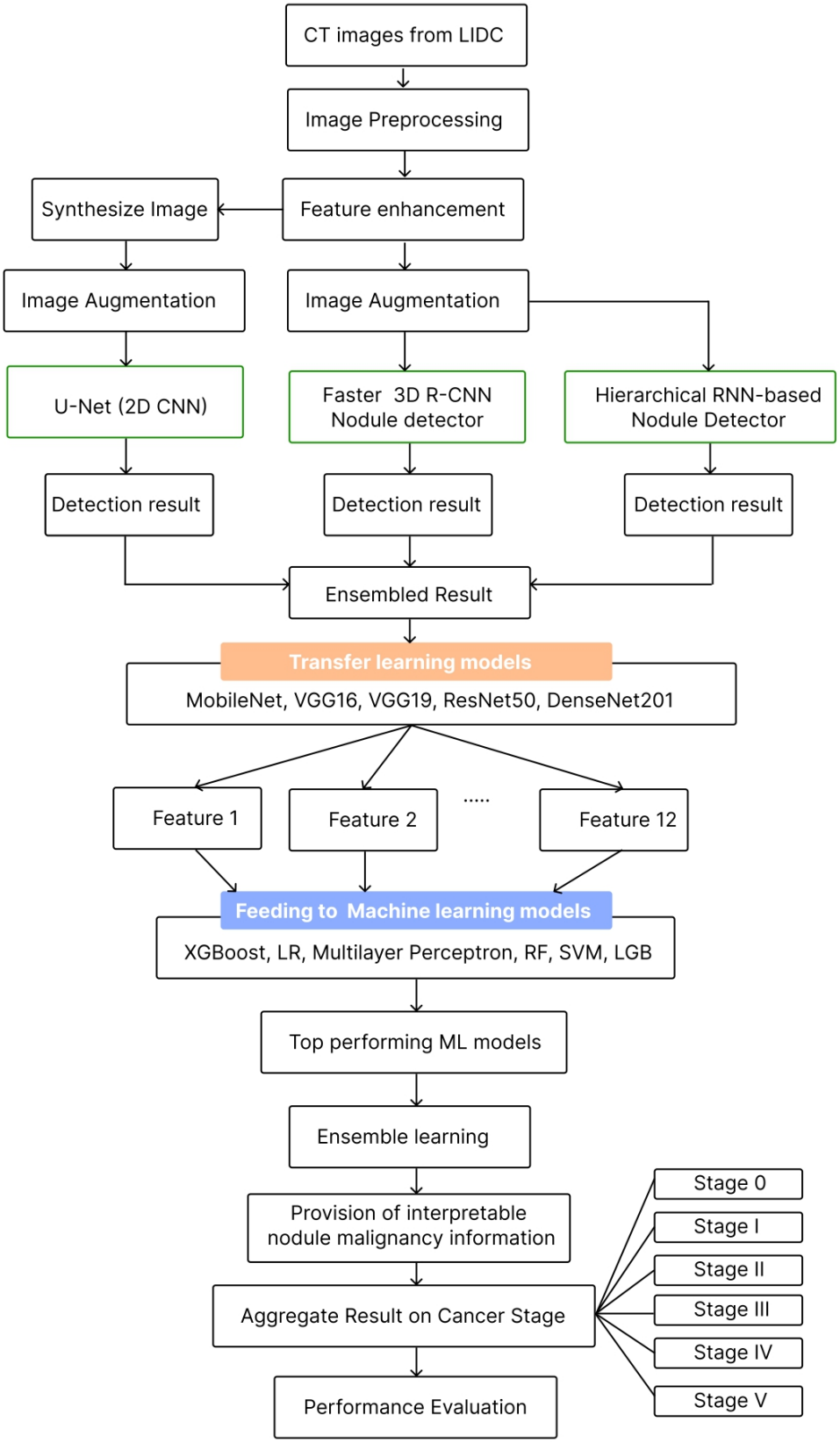

Figure 3.5: Proposed hybrid deep learning framework for the detection of lung nodule , classification with lung cancer stage determination.

3.4 Data Processing

3.4.1 Dataset

The chosen framework will utilize the Lung Image Database Consortium image collection (LIDC-IDRI). Described on its website[33], LIDC-IDRI comprises diagnostic and thoracic computed tomography (CT) scans for lung cancer screening, each annotated with marked-up lesions. It functions as an internationally accessible asset for developing, training, and evaluating computer-assisted diagnostic (CAD) systems' methods specifically designed for detecting and diagnosing lung cancer. In total, 1018 cases are contained in this dataset, the whole description of this dataset is depicted in 3.6. There are 10 features labeled in this dataset which are described in 3.7.

Dataset	LIDC-IDDRI
• CT Scan Cases	• 1018
• Candidate Scans	• 8249
• Scan with Nodules	• 5249
• Lebeled Features	• 10
• Augmented Images	• 365000

Figure 3.6: Description of LIDC-IDRI dataset.

3.4.2 Preprocessing

Performance of machine-learning based models depends on the quality and the volume of data. Hence, preprocessing low-dose computed tomography is important. The stages are depicted in 3.8. Preprocessing of input images are accomplished using seven steps, i.e. images resizing, color-space conversion, candidate image validation, lung area segmentation, feature enhancement, image synthesization and iamge augmentation. These steps are described in the following.

1. **Image resizing:** model has a certain criteria that it can take images only in a certain size. But, CT scans from different machines can have different resolutions with different sizes. Hence, resizing all the CT images is performed.The size of the input layer is 244x244x3.

2. **Color-space conversion:** single-channel images will be converted into three-channel RGB images using:

$$I_{\text{rgb}} = I_{(\text{gray, R})} + I_{(\text{gray, G})} + I_{(\text{gray, B})}$$

NO.	FEATURE	DESCRIPTION
1	Diameter	Nodule Diamer
2	Malignancy	Level of Malignancy
3	Sphericity	How spherical the nodule is
4	Margin	Clarity of nodule margin
5	Spiculation	Amount of needle-Lick structure from nodule boundary
6	Texture	How regular the texture of nodulle
7	Calcification	Calcification Degree
8	Internal-Structure	Internal Position of Nodule
9	Lobulation	Degree of non-uniform bulging
10	Subtlety	Contrast between nodule region & surrounding

Figure 3.7: Description of features labeled in LIDC-IDRI dataset.

3. **Candidate image validation:** Images that does not contain lung area where nodule can form are considered as non-candidate images. These images will be removed for efficient learning of deep-learning models.

4. **Lung area segmentation:** In this stage only the area of lung in each CT images will be segmented so that all computation for analysis will perform within this region, thus lessen complexity and time.

5. **Feature enhancement:** Following noise removal, the Genetic Algorithm-based Adaptive Histogram Equalization (GAAHE) algorithm, along with window and level contrast enhancement, will be applied.

6. **Image synthesization:** To input an image into the U-Net model, CT images are generated by converting the middle slice into the green channel, while the red and blue channels respectively contain a slice preceding and succeeding the middle slice.

7. **Image augmentation:** To address the challenge of overfitting, data augmentation is implemented. Given the substantial disparity between the quantities of positive and negative training images, the introduction of random spatial

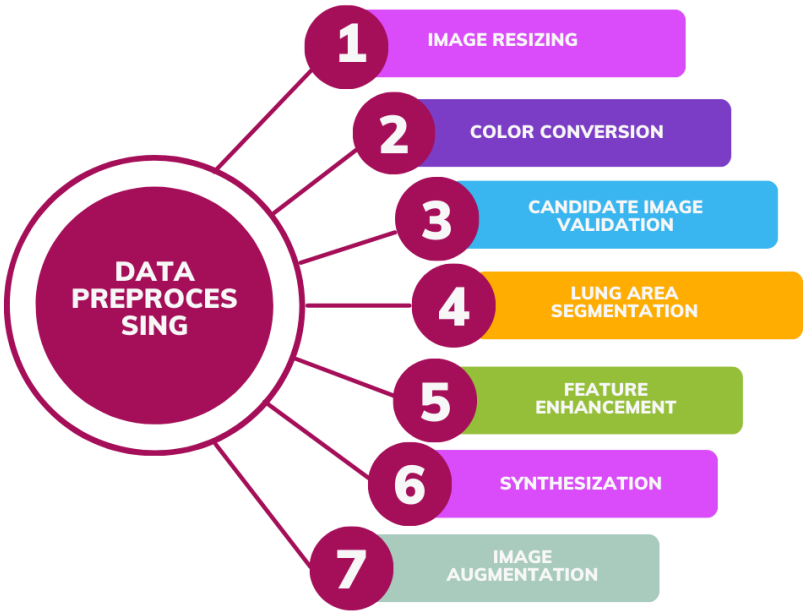

Figure 3.8: CT images of LIDC-IDRI datset is preprocessed before being sent to nodule detection models.

transformations facilitates the creation of a more balanced and extensive set of training samples.

3.5 Feature Extraction Techniques

Feature extraction is performed to extract the feature of all the detected lung nodules. These features includes diameter, malignancy, sphericity, margin, spiculation, texture, calcification, internal structure, lobulation, subtlety, probability of nodule, x-coordinate, y-coordinate, z-coordinate. To find the best feature extractor five TL model have been selected by now. Among these five TL models, the model with highest accuracy will be selected as the feature extractor for the proposed model and the evaluation will be carried on using it, in sha Allah.

3.5.1 MobileNet

With a reduced number of parameters, MobieNet is constructed based on depthwise separable convolutions, which were developed to minimize the computational effort in the initial layers. The architecture consists of two layers, namely depthwise and pointwise convolutions. Consequently, the pointwise convolution procedure, responsible for generating novel features, has been incorporated. Therefore, the term "depthwise separable convolution" was introduced to encapsulate the integration of these two layers. In a distinctive approach, this technique assigns an individual filter to each input stream using depthwise convolutions, followed by 1x1 convolutions (pointwise) to create a unified output distribution from the depthwise layer. Sub-

sequently, Batch Normalization (BN) is employed along with Rectified Linear Unit (ReLU) activation functions after each convolution. The efficiency of this approach is contingent on a balance between delay and accuracy. Given its compact nature, this method facilitates task execution on the CPU rather than necessitating reliance on the GPU.

3.5.2 VGG16

VGG16 attained the prestigious distinction of being crowned the victor of the 2014 ILSVRC object identification algorithm[34]. The VGG16 transfer learning approach encompasses three fundamental constituents, namely convolution, pooling, and fully connected layers. Within the convolution layer, the filters assume an integral role in the extraction of pertinent information from images, with the kernel and stride size serving as pivotal factors in this particular stratum. Conversely, the pooling layer is employed to effectively diminish network dimensions and computational burden by reducing the spatial scale. As for the fully connected layers, an extensive array of connections is established with the antecedent layers.A modified version of VGG16 is proposed is illustrated in 3.9. The last fully connected layer is cut-off so that the extracted feature can be passed to the next stage.

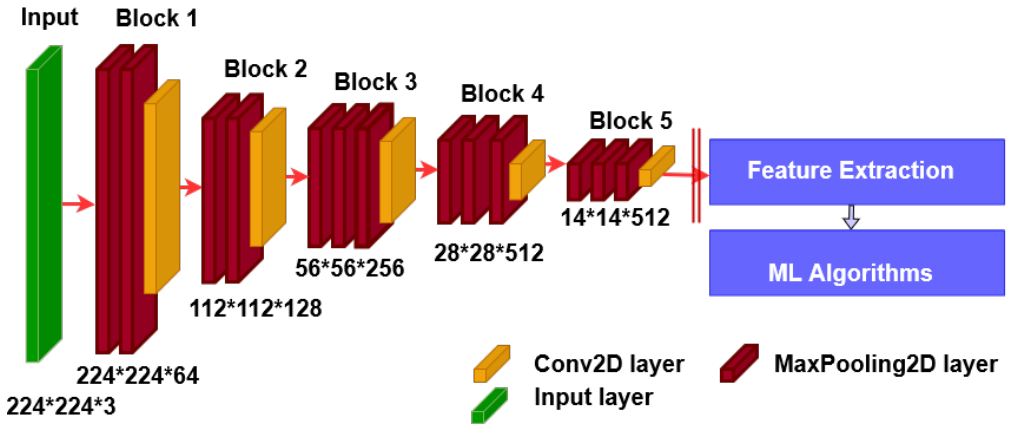

Figure 3.9: The proposed modified arcitechture of VGG16 for feature extraction.

3.5.3 VGG19

The VGG19 model, a CNN model, was introduced by [35]. It functions as a feature extractor with weighted ground characteristics comprising of 19 layers. These layers consist of a combination of 16 convolutional networks (convnets) and 3 dense layers, which are interconnected to classify images into 1000 categories. The model follows a sequence of 2 convolutional layers followed by 1 max pooling layer, repeated twice, then 4 convolutional layers followed by 1 max pooling layer, repeated thrice, and finally 3 fully connected layers. The extensive use of 3x3 filters in each convnet makes it a widely employed image prediction model.

3.5.4 ResNet50

The ResNet-50 architecture is a deep neural network used for image classification. It is built upon the concept of residual blocks, which include multiple convolutional layers and shortcut connections. These shortcut connections enable smoother gradient flow during training, making it feasible to construct very deep networks. ResNet-50 typically begins with an initial convolutional layer followed by max-pooling. It then consists of several residual blocks, which contain convolutional layers and shortcut connections. Periodic max-pooling layers are implemented to decrease the spatial dimensions of the feature maps.Towards the end of the network, fully connected layers are employed for making class predictions. The output layer often has 1,000 units, suitable for classifying images into 1,000 categories. The distinctive feature of ResNet-50 is the use of skip connections, which enhance the training of deep networks by addressing gradient vanishing issues.

3.5.5 DenseNet201

DenseNet, another tranfer learning model which is introduced by [36]. There is a mutual influence between each layer and all other layers in a forward-propagating manner. A DenseNet block comprises of five levels. The initial four layers are characterized by their density, whereas the last layer is a transition layer. The growth rate of each layer in this dense block is denoted as (k). If the growth rate is set to 4, a transition layer with a 2x2 average pool, a 1x1 convolution, and a stride of 2 is introduced. The dense layer consists of both 1x1 and 3x3 convolutions with a stride of 1. Our research experiment employs DenseNet201.

3.6 Classification Techniques

In these stage, the extracted features will be learned by ML models to classify whether or not a nodule is malignant or not. Six well-known ML models are decided to be used in this stage for their outstanding unique capability of classification. These models are discussed bellow.

3.6.1 Logistic Regression (LR)

Logistic Regression, often abbreviated as LR, is a technique that uses a special equation to make predictions. It helps us figure out which factors are the most important when trying to predict things. We can choose the most important factors by checking their importance one by one, making sure they are really important (with a p-value less than 0.05). It's like solving a puzzle, where we want to guess categories or labels based on some information. In the end, our predictions will also be in categories, like "yes" or "no," or in specific groups.

3.6.2 Random Forest (RF)

Random Forest is a meta-approach that leverages averaging for improvement of accuracy. It employs multiple decision tree classifiers on various subsets of the dataset to mitigate overfitting. This involves creating independent decision trees from various bootstrap samples extracted from the training data, each with a separately sampled attribute subset. The final classification model is formed by combining the outcomes of these trees through voting. By aggregating predictions from numerous decision trees, Random Forest achieves improved prediction accuracy.

3.6.3 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a classification method that starts by identifying a hyperplane and then using it to classify different classes. SVM is a widely adopted machine learning technique for classification tasks. It enhances generalization by minimizing the structural risk of the learning model. SVM is popular across various domains due to its high accuracy and quick training speed, especially when working with relatively small datasets. Its efficiency in classification tasks is well-established as it strives to maximize the model's ability to generalize. Moreover, since solving an SVM problem is essentially a convex optimization problem, the solution obtained internally is also globally optimal. Therefore, SVM is frequently utilized for training base classifiers in various applications.

3.6.4 Light Gradient Boosting (LGB)

Light Gradient Boosting (LGB) represents a novel approach to gradient learning inspired by the concept of decision trees[37]. It surpasses XGBoost and is characterized by its efficient use of memory, adopting a growth method which is leaf-wise with specified depth constraints, and employing an algorithm which is histogram-based to accelerate the learning procedure. LightGBM constructs histograms of width 'k' by discretizing continuous floating-point values into 'k' bins using the histogram algorithm mentioned above.

3.6.5 Extreme Gradient Boosting (XGBoost)

XGB is like a supercharged helper for making predictions with data. It's like having a group of smart decision-makers, and they work together really efficiently. XGB is faster than many other similar methods. To make the best predictions, it combines the wisdom of the group by learning from what was not perfect in previous guesses and making new and improved guesses. It's like a team effort where each member learns from the others and tries to get better each time. This method is great at getting closer to the right answer faster.

3.6.6 Multi-Layer Perceptron (MLP)

A Multi-Layer Perceptron (MLP) is a common type of artificial neural network (ANN) that works like a stack of building blocks. Each block has tiny decision-making

units, called neurons, that talk to each other. These neurons calculate and discuss the information they receive, and they decide whether to send a message to the next block. Think of it like a relay race where one runner passes the baton to the next. In the middle, there are hidden layers where they have their discussions, and it's like teamwork. The number of runners (neurons) in the starting block equals the things we want to know about, and the number of runners in the last block equals the choices we can make. It's a clever way to solve problems and make decisions based on data.

3.6.7 Ensemble learning (EL)

Ensemble learning involves the process of training and combining multiple individual learners to effectively address a specific problem[38]. In hard voting, the decision-making process relies on the predicted class labels provided by individual classifiers. Each classifier offers its prediction, and the final decision is determined by a majority vote based on the class labels predicted by the individual learners. This approach is particularly valuable when working with a diverse set of classifiers.

In **hard voting**, the majority rule is used to select the class label. The decision is made by counting the class labels predicted by individual classifiers. Each classifier provides its prediction, and the final choice is based on the majority vote. This method is useful when working with a diverse group of classifiers.

On the other hand, in **soft voting**, class labels are predicted based on the estimated probabilities provided by the classifiers, using the average class probability[39]. By default, the implementation calculates soft voting using a uniform value, resulting in a consistent weight applied in the weighted average process across all elements. This soft voting method is especially useful when dealing with a set of well-calibrated learners.

3.7 Validation Technique

3.7.1 K-fold Cross-Validation

K-fold Cross-Validation (CV) is a fundamental technique that involves partitioning a training dataset into k smaller subsets. In each of the k iterations, a fold is used as the training data for a model, while the remaining folds serve for validation. The process is repeated for each fold, and the average of the evaluation metrics obtained during these iterations is used as the final performance measure. In this specific case, the dataset is divided into 10 separate folds, with each one taking turns as the testing set while the others are used for training in a systematic manner.The procedure of k-fold cross-validation is depicted in 3.10.

3.7.2 Stratified K-fold Cross-Validation

Stratified Cross-Validation is a rigorous approach commonly employed in academic research to enhance the reliability of model evaluation, particularly when dealing with

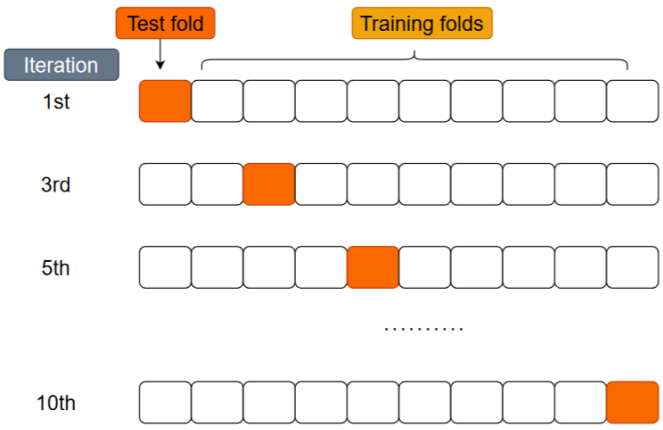

Figure 3.10: Demonstration of K-Fold Cross-validation.(For K = 10)

datasets containing imbalanced class distributions. In essence, it's a method that ensures each "fold" or subset of data used for testing during cross-validation maintains a proportion of classes similar to that of the entire dataset. The primary aim is to prevent any single fold from having a significantly different class distribution compared to the complete dataset. This approach is particularly crucial when dealing with classification problems where one class may be significantly smaller or more abundant than others. By maintaining this balance, Stratified Cross-Validation promotes fairness and accuracy in assessing a model's performance, providing a more robust and academically sound evaluation. The procedure of stratified k-fold cross-validation is depicted in 3.11.

Figure 3.11: Schematic diagram of Stratified K-fold cross validation.[2]

3.8 Evaluation Techniques

There are a bunch of ways to check how well our model is doing, like looking at confusion matrix, recall, accuracy, F1-score,ROC curve, precision , and various measures like MAE, RMSE, and MSE.The matrices to check how well the model holding up under all those scrutiny lenses are stated below:

- **Confusion Matrix** is a table-based tool which can assess a machine learning classification system's performance. It categorizes predictions into four outcomes: True Positive (correctly predicted positives), True Negative (accurately predicted negatives), False Positive (incorrectly forecasted positives), and False Negative (inaccurately predicted negatives). These outcomes aid in calculating metrics such as accuracy, precision, f1-score and recall and are instrumental in plotting the ROC curve.

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	TN	FN

Table 3.1: Description of Confusion matrix.

- **Accuracy** is a straightforward metric. It's the proportion of accurate predictions to the overall number of observations.

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

- **Precision** is like the standard for correctly predicted positive values. It's the ratio of how many times a model gets a positive prediction right to the total number of times it predicts something positive.

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

- **Recall** is like the thorough detective of predictions. It's the ratio of positive values which are correctly predicted to the total number of actual positive values.

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

- **F1-score** represents a harmonious synthesis of precision and recall in a classification problem, providing a balanced metric that considers both the accuracy of positive predictions (precision) and the model's ability to capture all relevant instances (recall).

$$F1 - score = 2 * \frac{recall * precision}{(recall + precision)}$$

- **MAE** is like the middle ground of errors. It calculates the average absolute difference between predicted and actual values, providing an indication of the average discrepancy in predictions.

$$MAE = \sum_{i=1}^{n} \frac{p(i) - a(i)}{n}$$

- **MSE** is like the squared version of errors. It calculates the mean of the squared discrepancies between predicted and actual values., assigning greater significance to larger errors. The catch is, MSE is always positive, which means it doesn't let errors cancel each other out.

$$MSE = \sum_{i=1}^{n} \frac{(p(i) - a(i))^2}{n}$$

- **RMSE** is a commonly used metric for assessing regression problems, assuming unbiased errors and a normal distribution. However, it's crucial to note that RMSE can be sensitive to outliers, which may impact its accuracy in the presence of extreme values.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(p(i) - a(i))^2}{n}}$$

  Here,

  - p(i) is the predicted value of i
  - a(i) is the actual value of i
  - n is the total number of values.

- **ROC curves** are effective tools for evaluating classifiers, providing a visual representation of a model's ability to differentiate between classes. The plot highlights the trade-off between sensitivity and specificity across various classification thresholds, which is especially useful for user requirements related to error costs and accuracy expectations. The AUC (area under the curve) adds another layer by measuring the model's capacity to distinguish between classes. A higher AUC indicates strong class separability, while a lower value suggests room for improvement in this aspect.

We have chosen to integrate the specified terms—such as accuracy, ROC curve, confusion matrix, MSE, MAE, and RMSE—into our system as key components for evaluating and assessing the future performance of our models. These metrics provide a comprehensive and nuanced analysis, allowing us to gain insights into various aspects of model behavior, from classification accuracy to error measurement in regression tasks.

CHAPTER IV

Conclusion Future Work

4.1 Conclusion

Lung cancer stands out as one of the most prevalent and unfortunately, deadliest forms of cancer, often remaining asymptomatic until reaching advanced stages. The accurate detection and segmentation of lung nodules from CT images are pivotal for early diagnosis. This research aims to identify and classify lung nodules as benign or malignant from low-dose CT images, with the ultimate goal of determining the stage of lung cancer. The study pursues this objective by showcasing a method that combines deep learning and machine learning models for the identification of lung cancer from CT images. For better performance preprocessing methods are implemented. Three DL models, 2D CNN based U-Net, 3D CNN, hierarchical RNN are considered to be used for nodule detention for their unique capability of learning from data. By now, the feature extractors are taken into account are five TL models, MobileNet, ResNet50, DenseNet201, VGG16, VGG19. And six machine learning classifiers, LR, SVM, RF, LGB, XGBoost, MLP , are considered to be used to find the best performing ML models which are considered for ensemble learning methods to determine whether nodules are malignant or not. The model considers malignancy of each detected malignant nodule for determining lung cancer stage. Along with this, the information about malignancy of each detected malignant nodule is provided for determining the validity of the decision on cancer stage.

4.2 Gantt Chart

Figure 4.1 shows a gantt chart showing the relative timing of research activities and their relationships.

Figure 4.1: Gantt Chart

References

[1] W. Wang and G. Chakraborty, "Evaluation of malignancy of lung nodules from ct image using recurrent neural network," in *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2992–2997, 2019.

[2] X. Duan, "Automatic identification of conodont species using fine-grained convolutional neural networks," 2023. Accessed: November 9, 2023.

[3] J. Ferlay, "Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2," *Int. J. cancer*, vol. 136, pp. .E359–86, 2015.

[4] R. Siegel, K. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.

[5] K. D. M. R. L. Siegel and A. Jemal, "Cancer statistics, 2," *CA. Cancer J. Clin*, vol. 68, p. 7–30, 2018.

[6] D. R. Baldwin, "Prediction of risk of lung cancer in populations and in pulmonary nodules: Significant progress to drive changes in paradigm," *Lung Cancer*, vol. 89, no. 1, pp. 1–3, 2015.

[7] J. Llert and L. Kreel, "The role of computed tomography in the initial staging and subsequent management of the lymphomas," *J Comput Assist Tomogr*, vol. 4, no. 3, pp. 368–391, 1980.

[8] P. Bach, M. Kelley, R. Tate, and D. McCrory, "Screening for lung cancer: a review of the current literature," *Chest*, vol. 123, no. 1, pp. 72–82, 2003.

[9] D. Aberle, A. Adams, C. Berg, W. Black, J. Clapp, R. Fagerstrom, I. Gareen, C. Gatsonis, P. Marcus, and J. Sicks, "Reduced lung-cancer mortality with low-dose computed tomographic screening," *N Engl J Med*, vol. 365, no. 5, pp. 395–409, 2011.

[10] S. Shen, S. Han, D. Aberle, A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," *Expert Systems with Applications*, vol. 128, pp. 84–95, 2019.

[11] S. Li and D. Liu, "Automated classification of solitary pulmonary nodules using convolutional neural network based on transfer learning strategy," *Journal of Mechanics in Medicine and Biology*, no. 2140002, 2021.

[12] S. Singh, D. Gierada, *et al.*, "Reader variability in identifying pulmonary nodules on chest radiographs from the national lung screening trial," *Journal of Thoracic Imaging*, vol. 27, no. 4, p. 249, 2012.

[13] M. Winkels and T. Cohen, "Pulmonary nodule detection in ct scans with equivariant cnns," *Medical Image Analysis*, vol. 55, pp. 15–26, 2019.

[14] G. Zhang, S. Jiang, Z. Yang, L. Gong, X. Ma, Z. Zhou, C. Bao, and Q. Liu, "Automatic nodule detection for lung cancer in ct images: A review," *Computers in Biology and Medicine*, vol. 103, pp. 287–300, 2018.

[15] A. Naik and D. Edla, "Lung nodule classification on computed tomography images using deep learning," *Wireless Personal Communications*, vol. 116, no. 1, pp. 655–690, 2021.

[16] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, and J. Wu, "Survey on deep learning for pulmonary medical imaging," *Frontiers in Medicine*, pp. 1–20, 2019.

[17] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, and P. Abreu, "Using deep learning techniques in medical imaging: A systematic review of applications on ct and pet," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4093–4160, 2020.

[18] W. Cao, R. Wu, G. Cao, and Z. He, "A comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans," *IEEE Access*, vol. 8, pp. 154007–154023, 2020.

[19] S. Sharma, M. Kaur, and D. Saini, "Lung cancer detection using convolutional neural network," *International journal of engineering and advanced technology*, vol. 8, no. 6, pp. 3256–3262, 2019.

[20] M. H. Hesamian, W. Jia, X. He, Q. Wang, and P. J. Kennedy, "Synthetic ct images for semi-sequential detection and segmentation of lung nodules," *Applied Intelligence*, vol. 51, pp. 1616–1628, 2021.

[21] J. Song, S.-C. Huang, B. Kelly, G. Liao, J. Shi, N. Wu, W. Li, Z. Liu, L. Cui, M. P. Lungre, *et al.*, "Automatic lung nodule segmentation and intra-nodular heterogeneity image generation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2570–2581, 2021.

[22] Y.-W. Wang, C.-J. Chen, H.-C. Huang, T.-C. Wang, H.-M. Chen, J.-Y. Shih, J.-S. Chen, Y.-S. Huang, Y.-C. Chang, and R.-F. Chang, "Dual energy ct image prediction on primary tumor of lung cancer for nodal metastasis using deep learning," *Computerized Medical Imaging and Graphics*, vol. 91, p. 101935, 2021.

[23] N. Faruqui, M. A. Yousuf, M. Whaiduzzaman, A. Azad, A. Barros, and M. A. Moni, "Lungnet: A hybrid deep-cnn model for lung cancer diagnosis using ct and wearable sensor-based medical iot data," *Computers in Biology and Medicine*, vol. 139, p. 104961, 2021.

[24] P. Sahu, D. Yu, M. Dasari, F. Hou, and H. Qin, "A lightweight multi-section cnn for lung nodule classification and malignancy estimation," *IEEE journal of biomedical and health informatics*, vol. 23, no. 3, pp. 960–968, 2018.

[25] M. A. Khan, S. Rubab, A. Kashif, M. I. Sharif, N. Muhammad, J. H. Shah, Y.-D. Zhang, and S. C. Satapathy, "Lungs cancer classification from ct images: An integrated design of contrast based classical features fusion and selection," *Pattern Recognition Letters*, vol. 129, pp. 77–85, 2020.

[26] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," *Expert systems with applications*, vol. 128, pp. 84–95, 2019.

[27] M. Nishio, C. Muramatsu, S. Noguchi, H. Nakai, K. Fujimoto, R. Sakamoto, and H. Fujita, "Attribute-guided image generation of three-dimensional computed tomography images of lung nodules using a generative adversarial network," *Computers in Biology and Medicine*, vol. 126, p. 104032, 2020.

[28] G. Kasinathan, S. Jayakumar, A. H. Gandomi, M. Ramachandran, S. J. Fong, and R. Patan, "Automated 3-d lung tumor detection and classification by an active contour model and cnn classifier," *Expert Systems with Applications*, vol. 134, pp. 112–119, 2019.

[29] N. Maleki, Y. Zeinali, and S. T. A. Niaki, "A k-nn method for lung cancer prognosis with the use of a genetic algorithm for feature selection," *Expert Systems with Applications*, vol. 164, p. 113981, 2021.

[30] W. Wang and G. Charkborty, "Automatic prognosis of lung cancer using heterogeneous deep learning models for nodule detection and eliciting its morphological features," *Applied Intelligence*, vol. 51, pp. 2471–2484, 2021.

[31] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning," *Expert Systems with Applications*, vol. 205, p. 117695, 2022.

[32] J. De Wit, "2nd place solution for the 2017 national datascience bowl,"

[33] B. Vendt, "The cancer imaging archive," 2023.

[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. Accessed: 2021-12-22.

[35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, 2014.

[36] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[37] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 3146–3154, 2017.

[38] Z. Zhou, "Machine learning," *China Merchants*, vol. 3, no. 3, pp. 30–35, 2016.

[39] S. Raschka, *Python Machine Learning*. Packt Publishing Ltd, 2015.