International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V71I1P204
Volume 71 Issue 1, 31-40, January 2023

Original Article

Computer-Aided Detection (CAD) system for the Detection and Classification of Pulmonary Nodules in Lung Cancer using CT images

Aparna M. Harale[1], Vinayak K. Bairagi[2]

[1,2]*Department of Electronics and Telecommunication, AISSMS Institute of Information Technology, India.*

[1]*Corresponding Author : aparna.harale@gmail.com*

Received: 23 October 2022 Revised: 13 December 2022 Accepted: 23 December 2022 Published: 24 January 2023

Abstract *- Lung cancer is one type of serious disorder around the globe. In comparison to other forms of cancer in both males and females, lung cancer records the highest number of cancer-related deaths. Pulmonary nodules are blob-like shapes that are potential manifestations of lung cancer and diameter between 3 to 30 mm. correct examinations of nodules are required for the lung cancer diagnosis and the subsequent treatment schedule. Lung cancer screening can significantly reduce the death rate. Novice radiologists rely on professional specialists to study lung CT images. The primary bottleneck for such a traditional learning system is a lack of time from professional radiologists. Hence, there is a requirement for the development of an Automatic Computer Aided Detection system (CAD) to assist radiologists and the analysis of lung cancer. Due to poor image quality that interferes with the segmentation process, traditional lung cancer prediction approaches could not maintain accuracy. In order to predict lung cancer, novel, improved image processing and machine learning technique is presented in this study. This paper aims to develop an Automatic CAD system to diagnose lung cancer. The Lung Image Database Consortium (LIDC) CT image was used. The detection of nodules is done using UNET architecture, and its malignancy is decided with an ensemble of three classifiers, SVM, KNN and LR are examined using Python. LR gives the highest accuracy, 97.92%. Performance check of different classifiers using Accuracy, Sensitivity, Specificity, Precision, and F1 Score. The ensemble of three classifiers model detects and classifies lung nodules with an accuracy of 83%.*

Keywords *- Computer tomography, Computer-aided diagnosis, Lung cancer, Nodule, Nodule detection.*

1. Introduction

One form of severe disease in the world is lung cancer. The lung cancer death rate is more comparable to any other cancer type, such as brain, prostate, and breast cancers, per year. For people aged 45 to 70 years, the Lung has the highest incidence rate of cancer-related death. [1]. Lung cancer is greatest frequently infected cancer in both genders. In the US, lung cancer is the largest cause of cancer-related death, in line with the statics of the American cancer society [1]. Estimated 228,820 (i.e. 116,300 cases of Men and 112,520 cases of female) new lung cancer cases are supposed to occur in 2022. The estimate accounts for almost 14% of all cancer diagnoses [1]. Approximately 135,720 deaths due to lung cancer will happen in 2022 [1]. Lung cancer records more deaths compared to any other cancer in both genders. To SEER Cancer Statistics Review, the Five-year Relative Survival Rates, when combined, are 19%, 16% in males and 23% in females [1]. Early analysis of the lung tumor is the solution to increase the survival rate by 57% for 1 to 5 years. It can be further increased from 65 to 80%. Hence major research efforts are made in the early detection of lung nodules to win the war against lung cancer [2]. Due to the decline in smoking prevalence, death rates have decreased by 51 percent each year in men and by 26 percent each year in women from 2008 to 2017 [1]. A pulmonary nodule in the CT images is identified with almost rounded opaqueness, comparatively proper margin and a maximum diameter of less than 3cm [5].

The purpose of the CAD system:

- To the development of an Automatic Computer Aided Detection (CAD) system for the diagnosis of lung cancer.
- Improve the accuracy of detection
- Decrease the estimation period of the radiotherapist.
- Detection of lung nodule cancer at the initial stage will improve the patient's survivability by over 5 years.
- Output from the CAD system is considered a second opinion and assists radiotherapists in concluding.
- To lower the price for several unnecessary health checks (avoid unnecessary biopsies through a lung).

The current paper aims to develop an Automatic Computer Aided Detection system (CAD) tool for lung cancer diagnosis. This paper is arranged into subsequent sections: section 1 provides an introduction, and section 2

discusses some related works. Section 3 is described the proposed methodologies and techniques to be used for the Lung Cancer detection CAD system; in Section 4, experimental results are described. Lastly, the conclusions and Section 5 discuss the future direction.

2. Related Work

The Block diagram of the lung nodule detection CAD system is shown in Fig.1. There are five phases of the CAD system. 1) Acquisition, 2) Pre-processing, 3) Segmentation, 4) Nodule Detection and false positive reduction, 5) Nodule Classification. The following subsections describe each block and existing methods relating to each component.

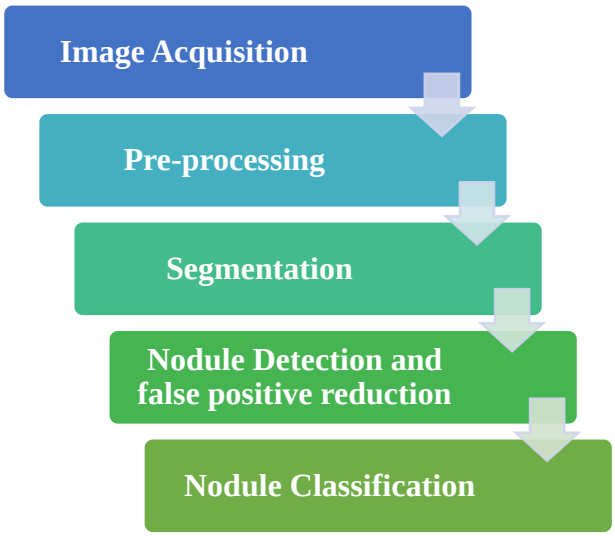


Fig. 1 Block diagram of lung nodule detection CAD system

2.1. Image Acquisition

Image acquisition is the procedure of obtaining CT scans in medicine. Lung CT pictures are available in numerous public and private databases. Early Lung Cancer Action Program (ELCAP) [3], Lung Image Database Consortium (LIDC) in National Imaging Archive [4], and Medical Image Database [5] are three popular public lung nodule databases. The LUNA16 dataset [8] is a subsection of the LIDC-IDRI dataset, in which various benchmarks sort the various images.

2.2. Pre-Processing

The Pre-processing image phase means image improvement to remove noise, which plans to make it easier to understand or sensitivity of knowledge in the image to supply enhanced input for additional programmed image processing methods and boundary identification. Various filtering methods in literature to eliminate these noises are wiener filter, median filter, Isotropic re-sampling, and adaptive median, selective enhancement filter. The literature found that the median filter gives the best de-noising performance over CT images without distorting the edge. Different preprocessing techniques proposed in the literature are compiled in Table 1.

Table 1. Preprocessing method reported in the literature

Author	Reported	Method of Pre-processing	Purpose
Obulesu et al.[20]	2021	Signed-Rank Gain	To obtain significant feature
Shakeel et al.[10]	2020	multilevel brightness-preserving approach	To remove noise
Reddy et al. [11]	2019	image securing, pre-handling, Binarization, thresholding,	To preprocess the image
Makaju et al. [12]	2018	median and Gaussian filter	To remove noise

2.3. Lung Segmentation

The segmentation method separates the image into its component regions or objects and outlines things and boundaries, such as lines, curves, etc., in images. Different segmentation techniques proposed in the literature are Region-Based Segmentation, model-based method, Thresholding, Morphological Operation etc. Different segmentation techniques were proposed in the survey and are presented in Table 1.

Table 2. Literature review of Segmentation Techniques of lung CT images

Author	Year	Dataset	Proposed technique	Performance (Accuracy)
Li et al.[13]	2020	Japanese Society of Radiological Technology	lung field segmentation and rib suppression	99%
Bhandary et al.[17]	2020	LIDC IDRI	Morphological and watershed segmentation	97.27%
Makaju et al. [12]	2018	LIDC IDRI	Watershed segmentation	92%.
Filho et al.[15]	2017	CT scans 40	Shape-based deformable	99.14%
Soliman et al.[14]	2017	105	Shape-based	98.4%
Shi et al. [16]	2016	CT scans 23	Thresholding	98%

Table 3. Different review methods of lung nodule Detection process

Study	Reported	Detection method	Dataset	Remark
Wenfa Jiang et al.[21]	2022	U-Net and RPN network	LIDC	Sensitivity 93%, Specificity 94%, ROC_AUC 93%
O. Obulesu et al.[20]	2021	Wilcoxon Signed Generative Deep Learning	470	Accuracy 86%
Li et al.[13]	2020	Multi-resolution patch-based CNNs were trained	Japanese Society of Radiological Technology	Accuracy 99%
Reddy et al.[11]	2019	fuzzy neural system use with ML	From UCI repository	Accuracy 96.67%.

Table 4. Different review methods of lung nodule Classification Techniques

Authors	Year	Technique/Method	Remarks	Database
Bhandary et al.[17]	2020	EFT is used to classify lung CT images.	Accuracy 97.27%	LIDCIDRI
P. Shakeel [9]	2019	Discrete AdaBoost optimized ensemble learning generalized	Accuracy 99.48%	--
Bhatia et al.[18]	2019	XGBoost and Random Forest	Accuracy 84%	LIDCIDRI

2.4. Lung Nodule Detection and False Positive Reduction

2.4.1. Lung Nodule Detection

The method of determining the probability of nodule patterns in an image is known as lung nodule detection, which also identifies the nodule location within the lung field. Lung nodule detection is an essential and especially important step in extracting Region of interest (ROI). Different detection techniques were proposed in the survey and are presented in Table 3.

2.4.2. False Positive Reduction

False positive reduction is the procedure of further eliminating the false positives (nodule similar structural formation through detection of nodule). Non each remaining system incorporates a false positives reduction section. Two-stage technique, multi-resolution-based technique etc., are used in literature to decrease the false positive occurrences.

2.5. Nodule Classification

Nodule Classification is the process of classifying if the detected nodule is benign or malignant. Nodules with a diameter of less than 10 mm are quite common, and the vast majority of them are harmless. Nodules with a diameter bigger than 10 mm are considered malignant and pose a significant risk of malignancy. In a high-risk population, the percentage of lung cancer with small nodules is below 10% [2]. Different classification techniques proposed in the literature are presented in Table 4.

3. Methodologies and Techniques to be Used

The proposed technique steps are Acquisition, preprocessing, Segmentation, Nodule Detection, feature extraction, and followed by Nodule Classification. We used Python to implement this design.

3.1. Implementation of lung nodule Detection with sample CT images

INPUT-

3.1.1. LIDC –IDRI Dataset

The Lung Image Database Consortium image collection (LIDC-IDRI) contains thoracic CT scans with marked-up labeled tumors for diagnosis and lung cancer screening. It is a 125GB dataset, with 1018 patients' information consisting of nodules and non-nodules.

3.1.2. List csv file

Consists of the slice no. where the nodule is located, scan no., volume, x and y locations, diameter, and nodule IDs. The LIDC dataset contains 3D CT volumes of patients along with XML files. The XML file includes annotations by radiologists developed in 2 phase process. The dataset has a csv file which contains x, y, and z coordinates and the radius of benign or malignant nodules. The labels for malignant and benign nodules are generated using the pylidc library. Pylidc library uses xml annotations to query data in an SQL-like fashion.

OUTPUT: Lung Nodule Detected with the help of classification techniques.

Flow Chart of Methodology

CT scans are represented in Hounsfield Units (HU). Hounsfield unit is a quantitative radio density measurement scale. Pixel data is converted into Hounsfield units using the formula:

$$\text{HU} = \text{Gray Value} * \text{slope} + \text{intercept}$$

The pixels outside CT scans are removed by setting them to 0. HU value for air is - 1024. Background and air gaps are removed by setting values to 0, as shown in fig. 2.

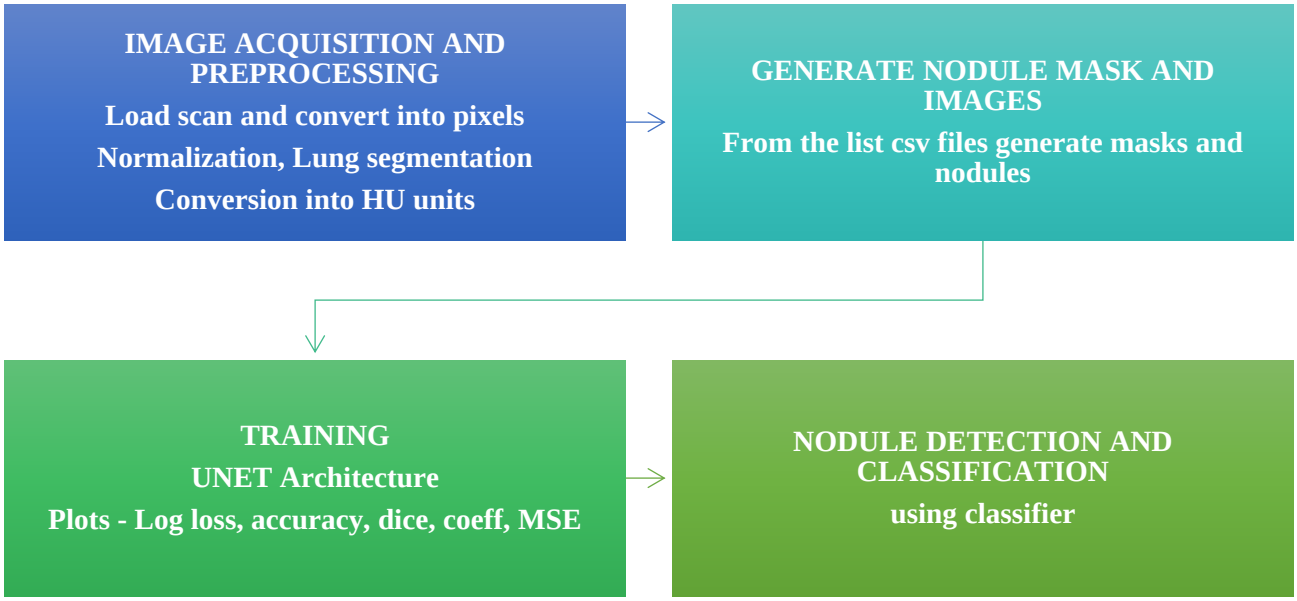


Fig. 2 Flow chart of Methodology

The scans are standardized to improve the model's performance and are centered on the mean with unit standard deviation. The regions whose CT values are not consistent with that of lung tissues are masked.

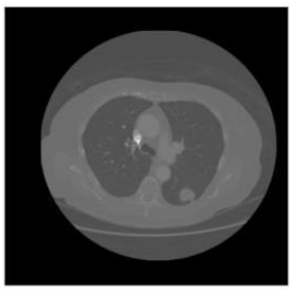


Fig. 3 Original CT scan background

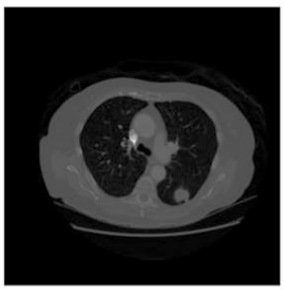


Fig. 4 CT scan after removing

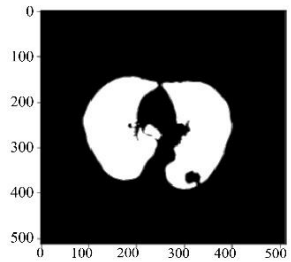


Fig. 5 Lung segmentation output

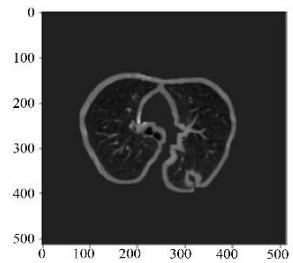


Fig. 6 Scans after preprocessing

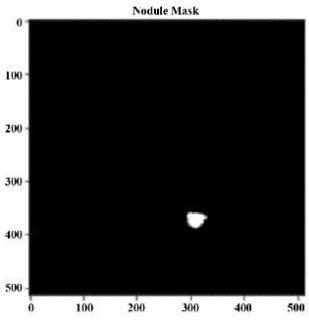


Fig. 7 Nodule Mask

K means clustering is implemented with the number of clusters equal to 2 for thresholding lung segment. The result is shown in fig. 5. Morphological erosion and dilation are used to add pixels into the lung wall to detect nodules confined to the wall. The final image after preprocessing is shown in fig. 6. These are used as input to the neural network for training.

Nodule locations for each patient are given in a csv file list3_2.csv. Cell magic wand package generates masks by using the x and y coordinates of the nodule at the z location.

Training the Network

a) UNet

UNet is a convolutional neural network developed for biomedical picture segmentation; its design has grown in size with minor changes to the CNN architecture. Convolutional Neural Networks (CNNs) performed admirably in simpler picture segmentation issues but not so well in more difficult ones. UNet enters the picture at this point. UNet was originally created to segment medical images. It had such positive outcomes that it was applied to various other fields. A recent graphics processing unit (GPU) takes less than a second to segment a 512*512 image after segmentation. U-Net is an architecture for semantic segmentation. The convolutional network's contracting path follows the standard architecture. UNet, multiple encoders (backbones) can be used here to generate strong features for the input image. Preprocessed 2D scans are passed into the network of UNET Architecture, which involves an expansive path (right side) and a contracting path (left side).

Two 3x3 convolution layers (unpadded convolutions) make up the contracting path, each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. Each stage doubles the amount of feature channels.

In the architecture's expansive path, the feature map is up-sampled, followed by a 2x2 convolution ("up-convolution") which the number of feature channels is cut in half. The feature maps from the contracting path are concatenated with feature maps in expanding path. Two 3x3 convolutions are applied, each followed by a ReLU. A 1x1 convolution is performed at the final layer to transfer each 64-component feature vector to the desired number of classes. There are 23 convolutional layers.
The Train Validation splitting ratio used is 80:20.

The pixel accuracy is insufficient to determine the segmentations model's performance due to class imbalance. The Dice coefficient tells how similar the prediction and ground truth are.

$$\text{Dice Coefficient} = 2 * \frac{\text{Area of Overlap in both pictures}}{\text{Total number of pixels}}$$

The second stage is classification:

The 2D scans are passed through UNet to get regions with the maximum possibility of the nodule.

i) Epoch

The epoch is the number of times the data has been passed over. The number of epochs isn't that important. The validation and training mistakes are more significant. Training should continue as long as these two errors continue to decrease. For example, if the validation error rises, this could indicate overfitting. (Overfitting is a statistical modelling error that arises when a function is too tightly fitted to a small number of data points. Attempting to make the model adhere to slightly erroneous data too closely will infect the model with significant flaws and diminish its predictive power).

The nodule masks are used to generate features for classifiers. For feature, extraction labeled array is generated, and region properties like diameter, area, major axis length, and solidity are obtained.

b) Mean HU

The Hounsfield unit (HU) is a quantitative scale describing radio density.

$$MeannoduleHU = Sum \sum \frac{(noduleimage[i] * mask)}{\text{np. sum (mask)}}$$

c) Area

Scalar quantity which provides the actual number of total nodule pixels.

d) Diameter

Scalar quantity which provides the average of the major axis and minor axis.

e) Diameter Major

Diameter Major Axis length

f) Eccentricity

i) Quantity Scalar

The eccentricity is the ratio of the distance between the ellipse's foci to the length of its major axis. The range is 0 to 1. A circle has an eccentricity of 0, and a line segment has an eccentricity of 1.

g) Spiculation

Solidity measures the density of an object. The area of an object is divided by the area of the object's convex hull. A number of 1 indicates a solid item, while a value less than 1 indicates an entity with uneven boundaries or holes.

Classification

Feature tables and corresponding labels are used for classification. The nodules are classified into 2 groups: Benign and malignant, where 0 stands for benign and 1 stand for malignant. Classifiers used are: Support Vector Machine, K nearest neighbors, Logistic Regression

a) Support Vector Machines (SVM)

The SVM is a classification algorithm that uses supervised machine learning. Parameters considered for tuning SVM are as follows: Kernel- Kernels are used according to the linear and nonlinear hyperplanes. The linear kernel is used. Regularization(C) –The regularization value used is 1.

b) K Nearest Neighbors (KNN)

The algorithm works on the assumption that related things occur in nearby proximity. KNN captures the idea of similarity with some mathematics of calculating the distance between points on a graph. The selected K is 5.

c) Logistic Regression (LR)

LR classification is used when output is categorical. It predicts the probability of categorical dependent variables based on independent input variables.

4. Results and Discussion

Table 5 shows the Result of Training and Testing based on loss, dice coefficient and accuracy.

Table 5. Result of Training and Testing

	Training	Testing
Loss	-0.7694	-0.6270
Dice Coefficient	0.7694	0.6225
Accuracy	0.7001	0.7114

Dice Coefficient

The Dice similarity coefficient is a statistical tool that determines how similar two sets of data are. The Dice coefficient and the IoU are quite close. They are positively correlated, which means that if one claims model A is better at segmenting a picture than model B, the other will agree. They both range from 0 to 1, with 1 denoting the greatest resemblance between expected and truth, just like the IoU (intersection over union). The dice coefficient shouldn't be greater than 1. If you are getting a coefficient greater than 1, maybe you need to check your implementation.

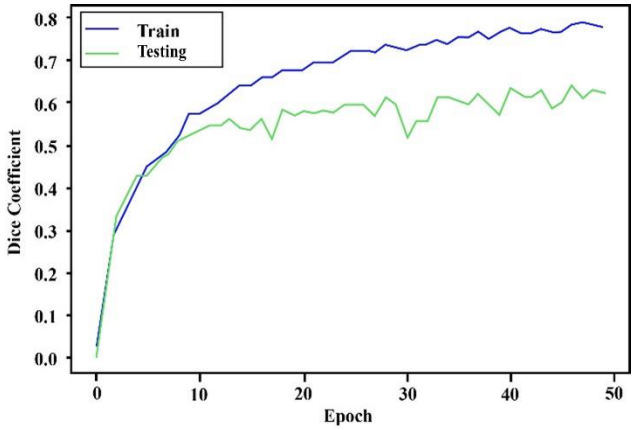


Fig. 8 Dice coefficient plot

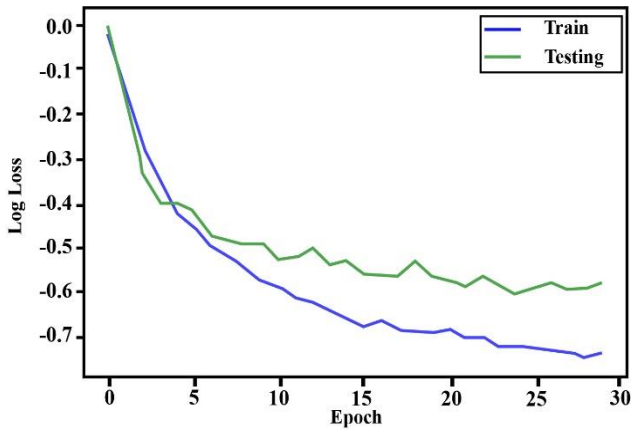


Fig. 9 Log loss coefficient plot

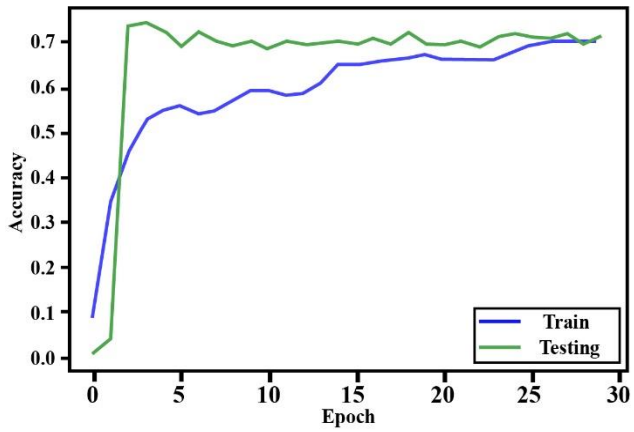


Fig. 10 Accuracy plot

a) Log Loss

The negative average of the log of corrected predicted probability for each case is called log loss. Logarithmic. By penalizing erroneous classifications, log loss quantifies a classifier's accuracy. The log-loss represents how close the forecast probability is to the actual/true number (0 or 1 in case of binary classification). The higher the log-loss number, the more the anticipated probability differs from the real value.

b) Negative Loss

Loss is multiplied by gradient when taking a step with gradient descent. So when the gradient becomes negative, gradient descent takes a step in the opposite direction. Such an idea is well captured when implementing gradient ascent, as it can simply be implemented by multiplying -1 by the loss.

The loss function is calculated across all data items during an epoch and guaranteed to deliver the quantitative loss measure at that time. However, showing the curve over iterations only provides the loss for a portion of the total dataset. The loss value describes how a model performs after each optimization cycle.

c) Accuracy

The accuracy curve indicates that the model, trained on cases from the training set, properly predicts instances from outside the training set.

Results of some sample images show in fig11, Fig. 12, and Fig. 13. It includes the original image, the image after the noise was removed, the Binary image, the segmented image, the final picture with segmentation and detected nodules, the image with segmented nodules.

i) Confusion Matrix

A confusion matrix is a table that explains a classifier's performance.

The sensitivity is its capability to identify those individuals with cancer.

The specificity is its capability to identify those individuals with non-cancer.

The Precision quality of a positive prediction made by the model.

F1-score is the harmonic mean of precision and recall. The real positive rate (recall) and precision are weighted and averaged.

The results of three classifiers, including Support Vector Machines, K Nearest Neighbors, and Logistic Regression, are shown bar graph in Fig. 14 in comparisons based on performance measures, including accuracy, sensitivity, specificity, precision, and F1 Score.

Results from three classifiers, including Support Vector Machines, K closest neighbours, and logistic regression based on TP, FP, FN, TN, Accuracy (%), Sensitivity (%), Specificity (%), Precision (%), F1 score (%) were compared in Table VI.

Test data

The masks are predicted on testing data of 105 slices, using the weights generated from this model. Malignancy is decided with the ensemble of three classifiers Support Vector Machines, K nearest Neighbors and Logistic regression.

Predicted Yknn, Ysvm, and Ylog for 105 test data.

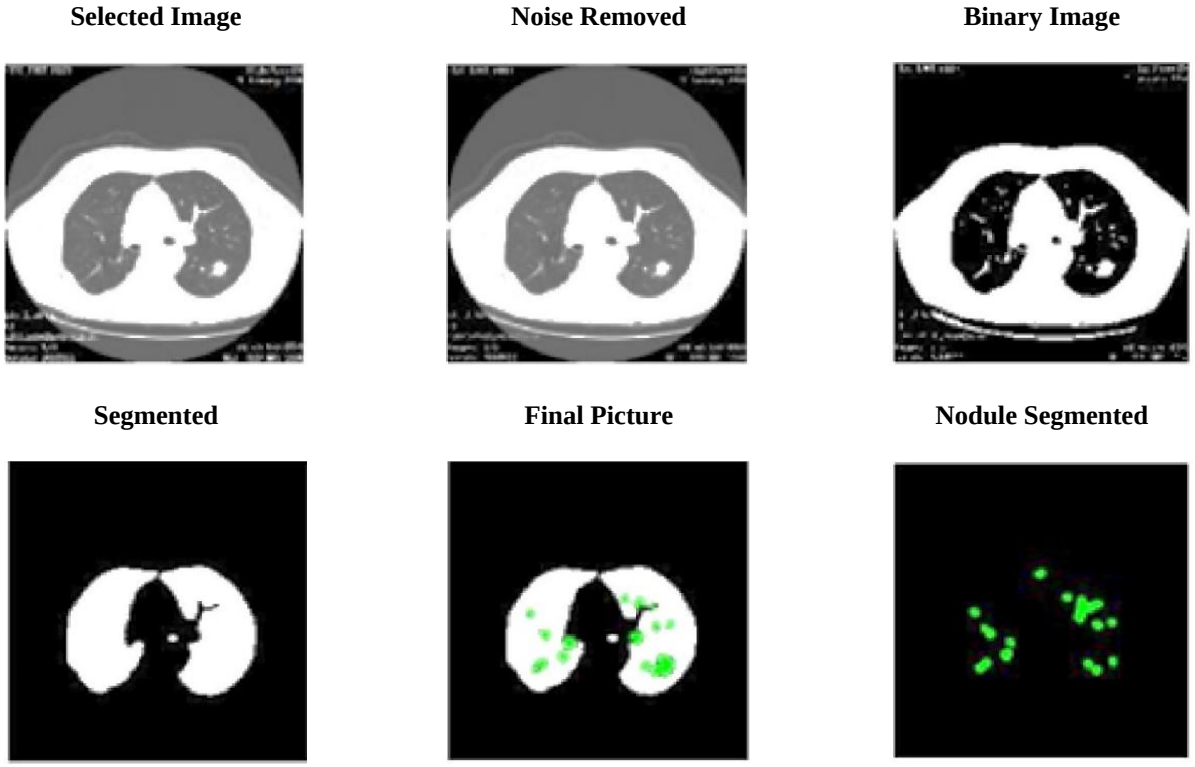


Fig. 11 Results of Input Image: 00075

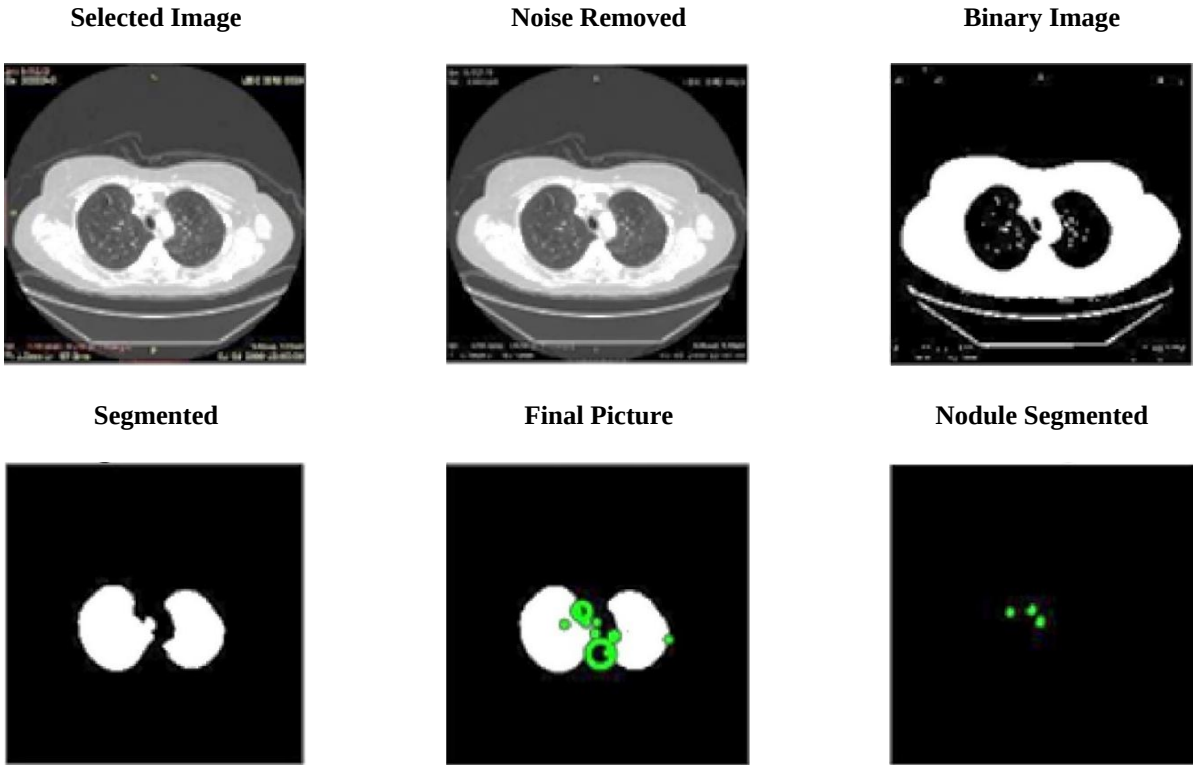


Fig. 12 Results of Input Image: 67_0029_00001

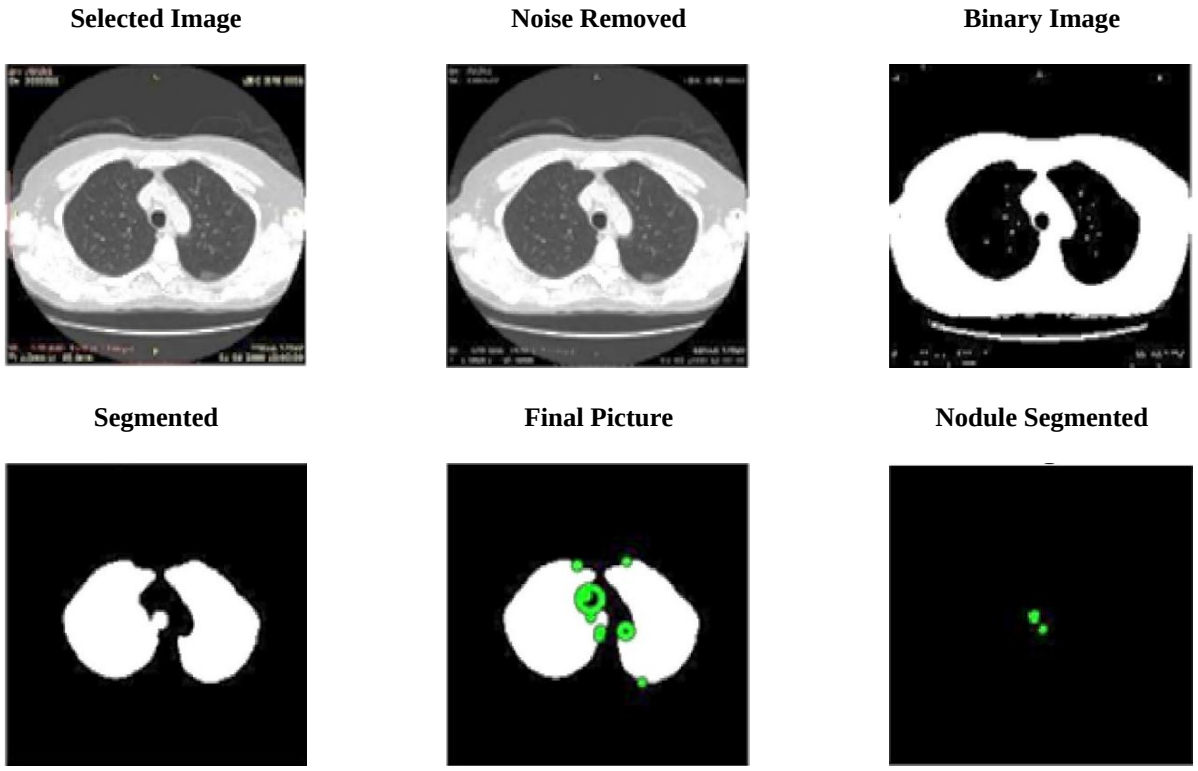
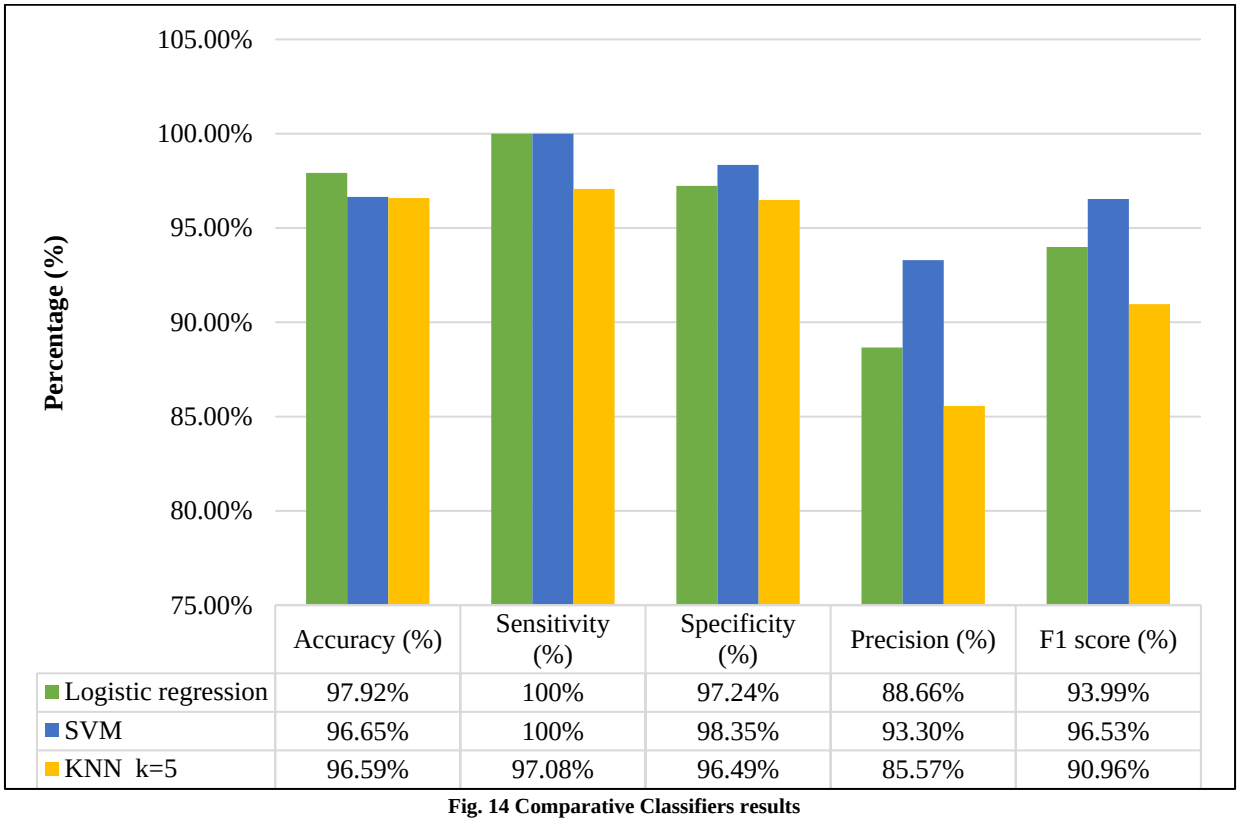


Fig. 13 Results of Input Image: RAD00001



	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)
Logistic regression	97.92%	100%	97.24%	88.66%	93.99%
SVM	96.65%	100%	98.35%	93.30%	96.53%
KNN k=5	96.59%	97.08%	96.49%	85.57%	90.96%

Fig. 14 Comparative Classifiers results

$$Avg\ mal\_val = \frac{y.KNFN[0] + y.SVM[0] + y.LogisticReg[0]}{3}$$

If avg mal_val is > 0.5, then verify if the nodule is malignant; otherwise, it is not.

Classification Results: Testing Accuracy=83%. The results of the proposed system are tabulated in Table VII.

Fig. 15 shows the Confusion matrix of Test data with all values.

Table 6. Comparative results of different Classifiers

Classifier	TP	FP	FN	TN	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)
Logistic regression	172	22	0	774	97.92%	100%	97.24%	88.66%	93.99%
SVM	181	13	0	774	96.65%	100%	98.35%	93.30%	96.53%
KNN k=5	166	28	5	769	96.59%	97.08%	96.49%	85.57%	90.96%

Table 7. Model Classifiers results

Classifier	TP	FP	FN	TN	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)
Proposed system	71	8	9	17	83.8	88.75	68.00	89.87	89.31

		Predicted Class		
		Positive	**Negative**	**%**
Actual Class	**Positive**	**True Positive (TP) 71**	**False Negative (FN) Type II Error 8**	**Sensitivity** $\frac{TP}{(TP+FN)}$ **88.75**
	Negative	**False Positive (FP) Type I Error 9**	**True Negative (TN) 17**	**Specificity** $\frac{TN}{(TN+FP)}$ **68.00**
	%	**Precision** $\frac{TP}{(TP+FP)}$ **89.87**	**Negative Predictive** $\frac{TN}{(TN+FN)}$ **68.00**	**Accuracy** $\frac{TP+TN}{(TP+TN+FP+FN)}$ **83.80**

Fig. 15 Confusion matrix of Test data

5. Conclusion

Implementation of CAD system (includes Pre-processing, Segmentation, Detection, feature extractions and Classification) for Lung CT images using Python. The detection of candidate nodules is done using UNET architecture. Biomedical image segmentation was the first application for which U-net was created and implemented. U-Net boasted an image segmentation method that surpassed its rival at the time to boost the network's learning capacity. The malignancy of the detected nodule is decided with an ensemble of three classifiers Support Vector Machines, K nearest Neighbours and Logistic regression. Performance check of different classifiers using Accuracy, Sensitivity, Specificity, Precision, and F1 Score. Logistic regression gives the highest accuracy, 97.92%. The ensemble of three classifiers model detects and classifies lung nodules with an accuracy of 83%. The diameter of the lung nodule and spiculation plays an important role in deciding the malignancy of the lung nodule. Nodules with more diameter and spiculation have a high probability of being malignant.

However, there is room to improve the accuracy of the model. Detection of less dense nodules surrounded by fats is difficult to detect. More information about the nodule can be studied using its connection with neighbouring slices to improve accuracy.

Acknowledgments

The authors are grateful to AISSMS IOIT Pune for providing resources for this paper.

References

[1] Cancer Facts and Figure 2022 by American Cancer Society. [Online]. Available : http://www.cancer.org

[2] Ashis Kumar Dhara, Sudipta Mukhopadhyay, and Niranjan Khandelwal, "Computer-Aided Detection and Analysis of Pulmonary Nodule from CT Images: A Survey," *IETE Technical Review*, vol. 29, no. 4, pp. 265-275, 2012. *Crossref*, https://doi.org/10.4103/0256-4602.101306

[3] Early Lung Cancer Action Program (ELCAP), [Online]. Available: http://www.via.cornell.edu/lungdb.html.[last cited on 2022].

[4] Lung Imaging Database Consortium (LIDC), 2022. [Online]. Available: https://imaging.nci.nih.gov/ncia/login.jsf./ http://www.cancerimagingarchive.net

[5] Medical Image Database, Medpix, 2022. [Online]. Available: http://rad.usuhs.edu/medpixlindex.html.

[6] Martin Dolejsi et al., "The Lung TIME: Annotated Lung Nodule Dataset and Nodule Detection Framework," *Proceedings Medical Imaging 2009: Computer-Aided Diagnosis*, vol. 7260, 2009. *Crossref*, https://doi.org/10.1117/12.811645

[7] Aparna Rajesh Lokhande, and Dr. Vinayak K. Bairagi, "Computer Aided Detection of Lung Cancer Pulmonary Nodules From CT Images: A Recent Survey," *International Journal of Advanced Information Science and Technology*, vol. 6, no. 6, pp. 20-32, 2017.

[8] Salsabil A. El-Regaily et al., "Survey of Computer Aided Detection Systems for Lung Cancer in Computed Tomography," *Current Medical Imaging Reviews*, vol. 14, no. 1, pp. 3-18, 2018. *Crossref*, http://dx.doi.org/10.2174/1573405613666170602123329

[9] P. Mohamed Shakeel et al., "Automatic Detection of Lung Cancer From Biomedical Data Set Using Discrete AdaBoost Optimized Ensemble Learning Generalized Neural Networks," *Neural Computing and Applications*, vol. 32, pp. 777-790, 2020. *Crossref*, https://doi.org/10.1007/s00521-018-03972-2

[10] P. Mohamed Shakeel, M. A. Burhanuddin, and Mohammad Ishak Desa, "Automatic Lung Cancer Detection from CT Image Using Improved Deep Neural Network and Ensemble Classifier," *Neural Computing and Applications*, vol. 34, pp. 9579–9592, 2022. *Crossref*, https://doi.org/10.1007/s00521-020-04842-6

[11] Ummadi Janardhan Reddy et al., "Recognition of Lung Cancer Using Machine Learning Mechanisms with Fuzzy Neural Networks," *Traitement Du Signal*, vol. 36, no. 1, pp. 87-91, 2019.

[12] Suren Makaju et al., "Lung Cancer Detection Using CT Scan Images," *Procedia Computer Science*, vol. 125, pp. 107-114, 2018. *Crossref*, https://doi.org/10.1016/j.procs.2017.12.016

[13] Xuechen Li, "Multi-Resolution Convolutional Networks for Chest X-Ray Radiograph-Based Lung Nodule Detection," *Artificial Intelligence in Medicine*, vol. 103, 2020. *Crossref*, https://doi.org/10.1016/j.artmed.2019.101744

[14] Ahmed Soliman et al., "Accurate Lungs Segmentation on CT Chest Images by Adaptive Appearance-Guided Shape Modeling," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 263–276, 2017.

[15] Pedro Pedrosa Rebouças Filho et al., "Novel and Powerful 3D Adaptive Crisp Active Contour Method Applied in the Segmentation of CT Lung Images," *Medical Image Analysis*, vol. 35, pp. 503–516, 2017. *Crossref*. https://doi.org/10.1016/j.media.2016.09.002

[16] Z. Shi et al., "Many is Better Than One: An Integration of Multiple Simple Strategies for Accurate Lung Segmentation in CT Images," *Biomed Research International*, pp. 1-13, 2016. *Crossref*, https://doi.org/10.1155/2016/1480423

[17] Abhir Bhandary et al., "Deep-Learning Framework to Detect Lung Abnormality–A Study with Chest X-Ray and Lung CT Scan Images," *Pattern Recognition Letters*, vol. 129, pp. 271-278, 2020. *Crossref*, https://doi.org/10.1016/j.patrec.2019.11.013

[18] Siddharth Bhatia, Yash Sinha, and Lavika Goel, "Lung Cancer Detection: A Deep Learning Approach," *Soft Computing for Problem Solving, Springer*, vol. 817, pp. 699-705, 2019. *Crossref*, https://doi.org/10.1007/978-981-13-1595-4_55

[19] João Rodrigo Ferreirada Silva Sousa et al., "Methodology for Automatic Detection of Lung Nodules in Computerized Tomography Images," *Computer Methods and Programs in Biomedicine*, vol. 98, no. 1, pp. 1-14, 2018. *Crossref*, https://doi.org/10.1016/j.cmpb.2009.07.006

[20] O. Obulesu et al, "Adaptive Diagnosis of Lung Cancer By Deep Learning Classification Using Wilcoxon Gain and Generator," *Hindawi Journal of Healthcare Engineering*, vol. 2021, 2021. *Crossref*, https://doi.org/10.1155/2021/5912051

[21] Wenfa Jiang et al., "Application of Deep Learning in Lung Cancer Imaging Diagnosis," *Hindawi Journal of Healthcare Engineering*, vol. 3, 2022. *Crossref*, https://doi.org/10.1155/2022/6107940

[22] Nasrullah Nasrullah et al., "Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies," *Sensors*, vol. 19, no. 17, p. 3722, 2022. *Crossref*, https://doi.org/10.3390/s19173722

[23] Ebanesar.C et al., "Computer Aided System for Automated Heterogeneous Cancer Recognition Using Google Cloud Platform," *SSRG International Journal of Computer Science and Engineering*, vol. 5, no. 5, pp. 11-16, 2018. *Crossref*, https://doi.org/10.14445/23488387/ijcse-v5i5p103

[24] Ashis Kumar Dhara et al., "A Combination of Shape and Texture Features for Classification of Pulmonary Nodules in Lung CT Images," *Journal of Digital Imaging*, vol. 29, no. 4, pp. 466-475, 2016. *Crossref*, https://doi.org/10.1007/s10278-015-9857-6

[25] V. Shalini, and K. S. Angel Viji, "Integration of Convolutional Features and Residual Neural Network for the Detection and Classification of Leukemia From Blood Smear Images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 176-184, 2022. *Crossref*, https://doi.org/10.14445/22315381/ijett-v70i9p218

[26] Ivan William Harsono et al., "Lung Nodule Detection and Classification From Thorax CT-Scan Using Retinanet with Transfer Learning," *Journal of King Saud University – Computer and Information Sciences, Elsevier*, vol. 34, no. 4, pp. 567-577, 2022. *Crossref*, https://doi.org/10.1016/j.jksuci.2020.03.013

[27] Akila Victor et al., "Detection and Classification of Breast Cancer Using Machine Learning Techniques for Ultrasound Images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 3, pp. 170-178, 2022. *Crossref*, https://doi.org/10.14445/22315381/ijett-v70i3p219

[28] Ruchika, and Archna Sharma, "CAD Implementation for Detection of Lung Cancerous Nodules," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 6, no. 4, pp. 804-806, 2016.