

# Lung Cancer Detection using Segmented 3D Tensors and Support Vector Machines

Zaib un Nisa<sup>1</sup>, Arfan Jaffar<sup>2</sup>, Sohail Masood Bhatti<sup>3</sup>, Umair Muneer Butt<sup>4</sup>

Dept. of Computer Science and Information Technology, Superior University, Lahore, Pakistan<sup>1, 2, 3</sup>

Dept. of Computer Science, The Chenab University, Gujrat, Pakistan<sup>4</sup>

**Abstract**—Tumor is currently the second most prevalent cause of mortality, and its prevalence is expanding rapidly. The development of pulmonary nodules inside the lungs is suggestive of the existence of lung cancer. The detection of cancer is achieved using nodules detected in computer tomography (CT) images obtained from the LUNA 16 dataset. This study uses the Python library "PyTorch" for this purpose. A three-dimensional model has been used to train and extract the nodular segments from CT-Scan images, referred to as CT-scan chunks. It is done due to the impracticality of handling the whole CT scan image due to its vast size. The previously mentioned chunks are then transformed into PyTorch tensors. The tensors are subsequently input into a deep learning model to extract features, which are then passed through a sequence of machine learning classifiers for the purpose of classification. These classifiers include Support Vector Machines, Multi-layer Perceptron, Random Forest Classifier, Logistic Regression, K Nearest Neighbor, and Linear Discriminant Analysis. Our research has shown that the use of chunk extraction from CT-Scan images, coupled with the creation of tensors using segmented CT scans, has significantly enhanced the precision of various machine learning algorithms. Additionally, this approach has the advantage of reducing the computational time during runtime. In our study, the use of Support Vector Machines yielded the best degree of accuracy, reaching 99.68%. The findings of this study have the potential to be valuable in the practical implementation of real-time lung nodule identification applications.

**Keywords**—Deep learning; lung cancer; LUNA16; machine learning; nodules; PyTorch

## I. INTRODUCTION

Cancer is a fatal illness that kills thousands and millions of people worldwide. Lung cancer is responsible for around 350 daily fatalities, making it the second leading cause of cancer-related deaths. Roundabout 103,00 of 127 lakh lung cancer fatalities in 2023 (81%) are directly attributable to smoking. The remaining approximately 20,500 lung cancer fatalities unrelated to smoking would be classified as the eighth leading cause of cancer mortality [1]. The process of diagnosing pulmonary nodules is valuable for the identification and treatment of cancer in its initial stages, therefore significantly contributing to the preservation of many human lives [2].

Doctors have significant challenges when dealing with a substantial volume of CT-scan data. Computer-aided detection (CAD) systems have been shown to alleviate the burden of radiologists while enhancing diagnosis accuracy in the context of lung cancer [3]. The CDAM technique was used in order to visually represent the characteristics and highlight the

decision-making process inside Zhang's unique CAD (Computer-Aided diagnosis) system for lung cancer diagnosis [4]. Recent research on lung nodule identification demonstrates the exceptional performance of Deep Learning (DL)-based CAD systems [5].

The use of DL techniques has shown exceptional efficacy in the domain of lung cancer diagnosis, namely in the identification of pulmonary nodules. A study used the Faster R-CNN and U-Net-like encoder-decoder models to analyze the LUNA16 dataset, resulting in an accuracy rate of 0.842 [6]. The 3D CNN model is employed by Duo et al. [7] for Lung cancer detection. The researchers partitioned their approach into two distinct components, namely candidate screening and false positive deduction. The researchers attained a sensitivity rate of 90.06% and a FROC score of 0.839. Khosravan et al. obtained a FROC score of 0.897. The researchers put out a computational procedure known as S4ND with the purpose of facilitating end-to-end learning. The architecture is comprised of densely connected CNN units [8]. Wang et al. used three slices to make a 3D-RGB image for nodule detection and achieved an accuracy of 0.968 [9].

There was a need to improve the nodule detection accuracy, so it can be effectively used in real time scenarios. Moreover, the use of three-dimensional convolutional neural network (3D-CNN) architectures has more potential in efficiently capturing the unique attributes of malignant nodules in contrast to two-dimensional CNN (2D CNN) designs. Thus far, only a limited number of publicly available research publications have been published on the use of 3D CNNs for the identification of lung cancer. This research study used 3-dimensional CT-Scan results, 3D CNN features and machine learning classifiers, using specialized Python libraries to leverage the capabilities of Python and the effectiveness of machine learning methods. Various state-of-the-art ML techniques have been proposed for the purpose of identifying lung nodules in the context of lung cancer detection in literature. However, there is ongoing debate over the efficacy of current methods for the timely identification of lung cancer. Therefore, we advocate for the adoption of a comprehensive methodology in the detection of lung cancer. The research makes many noteworthy contributions, which are outlined below:

- Nodules have been identified in the photos obtained from the publicly accessible dataset LUNA16.
- Images were preprocessed by isolating the lung region, creating slices, and then segmenting the CT-Scan

slices. Additionally, the nodule segment from the segmented slices was transformed into *CT-Scan chunks and a 3D tensor was created.*

- In our study, several machine learning classifiers for the purpose of classification are used after extracting features using 3D deep learning algorithm. These ML classifiers include Support Vector Machine, Random Forest, Machine Learning Perceptron, Logistic Regression, Linear discriminant analysis, and K Nearest Neighbors.

After this introductory section, a comprehensive assessment of the existing literature pertaining to the detection of lung cancer is presented in Section II. Subsequently, our proposed technique is outlined in Section III. The findings of the proposed methodology are provided and compared with the outcomes of state-of-the-art algorithms in Section IV along with detailed analysis and examination of the methodology used in the study, as well as a comprehensive presentation of the obtained findings. The concluding section encapsulates the research effort and its novel contributions.

## II. BACKGROUND STUDY

According to the data shown in Fig. 1, lung cancer has emerged as the second most prominent cause of mortality. Lung cancer can be diagnosed efficiently by detecting nodules in the CT scan images as shown in Fig. 2. Currently, pulmonary nodules are mostly diagnosed using various forms of chest imaging examination, such as MRI, X-ray, CT-Scan, and others. Because of its great accuracy and cheap cost, CT scan has become the gold standard in diagnostics [10]. Due to its high level of sensitivity, CT scans have become the preferred method in screening for lung cancer. Classifying pulmonary nodules, however, remains a laborious task. Furthermore, the size of CT images is continually expanding, adding more labor to the normal diagnosis procedure for radiologists [11].

Manual feature engineering is often employed in conjunction with machine learning for medical picture categorization. For both binary and multi-class classification, support vector machines (SVM) are a well-known machine learning approach. Nodule detection was performed using an adaptive morphology-based strategy and SVM [12]. Manifold classification regularization was applied by Ren et al. for nodule identification along with the Dense network of neurons (DNN) [13]. Image categorization and object recognition both benefited greatly from DNN's automatic feature extraction. In a research study, a convolutional neural network (CNN) was used for the purpose of detecting and classifying lung nodules. [14]. Classifying nodules in CT scan pictures is another application for CNN-based Generative Adversarial Networks [15]. A pulmonary nodule classification system was designed by Jiang et al. The authors used improved deep features by using contextual attention and located region of interest (ROI) using spatial attention. For enhancing the performance of the classifier an ensemble was employed [16]. CNN was used along with kernel-based non-Gaussian for cancer detection. The authors applied an adaptive histogram equalization technique in the preprocessing step for segmenting the region of interest. They have achieved an accuracy of 87% [17].

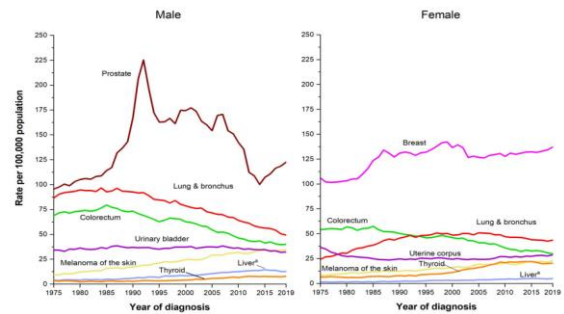


Fig. 1. Cancer incidence rates by gender in the USA from 1975 to 2019. Rates are adjusted for reporting lag times and age standardization to the US standard population of 2000 [1].

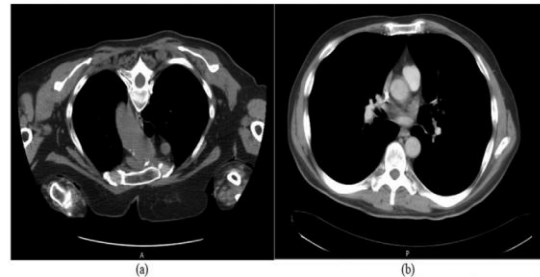


Fig. 2. Lungs images: (a) Normal lung CT image, (b) Cancerous lung CT image [49].

In a study, the deep learning model employed for nodule identification is a receptive field regularized V-net [18]. The classification of lung nodules is accomplished by the use of a hybrid approach including the SqueezeNet and ResNet models. In addition to convolutional neural networks (CNN), authors used a bio-inspired approach called the “Whale optimization algorithm with adaptive particle swarm optimization” (WOA-APSO) for detecting lung cancer [19]. They achieved an accuracy of 97.18%. Histogram equalization, the Tophat approach for noise removal, and the Boosted deep convolutional neural network (BDCNN) were utilized by Rani et al. [20] while classifying lung cancer. The noise was filtered out using weighted mean histogram analysis, and features were extracted using a hybrid dual-tree complex. Ultimately, a deep CNN was used for the purpose of lung cancer detection [21].

A sequential technique for identifying lung cancer in CT lung images was devised by Guo et al. [22]. The used approach included the utilization of both a CNN-based classifier and a feature-based classifier. The first use of a CNN-based classifier was seen in the analysis of a dataset pertaining to the detection of lung cancer. Subsequently, a classifier based on features was employed for further investigation. [22]. Irregular nodule shapes were segmented by Hesamian et al. [23]. To tackle the issue of diminished contrast, scholars have conducted investigations including the use of deep-learning methodologies on artificially created images derived from innovative color schemes. In this work, a modified version of the DL-based U-Net model was used. Chen et al. used the dense neural network methodology [24].

The usage of batch normalization and dropout was employed in conjunction with this dense architecture. The

performance on the LUNA16 dataset has been commendable. Basal et al. and Xu et al. used ensemble learning to identify lung cancer [25], [26] and combined YOLOv3 with CNN for malignant nodule detection. Fractal networks are employed in research to categorize lung nodules [27]. The performance of the system was evaluated and verified using the LUNA16 dataset, using the Fractal net model. The resulting accuracy was 94.7%. With the use of ensemble learning, Muzammil et al. improved accuracy to 96.89%. In their work [28], they integrated the SVM and AdaBoostM2 algorithms with the Linear discriminant analysis (LDA). Another research [29] used deep learning to build a segmentation technique for identifying lung cancer. Numerous more cutting-edge research [30]–[33] covered a wide range of topics and highlighted additional real-world issues.

TABLE I. SUMMARY OF LUNG CANCER DETECTION ARTICLES

References	Year	Method	Accuracy
[34]	2015	Multi-scale CNN	86.84%
[35]	2017	Multi-crop CNN	87.14%
[35]	2017	Deep 3D DPN	88.74%
[35]	2017	Deep 3D DPN+ GBM	90.44%
[36]	2018	3D-CNN with transfer learning	71%
[37]	2018	FR Using DAN	93.9%
[38]	2018	Wavelet-Based CNN	91.9%
[17]	2019	3D Ensemble	90.24%
[39]	2019	CNN with feature fusion	99.2%
[40]	2020	Multi-scale CNN (MCNN)	93.7%
[41]	2020	Kernel-based non-gaussian Convolutional Neural Network	87.3 %
[19]	2020	WOA	91.18%
[20]	2020	Boosted Deep Convolutional Neural Network	97.30%
[42]	2020	TL with VGG16, VGG19	95.0%
[26]	2020	Deep 3D-Scan	92.7%
[28]	2021	EL	96.9%
[22]	2021	FB-and OCNN	95.96%
[43]	2021	3D attention U-Net	94.43%
[44]	2021	Res BCDU-Net	97.58%
[45]	2021	HMB-HCF	99.30%
[46]	2022	YOLOv3	95.17%
[47]	2022	ProCAN	95.28%
[18]	2022	Squeeze net. Res Net	94.87%
[2]	2023	LungNet-SVM	96.37%

Table I provides a review of recent research that has advanced our understanding. Literature Table I illustrates the substantial amount of scholarly literature dedicated to the examination of lung cancer diagnosis, nodule detection, and

categorization. The main objective of this research is to investigate lung cancer screening. The CT scans were segmented into chunks and then subjected to analysis using a three-dimensional convolutional neural network (3D-CNN) model. Deep learning-based imaging algorithms have significantly enhanced the accuracy and efficacy of lung nodule segmentation, identification, and classification. These algorithms use old medical images to achieve superior performance in these tasks. The system's ability to engage in reinforcement learning facilitated the attainment of this outcome.

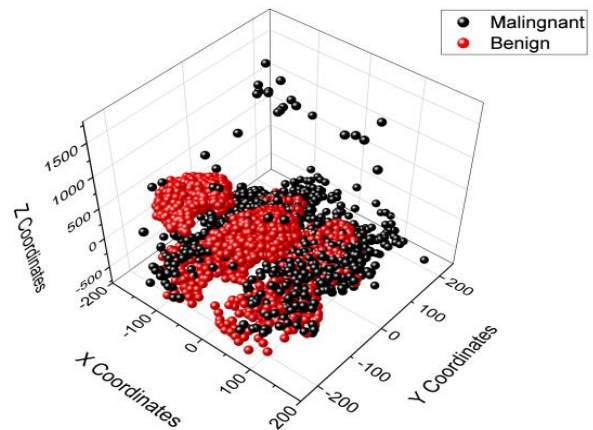


Fig. 3. The spatial coordinates of the lesions are situated in the x, y, and z dimensions. Red colour is used to indicate benign lesions, whereas black colour is used to indicate malignant tumours [26].

An assortment of unsupervised deep learning algorithms, such as CNN, Faster R-CNN, Mask R-CNN, and U-Net have been used to create convolutional networks that aim to identify lung cancer and minimize the occurrence of false positives. This is in contradiction to the conventional approach of using supervised and reinforced learning techniques. Previous researches have shown that computed tomography (CT) is the most often used imaging modality in computer-aided detection (CAD) systems for the detection of lung cancer. The use of 3D-CNN architectures shows more promise in effectively capturing the distinctive characteristics of malignant nodules compared to 2D CNN designs. So far, only a restricted number of research articles using 3D convolutional neural networks (CNNs) for the purpose of lung cancer detection have been made accessible to the public.

### III. METHOD

In the subsequent part, we will elucidate the strategy that has been put forward. In Fig. 8 a block diagram is provided, illustrating the suggested technique. This demonstrates the detection of nodules in the images, followed by the segmentation of the CT-Scan image into distinct sections, as seen in Fig. 7. The segmentation and classification tasks were effectively accomplished by using deep learning techniques. Within the scope of this research activity, the process involves the concatenation of a 3D tensor, which is afterwards used in conjunction with machine learning algorithms, with the aim of facilitating the identification of lung cancer.

### A. Dataset

The data used in this research was obtained from Kaggle, a reputable and well-known data repository. The Lung Nodule Analysis 2016 Challenge dataset, often known as Luna16, has been released to the public. It was updated in 2020. The dataset comprises CT scans from a total of 888 patients. Every computed tomography (CT) scan has a thickness that exceeds 2.5 millimeters. The data is saved in the MetaImage format, more especially in the mhd/raw file format. Every .mhd file is accompanied by a separate binary file that contains the raw pixel data. The dataset has a total of 754,975 candidates, whose information is categorized into two distinct groups: malignant and benign. Malignant nodules are classified as "1" whereas benign nodules are classified as "0". Fig. 3 depicts the distribution of lesion sites among the patients.

### B. Preprocess

During the preprocessing phase, the following steps were conducted:

- 1) *Extract CT-Scan images:* The CT-Scan files were retrieved from the dataset using the "SimpleITK" package.
- 2) *Unique Ct scans:* We found that there are total 888 unique CT-Scans for 754975 candidates.
- 3) *Missing values:* The missing values in the data were examined. A total of 443 missing data were identified, indicating the absence of unique CT-Scans. Consequently, the aforementioned values were eliminated.

4) *Histogram:* The pixel values undergo a transformation to Hounsfield units (HU), and afterwards, their histograms are shown, as seen in Fig. 4. The histogram has a range of -2000 to 1000 Hounsfield Units (HU). The histograms of all the CT-scan pictures exhibit a similar distribution.

5) *Pulmonary region:* The pulmonary area is marked out by using several threshold values, as seen in Fig. 5.

### C. Process

The data is divided into training and validation. The input images are transformed in the data preprocessing step, and noise and outliers are filtered away. Data is analyzed and it is found that the candidate file has a flag for each mass in the CT scans. "seriesuid" is a unique identifier for each CT Scan. The class label is 0 if the CT-Scan doesn't have a nodule. The class label is 1, if CT scans have a nodule (both malignant and benign). There are over 750k candidates. The candidates belonging to the same CT scan are categorized into groups, after which the diameters dictionary is used to get the diameter of each candidate. The order of the applicants was flipped, resulting in all individuals with nodules being positioned at the top of the list. The candidates without comparable CT scans were excluded from the dataset. The entire number of applicants is 750,000, with just 37,000 having comparable CT scans. A chunk was obtained by generating a list consisting of three slices, with each slice corresponding to a certain direction. The segments are then transformed into PyTorch Tensors for CT-Scan analysis. Every tensor in the dataset has a shape of 1x10x18x18. These PyTorch tensors are fed into deep-learning models for feature extraction. Next, a training and testing dataset is created.

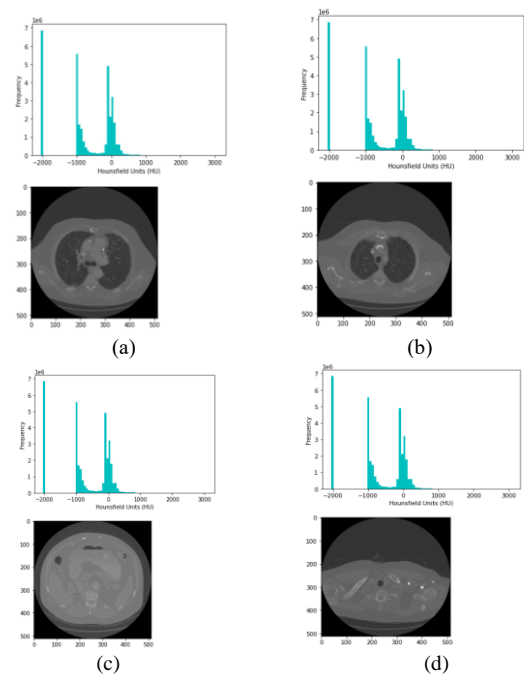


Fig. 4. CT-Scans with their corresponding histograms in HU units. (a) CT-scan number 80, (b) CT-scan number 100, (c) CT-scan number 2, (d) CT scan number 120.

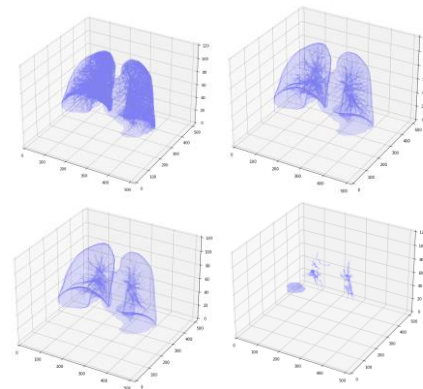


Fig. 5. 3D view of the lungs and bronchioles, for different threshold values.

The step-by-step process for carrying out experiments is as follows:

- 1) *Convert* the image into the fixed slices.
- 2) *Segment* the CT-Scan slices.
- 3) *Convert* the segmented CT-scan slices into 3D tensors.
- 4) *Split* the data into training and testing.
- 5) *Use* a 3D Deep learning algorithm for feature extraction.
- 6) *Use* machine learning classifiers for classification.

The 3D CNN model that is used for feature extraction is elaborated in Table III. The use of flattened layers involves the conversion of the output of a maxpooling layer, which is a multidimensional tensor, into a one-dimensional representation. The neural network architecture is composed

of two fully connected layers, specifically referred to as fc1 and fc2, which are positioned after the flattened layer. Fully connected layers are often used for the purpose of feature extraction. In this instance, the fc1 layer is utilized.

TABLE II. COMPARISON OF PROPOSED MODEL WITH STATE-OF-THE-ART ALGORITHMS

Table with 5 columns: Authors, Year, Dataset, Method, Accuracy. Rows include Shah et al. [42], Bansal et al [26], Naik et al.[27], Muzammil et al. [28], Dodia et al.[18], Naseer et al.[2], and Proposed method.

TABLE III. 3D-CNN MODEL'S PARAMETERS

Table with 6 columns: Sr No., Layer, Input Channel, Kernels, size, Parameters, Output shape. Rows 1-7 describe layers from Conv3D-1 to Flatten-7.

TABLE IV. ABBREVIATIONS AND THEIR EXPLANATIONS USED IN THIS STUDY

Table with 2 columns: Abbreviation, Explanation. Lists terms like CT-Scans, CDAM, CAD, DL, DNN, FROC, ML, MRI, SVM, ROI, BDCNN, YOLO, LUNA, LDA, Seg3DT\_SVM.

IV. RESULTS AND DISCUSSION

In this research study, a deep learning model is used for the purpose of feature extraction. Prior to this, the data is subjected to preprocessing, whereby it is transformed into three-dimensional CT-scan chunks to serve as input.

Accuracy is defined as:

Accuracy = (TN + TP) / (TN + FN + FP + TP) (1)

The variables TN and TP in equation (1) represent the rates of true negative and true positive, respectively, in the confusion matrix. A confusion matrix is shown in Fig. 9. In this research work, data is pre-processed, features are extracted and then machine learning algorithms are applied at the end of pipeline.

The findings clearly demonstrate that the suggested model has surpassed the performance of existing algorithms by attaining an accuracy rate of 99.6%. Hence, this model has the capability to be used in real-time situations for the purpose of identifying lung malignancies.

This research study presents a unique methodology for the identification of lung cancer. The task of lung cancer detection is accomplished by the use of data collection that is accessible to the general public. The data used in this study was acquired from Kaggle, a well-known and esteemed data source.

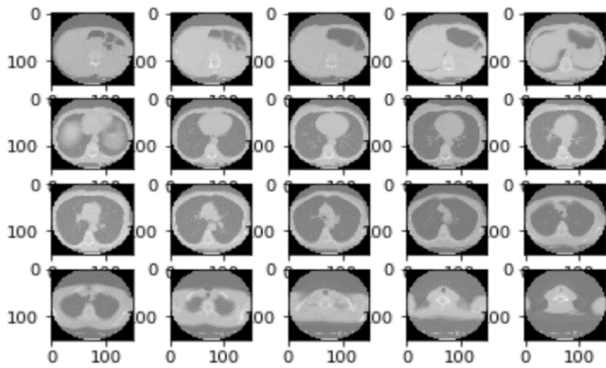


Fig. 6. CT-Scan slices.

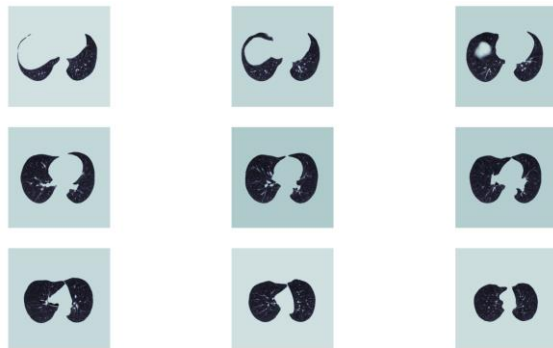


Fig. 7. CT-Scan segmented slices.

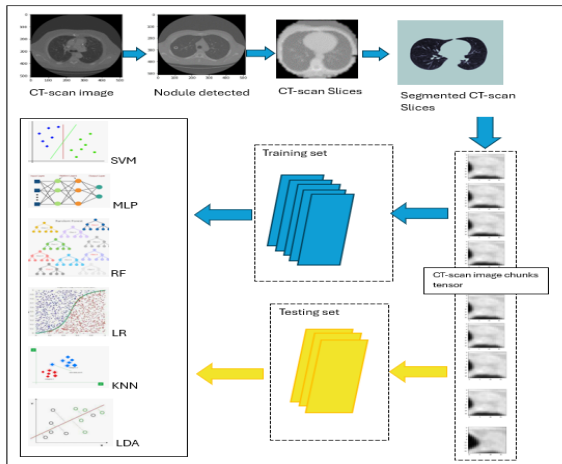


Fig. 8. Block diagram of this research work.

		Prediction	
		0	1
Actual	0	TN	FP
	1	FN	TP

Fig. 9. Confusion matrix [50].

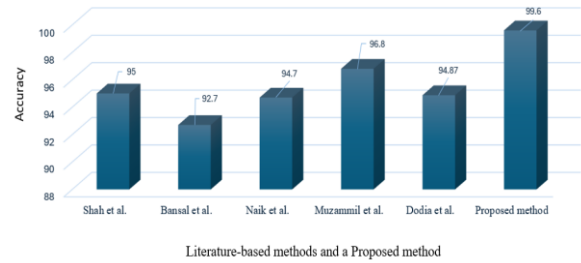


Fig. 10. Comparison of state-of-the-art accuracies with the proposed method.

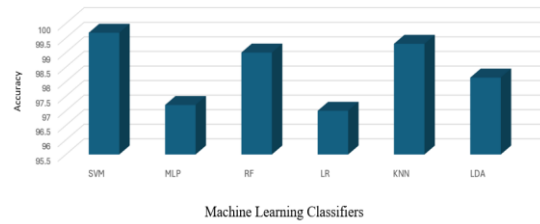


Fig. 11. Accuracies of various classifiers, using the proposed preprocessing method.

The examination of missing values in the dataset was conducted. A total of 443 instances of missing data were detected, suggesting the lack of distinct CT-Scans. As a result, the previously indicated values were eradicated. The pixel values are subjected to a conversion process into Hounsfield units (HU), and subsequently, their histograms are shown, as seen in Fig. 4. The histogram exhibits a range spanning from -2000 to 1000 Hounsfield Units (HU). The histograms of all the CT-scan images show a comparable distribution. The delineation of the pulmonary region is achieved by the use of several threshold values, as seen in Fig. 5. The dataset is partitioned into two subsets, namely the training set and the validation set. During the data preparation stage, the input pictures undergo a transformation process, which includes the removal of noise and outliers. Upon analysis of the data, it has been determined that the candidate file contains a flag for each mass seen in the CT images that serves as a unique identifier for each CT-Scan. The assigned class label is 0 in cases when the CT-Scan does not reveal the presence of a nodule. The assigned class label is denoted as 1, indicating the presence of nodules in CT scans, including both malignant and benign cases. There are a total of more than 750,000 candidates. The candidates belonging to the same CT scan are categorized into groups, after which the diameters dictionary is used to get the diameter of each candidate. The order of the applicants was reversed, resulting in all individuals with nodules being placed at the top of the list. The individuals who did not have CT scans that could be compared were eliminated from the dataset. The procedure of segmentation has considerable significance within the realm of diagnostics.

The extraction of CT scan slices from images is performed in order to eliminate any unwanted portions. The lung pictures were then subjected to segmentation. The extraction of three-dimensional segments includes nodules from the segmented slices. The parts are put into our model. The chunk's measurements are 10 units in length, 18 units in breadth, and

18 units in height. A chunk was acquired by constructing a list including three slices, where each slice corresponds to a certain direction. The segments are then converted into PyTorch Tensors to facilitate CT-Scan processing. The PyTorch tensors are used as inputs in 3D deep learning models for feature extraction. Following this, a dataset is constructed for the purpose of training and testing, and a variety of machine learning classifiers are used to choose the classifier that exhibits the highest performance. The results unequivocally indicate that the proposed model has outperformed current algorithms, achieving a remarkable accuracy rate of 99.68%. Therefore, this model has the capacity to be used in real-time scenarios with the objective of detecting lung cancers.

## V. CONCLUSION

This study presents a novel approach that utilizes segmented 3D tensors, 3D convolutional neural network and SVM to identify lung nodules in computed tomography (CT) data. The proposed model is abbreviated as Seg3DT\_SVM. The model being discussed is specifically built to accept data chunks in the format of PyTorch tensors. The discrete components, referred to as chunks, that comprise the nodule are the individual segments of computed tomography (CT) slices used. The process of obtaining cross-sectional images is an essential element of computed tomography (CT) scans. The proposed technique includes an initial phase of data preparation, as previously elucidated, succeeded by the use of machine learning algorithms. The experimental investigation used the publicly available LUNA16 dataset and was conducted using the PyTorch framework. The analyzed model demonstrates a much greater degree of accuracy (99.68%) compared to the existing state-of-the-art approaches used for the same dataset. The method demonstrates the capability to assist healthcare workers in promptly identifying cases of lung cancer in real-time situations.

## REFERENCES

- [1] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023," *CA Cancer J Clin*, vol. 73, no. 1, pp. 17–48, Jan. 2023, doi: 10.3322/caac.21763.
- [2] I. Naseer, T. Masood, S. Akram, A. Jaffar, M. Rashid, and M. A. Iqbal, "Lung Cancer Detection Using Modified AlexNet Architecture and Support Vector Machine," *Computers, Materials and Continua*, vol. 74, no. 1, pp. 2039–2054, 2023, doi: 10.32604/cmc.2023.032927.
- [3] D. Liu, F. Liu, Y. Tie, L. Qi, and F. Wang, "Res-trans networks for lung nodule classification," *Int J Comput Assist Radiol Surg*, vol. 17, no. 6, pp. 1059–1068, Jun. 2022, doi: 10.1007/S11548-022-02576-5.
- [4] S. Zheng et al., "Interpretative computer-aided lung cancer diagnosis," *Comput Methods Programs Biomed*, vol. 210, Oct. 2021, doi: 10.1016/J.CMPB.2021.106363.
- [5] H. Jiang, F. Shen, F. Gao, and W. Han, "Learning Efficient, Explainable and Discriminative Representations for Pulmonary Nodules Classification," Jan. 2021, [Online]. Available: <http://arxiv.org/abs/2101.07429>
- [6] A. A. A. Setio et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Med Image Anal*, vol. 42, pp. 1–13, Dec. 2017, doi: 10.1016/j.media.2017.06.015.
- [7] Q. Dou, H. Chen, Y. Jin, H. Lin, J. Qin, and P.-A. Heng, "Automated Pulmonary Nodule Detection via 3D ConvNets with Online Sample Filtering and Hybrid-Loss Residual Learning," Aug. 2017, [Online]. Available: <http://arxiv.org/abs/1708.03867>
- [8] N. Khosravan and U. Bagci, "S4ND: Single-Shot Single-Scale Lung Nodule Detection," May 2018, [Online]. Available: <http://arxiv.org/abs/1805.02279>
- [9] J. Wang et al., "Pulmonary Nodule Detection in Volumetric Chest CT Scans Using CNNs-Based Nodule-Size-Adaptive Detection and Classification," *IEEE Access*, vol. 7, pp. 46033–46044, 2019, doi: 10.1109/ACCESS.2019.2908195.
- [10] A. Pezeshk, S. Hamidian, N. Petrick, and B. Sahiner, "3-D Convolutional Neural Networks for Automatic Detection of Pulmonary Nodules in Chest CT," *IEEE J Biomed Health Inform*, vol. 23, no. 5, pp. 2080–2090, Sep. 2019, doi: 10.1109/JBHI.2018.2879449.
- [11] S. P. Pawar and S. N. Talbar, "LungSeg-Net: Lung field segmentation using generative adversarial network," *Biomed Signal Process Control*, vol. 64, Feb. 2021, doi: 10.1016/j.bspc.2020.102296.
- [12] A. Halder, S. Chatterjee, D. Dey, S. Kole, and S. Munshi, "An adaptive morphology based segmentation technique for lung nodule detection in thoracic CT image," *Comput Methods Programs Biomed*, vol. 197, p. 105720, Dec. 2020, doi: 10.1016/J.CMPB.2020.105720.
- [13] Y. Ren et al., "A manifold learning regularization approach to enhance 3D CT image-based lung nodule classification," *Int J Comput Assist Radiol Surg*, vol. 15, no. 2, pp. 287–295, Feb. 2020, doi: 10.1007/S11548-019-02097-8.
- [14] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00444-8.
- [15] Y. Onishi et al., "Multiplanar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks," *Int J Comput Assist Radiol Surg*, vol. 15, no. 1, pp. 173–178, Jan. 2020, doi: 10.1007/S11548-019-02092-Z.
- [16] K. Xia, J. Chi, Y. Gao, Y. Jiang, and C. Wu, "Adaptive aggregated attention network for pulmonary nodule classification," *Applied Sciences (Switzerland)*, vol. 11, no. 2, pp. 1–15, Jan. 2021, doi: 10.3390/app11020610.
- [17] H. Jiang, F. Gao, X. Xu, F. Huang, and S. Zhu, "Attentive and ensemble 3D dual path networks for pulmonary nodules classification," *Neurocomputing*, vol. 398, pp. 422–430, Jul. 2020, doi: 10.1016/J.NEUCOM.2019.03.103.
- [18] S. Dodia, A. Basava, and M. Padukudru Anand, "A novel receptive field-regularized V-net and nodule classification network for lung nodule detection," *Int J Imaging Syst Technol*, vol. 32, no. 1, pp. 88–101, Jan. 2022, doi: 10.1002/IMA.22636.
- [19] S. Vijh, P. Gaurav, and H. M. Pandey, "Hybrid bio-inspired algorithm and convolutional neural network for automatic lung tumor detection," *Neural Comput Appl*, 2020, doi: 10.1007/s00521-020-05362-z.
- [20] K. V. Rani and S. J. Jawhar, "Superpixel with nanoscale imaging and boosted deep convolutional neural network concept for lung tumor classification," *Int J Imaging Syst Technol*, vol. 30, no. 4, pp. 899–915, Dec. 2020, doi: 10.1002/IMA.22422.
- [21] A. D. Ananth and C. Palanisamy, "Extended and optimized deep convolutional neural network-based lung tumor identification in big data," *Int J Imaging Syst Technol*, vol. 32, no. 3, pp. 918–934, May 2022, doi: 10.1002/IMA.22667.
- [22] Z. Guo, L. Xu, Y. Si, and N. Razmjoooy, "Novel computer-aided lung cancer detection based on convolutional neural network-based and feature-based classifiers using metaheuristics," *Int J Imaging Syst Technol*, vol. 31, no. 4, pp. 1954–1969, Dec. 2021, doi: 10.1002/IMA.22608.
- [23] 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, no. 3, pp. 1616–1628, Mar. 2021, doi: 10.1007/s10489-020-01914-x.
- [24] Y. Chen, Y. Wang, F. Hu, L. Feng, T. Zhou, and C. Zheng, "Ldnnet: Towards robust classification of lung nodule and cancer using lung dense neural network," *IEEE Access*, vol. 9, pp. 50301–50320, 2021, doi: 10.1109/ACCESS.2021.3068896.
- [25] Y. Xu et al., "Identification of Benign and Malignant Lung Nodules in CT Images Based on Ensemble Learning Method," *Interdiscip Sci*, vol. 14, no. 1, pp. 130–140, Mar. 2022, doi: 10.1007/S12539-021-00472-1/METRICS.

- [26] G. Bansal, V. Chamola, P. Narang, S. Kumar, and S. Raman, "Deep3DScan: Deep residual network and morphological descriptor based framework for lung cancer classification and 3D segmentation," *IET Image Process*, vol. 14, no. 7, pp. 1316–1326, May 2020, doi: 10.1049/iet-ipr.2019.1164.
- [27] A. Naik, D. R. Edla, and V. Kuppili, "Lung Nodule Classification on Computed Tomography Images Using Fractalnet," *Wirel Pers Commun*, vol. 119, no. 2, pp. 1209–1229, Jul. 2021, doi: 10.1007/S11277-021-08258-W/METRICS.
- [28] M. Muzammil, I. Ali, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Pulmonary nodule classification using feature and ensemble learning-based fusion techniques," *IEEE Access*, vol. 9, pp. 113415–113427, 2021, doi: 10.1109/ACCESS.2021.3102707.
- [29] A. Shimazaki et al., "Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-021-04667-w.
- [30] S. Su, D. Yuan, Y. Wang, and M. Ding, "Fine Grained Feature Extraction Model of Riot-related Images Based on YOLOv5," *Computer Systems Science and Engineering*, vol. 45, no. 1, pp. 85–97, 2023, doi: 10.32604/csse.2023.030849.
- [31] H. Sun and R. Grishman, "Lexicalized Dependency Paths Based Supervised Learning for Relation Extraction," *Computer Systems Science and Engineering*, vol. 43, no. 3, pp. 861–870, 2022, doi: 10.32604/csse.2022.030759.
- [32] W. Sun, Y. Du, X. Zhang, and G. Zhang, "Detection and recognition of text traffic signs above the road," *International Journal of Sensor Networks*, vol. 35, no. 2, pp. 69–78, 2021, doi: 10.1504/IJSNET.2021.113626.
- [33] H. Sun and R. Grishman, "Employing Lexicalized Dependency Paths for Active Learning of Relation Extraction," *Intelligent Automation and Soft Computing*, vol. 34, no. 3, pp. 1415–1423, 2022, doi: 10.32604/iasec.2022.030794.
- [34] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale Convolutional Neural Networks for Lung Nodule Classification," *Inf Process Med Imaging*, vol. 24, pp. 588–599, 2015, doi: 10.1007/978-3-319-19992-4\_46.
- [35] W. Shen et al., "Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification," *Pattern Recognit*, vol. 61, pp. 663–673, Jan. 2017, doi: 10.1016/j.patcog.2016.05.029.
- [36] W. Lindsay, J. Wang, N. Sachs, E. Barbosa, and J. Gee, "Transfer Learning Approach to Predict Biopsy-Confirmed Malignancy of Lung Nodules from Imaging Data: A Pilot Study," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11040 LNCS, pp. 295–301, 2018, doi: 10.1007/978-3-030-00946-5\_29.
- [37] K. Mao, R. Tang, X. Wang, W. Zhang, and H. Wu, "Feature Representation Using Deep Autoencoder for Lung Nodule Image Classification," *Complexity*, vol. 2018, 2018, doi: 10.1155/2018/3078374.
- [38] E. Matsuyama and D.-Y. Tsai, "Automated Classification of Lung Diseases in Computed Tomography Images Using a Wavelet Based Convolutional Neural Network," *J Biomed Sci Eng*, vol. 11, no. 10, pp. 263–274, 2018, doi: 10.4236/jbise.2018.1110022.
- [39] A. Kumar, M. Fulham, D. Feng, and J. Kim, "Co-Learning Feature Fusion Maps from PET-CT Images of Lung Cancer," 2019. [Online]. Available: <http://ieeexplore.ieee.org>.
- [40] H. Yektaei and M. Manthouri, "DIAGNOSIS of LUNG CANCER USING MULTISCALE CONVOLUTIONAL NEURAL NETWORK," *Biomed Eng (Singapore)*, vol. 32, no. 5, Oct. 2020, doi: 10.4015/S1016237220500301.
- [41] S. R. Jena and S. T. George, "Morphological feature extraction and KNG-CNN classification of CT images for early lung cancer detection," *Int J Imaging Syst Technol*, vol. 30, no. 4, pp. 1324–1336, Dec. 2020, doi: 10.1002/IMA.22445.
- [42] G. Shah, R. Thammasudjarit, A. Thakkinstian, and T. Suwatanapongched, "NoduleNet: A Lung Nodule Classification Using Deep Learning," *Ramathibodi Medical Journal*, vol. 43, no. 4, pp. 11–19, Dec. 2020, doi: 10.33165/RMJ.2020.43.4.241727.
- [43] C. Chen et al., "An Effective Deep Neural Network for Lung Lesions Segmentation from COVID-19 CT Images," *IEEE Trans Industr Inform*, vol. 17, no. 9, pp. 6528–6538, Sep. 2021, doi: 10.1109/TII.2021.3059023.
- [44] Y. Jalali, M. Fateh, M. Rezvani, V. Abolghasemi, and M. H. Anisi, "ResBCDU-net: A deep learning framework for lung CT image segmentation," *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–24, Jan. 2021, doi: 10.3390/s21010268.
- [45] H. M. Balaha, M. H. Balaha, and H. A. Ali, "Hybrid COVID-19 segmentation and recognition framework (HMB-HCF) using deep learning and genetic algorithms," *Artif Intell Med*, vol. 119, Sep. 2021, doi: 10.1016/j.artmed.2021.102156.
- [46] Z. Bu et al., "Lung Nodule Detection Based on YOLOv3 Deep Learning with Limited Datasets," *MCB Molecular and Cellular Biomechanics*, vol. 19, no. 1, pp. 17–28, 2022, doi: 10.32604/mcb.2022.018318.
- [47] M. Al-Shabi, K. Shak, and M. Tan, "ProCAN: Progressive growing channel attentive non-local network for lung nodule classification," *Pattern Recognit*, vol. 122, p. 108309, Feb. 2022, doi: 10.1016/J.PATCOG.2021.108309.
- [48] "Luna16 | Kaggle." Accessed: Sep. 03, 2023. [Online]. Available: <https://www.kaggle.com/datasets/avc0706/luna16>
- [49] M. Toğaçar, B. Ergen, and Z. Cömert, "Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks," *Biocybern Biomed Eng*, vol. 40, no. 1, pp. 23–39, Jan. 2020, doi: 10.1016/J.BBE.2019.11.004.
- [50] A. Suresh, "What is a confusion matrix?. Everything you Should Know about... | by Anuganti Suresh | Analytics Vidhya | Medium." Accessed: Sep. 03, 2023. [Online]. Available: <https://medium.com/analytics-vidhya/what-is-a-confusion-matrix-d1c0f8feda5>