YOLOv8-Based Lung Nodule Detection: A Novel Hybrid Deep Learning Model Proposal

CİHAT ŞAMAN, ŞERİFE ÇELİKBAŞ

[1] BIOMEDİCAL ENGINEER , SMNWAY INC & cihatsaman@smnway.com, TURKEY
[2]LECTURER, İSTANBUL AYDIN UNIVERSITY & serifecelikbas@aydin.edu.tr, TURKEY

Abstract - *Small oval or circular masses identified in the lungs are known as lung nodules. When these nodules are smaller than 3 cm, they are often considered less concerning; however, over time, they may increase in size, potentially leading to more serious consequences. Early detection of these nodules and timely preventive measures are crucial to impede their progression to malignancy. Conventional diagnostic methods like computerized tomography (CT) and radiographic imaging techniques are utilized for this purpose. Nevertheless, these approaches can either subject patients to excessive radiation or prove inadequate in detecting small nodules. As a result, various deep learning-based image processing techniques are being explored for lung nodule detection. In this study, a novel deep-learning model is proposed for the automated real-time detection of lung nodules. The proposed model exhibits a remarkable accuracy of 92.3% in nodule detection, along with a sensitivity of 88.5% and a mean average precision (mAP) of 53.5%. The model is built using the YOLOv8 architecture, with the YOLOv8m configuration yielding the best results. Additionally, graphical comparisons with existing studies in the literature demonstrate the effectiveness of the training model.*

Key Words: deep learning, lung cancer detection, yolov8x, yolov8l, yolov8m, yolov8s, yolov8n

1. INTRODUCTION

Pulmonary nodules are small, round, or oval-shaped growths in the lungs that can be benign or malignant. They are often detected incidentally on chest imaging or through dedicated screening programs for lung cancer. The presence of pulmonary nodules may cause symptoms such as cough, chest pain, shortness of breath, or coughing up blood, especially if the nodules are large, multiple, or cancerous. Some nodules may also be associated with infections, inflammation, or autoimmune diseases that can affect other organs and systems. Therefore, determining the nature and cause of pulmonary nodules is essential for providing appropriate treatment and preventing complications. The epidemiology of pulmonary nodules depends on various factors, such as the prevalence of smoking, environmental exposures, infectious diseases, and genetic susceptibility. According to the latest guidelines and evidence, the management of pulmonary nodules should be based on the risk of malignancy, which can be estimated by using validated models that incorporate radiographic and clinical features. The size, shape, density, location, and growth rate of the nodules are important radiographic characteristics that influence risk assessment and follow-up recommendations. The use of low-dose computed tomography (LDCT) for lung cancer screening has increased the detection of pulmonary nodules, especially those that are sub-solid or ground-glass in appearance, which have a higher risk of cancer than solid nodules. New criteria for the classification and management of atypical pulmonary cysts, juxta pleural nodules, and inflammatory or infectious findings have been introduced in Lung-RADS® v2022, a standardized reporting system for LDCT screening. The goal of pulmonary nodule evaluation is to identify and treat lung cancer at an early stage while minimizing unnecessary interventions and harm for benign nodules [1-3].

Lung cancer is one of the most common and deadliest types of cancer in the world [4]. Early diagnosis and treatment of lung cancer is critical to improve survival. For this purpose, developed computational image processing studies are carried out. However, lung nodule detection and diagnosis poses challenges for image processing and analysis. Lung nodules can vary greatly in size, shape, density, location, and number. In addition, factors such as noise, artifact, and lack of contrast in images may complicate the detection and characterization of lung nodules. Lung nodule detection using the deep learning method is an important step for early detection of lung cancer. Deep learning techniques for lung nodule detection provide higher performance and accuracy than traditional computer-aided diagnostic systems. Zhang et al. [5] utilized 3D DenseNet and 3D FPN models for detecting lung nodules. They designed a dense feature pyramid network to extract multi-scale features from different layers of the DenseNet backbone. In their study, they achieved an impressive competition performance metric (CPM) value of 0.8934 on the LUNA16 dataset, showcasing compelling results. Furthermore, the detection performance was improved by approximately 2% compared to other methods Zhang et al. [6] proposed an integrated active contour model (IACM_MRFEBPD) for small ground glass opacity (GGO) pulmonary nodule segmentation. The method combines Markov random field energy and Bayesian probability difference, achieving an average IOU of 0.7444, 0.7503, and 0.7450 for LIDC-IDRI

test set, clinical test set, and all test sets, respectively. The approach enables accurate and robust segmentation, aiding in early lung cancer detection through medical imaging evaluation. Ye et al. [7] proposed an automated pulmonary nodule detection model using modified V-Nets and an advanced SVM classifier. The model achieved 66.7% accuracy, 81.1% true positive rate, 78.1% positive predictive value, and 78.7% F-score. Experimental results on the LUNA16 dataset demonstrate its effectiveness in early lung cancer detection. Huang et al. [8] propose a fast and fully automated end-to-end system for precise pulmonary nodule analysis. Their approach achieves high accuracy, with 91.4% and 94.6% detection rates and an average of 1 and 4 false positives (FPs) per scan. The nodule segmentation exhibits an average dice coefficient of 0.793, demonstrating its effectiveness in automated lung nodule analysis. Luo et al. [9] present an anchor-free 3D sphere-based lung nodule detection network. The approach automatically predicts nodule position, radius, and offset without predefined anchor parameters. The model effectively performs existing methods on the LUNA16 dataset, achieving high true positive rates (TPR) of 92.2 (1 FPs/image), 93.9 (2 FPs/image), and 96.4 (8 FPs/image) for lung nodule detection. Vipparla et al. [10] propose three diverse 3D Attention-based CNN architectures for lung nodule detection. The models use attention mechanisms and contextual dependencies to achieve significant improvements in classification performance. Fusion of their predictions leads to more effective results with a CPM score of 0.931 on the LUNA16 dataset. Han et al. [11] propose a CAD system, Pulmonary Nodules Detection Assistant Platform, for early detection and classification of pulmonary nodules in physical examination LDCT images. Their 3D CNN-based model achieves an accuracy of 0.879 with an average of 1 false positive per CT in LNPE1000 dataset. The system shows efficiency in detecting smaller and larger nodules, offering potential benefits for early lung cancer detection in physical examination applications. Alakwaa et al. [12] propose a computer-aided diagnosis (CAD) system for lung cancer classification using CT scans with unmarked nodules. The system achieves an accuracy of 86.6% and requires less labeled data, offering a more efficient and generalized approach to lung cancer diagnosis.

Lung nodule detection was addressed in this study using a two-stage deep learning model. In the first stage, image processing filters were employed to enhance learning performance and highlight nodule regions. Subsequently, a Convolutional Neural Network (CNN) was trained, achieving an impressive 94% accuracy. The resulting weight model was then utilized to detect cancerous regions. For the next step, the nodule regions identified by the CNN and those marked by radiology experts were prepared as input windows and trained with YOLOv8 models. The performance metrics, including accuracy (92.3%), precision (87.3%), and mean Average Precision (mAP - 53.5%), were obtained based on the training data. The study results were presented in tables and compared with findings from the existing literature.

The other sections of the paper are organized as follows: Section 2 introduces the dataset and presents the explanations of the methods and techniques used. Section 3 explains the general workflow of the proposed convolutional model. Section 4 presents experimental results and comparative analyses demonstrating the effectiveness of the proposed approach. Finally, Section 5 includes evaluations related to the obtained results.

2. MATERIAL METOD

2.1. Dataset

The LUNA16 dataset is derived from the Lung Image Database Consortium (LIDC-IDRI) database, which has been publicly released. The LIDC-IDRI database contains CT images and expert radiologists' annotations. The LUNA16 dataset consists of 888 chest CT scans in DICOM format. The dataset provides DICOM files along withs .mhd and..zraw extensions. The annotations for nodule regions were determined by expert radiologists. The dataset includes annotations for both malignant (cancerous) and benign (non-cancerous) nodules. Nodules vary in terms of their sizes, shapes, and locations within the lungs. In this study, the LUNA16 dataset was processed using the Python-3.11.1 torch-2.0.1+cpu CPU (12th Gen Intel Core (TM) i7-12700H) hardware in the PyCharm IDE, and a deep learning model was proposed for nodule detection [13].

2.2. Convolutional Neural Network

The increasing advancement and accessibility of hardware technologies, coupled with the abundance of data in various fields, have positively influenced the proliferation and success of deep learning studies [15]. Convolutional Neural Network (CNN) is a deep learning method that processes its inputs by analyzing specific features through different layers for classification tasks. Its ability to determine features on its own provides ease of use. As a result, it is effectively employed in various medical applications such as cancer diagnosis and nodule detection [16-18]. CNN, as shown in Figure 1, consists of a convolutional layer, pooling layer, activation layer, and fully connected layer. The convolutional layer works based on the logic of applying filters of a predetermined size to perform operations on the image and extract overlapping regions. It demonstrates efficiency in detecting low-level features such as edge detection, gradient orientation, and color.

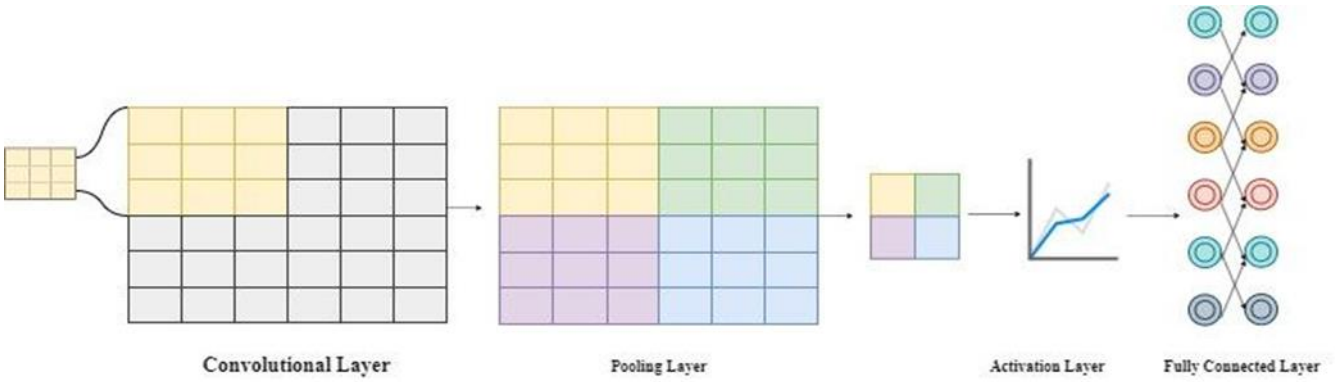

Fig -1: CNN architecture [19]

After the convolutional layer, the pooling layer contributes to reducing the dimensionality of the data, thereby reducing the computational power required to process it. It also aids in extracting spatially invariant features and helps prevent overfitting. An activation function is applied to the computation results to finalize the output and create a mathematical value. Subsequently, the obtained outputs are passed to the fully connected layer to be transformed into a single vector and then forwarded to classification functions [19].

In this study, images obtained from .mhd extension DICOM files were processed with the CNN architecture shown in Figure 4 after the preprocessing steps. In addition to the general layers, the dropout layer is also used in the architecture. In order to increase the learning performance, the learning rate was reduced after 10 epochs and satisfactory results were obtained. The regions marked by the weight model formed as a result of the training were recorded in .txt format. In these txt files, png coordinates are retained and the corresponding coordinates are determined when the coordinates shared in the LUNA16 dataset are converted to 2D form. For each png, the positions determined by both the CNN architecture and the expert radiologists in the data set were taken as the YOLO input. The results obtained are presented in Chapter 4.

2.3. You Only Look Once (YOLO)

YOLO (You Only Look Once) is an object detection algorithm that has become increasingly popular in recent years, reducing the need for high-performance hardware. YOLO is fundamentally a deep learning algorithm based on CNN and is commonly used in applications such as segmentation, classification, object detection, and real-time object tracking [20, 21]. Its initial version was published in 2015, and since then, various versions have been developed to maintain a balance between speed and high accuracy.

Real-time nodule detection in medical data is crucial for achieving fast and high-performance results with deep learning [13]. Therefore, this study utilizes YOLOv8 for this purpose Accordingly, this study proposes a two-stage learning system for medical data. In the first stage, an abundant and sufficient amount of data is preprocessed and trained with the CNN architecture. After the training and nodule identification, in the second stage, YOLOv8 is fine-tuned for real-time applications to achieve fast and convenient integration.

YOLOv8 is created in 5 different scale types. It can learn features such as object detection, segmentation, pose estimation, tracking, as well as multiple views in different dimensions. These scales; YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra-large) [22]. This study also presents the results of these five different scales for the same epoch comparatively.

3. PROPOSED MODEL

A hybrid model is being proposed to improve the accuracy, efficiency, and real-time capability of lung nodule detection and enhance the differentiation between benign and malignant nodules. The overall workflow of the study is illustrated in Figure 2.

Initially, the data from the LUNA16 dataset was preprocessed to determine the lung parenchyma by adjusting the values to the Hounsfield unit scale [-1000, 400][13]. Preprocessing steps were then applied to each DICOM slice, followed by combining each slice with other images from the corresponding .mhd extension file. This process yielded a total of 888 .png files. These files were subsequently divided into training, testing, and validation sets using the CNN model, and then trained with the CNN model, resulting in the creation of a weight file (.h5).

Using the obtained weight file, nodule detection was performed on all 888 images, and the predicted regions containing nodules were annotated. The coordinates of these regions were recorded and compared with the average coordinates provided by expert radiologists in the LUNA16 dataset using an .mhd file. In cases where there

were significant differences in the coordinates, two sets of coordinates were generated. Coordinates that were in close proximity to each other were merged to form a single rectangle, and the data was saved.

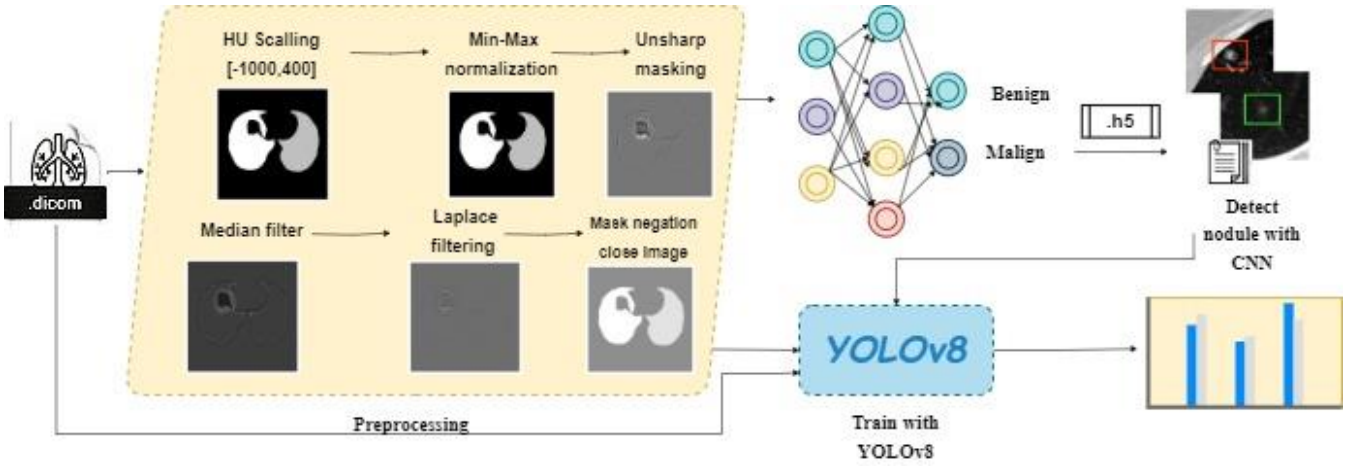

Fig -2: Workflow diagram of the proposed model

Subsequently, the acquired data and image inputs were trained using YOLOv8, which offers fast training and real-time processing capabilities. The learning outputs were presented along with performance metrics. Finally, the trained weight model was validated for each DICOM slice.

3.1. İmage Preprocess

After extracting the files with the .mhd extension from the segmentation folder containing 888 CT images in the LUNA16 dataset, the first step involved extracting the lung parenchyma based on the Hounsfield Unit scale. Subsequently, a series of image-processing steps were applied to enhance the lung parenchyma and facilitate subsequent analysis. These steps included min-max normalization to ensure consistent data scaling and prevent errors due to unusual data values. Furthermore, the following procedures were performed: Unsharp Masking, Median Filtering, Laplacian Filtering, Binarization, Mask Negation, Morphological Closure, Mask Negation (Repeated), Hollow Filling, and Mask Subtraction. The accompanying Figure 3 illustrates the progressive changes in the parenchymal image as each procedure was applied.

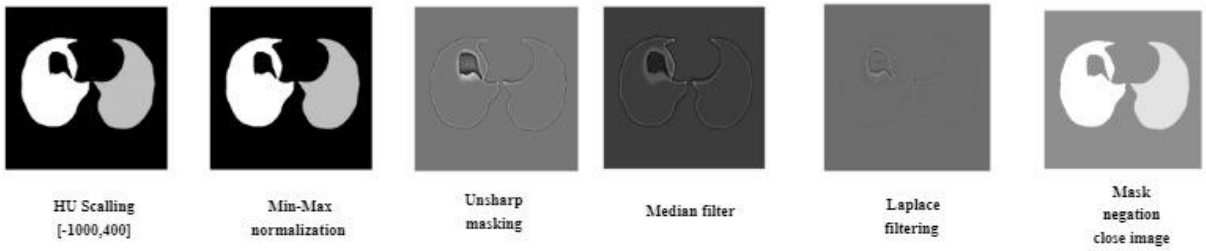

Fig -3: Image preprocessing pipelines

Hounsfield Units are characterized as a quantitative measurement of radio intensity. It is used to make regional comments on CT and MR images and to analyze the existing tissue type. For this study, HU scaling was performed in the range of [-1000, 400]. These values were chosen because they highlight the lung parenchyma tissue [13]. After determining the pakanchyme texture, min-max normalization was applied to scale the intensity values of each pixel in the image to the [0, 1] range. The normalization formula is as follows:

$$X' = (X - min\_density) / (max\_density - min\_density) \quad (1)$$

Here X represents the original pixel intensity value, X' represents the normalized pixel intensity value, min_density, and max_density represent the minimum and maximum intensity values in the original range. In the second stage, unsharp masking was applied. It is used to highlight edges and details in the image by removing a blurred version of the original image from the original image. A median filter, which is a noise reduction technique that replaces the pixel value with the median

value, has been applied in order to eliminate the noise generated after this process. It is the Laplace filtering technique, which is an edge detection technique that emphasizes rapid density changes in the image in order to more precisely detect the nodules in the image. After this filtering step, a binarization process determined according to a certain threshold was applied in order to analyze the nodules more easily. And in the final stage, mask negation process was applied in order to emphasize the nodules more and reduce the noise from the background.

After the creation of the data, the images were adjusted to 512x512 for training purposes and passed through the CNN model given in Figure 4. A convolutional layer is used, which creates feature maps by scanning with various filters and is effective in detecting edges, color changes and other important patterns in the image. 32 3x3 filters and relu layers as activation functions, 64 3x3 filters and relu layers as activation functions are used, respectively. In order to reduce the computational load and reduce the risk of overfitting, maximum pooling is applied in size (2, 2). In order to provide input to the dense layer, it flattens the output of Conv2D and MaxPooling2D layers into a single vector. In the Dense layer, important features are learned for classification by combining their features. In the last stage, the SoftMax function is used to determine the class prediction. The weight model with .h5 extension obtained at the end of the process was recorded.

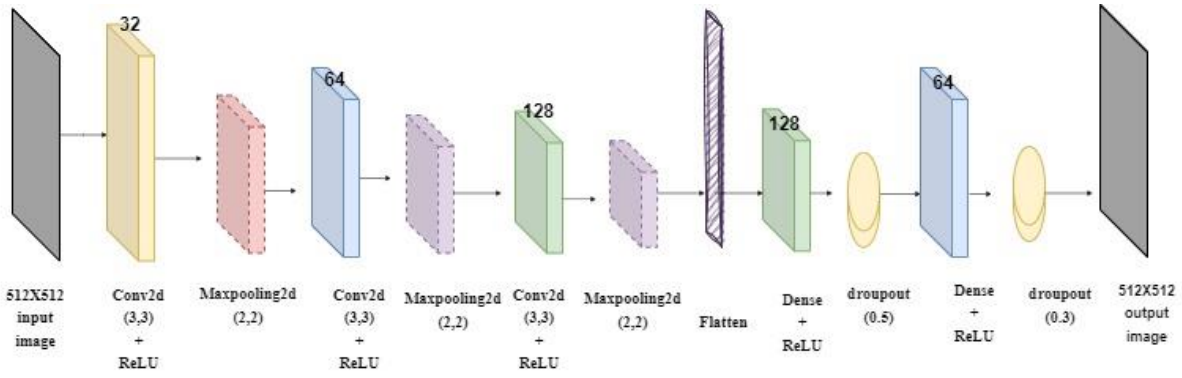

Fig -4: Proposed CNN architecture

After 150 epochs of training, we achieved a satisfactory training success rate of 94%. Detailed training outcomes are extensively documented in Chapter 4. The obtained model was employed for detecting nodules, and each nodule's boundary was enclosed within a rectangle. The resulting data from this process were stored in separate .txt files, which serve as input data for YOLOv8. Additionally, the data information annotated by expert radiologists in .mhd files was also incorporated into these .txt files. Subsequently, we conducted training with 5000 epochs using the ultrasonic library on the yolov8.pt model. The results obtained were highly satisfactory and demonstrated comparability with findings reported in the literature. Section IV provides a comprehensive account of the experimental results.

4. EXPERİMENTAL RESULTS

In this study, the evaluation metrics employed for lung nodule detection include accuracy, sensitivity, recall, precision, mAP (mean Average Precision), and mAP50-95 (mean Average Precision between 50 and 95). The obtained results are based on the outputs of TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) generated after the training process.

Accuracy represents the percentage of correctly classified nodules relative to the total number of detected nodules. It is calculated using the formula:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

Sensitivity measures the percentage of true positive cases relative to the sum of true positive and false negative cases. The formula is as follows:

$$Sensivity = TP / (TP + FN) \quad (3)$$

Recall denotes the sensitivity percentage relative to the sum of true positive and false negative cases. The formula is given by:

$$Recall = TP / (TP + FN) \quad (4)$$

Precision indicates the percentage of true positive cases relative to the sum of true positive and false positive cases. The formula is as follows:

$$Precision = TP / (TP + FP) \quad (5)$$

mAP(mean Average Precision) represents the average value of sensitivity at different recall levels. The formula is given by:

$$mAP = (1 / N) * \Sigma(Recall) \quad (6)$$

mAP50-95(mean Average Precision between 50 and 95 reflects the average sensitivity between recall levels of 50% to 95%. The formula is given by:

$$mAP50\text{-}95 = (1 / (95 - 50 + 1)) * \Sigma(Recall) \text{ (for i from 50 to 95)} \quad (7)$$

In this study, the primary objective is to enable real-time and rapid detection of benign or malignant nodules within the region captured by CT images. Lung nodule detection is achieved by training a Convolutional Neural Network (CNN) using information from nodule regions based on radiologist annotations. The dataset consists of 888 CT images in DICOM format, each containing corresponding .mhd and .zraw files. Through image processing and preprocessing steps, 551,065 PNG images were generated. Notably, the 93rd and 94th PNG layers, which generally display the lung parenchyma for each CT, were included in the training process. The training achieved 94% accuracy over 150 epochs.

Subsequently, an .h5 model was created after training. Using this model, nodules were detected in the images and saved to a .txt file. Additionally, the coordinates marked by radiology experts in the candidates.csv file of the LUNA16 dataset were converted into 2D .png planes. The resulting coordinate information was also saved in the relevant .txt file. Both the generated .txt files and images were utilized as input for YOLOv8 and evaluated for five different scales over 30 epochs. The obtained results have been summarized in Table 1. According to these results, when evaluated at different scales, the YOLOv8x model demonstrated the highest precision value. Furthermore, the mAP (mean average precision) and mAP50-95 (mean average precision at IoU thresholds of 50 to 95) values showed the best performance for the YOLOv8m model. These findings indicate that the YOLOv8 model achieved high success in the object detection task, and particularly, the YOLOv8x and YOLOv8m models outperformed others when compared across different scales.

Simultaneously, these findings provide evidence that the proposed methodology is suited and confirm its practicality, particularly in the context of real-time applications. The results underscore the methodology's potential to excel when supplied with more comprehensive training data and deployed on robust hardware.

Table 1. Performance Evaluation of YOLOv8 Scales Trained over 30 Epochs

YOLO scales	Precision	mAP50	mAP50-95
Yolov8n	0.291	0.356	0.296
Yolov8s	0.775	0.505	0.394
Yolov8m	0.873	**0.535**	**0.4**
Yolov8l	0.626	0.523	0.303
Yolov8x	**0.894**	0.482	0.36

Moreover, the study results exhibit comparability with similar previous works, as presented in Table 2, showcasing the performance outcomes in a comparative manner. The findings highlight the efficacy and potential of the proposed approach for lung nodule detection, contributing to the advancement of medical imaging and early lung nodule diagnosis.

Table 2. Performance Evaluation of Studies Utilizing the LUNA16 Dataset

Study	Acc.	Sens.	Method Used
Ye et al. [7]	66.7 %	81.1%	Modified V-Nets and Advanced SVM Classifier
Huang et al. [8]	91.4 %	-	Fast and Fully Automated End-to-End System
Han et al. [11]	87.9 %	-	3D CNN-Based Model for Physical Examination LDCT Images
Alakwaa et al. [12]	86.6 %	-	Computer-Aided Diagnosis (CAD) System for Lung Cancer Classification Using CT Scans with Unmarked Nodules
Proposed	**92.3 %**	**88.5 %**	**The model trained using 2D Convolutional Neural Networks (2DCNN) and YOLOv8**

5. CONCLUSIONS

A deep learning model is proposed in this research to facilitate disease diagnosis with reduced radiation exposure. The detection of nodules and cancerous tissues is targeted using real-time imaging devices, such as C-arm fluoroscopy, which involve lower radiation doses. A novel

deep learning model, specifically tailored for the detection of lung nodules, enables real-time automated detection of predefined nodules. Lung nodule detection is achieved with an accuracy of 92.3%, a sensitivity of 88.5%, and a mean average precision (mAP) of 53.5%. When compared to the YOLOv8 architecture, superior results are obtained with the YOLOv8m architecture. Furthermore, graphical comparisons with related studies in the literature effectively demonstrate the effectiveness of the training model. This study represents a significant stride towards early lung nodule detection and the implementation of preventive measures, offering valuable insights for future research. Deeper networks and larger datasets can be explored in future studies to further enhance the robustness and accuracy of the results.

REFERENCES

[1] Callister, M.E. et al. (2015) 'British Thoracic Society Guidelines for the investigation and management of pulmonary nodules: Accredited by nice', Thorax, 70(Suppl 2), pp. ii1–ii54. doi:10.1136/thoraxjnl-2015-207168.

[2] Larici, A.R. et al. (2017) 'Lung nodules: Size still matters', European Respiratory Review, 26(146), p. 170025. doi:10.1183/16000617.0025-2017.

[3] Mazzone, P.J.; L. (2022) Evaluating the patient with a pulmonary nodule, jama, DeepDyve. Available at: https://www.deepdyve.com/lp/american-medical-association/evaluating-the-patient-with-a-pulmonary-nodule-Z1JwqOiXDs (Accessed: 06 August 2023).

[4] Chen, X., Mo, S. and Yi, B. (2022) 'The spatiotemporal dynamics of lung cancer: 30-year trends of epidemiology across 204 countries and Territories', BMC Public Health, 22(1). doi:10.1186/s12889-022-13281-y.

[5] Zhang, M. et al. (2021) 'Pulmonary nodule detection based on 3D feature Pyramid Network with incorporated squeeze-and-excitation-attention mechanism', Concurrency and Computation: Practice and Experience, 35(16). doi:10.1002/cpe.6237.

[6] Zhang, S. et al. (2020) 'Segmentation of small ground glass opacity pulmonary nodules based on Markov random field energy and bayesian probability difference', BioMedical Engineering OnLine, 19(1). doi:10.1186/s12938-020-00793-0.

[7] Ye, Y. et al. (2020) 'Pulmonary nodule detection using V-net and high-level descriptor based SVM classifier', IEEE Access, 8, pp. 176033–176041. doi:10.1109/access.2020.3026168.

[8] Huang, X. et al. (2019) 'Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks', Computerized Medical Imaging and Graphics, 74, pp. 25–36. doi:10.1016/j.compmedimag.2019.02.003.

[9] Luo, X. et al. (2022) 'SCPM-net: An anchor-free 3D lung nodule detection network using sphere representation and center points matching', Medical Image Analysis, 75, p. 102287. doi:10.1016/j.media.2021.102287.

[10] Vipparla, V.K., Chilukuri, P.K. and Kande, G.B. (2021) 'Attention based multi-patched 3D-cnns with hybrid fusion architecture for reducing false positives during lung nodule detection', Journal of Computer and Communications, 09(04), pp. 1–26. doi:10.4236/jcc.2021.94001.

[11] Han, Y. et al. (2022) 'Pulmonary nodules detection assistant platform: An effective computer aided system for early pulmonary nodules detection in physical examination', Computer Methods and Programs in Biomedicine, 217, p. 106680. doi:10.1016/j.cmpb.2022.106680.

[12] Alakwaa, W., Nassef, M. and Badr, A. (2017) 'Lung cancer detection and classification with 3D Convolutional Neural Network (3D-CNN)', International Journal of Advanced Computer Science and Applications, 8(8). doi:10.14569/ijacsa.2017.080853.

[13] Terven, J. and Cordova-Esparza, D. (2023) A comprehensive review of Yolo: From Yolov1 and beyond, arXiv.org. Available at: https://doi.org/10.48550/arXiv.2304.00501 (Accessed: 03 August 2023).

[14] Chi, J. et al. (2020) 'A novel pulmonary nodule detection model based on multi-step cascaded networks', Sensors, 20(15), p. 4301. doi:10.3390/s20154301.

[15] Potok, T.E. et al. (2018) 'A study of complex deep learning networks on high-performance, neuromorphic, and quantum computers', ACM Journal on Emerging Technologies in Computing Systems, 14(2), pp. 1–21. doi:10.1145/3178454.

[16] Jin, H. et al. (2023) 'Machine learning techniques for pulmonary nodule computer-aided diagnosis using CT images: A systematic review', Biomedical Signal Processing and Control, 79, p. 104104. doi:10.1016/j.bspc.2022.104104.

[17] Alzubaidi, L. et al. (2021) 'Review of Deep Learning: Concepts, CNN Architectures, challenges, applications, Future Directions', Journal of Big Data, 8(1). doi:10.1186/s40537-021-00444-8.

[18] Ker, J. et al. (2018) 'Deep learning applications in medical image analysis', IEEE Access, 6, pp. 9375–9389. doi:10.1109/access.2017.2788044.

[19] Bhatt, D. et al. (2021) 'CNN variants for Computer Vision: History, architecture, application, challenges and future scope', Electronics, 10(20), p. 2470. doi:10.3390/electronics10202470.

[20] Ünver, H.M. and Ayan, E. (2019) 'Skin lesion segmentation in dermoscopic images with combination of Yolo and GrabCut algorithm', Diagnostics, 9(3), p. 72. doi:10.3390/diagnostics9030072.

[21] Tan, L. et al. (2021) 'Comparison of RetinaNet, SSD, and Yolo V3 for real-time pill identification', BMC Medical Informatics and Decision Making, 21(1). doi:10.1186/s12911-021-01691-8.

[22] Ultralytics, Ultralytics/ultralytics: New - yolov8 in PyTorch > ONNX > OpenVINO > CoreML > TFLite, GitHub. Available at: https://github.com/ultralytics/ultralytics (Accessed: 06 August 2023).