ORIGINAL RESEARCH

DETERMINATION OF LOCATION AND SEVERITY OF NODULES ON LUNG CANCER CT IMAGE USING YOLO METHODS

Hanun Masitha Ramadhani | Chastine Fatichah*

[1]Dept. of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Correspondence

*Chastine Fatichah, Dept of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Email: chastine@if.its.ac.id

Present Address

Gedung Teknik Informatika, Jl. Teknik Kimia, Surabaya 60111, Indonesia

Abstract

The severity of lung cancer can be used to determine appropriate treatment measures and reduce the risk of death. The severity identification is monitored based on the size and location of the nodule. However, previous studies still focused on determining the location of nodules without identifying their severity. In this study, the severity of lung cancer is detected based on the size of its nodules. This research contributes to the annotation of severity to the Lung Image Database Consortium image collection (LIDC-IDRI) dataset and the development of automatic severity detection using You Only Look Once (YOLO) methods. The data is given a severity level based on the nodule size calculated based on the number of pixels in the nodule length. Automatic detection is done using YOLO methods, which consist of several versions, namely YOLOv5, YOLOv7, and YOLOv8. YOLO methods can properly detect the location and severity of cancer nodules with the IoU evaluation results obtained using YOLOv5, YOLOv7, and YOLOv8, which are 0.86, 0.6, and 0.87, respectively. From the experiment, it can be concluded that determining the location and severity of cancer based on nodule size using YOLO methods is proven effective and can be done in real-time.

KEYWORDS:

Lung Cancer, Nodule's Severity, Object Detection, YOLO

1 | INTRODUCTION

Lung cancer is abnormal tissue that grows in the lung organs[1]. Based on data from the Global Cancer Observatory (Globocan), about 2.2 million new patients have been diagnosed with lung cancer, which ranks second after breast cancer. Meanwhile, the number of patients who died was 1.7 million throughout 2020, which ranks first in the world[2]. From the following data, it can be concluded how deadly lung cancer is if it is not detected properly and treated immediately. Cancer can be detected by

utilizing various techniques in the field of radiology, such as X-rays, Computed Tomography (CT), and Magnetic Resonance imaging (MRI).

The output of these techniques is usually in the form of images that are generally grouped under the term medical image. CT is the most popular method because it provides more accurate results and is cheaper than X-rays and MRI[3]. The results of CT images can show lung segments to show the location and size of lung nodules[1]. Nodule size is one of the indicators used to assess the malignancy of a nodule and can also determine health procedures as early as possible by international guidelines[4]. So, in the early treatment of lung cancer, precise measurements of nodules are needed[5]. Although CT can describe the condition of the lungs with high accuracy, the results of CT can also be misdiagnosed. One of the main reasons for misdiagnosis is the misinterpretation of the image by the attending physician[6]. The interpretation of CT images is still done manually by a radiologist. Radiologists take 15 to 20 minutes to diagnose the disease; each day, there are 45 images to analyze; this affects the radiology performance itself[6]. Errors in interpreting this image count as having a 25% - 45% chance of error. Factors influencing this error rate are the level of vigilance and perspective of radiologists in looking for lung nodules, working conditions, human factors such as fatigue and bias, and suspect indices of patients such as patient age and clinical history[7].

In this situation, a tool is needed to help experts by reducing reading time and detecting nodules with high accuracy. This can be helped by creating Computer-aided Detection (CAD)[8]. CAD itself is a technology designed to help experts detect diseases. CAD has been proven to help doctors detect cancerous nodules[9]. Making CAD is based on several machine learning methods to read and process data to help detect a disease.

One example of machine learning in CAD is research that uses Faster DCNN[10] and YOLO[11] in making nodule location detection devices. In research using YOLO, it can be concluded that the architecture can detect nodules quickly and precisely. In previous studies, no detection was found based on nodule severity. The severity of the nodule is one of the elements to consider while diagnosing and deciding the next medical procedure to be conducted[4]. This study proposes a technique based on lung CT scan pictures to determine the location of nodules and the severity of lung cancer. The contribution is a new annotation of the dataset based on the severity as measured by the length of the nodule. The YOLO technique detects nodule location and severity automatically depending on nodule size. This study analyzed different YOLO variations to determine the optimum YOLO variant model performance.

2 | PREVIOUS RESEARCHES

One example of detecting malignant nodules is using the object detection approach to display the nodule's location[9]. In previous research, researchers used the Lung Nodule Analysis 16 (LUNA-16) dataset and the Lung Image Database Consortium (LIDC) to obtain a CT image to see the Region of Interest (ROI). This study divided it into three parts: identifying the predicted ROI nodule part, minimizing false positives, and identifying nodules in advance. Every layer in the first section identifies its ROI. After each ROI has been calculated, it can be concatenated to provide a 3D image. Following that, the size of each nodule will be used to estimate each ROI. The size of each nodule is compared to a threshold. The data will be removed if the size is smaller than the threshold (false positive). After that, InceptionV4 is used to classify the image and identify nodules and non-nodules. A Competition Performance Metric (CPM) was used to calculate the experiment's findings, and the outcome was 90%[10].

In another study, the researchers attempted to use various YOLO methods on the Lung Image Database Consortium (LIDC) dataset, namely YOLOV3, YOLOV5, and Improved YOLOV5. The YOLO approach is recommended because its algorithm can learn new data quickly and accurately. In the Improved YOLOV5 model, researchers added an SE Net layer to determine different weights for each input feature map. Then, add the F-Concat Block layer, which functions to combine several features at different levels. The results of this study will measure precision, recall, and AP. The improved YOLOV5 produced the best results from the experiment with 89.5% precision, 84.6% recall, and 90% AP[11].

Nodule detection has been done before using the YOLO method, which shows the nodule's location[10, 11]. According to this research, YOLO has a faster detection speed and improved system performance. However, previous studies only focused on detecting the location of nodules and did not classify nodule severity. One of the factors in deciding the disease's diagnosis and the subsequent medical therapy is considering the severity. Therefore, this study proposes a system to detect the location of nodules and the severity of lung cancer based on lung CT scan images.
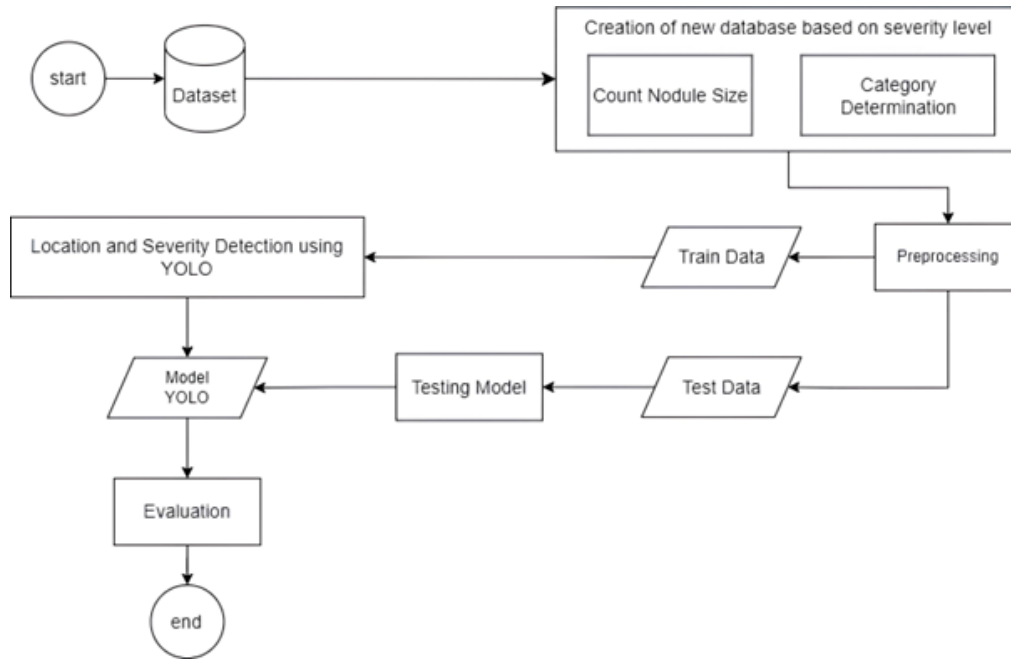
FIGURE 1 The proposed system for determining the location and severity of lung cancer.

3 | MATERIAL AND METHOD

In the study of Determining the Location and Severity of Nodules on CT Images of Lung Cancer Using the YOLO Method, there are three research steps, namely the development of a new dataset based on severity, dataset training, and dataset testing. The research design for this study includes information on the actions taken to complete the task by the research framework shown in Fig. 1 .

3.1 | Dataset

The dataset used in this study is taken from the Lung Image Database Consortium image collection (LIDC-IDRI)[12]. The dataset was obtained from 7 academic centers and eight medical imaging companies. The dataset consists of 1018 cases of lung cancer. The image from the dataset itself is a clinical thoracic (CT) image with a size of 512 x 512. The dataset is stored in several types: CT scan image files, nodule annotation images, and data records of the cancer case.

Each case is classified into nodule cases and clean nodule cases based on the diagnoses of five professional doctors. If just two doctors find a nodule with a size of more than 3 mm, the patient is not diagnosed as having a pulmonary nodule or clean nodule case. If more than three doctors find nodules with a size of more than 3 mm, the patient is diagnosed as having pulmonary nodules or nodule cases.

3.2 | Creation of Severity Level Database

We divided the new database based on severity level into two parts: calculating the nodule's size and determining the severity.

3.2.1 | Count Nodule Size

The nodule measurement process starts with searching the contour lines of the image. Contour lines can be detected due to differences in image color intensity. These color differences cause the dots between colors to become contour lines. We look for the four outermost points from the contour lines, which later form a square of the nodule. The 4 points are named Top Left (TL), Top Right (TR), Bottom Left (BL), and Bottom Right (BR). From each TL and TR, the midpoint will be searched, which
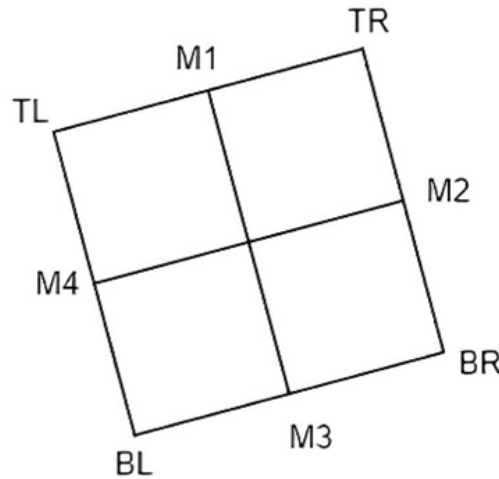
FIGURE 2 The illustration of dotting in nodule images.

TABLE 1 The nodule category determination rules.

Level	Range
1	6 mm < Nodule Size ≤ 10 mm
2	10 mm < Nodule Size ≤ 20 mm
3	20 mm < Nodule Size ≤ 30 mm
4	Nodule Size > 30 mm

is named M1. The TR and BR points will be searched for the midpoint and named M2. The midpoint between BR and BL is M3, and the midpoint between TL and BL is M4. This point will be illustrated in Fig. 2 .

After finding all the points in the image, the next step is to find their length. The Euclidean distance measurement formula (see equation 1) assists in measuring length, where (x_1, y_1) is the coordinate of the first point and (x_2, y_2) is the coordinate of the second point.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

We must find two distances between points in (M1 and M3) and (M2 and M4) with the Euclidean distance measurement formula. After that, we convert the unit from pixels to millimeters (mm). 1 pixel in a 512 x 512 lung CT image has a size of 0.617 mm. In two distances, the longest distance will represent the distance in the nodule in the picture.

3.2.2 | Category Determination

The next stage is labeling each nodule according to the severity of malignancy identifiable from the nodule's size. This classification is based on the Nelson test[13] and the Fleischener Society Guidelines[14] regarding the likelihood of nodule malignancy based on nodule size. According to Table 1 , the classification is separated into four stages, labeled from level 1 to level 4. Level 1 is a level with nodules ranging from 6 to 10 mm. They have a less than 10% malignancy rate. Level 2 is an additional option, with nodules between 10 and 20 mm in size and a 10% to 20% chance of malignancy. Level 3 contains nodules between 20 and 30 mm in size with a chance of spite of about 25%, while level 4 has nodules larger than 30 mm and a probability of malignancy greater than 30%. Roboflow Web assists with labeling. Nodule labeling is based on the nodule category determination rules found in Table 11 . Roboflow also helps divide data into training, testing, and validation data.

TABLE 2 The differences from the architecture of YOLO.

Architecture	Year	Note
YOLO	2015	The first time the YOLO system was published.
YOLOv2	2016	Improves batch normalization, anchor boxes, and dimension cluster functions.
YOLOv3	2018	Improves efficient backbone function. Dn applies multiple anchors and spartial pytamid pooling.
YOLOv4	2020	Introducing mosaic data and anchor-free detection head.
YOLOv5	2020	Added hyperparameter optimizer, combining experiment tracking and automatic export.
YOLOv6	2022	Introducing finetuned with the help of the proposed hidden layer pruning algorithm.
YOLOv7	2022	Add an estimation pose.
YOLOv8	2023	Introducing C2F as a lighter CSP replacement and improving the performance of PAN-FPN.

3.3 | Preprocessing

In the preprocessing process, we are processing an image augmentation process. Data augmentation is a way to increase the diversity of training data that utilizes transformation inputs. In computer vision, image augmentation becomes a regulatory technique to reduce overfitting in deep learning processes[15]. Rotation is one technique for augmentation. A technique for data augmentation called rotation involves tilting an image or object. This method is used in image processing and machine learning to broaden the variety of datasets and make trained models more tolerant of rotated or slanted object recognition. In this research, rotations at 15°, 30°, 180°, 330°, and 340° degrees were used.

3.4 | YOLO

You Only Look Once (YOLO) algorithm detects an object directly. YOLO implements an architecture similar to a Convolutional Neural Network (CNN). YOLO applies convolution and pooling layers to its architecture[16]. The difference between YOLO and classification using CNN is the addition of localization to the classification of YOLO. Localization is the location of objects in the form of bounding boxes (bx, by, bh, bw)[17]. Because of the need for location on the intended object or label, it is necessary to annotate the label before entering the YOLO training process. In the annotation process, each image stores information on the label name, x label coordinate points, y label coordinate points, bounding box length, and bounding box width [18].

The main advantages of YOLO are small model sizes and fast computational speed when detecting objects in real-time. YOLO itself has several versions, from version one to version 8. The differences between each version of YOLO are shown in Table 2 .

3.4.1 | YOLOv5

YOLOV5 focuses on the application of data enhancement, which plays an important role in detecting small image objects[11]. Data enhancement is the process of improving data by adding external context. The data enhancement processes used in YOLOV5 are scaling, color space adjustment, and mosaic enhancement. The advantage of YOLOV5 is that it is also simpler to use and faster due to its lightweight model[18].

The architecture of the YOLOV5 is divided into four parts: the input, backbone, neck, and head. In the input process, data enhancement processes prepare data before entering the process in YOLOV5. The next step is the backbone, which focuses on extracting the features in the image[19]. The backbone of YOLOV5 has a focused structure, multiple Conv, C3 modules, and SSP modules. The focus structure ensures information from the incoming image and reduces the size of the input data. After that, the data was extracted using multiple Conv, C3 modules, and SPP modules to form a new feature map[11]. Then, the feature map enters the neck section. The neck is often considered a decoder layer that combines feature maps from the previous step. YOLOV5 developed the Path Aggregation Network (PANet), which enhances information by shortening low-level and high-level information pathways through bottom-up path enhancement[19]. The final step on YOLOV5 is the head section. This head predicts targets and forms a target bounding box[11]. The illustration of the architect YOLOV5 is shown in Fig. 3 .

Each of YOLOv5's versions has a distinct number of parameters and input data size. YOLOv5 versions include YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv5n6, YOLOv5s6, YOLOv5m6, YOLOv5l6, and YOLOv5x6. Details of the differences between each version of YOLOv5 are in Table 3 .
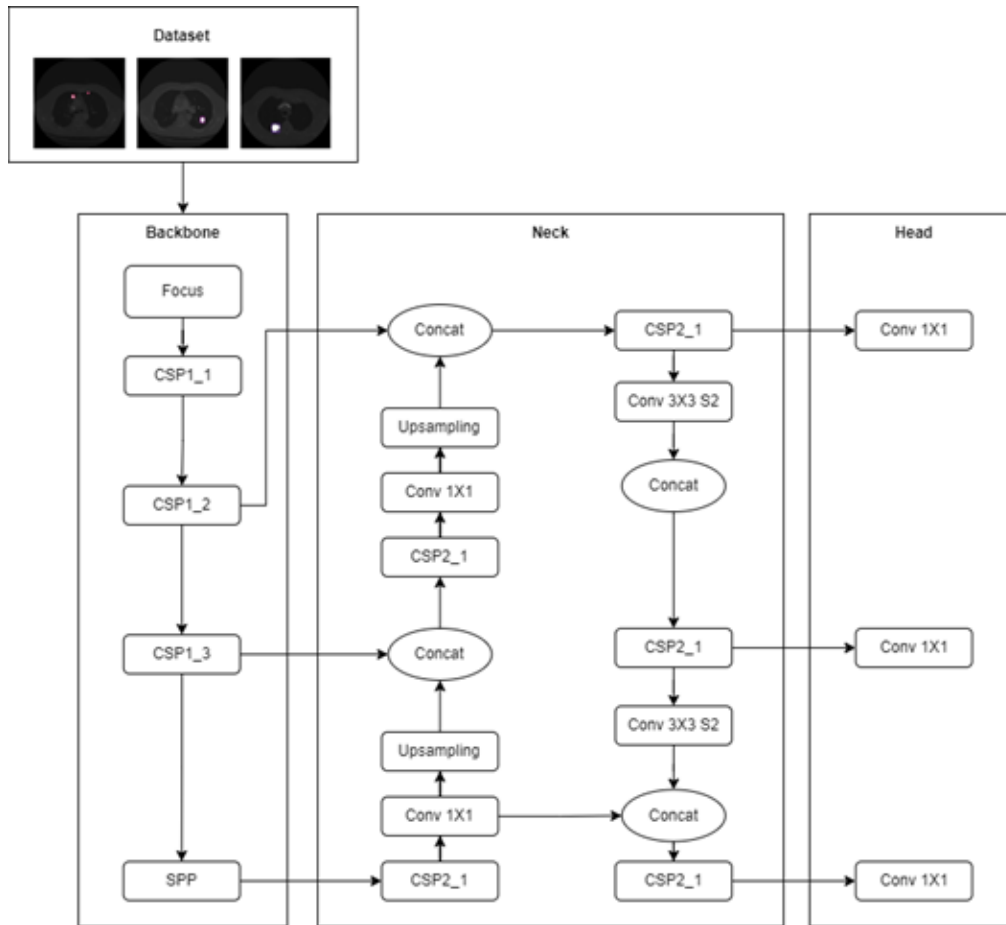
FIGURE 3 The YOLOv5 architecture for determining location and severity of lung cancer.

TABLE 3 The differences between various YOLO5 versions.

Version	Input Size (pixel)	Sum of Param (s)
5n	640	1.9
YOLOv5s	640	7.2
YOLOv5m	640	21.2
YOLOv5l	640	46.5
YOLOv5x	640	86.7
YOLOv5n6	1280	3.2
YOLOv5s6	1280	12.6
YOLOv5m6	1280	35.7
YOLOv5l6	1280	76.8
YOLOv5x6	1280	140.7

3.4.2 | YOLOv7

Developments in YOLOv7 focus more on module optimization by improving the accuracy of object detection in real time without improving computation. This YOLOv7 optimization result can reduce parameters by 40% and 50% in compute time. Optimization on YOLOv7 is called the trainable bag-of-freebies method.

Fig. 4 represents the architecture of YOLOv7. In this architecture, the backbone architecture is optimized using the E-ELAN (Extended Efficient Layer Aggregation Network). In this E-ELAN architecture, the model can continue to improve the network's learning ability without destroying the existing gradient path. Therefore, it can improve accuracy without worrying about a
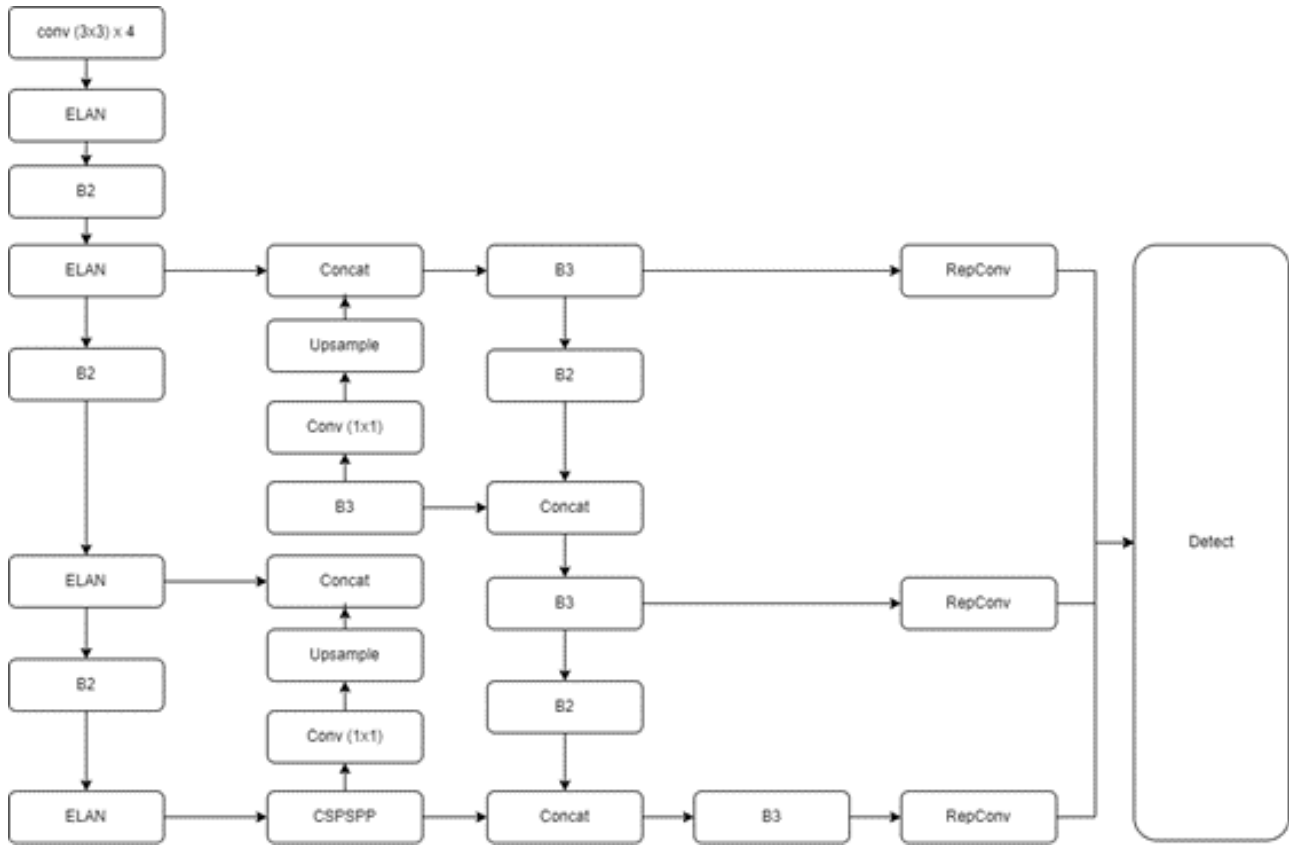
FIGURE 4 The YOLOv7 architecture for determining location and severity of lung cancer.

TABLE 4 The differences between various YOLO7 versions.

Version	Input Size (pixel)	Sum of Param (s)
YOLOv7	640	36.9
YOLOv7x	640	96.9
YOLOv7-W6	1280	70.4
YOLOv7-E6	1280	97.2
YOLOv7-D6	1280	154.7
YOLOv7-E6E	1280	151.7

gradient exploding. In addition, YOLOv7 also introduces compound model scaling. It can maintain model properties to remain optimal in the data training process[20].

Yolov7 can also be distinguished based on the version. Table 4 provides information on the differences between each version of YOLOv7. Each version has a different input size and a different number of parameters. For smaller input sizes (640), YOLOv7 can be divided into YOLOv7 and YOLOv7x. The difference between the two versions is in the number of parameters. YOLOv7, there are 36.9 million parameters, and for YOLOv7x, there are 96.9 million parameters. While for the version with an input image of 1280, there are four different versions, namely YOLOv7-W6, YOLOv7-E6, YOLOv7-D6, and YOLOv7-E6E. Each version also has a different number of parameters.

3.4.3 | YOLOv8

YOLOv8 is the latest YOLO release that focuses on improving enhanced performance. It has applied several methods in AI, such as detection, segmentation, pose estimation, tracking, and classification, So It can be applied to various projects[21].
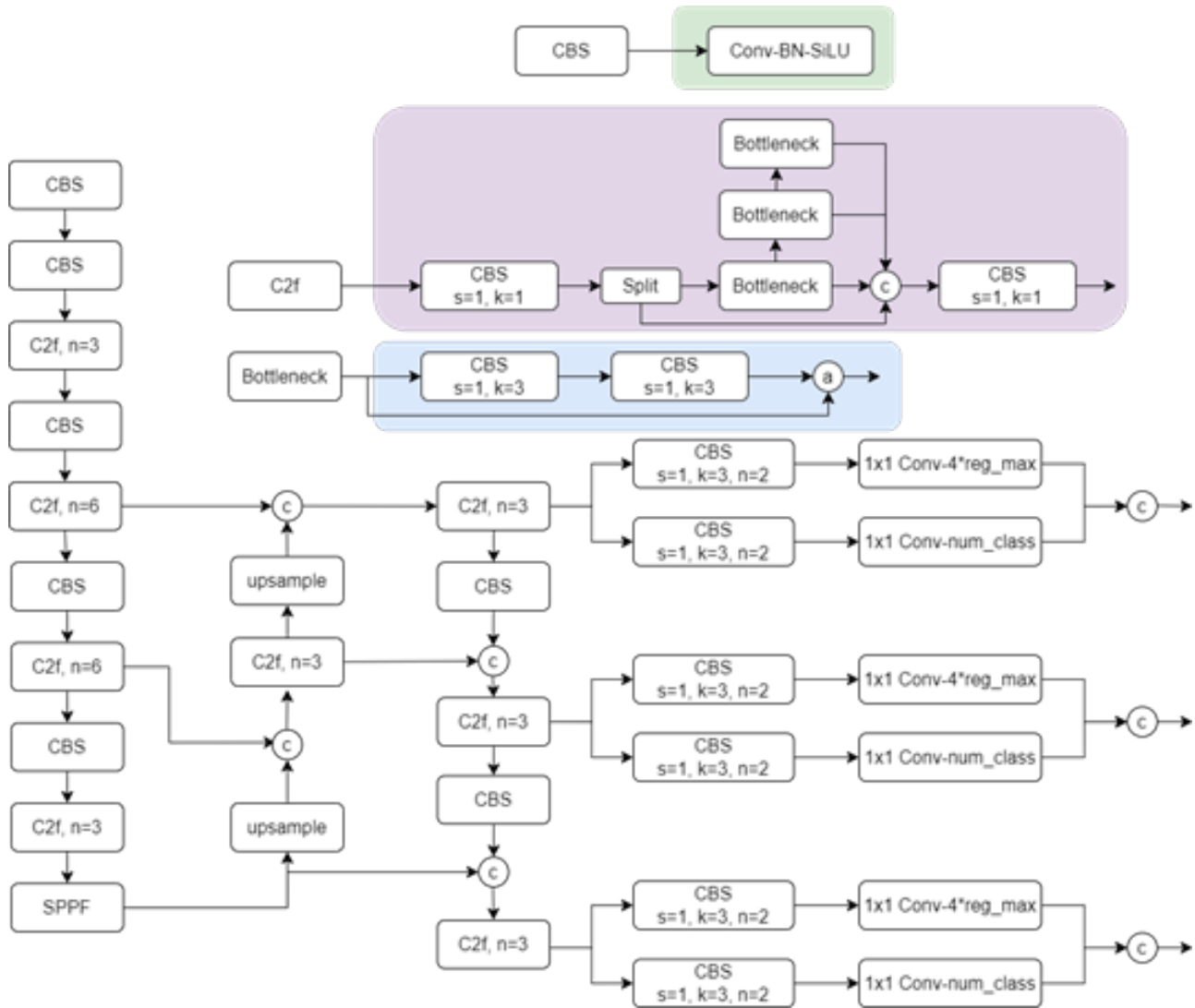
FIGURE 5 The YOLOv8 architecture for determining location and severity of lung cancer.

TABLE 5 The differences between various YOLO8 versions.

Version	Input Size (pixel)	Sum of Param (s)
YOLOv8n	640	3.2
YOLOv8s	640	11.2
YOLOv8m	640	25.9
YOLOv8l	640	43.7
YOLOv8x	640	68.2

Fig. 5 is an architectural drawing of YOLOv8. In the YOLOv8 architecture, there is a change from the CSP layer to the C2f layer. This C2f layer is lighter than the CSP layer because of the Bottleneck application, which increases gradient flow information with a lightweight. In addition, there is an improvement in PAN-FPN compared to YOLOv5 and YOLOv6. In this process, the C3 module and RepBlock are replaced with C2f, and the conv layer is removed before the upslung in the backbone. By changing this layer, the extracted feature goes directly into the upsampling process[22].

In Table 5 , there are several versions of YOLOv8: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. Each version has the same input image, 640, but different parameters.

Actual Class

		1	0
Predicted Class	1	True Positive	False Positive
	0	False Negative	True Negative

FIGURE 6 The confusion matrix.

Based on several YOLO architectures with different versions, this study used YOLOv5: YOLOv5s, YOLOv5m, and YOLOv5l; YOLOv7; YOLOv8: YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. Each version is not tested in its entirety because it considers the size of the input data and the number of parameters used. The size of the input data sought is close to the size of the dataset image, which is 640. Meanwhile, based on the number of parameters, we look for at most 80 million parameters. This is because the greater the number of parameters, the higher the computational level.

3.5 | Testing

3.5.1 | Confusion Matrix

A confusion matrix is a matrix that describes the performance of a classification model using testing data. In the confusion matrix, there are four variables, namely True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). True Positive describes the original data as having a label of 1, and the predicted result is 1. False Positive describes the original data as having a label of 0, and the predicted result is 1. True Negative describes the original data as having a label of 0, and the predicted result is 0. False Negative describes the original data as having a label of 0, and the predicted result is 1 [23]. An illustration of the confusion matrix can be seen in Fig. 6 .

After the matrix is formed, several performance calculations can be performed, such as accuracy, recall, precision, F1 score, etc. Accuracy calculations are used to calculate the efficiency of the method. The accuracy calculation is calculated by dividing the total number of correct prediction results by the amount of data [25]. The calculation of the accuracy equation as in equation 2. The recall calculation is calculated by dividing the TN value by the sum of the TN and FP values (see equation 3). Meanwhile, precision is determined by dividing the TN value by the total TN and FN values (see equation 4) [24].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Recall = \frac{TN}{TN + FP} \tag{3}$$

$$Precision = \frac{TN}{TN + FN} \tag{4}$$

3.5.2 | Mean Average Precision

Mean Average Precision (mAP) is a calculation matrix commonly used to evaluate object detection models. The map is calculated using the Average Precision (AP) values (see equation 5).

$$mAP = \frac{1}{n} \sum_{i=0}^{n} AP_i, \qquad n \ is \ sum \ of \ class \tag{5}$$
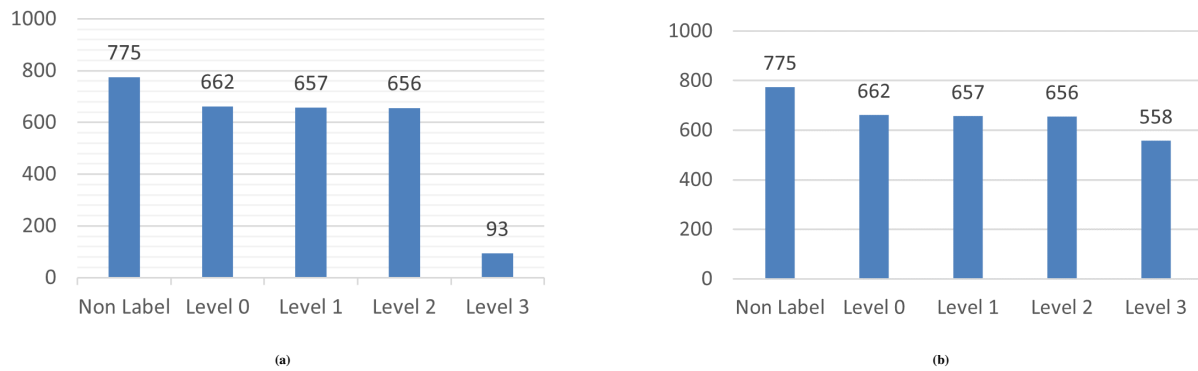
FIGURE 7 The data distribution (a) before augmentation and (b) after augmentation.

TABLE 6 The number of Splitting Data.

Image Type	Num of Nodule Images
Training Data	1798
Validation Data	506
Testing Data	247

The Average Precision (AP) calculation is calculated using Weighted Average Precision. Calculations using Weighted Average Precision can consider the precision and recall values equally. The AP calculation is obtained by calculating the formula for average precision by multiplying each recall by the realized recall gain.

4 | RESULTS AND DISCUSSION

This study uses data from 439 patients with 4643 images labeled according to nodule size. The results of data labeling found that Level One, which has a size between 6-10 mm, consists of 775 images; Level Two, which has a size between 10-20 mm, consists of 662 images; Level 3, which has a nodule size between 20-30 mm consists of 656 images, Level 4 which has a size of more than 30 mm consists of 93 images, and 2457 images did not go into all levels because the size of the nodules in the images was less than 6 mm. If the nodule is less than 6 mm, it has a percentage of becoming cancerous of 0-1%, and no follow-up is done shortly because it is still considered a harmless nodule[14]. The graph of the amount of data is in Fig. 7 a.

Data is smaller at level 4 than at levels 1 and 2. Therefore, at levels 3 and 4, augmentation will be carried out to multiply the data. The augmentation process is carried out by rotating the image by degrees 15, 30, 180, 330, and 345—the augmentation process increases data at level 4 with 558 images. The graph of the amount of data is in Fig. 7 b.

After that, the data was divided into 70% training data, 20% validation data, and 10% testing data. The data sharing resulted in training data totaling 1798 images, validation data 506 images, and testing data 247 images. The amount of data distribution is in Table 6 .

The device used for research has an Intel i7-8700 CPU with a 6 GB NVIDIA GeForce GTX 1060 GPU. For the YOLOv5 architecture, use YOLOv5s, YOLOv5m, and YOLOv5l versions. for YOLOv7 only use YOLOv7 version only. As for YOLOv8, it use YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x versions. In the training process, the data use epoch = 200, batch size = 8, and image size = 512. The training process results are tested again and calculated using mAP calculations with IoU threshold = 0.5.

4.1 | Experimentation of YOLOv5

In the experiment using YOLOv5, we used three versions, namely YOLOv5s, YOLOv5m, and YOLOv5l. The results of each version will be broken down based on data levels.
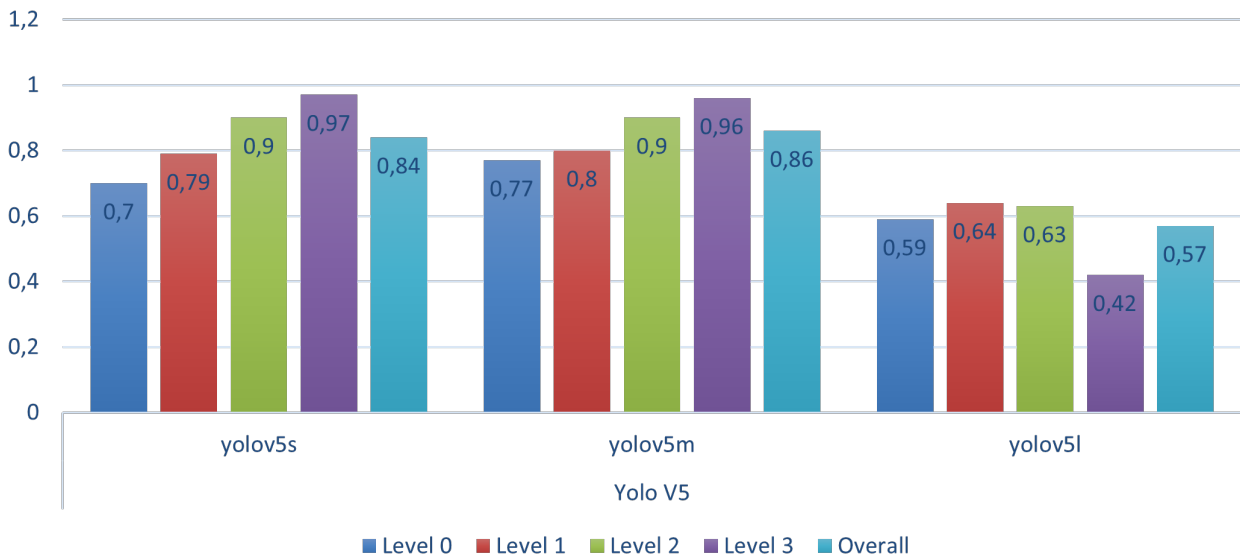
FIGURE 8 The experimental results using YOLOv5 versions, i.e. YOLOv5s, YOLOv5m, and YOLOv5l.
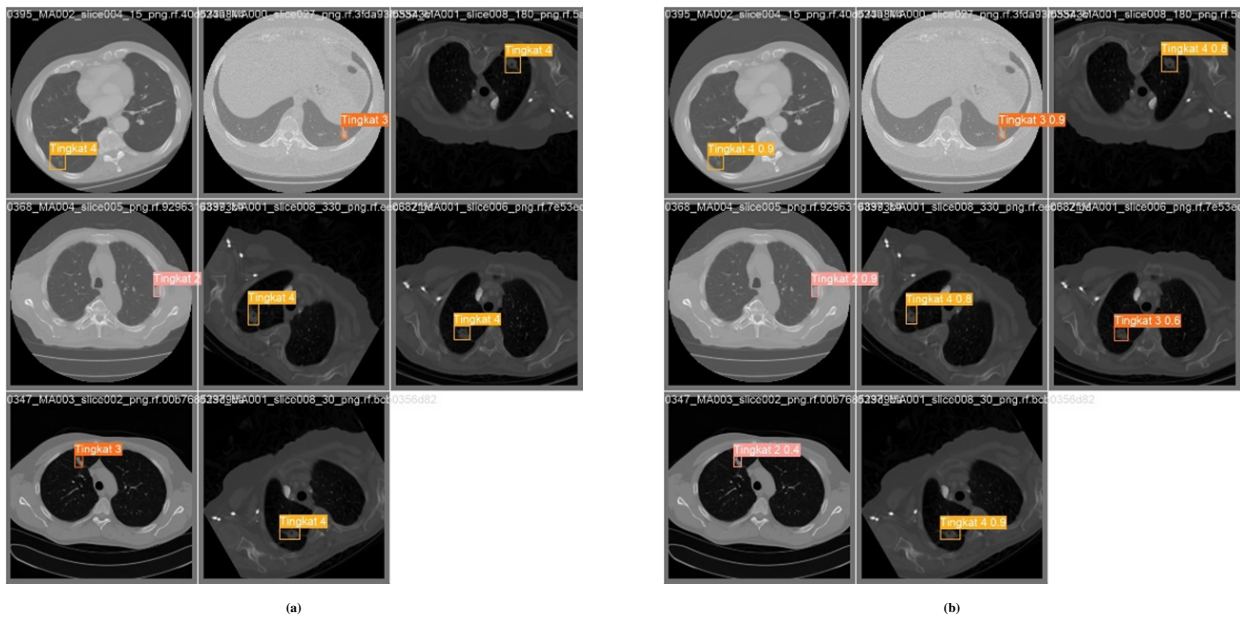

FIGURE 9 The prediction score (a) before augmentation and (b) after augmentation.

Based on Fig. 9 , Level 1 averages 0.68, and YOLOv5m has the best results. At Level 2, it averages 0.73, with YOLOv5m having the most favorable results. For Level 3, it has an average of 0.81, with YOLOv5 and YOLOv5m having the best results. Lastly, Level 4 averages 0.78, with YOLOv5s having the most excellent results. Based on the following results, level 2 data has worse results than other levels. In contrast, Level 3 data has the best average compared to the others. YOLOvm obtained the best YOLOv5 model with an average MAP value of 0.86.

The prediction results using YOLOv5 are shown in Fig. 9 . Based on the picture, some nodules are still wrong predictions and have a small confidence value.
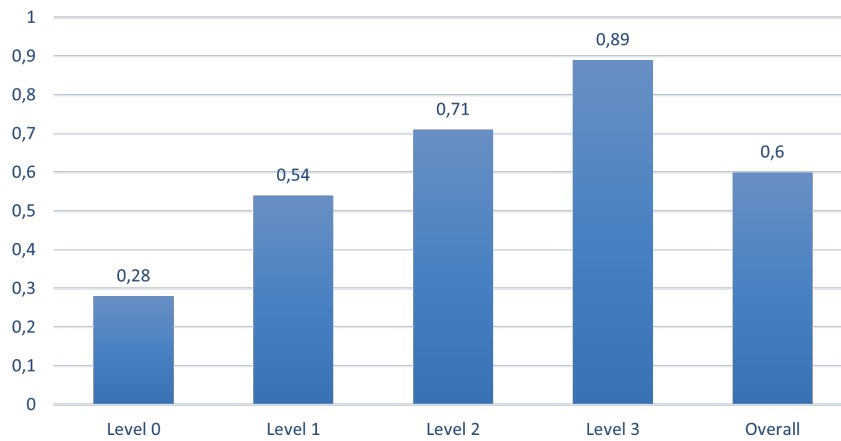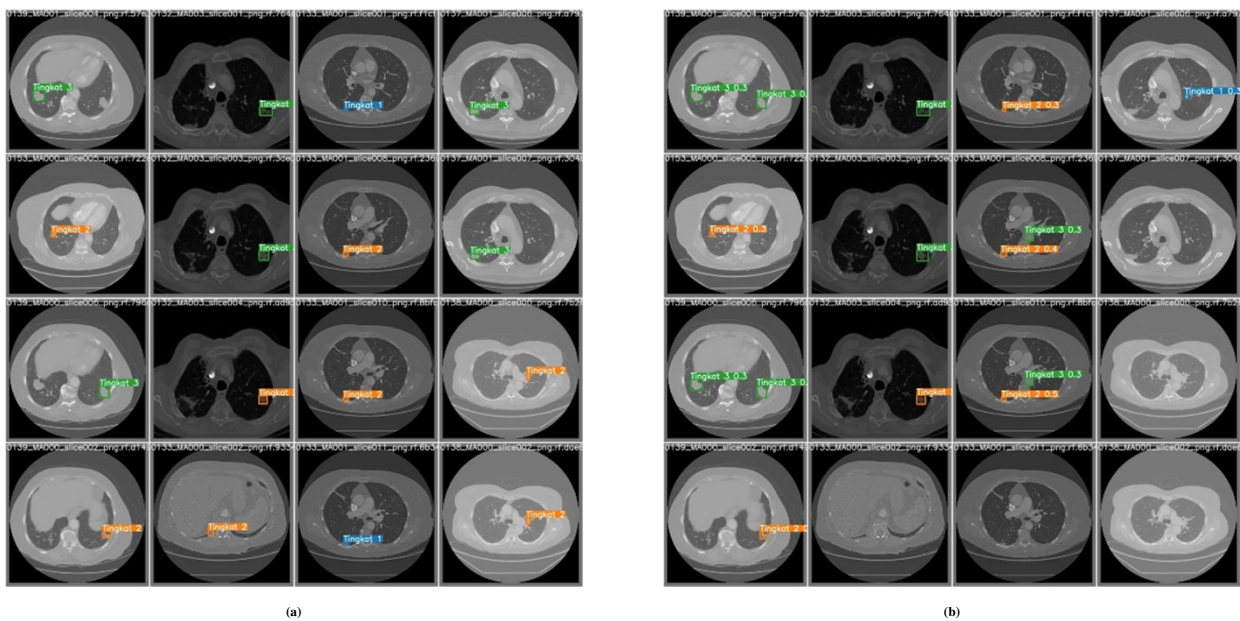
FIGURE 10 The experimental results using YOLOv7.

FIGURE 11 The prediction score (a) before augmentation and (b) after augmentation.

4.2 | Experimentation of YOLOv7

We only used one version of YOLOv7 in the experimentation. YOLOv7 results can be broken down by levels.

Based on Fig. 10 , Level 1 has the smallest MAP value with 0.28. Level 2 has a MAP value of 0.54. Level 3 has a MAP value of 0.7; finally, Level 4, which has the highest MAP value, gets a value of 0.89. If the results of YOLOv7 are compared to the previous YOLOv5, then there is a decrease in performance based on the size of the MAP. At each Level, YOLOv7 decreased by 0.4 at Level 1, 0.19 at Level 2, 0.11 at Level 3, and 0.11 at Level 4.

Fig. 11 is an example of prediction results using YOLOv7. It can also be seen from the image that several nodules cannot be detected. Not only that but there are also some predictions. So, using YOLOv7 is still not optimal with the data we use.
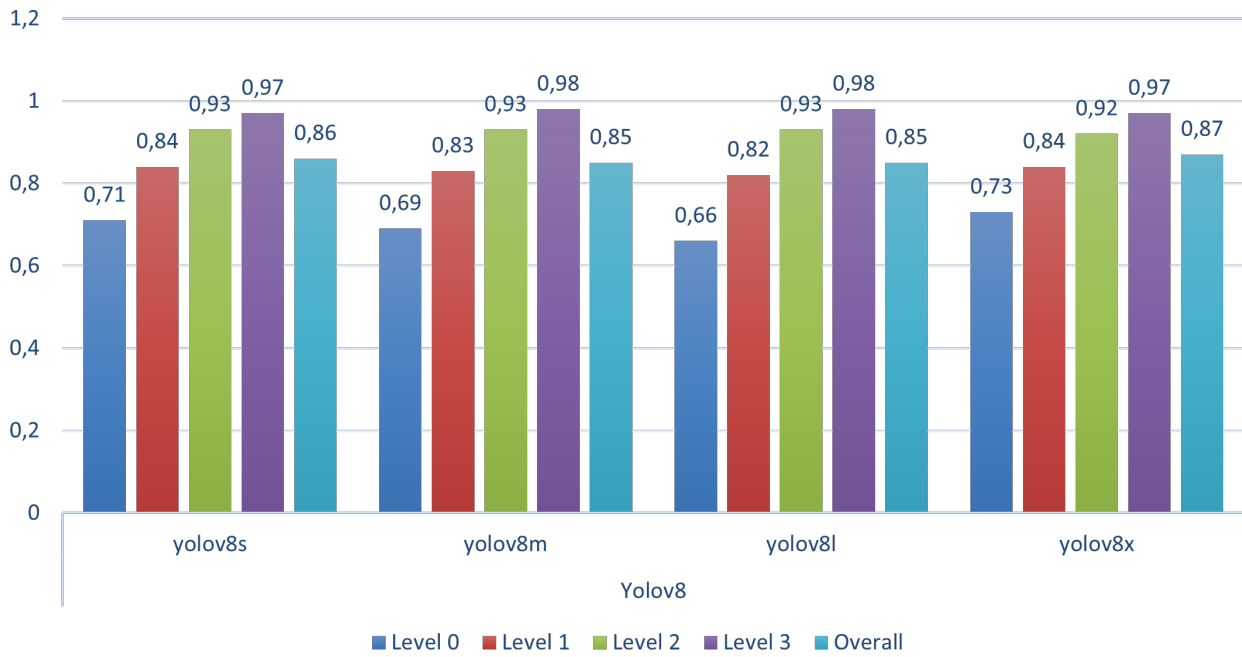
FIGURE 12 The experimental results using YOLOv8 versions, i.e. YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x.

4.3 | Experimentation of YOLOv8

In the experiment using YOLOv8, we used four versions, namely YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. The results of each version will be broken down based on data levels.

Based on the results from Fig. 12 , it can be seen that the smallest MAP value is recovered by Level 1 with an average of 0.69, with YOLOv8x having the largest MAP value at Level 1. Then, at Level 2, get a MAP value of 0.83 with YOLOv8s, and YOLOv8x has the largest MAP value at Level 2. At Level 3, get a MAP value of 0.92 with YOLOv8m, and YOLOv8l is the version that has the best MAP value at this Level. Finally, Level 4 has an average MAP value of 0.97, and YOLOv8m and YOLOv8l are the versions that have the best MAP values at Level 4. If the results of YOLOv8 are compared with YOLOv5, there is an increase in performance results when viewed from the average MAP at each level. Increases at each Level are 0.01 at Level 1, 0.1 at Level 2, 0.11 at Level 3, and 0.21 at Level 4.

Fig. 13 is an example of the results of predictions made with YOLOv8. It can be seen that there are still some pictures that have wrong predictions.

Based on Fig. 14 , which is the best result graph of each architecture, it can be concluded that YOLOv8 is the most optimal YOLO between YOLOv8, YOLOv7, and YOLOv5. This can be seen from the MAP value on YOLOv8x, which has the largest average MAP value in other experiments.

Fig. 15 is an accuracy matrix based on predicted results from the testing data. Level 1 is the most difficult to predict, and 0.18 prediction becomes background. This is because the image is small and difficult to distinguish from the lung organs that are small in size as well. Compared to Level 4, which has the best correct prediction results than the other Levels. This is because nodules in Grade 4 are visible and large, so they can easily distinguish nodules from organs in the lungs.
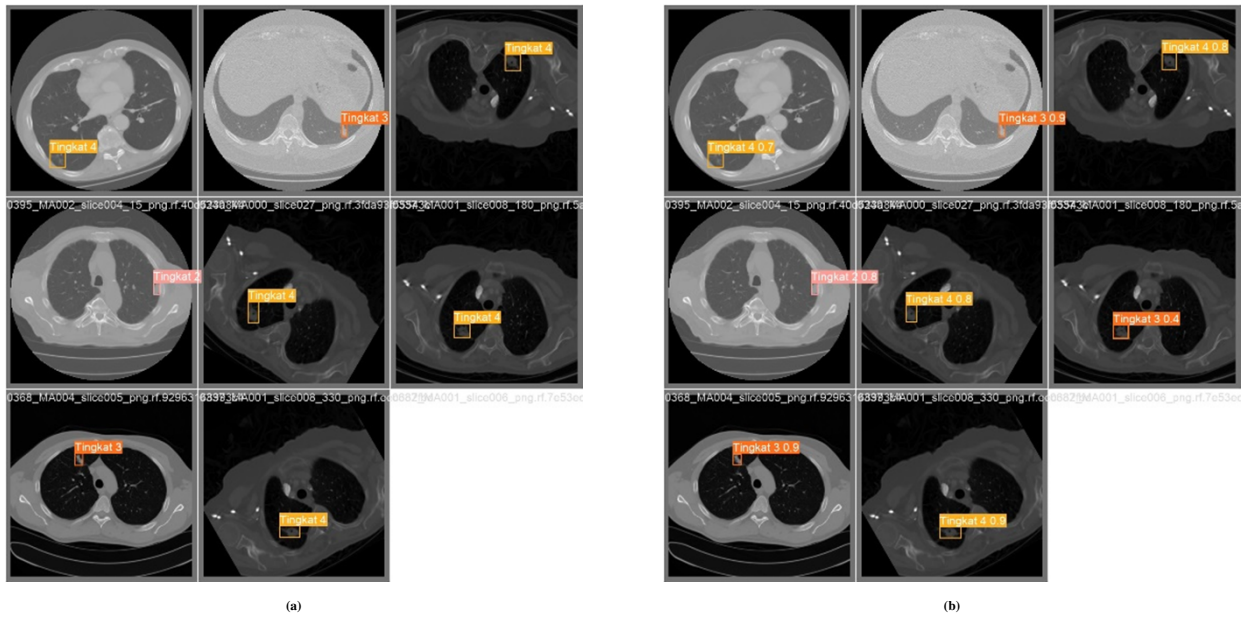
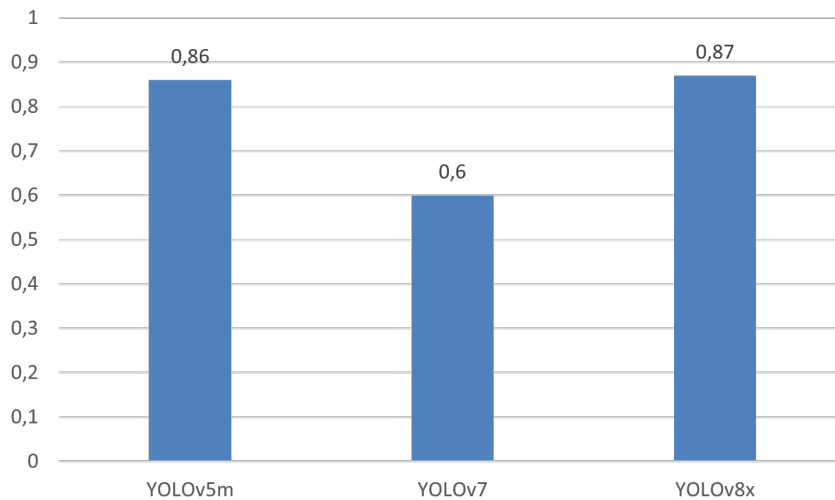**FIGURE 13** The MAP value (a) before augmentation and (b) after augmentation.

FIGURE 14 The comparison results of all best YOLO versions.

predict		Level 1	Level 2	Level 3	Level 4	Background
	Level 1	0.79	0.06	0	0	0.2
	Level 2	0.02	0.81	0.07	0	0.25
	Level 3	0	0.05	0.87	0.02	0.26
	Level 4	0	0	0.04	0.98	0.09
	Background	0.18	0.09	0.03	0	0
		Level 1	Level 2	Level 3	Level 4	Background
				True		

FIGURE 15 The accuracy matrix.

5 | CONCLUSION

This study detects lung cancer nodules' location and severity with several YOLO models, such as YOLOv5, YOLOv7, and YOLOv8. To detect the severity level, research also builds a new classification dataset based on the severity level. The severity level is obtained from the nodule size classification into four different levels.

Datasets are created based on the severity level seen from the nodule's size. The dataset is divided into four classes, namely Level 1, Level 2, Level 3, and Level 4. Level 1 has 775 images, Level 2 has 662 images, Level 3 has 656 images, and Level 4 has 93 images. The dataset is used for the training process of detecting location and severity using YOLO. In the detection training process, three different methods were used, namely YOLOv5, YOLOv7, and YOLOv8. Each YOLO model result was tested again with 247 images using mAP 0.5. The best results from the area detection and severity detection models are combined to produce an area and severity detection system. The YOLOv8 model produces the best results for detecting severity, with mAP values ranging from 0.5 to 0.87.

Based on the experiments, the class labels that were most difficult to detect were data labeled Level 1 or images with nodule sizes between 8 mm and 10 mm. The data on this label is the data with the smallest size, so it is more difficult to detect than images at other levels. For further research, optimization of parameters in architecture can also be done to improve the model. Optimization can be done by changing the learning rate parameter to be able to detail the data learning. Apart from that, optimization can also be done by modifying the anchor, which functions as the expected parameter for the size of the detected object's bounding box.

CREDIT

Hanun Masitha Ramadhani: Conceptualization, Methodology, Validation, Resources, Data Curation, Writing - Original Draft, and Visualization. **Chastine Fatichah:** Methodology, Validation, Writing - Review & Editing, and Supervision.

References

1. Joseph J, Rotty LWA. Kanker Paru: Laporan Kasus. Medical Scope Journal 2020;2.

2. Cao W, Chen HD, Yu YW, Li N, Chen WQ. Changing profiles of cancer burden worldwide and in China: A secondary analysis of the global cancer statistics 2020. Chinese Medical Journal 2021;134.

3. Ramadhaniah F, Syarif S. Studi Tinjauan Pustaka: Risiko Kejadian Kanker Paru pada Penderita Tuberkulosis Paru. Jurnal Epidemiologi Kesehatan Indonesia 2020;4.

4. Gregson CL, Armstrong DJ, Bowden J, Cooper C, Edwards J, Gittoes NJL, et al. UK clinical guideline for the prevention and treatment of osteoporosis. Archives of Osteoporosis 2022;17.

5. Larici AR, Farchione A, Franchi P, Ciliberto M, Cicchetti G, Calandriello L, et al. Lung nodules: Size still matters. European Respiratory Review 2017;26.

6. Vaghasiya K, Sharma A, Verma RK. Misdiagnosis Murder: Disguised TB or Lung Cancer? Pulmonary Research and Respiratory Medicine - Open Journal 2016 9;3:e5–e6.

7. Ciello Ad, Franchi P, Contegiacomo A, Cicchetti G, Bonomo L, Larici AR. Missed lung cancer: When, where, and why? Diagnostic and Interventional Radiology 2017;23.

8. Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, et al. Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation. IEEE Transactions on Medical Imaging 2016;35(5):1170–1181.

9. Liu C, Hu SC, Wang C, Lafata K, Yin FF. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. Quantitative Imaging in Medicine and Surgery 2020;10.

10. Wang J, Wang J, Wen Y, Lu H, Niu T, Pan J, et al. Pulmonary Nodule Detection in Volumetric Chest CT Scans Using CNNs-Based Nodule-Size-Adaptive Detection and Classification. IEEE Access 2019;7.

11. Wang Z, Liu H, Zhang G, Yang X, Wen L, Zhao W. Diseased Fish Detection in the Underwater Environment Using an Improved YOLOV5 Network for Intensive Aquaculture. Fishes 2023;8.

12. Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans.

Medical Physics 2011;38.

13. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: A prespecified analysis of data from the NELSON trial of low-dose CT screening. The Lancet Oncology 2014;15.

14. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner Society 2017. Radiology 2017;284.

15. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: Fast and flexible image augmentations. Information (Switzerland) 2020;11.

16. Geraldy C, Lubis C. Detecting and Identifying Vehiche Type Using YOLO and CNN (Pendeteksian dan pengenalan jenis mobil menggunakan algoritma You Only Look Once dan Convolutional Neural Network. Jurnal Ilmu Komputer dan Sistem Informasi 2020;8.

17. Jupiyandi S, Saniputra FR, Pratama Y, Dharmawan MR, Cholissodin I. Development of vehicle detection for counting utility in parking space using CUDA and Modified-YOLO (Pengembangan Deteksi Citra Mobil untuk Mengetahui Jumlah Tempat Parkir Menggunakan CUDA dan Modified YOLO). Jurnal Teknologi Informasi dan Ilmu Komputer 2019;6.

18. Jiang P, Ergu D, Liu F, Cai Y, Ma B. A Review of Yolo Algorithm Developments. In: Procedia Computer Science, vol. 199; 2021. p. 733–736.

19. Chen S, Duan J, Wang H, Wang R, Li J, Qi M, et al. Automatic detection of stroke lesion from diffusion-weighted imaging via the improved YOLOv5. Computers in Biology and Medicine 2022;150.

20. Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2023. p. 7464–7475.

21. Qadri SAA, Huang NF, Wani TM, Bhat SA. Plant Disease Detection and Segmentation using End-to-End YOLOv8: A Comprehensive Approach. In: 2023 IEEE 13th International Conference on Control System, Computing and Engineering (ICCSCE); 2023. p. 155–160.

22. Wang X, Gao H, Jia Z, Li Z. BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8. Sensors (Basel, Switzerland) 2023;23.

23. Mishra S, Dash A, Jena L. 10. In: Bhoi AK, Mallick PK, Liu CM, Balas VE, editors. Use of Deep Learning for Disease Detection and Diagnosis Singapore: Springer Singapore; 2021. p. 181–201.

24. Rachh R, Allagi S, Shravan BK. 11. In: N P, Kautish S, Peng SL, editors. Chapter 11 - Machine learning algorithms for prediction of heart disease Academic Press; 2021. p. 247–275. https://www.sciencedirect.com/science/article/pii/B9780128216330000106.