Identify Tumors on Lung CT Images

Phong Thanh Le[1,2], Thai Hoang Le[1,2(✉)], and Hieu Duc Thai Tran[1,2]

[1] Faculty of Information Technology, University of Science, Ho Chi Minh City, Vietnam
lhthai@fit.hcmus.edu.vn
[2] Vietnam National University, Ho Chi Minh City, Vietnam

Abstract. This paper introduces You Only Look Once (YOLO) model to identify tumors on computed tomography (CT) lung images. The model uses a variant of the YOLO algorithm called YOLOv5 [1], which is known for its accuracy and speed in object detection tasks. To train and evaluate the YOLO model, we use the Lung Nodule Analysis 2016 dataset (LUNA16) [2]. This dataset contains a set of lung CT scans with annotations indicating the location of the tumors. We preprocess the CT images and annotations to prepare data for model training and testing. During the training phase, the YOLO model uses a loss function named Generalized Intersection over Union (GIoU) loss [3], which provides a more accurate measure of box overlap between the predicted objects and the ground truth. The combination of the YOLO architecture and the GIoU loss enables accurate and fast detection, making the proposed model a promising tool to aid physicians in diagnosing lung cancer.

Keywords: You Only Look Once · Tumor identification · Lung Computed Tomography (CT) images · Lung Nodule Analysis 2016 (LUNA16) dataset · Generalized Intersection over Union (GIoU) loss · Diagnosis Lung cancer

1 Introduction

Lung cancer remains a major global health concern as it continues to be the most common form of cancer and is associated with the highest number of cancer-related deaths worldwide [4]. Finding and diagnosing lung cancer at an early stage is critical to improving patient outcomes and improving the effectiveness of treatment options. To assist doctors in the diagnostic process, CAD (Computer Aided Detection) systems [5, 6] have been developed. These systems use advanced image processing algorithms and techniques to assist radiologists and other medical professionals in the analysis and interpretation of medical images, especially those Images are obtained through techniques such as Computed Tomography (CT). The CAD system for lung cancer typically consists of two main stages. The main stage is to detect candidate nodules, which are potential tumor regions within the lung. This phase involves applying complex algorithms to segment lung images and identify suspicious areas that may be indicative of nodules. Once candidate nodules are identified, the CAD system proceeds to classify them as

P. Cong Vinh and N. Thanh Tung (Eds.): ICCASA 2023, LNICST 579, pp. 161–174, 2024.
https://doi.org/10.1007/978-3-031-58878-5_12

positive, indicating the presence of a tumor. This classification stage is crucial in distinguishing between cancerous and noncancerous nodules. The CAD system not only detects and classifies nodules, but can also provide additional information about their characteristics. For example, it can help determine the size, location, shape irregularities, and other characteristics of identified nodules. This information is valuable in assisting physicians with treatment planning and decision making. Furthermore, CAD systems can assist in monitoring the growth or change of nodules over time, providing valuable insights into disease progression and treatment response. Machine learning techniques, especially deep learning algorithms, have revolutionized the analysis of medical images, including the detection and classification of lung nodules for the diagnosis of lung cancer. These algorithms are ideally suited for tasks involving pattern recognition in large datasets, making them highly effective in medical imaging applications. By leveraging the power of CAD systems and machine learning algorithms, medical professionals can improve the detection and diagnosis of lung cancer in its early stages, leading to more effective treatment strategies. And better outcomes for patients. As research and technology continue to advance, CAD systems are expected to play an increasingly important role in the fight against lung cancer and other medical conditions. In the field of tumor detection and identification on lung CT images, there are numerous methods available, including Computer-Aided Diagnosis (CAD) systems. These systems leverage advanced algorithms to assist radiologists and medical professionals in the accurate and efficient detection of tumors and abnormalities in medical images. For the specific task at hand, we have chosen to implement a new deep learning model as part of our CAD system. Deep learning has shown significant promise in various medical imaging tasks, including tumor detection, due to its ability to automatically learn and extract relevant features from the data. In the next section of our paper, we plan to provide a detailed description of our new deep learning model. The rest of the paper is organized as follows. Section 2 discusses an overview of the related works in Sect. 2. A description of our approach for is presented in Sect. 2. Section 3 details the experimentation carried out on. Finally, conclusions are given in Sect. 4

2 Related Work

2.1 Object-Detection [7] Algorithms

Deep learning models have demonstrated outstanding potential in detecting lung tumors, offering promising avenues for advancing medical imaging analysis. However, to understand their true performance in clinical settings, insights from medical professionals and researchers are indispensable. Medical professionals possess invaluable domain knowledge and practical experience in diagnosing lung tumors from CT images, making their perspectives important in evaluating the effectiveness of deep learning models. In this context, challenges in tumor identification arise in specific situations, such as distinguishing between benign and malignant tumors or detecting nodules in complex cases. Each method can excel under different conditions, requiring evaluation based on specific use cases and datasets. Swetha Subramanian [8] proposed a 50 x 50 pixel image-based tumor identification procedure, focusing on tumor identification using small arrays of images. This approach is capable of using a Convolutional Neural Network (CNN) architecture

specifically designed for image analysis tasks. On the other hand, Ding et al. [9] proposed a 3-D CNN-based lung nodule detection system, using the Pulmonary Nodule Analysis Challenge (LUNA16) dataset, containing CT scans with annotations for lung nodules (potential tumors). It is important to note that these models target different aspects of tumor identification, with Swetha Subramanian's model focusing on smaller arrays of images, while Ding et al. using 3-D CNN to analyze the entire volumetric CT image. To further compare the performance of the object detection models, we consider the Mean Accuracy (mAP) (Fig. 1). Faster algorithms like Single Shot Multibox Detector (SSD) may struggle with the complexity of the tumor identification problem. On the other hand, Faster R-CNN [10] prioritizes accuracy over speed and can achieve six times better performance than SSDs with the same feature extractor. Even though YOLO v5 exhibits improved performance over SSDs and Fully Area-Based Coherent Network (RFCN), Faster R-CNN maintains a small advantage of 0.1 mAP at 4,000 training steps. Given the diverse strengths and challenges associated with different deep learning models for tumor recognition, the decision was made to use YOLO v5x as the official technique for testing and evaluation to demonstrate its effectiveness. Superior performance compared to other object detection models such as SSD and RFCN. By using YOLO v5x, we aim to take advantage of its ability to efficiently detect lung tumors from CT images. We recognize the importance of accurate tumor identification in the clinical setting, especially when distinguishing between benign and malignant tumors or dealing with complex cases. The perspective and field knowledge of medical professionals informed our decision to select a model that could excel at detecting tumors from small arrays of images, which could be particularly beneficial in identifying subtle abnormalities and improving early detection rates.

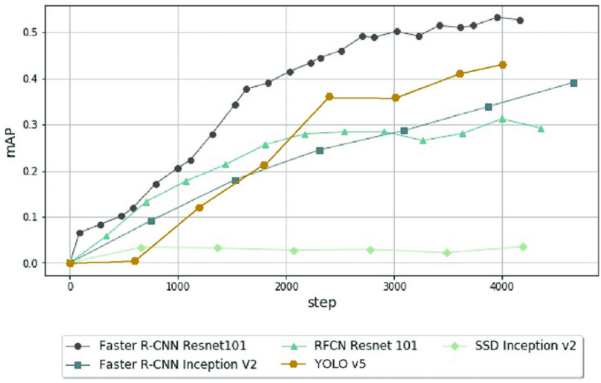

Fig. 1. Performance comparison using the mAP of the object-detection models

2.2 YOLO v5x

2.2.1 Architecture

YOLO v5x (You Only Look Once version 5) is an advanced object detection algorithm that builds on the success of its predecessors, most notably YOLOv3. It represents a significant improvement in both speed and accuracy. The key innovation of YOLO v5x lies in its single-stage approach, which allows it to predict bounding boxes and class probabilities for objects directly in a single neural network pass. This design not only reduces computational complexity, but also enables real-time object detection with impressive accuracy, making it well-suited to a wide range of applications, including mission critical is to detect lung nodules in medical imaging. The architecture of YOLO v5x consists of three essential components (Fig. 2), each serving a specific purpose.

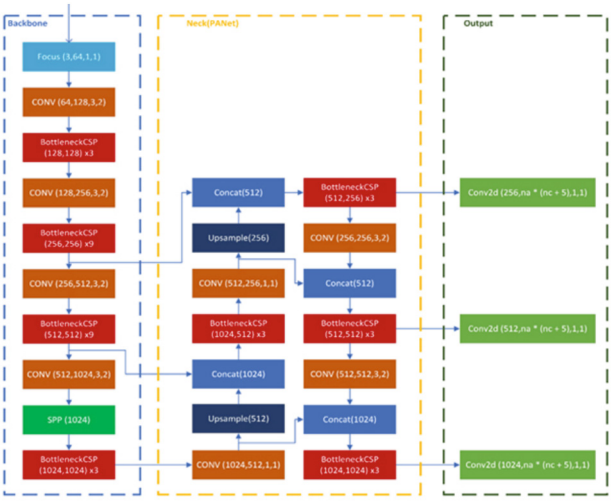

Fig. 2. Architecture of YOLO v5x

Backbone model is a fundamental part of YOLO v5x, responsible for extracting rich and informative features from input image. This process involves passing the image through several convolutional layers, capturing hierarchical features of increasing complexity. By learning relevant patterns and characteristics from images, Backbone models provide the foundation for accurate object detection. The Neck model is an intermediate component responsible for generating the characteristic pyramids. Feature pyramids are multi-scale representations of images that help identify objects of different sizes and scales in the input. This aspect is especially important in medical imaging, as lung nodules can vary considerably in size and shape. The Neck model's ability to generate pyramidal pyramids enhances YOLO v5x's ability to efficiently detect lung nodules of different sizes. The Head model is the final component of YOLO v5x and is responsible for making the final output predictions. It takes the features extracted by the Backbone model and further refines them based on the feature pyramids generated by the Neck model. The Head model then predicts bounding boxes consisting of detected objects,

along with corresponding class probabilities for each feature type. In the context of lung nodule detection, the Head model's predictions pinpoint the location and type of nodule in the CT scan. By integrating information from the Backbone and Neck models, YOLO v5x can efficiently and accurately detect objects, including lung nodules, of various sizes and scales within an input CT image. The use of feature pyramids generated by the Neck model contributes to the algorithm's robust object identification capabilities. Moreover, YOLO v5x leverages the concept of anchor boxes, combined with class probabilities, in the Head model to further enhance the precision and reliability of the final predictions. These anchor boxes act as reference templates for the algorithm to detect and localize objects more accurately. The end-to-end design of YOLO v5x enables both speed and accuracy in tumor detection, making it an incredibly promising choice for medical image analysis tasks, particularly tumor identification in lung CT images. Its real-time capabilities facilitate rapid analysis of large volumes of medical data, enabling timely diagnosis and treatment planning, which is crucial for ensuring the best possible patient outcomes.

2.2.2 Advantages

YOLO v5x offers several advantages that make it a powerful and efficient object detection algorithm, particularly for applications like tumor identification in lung CT images. YOLO v5x is a remarkable advancement in object detection algorithms, building upon the success of its predecessors and introducing several improvements to enhance its overall performance. One of its key strengths is its ability to achieve higher accuracy and better detection results compared to previous versions, such as YOLOv3. This enhanced performance is particularly critical for accurate tumor identification in lung CT images and can significantly impact the effectiveness of lung cancer diagnosis. One of the standout features of YOLO v5x is its optimization for speed, enabling real-time object detection even on resource-constrained devices. Its lightweight architecture allows for fast inference without compromising accuracy, making it well-suited for applications that require quick responses. In the context of lung CT image analysis, real-time inference can expedite the tumor identification process, leading to faster diagnosis and more prompt treatment planning. Furthermore, YOLO v5x's high level of customizability empowers researchers and developers to fine-tune the model to suit specific use cases. They can modify the architecture, backbone, and other components to adapt the algorithm to different tasks and datasets. This flexibility is especially advantageous in medical imaging, where customizing the algorithm to the unique characteristics of lung CT images can lead to improved performance and better tumor identification. Another crucial advantage of YOLO v5x is its ability to handle images of varying resolutions and aspect ratios without requiring fixed input image sizes. This adaptability is particularly beneficial when dealing with medical images that may have different dimensions or varying pixel densities. The algorithm's capability to handle diverse image sizes enhances its applicability to lung CT images, which often come in various resolutions. In challenging and cluttered scenes, YOLO v5x excels in detecting objects, including tumors. It can effectively handle overlapping objects and accurately identify small or distant tumors, a crucial feature in medical imaging applications. This robustness significantly contributes to accurate tumor identification in diverse clinical scenarios. The utilization of anchor

boxes and efficient regression techniques in YOLO v5x results in more accurate and precise bounding box predictions around detected objects. This level of object localization is paramount in medical applications, where accurate tumor boundaries are vital for accurate diagnosis and effective treatment planning. Precise localization empowers medical professionals to confidently assess tumor characteristics for making well-informed treatment decisions. YOLO v5x follows a single-stage detection approach, eliminating the need for proposal generation and subsequent refinement stages. This simplicity reduces computational overhead and enables faster inference. In the context of lung tumor identification, this streamlined process allows for rapid analysis of CT scans, making YOLO v5x an excellent choice for time-sensitive medical applications. The fact that YOLO v5x is open-source and comes with pre-trained models and extensive documentation makes it accessible and user-friendly for both researchers and developers. Its straightforward interface simplifies the training and deployment process, facilitating seamless integration into existing medical imaging workflows.

2.3 Our Contribution

To demonstrate the effectiveness of YOLO v5x in the domain of tumor identification, we propose a novel model that leverages the advanced capabilities of YOLO v5x in object detection. Our primary objective is to develop a system that can accurately and efficiently detect tumors on lung CT images, contributing to early tumor detection and improved patient outcomes in the context of lung cancer diagnosis. The development of our tumor identification model involved a series of well-defined steps, illustrated in Fig. 3.

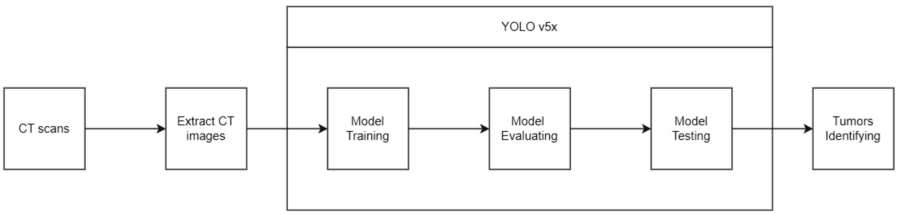

Fig. 3. Our model identifies tumors on lung CT images.

In Fig. 3, we began our tumor identification model by meticulously sorting and selecting an extensive dataset that included a diverse array of lung CT scans with annotated tumor regions. The quality, size, and diversity of the dataset are of prime importance in ensuring the robustness and generalizability of our model. To prepare the dataset for training, we used the necessary data preprocessing steps, including image resizing, normalization, and enhancement. These steps serve to normalize the input data, improve the performance of the model, and enhance the training data, thus increasing its diversity. Then, the custom YOLO v5x model underwent exhaustive training on the selected dataset, where we optimized its parameters to maximize the tumors identification accuracy. The training process involves iterative updates to the model based on the calculated

loss, which is mainly driven by the GIoU loss function. This loss function ensures precise bounding box positioning, allowing the model to accurately and accurately delineate tumor regions on lung CT images. An important aspect of YOLO v5x is its optimized speed for real-time object detection. To showcase the real-time inference capabilities of our model, we demonstrated the ability to rapidly process and identify tumors on lung CT images. This property makes our model well-suited to time-sensitive medical situations where rapid tumors identification is critical for prompt patient care. Furthermore, to ensure the model's adaptability and generalization to real-world data, we incorporated a fine-tuning and adaptability process. This allows our model to handle variations in the data set, such as different lung nodule types, sizes, and image quality, thus improving the model's performance in different cases. Different .

When using our model (Fig. 3) for tumor identification on lung CT images with the LUNA16 dataset, several significant contributions can be observed in the context of medical imaging and lung cancer diagnosis. Our model, with its improved performance and architecture, can achieve high accuracy in identifying lung tumors within the LUNA16 dataset. The algorithm's ability to handle various image sizes and complex scenarios in CT scans leads to more precise tumor localization and reduced false positives and false negatives. Our model optimized speed allows for real-time tumor detection, making it well-suited for analyzing large volumes of CT scans from the LUNA16 dataset efficiently. This real-time capability ensures rapid screening and diagnosis, which is crucial in medical settings for timely patient care. By accurately identifying small and subtle tumors in lung CT images from LUNA16, our model contributes to early detection. Early identification of lung tumors can lead to improved treatment outcomes and potentially save lung. Our model serves as a valuable tool to assist radiologists and healthcare professionals in the tumor identification process. Its accurate predictions and automated detection can speed up the screening process and provide valuable insights, allowing medical experts to focus on critical decision-making and patient care. Our model can be fine-tuned and customized to suit the specific characteristics of the LUNA16 dataset, such as the nature of lung nodules and variations in imaging quality. This adaptability ensures better generalization to the LUNA16 data, leading to improved performance on this specific dataset. Our model precise tumor localization in lung CT scans from LUNA16 provides crucial information for treatment planning. Accurate tumor boundaries aid in developing targeted treatment strategies, optimizing therapy delivery, and minimizing damage to healthy tissues. Integrating our model into clinical workflows with the LUNA16 dataset can streamline tumor identification processes. The algorithm's user-friendly interface and real-time inference allow for efficient analysis and diagnosis, saving time and resources for medical professionals. Our model's application to the LUNA16 dataset contributes to the advancement of research in lung cancer detection and medical image analysis. The insights gained from this study can lead to further improvements in lung nodule detection algorithms and drive innovations in the field. Our model's efficient and accurate tumor detection capabilities make it a valuable asset for large-scale lung cancer screening programs. By processing a significant number of CT scans, the algorithm can support population-based cancer screening initiatives, leading to early tumor detection and improved public health outcomes. The results of our experiments demonstrate the real-world efficacy of our model in tumor detection. During

the evaluation process, we utilized a carefully curated and diverse dataset, representative of the complexity and variety of lung nodules encountered in clinical practice.

3 Experiments Results

3.1 Datasets

Lung Nodule Analysis 2016 (LUNA16) is a significant dataset used in the field of lung nodule detection and analysis. It comprises 888 CT scans, which are three-dimensional medical imaging data obtained from computed tomography machines. These CT scans are crucial in identifying and diagnosing lung nodules, which are potential indicators of lung cancer. In the LUNA16 dataset, a file is provided that contains an extended set of candidate locations for the 'false positive reduction' track. False positives are instances where the algorithm incorrectly identifies a region as a nodule when it is not actually present. To address this, the dataset includes a set of candidate locations to help improve the accuracy of nodule detection algorithms by filtering out false positives. It is important to note that a single nodule can be detected by multiple candidates. This is because lung nodules can vary in size, shape, and appearance, leading to multiple possible detection points within a single nodule. In the LUNA16 dataset, a total of 1186 nodules have been detected, and the number of candidates generated for these nodules amounts to 754,975. This large number of candidates provides a diverse set of potential nodule locations for the algorithm to analyze and evaluate. Each CT scan in the LUNA16 dataset consists of a series of 200 to 400 individual images, known as slices. These slices are two-dimensional representations of the three-dimensional CT scan data. Each slice has a size of 512 x 512 pixels, which represents the image resolution and dimensions.

Based on the author [8], we used bounding boxes with a fixed size of 50 x 50 pixels to train, validate, and test our model for lung nodule detection. Bounding boxes are rectangular regions that enclose the regions of interest in an image, in this case, the lung nodules. The size of 50 x 50 pixels was chosen as a standard region to focus on the nodules and provide a consistent input for the model. The dataset used for training and evaluation consists of a total of 5486 lung nodules with positive labels. These positive labels indicate the presence of nodules in the corresponding images. The dataset is divided into three subsets: 4149 images for training, 977 images for validation, and 1418 images for testing. During the training phase, our model is presented with the 50 x 50 pixel gray-scale images of lung nodules along with their corresponding bounding box annotations. The model learns to detect and localize nodules within the images, effectively predicting the bounding boxes that enclose the nodules. The validation subset, comprising 977 images, is used during the training process to monitor the model's performance and prevent overfitting. This process helps ensure that the model generalizes well to new, unseen data, rather than memorizing specific examples from the training set. Finally, the testing subset, containing 1418 images, is employed to assess our model's performance on the data. The model's predictions on this test set are evaluated against ground-truth annotations to measure the algorithm's accuracy, precision, recall, F1-score, and other performance metrics. This step helps validate the model's effectiveness in detecting lung nodules in real-world scenarios. By using gray-scale images and bounding box annotations of fixed size, we achieve consistency in data representation, making the training process more

manageable and efficient. The use of a large dataset with a significant number of nodules ensures that our model can learn from diverse examples, leading to robust and accurate nodule detection.

3.2 Environment for Experiments

We conducted the implementation and execution of our model in a consistent environment using the Google Colab platform. The platform provided a 16 GB GPU A100 and 12 GB RAM, which was crucial for efficient training and inference of the deep learning model. To initialize the model, we used a set of weights based on the YOLO v5x architecture. Weight initialization is essential as it helps the model start with reasonable parameter values, which can expedite the training process and lead to more stable convergence. During the training phase, we employed the Adam optimizer with a learning rate of 1e-2 and a weight decay of 5e-4. Adam is a popular optimization algorithm that efficiently adapts the learning rate for each parameter, making it well-suited for training deep neural networks like YOLO v5x. During the training phase, we employed the Adam optimizer with a learning rate of 1e-2 and a weight decay of 5e-4. Adam is a popular optimization algorithm that efficiently adapts the learning rate for each parameter, making it well-suited for training deep neural networks like YOLO v5x. To improve the model's generalization and robustness, we applied data augmentation techniques during training. We adjusted the hue, saturation, and value (brightness) of the images with hyperparameters set at 0.015 for hsv_h, 0.7 for hsv_s, and 0.4 for hsv_v. This technique introduces variations in color to make the model more resilient to changes in illumination and contrast. We performed random translations of the images with a hyperparameter set to 0.1. This technique shifts the objects' positions within the images, simulating different viewpoints and object locations. We applied random scaling to the images with a hyperparameter set to 0.5. This technique allows the model to detect objects at various sizes, which is essential for handling lung nodules of different dimensions. We used mosaic augmentation with a hyperparameter set to 1. Mosaic augmentation combines four random images into a single mosaic image, providing the model with more complex and diverse examples for training. We trained the model for 100 epochs, with a batch size of 16 images. The number of epochs determines the number of times the entire dataset is processed during training. Training for 100 epochs allows the model to learn from the data effectively and potentially achieve convergence. To prevent overfitting and avoid wasting training time once the model has converged, we employed the early stop technique. Early stop enables the model to stop training when its performance on the validation set starts deteriorating, indicating that further training may lead to overfitting. By implementing the YOLO v5x models in the specified environment and applying appropriate hyperparameters, data augmentation, and early stop technique, we aimed to achieve robust and accurate lung nodule detection in lung CT images using the LUNA16 dataset.

3.3 Evaluation

We developed a two-step approach to train and evaluate our model for lung nodule detection in lung CT images using the YOLO v5x architecture. In the first step, we used a

pre-training model and the training data to find the initial parameters for our model. The pre-training model might have been a pre-trained YOLO v5x model on a large dataset, such as the COCO dataset, to capture general object detection capabilities. We fine-tuned this pre-training model on our specific dataset, which consists of lung CT images with labeled lung nodules. During training, we utilized the back-propagation method, a standard optimization technique for training neural networks. Back-propagation calculates gradients with respect to the model parameters, and the model iteratively adjusts these parameters to minimize the difference between predicted bounding boxes and ground-truth annotations of lung nodules. This process allows the model to learn the patterns and features relevant to lung nodule detection. After training, we used the cross-validation method to assess the model's accuracy and generalization performance. Cross-validation involves splitting the training data into multiple subsets, or "folds," and iteratively using each fold as a validation set while training the model on the remaining folds. This process is repeated to obtain more robust estimates of the model's performance. During cross-validation, we measured several performance metrics, including Precision, Recall, and F1-score. Precision represents the ratio of true positive detections (correctly identified nodules) to all positive detections (both true and false positives). Recall, also known as Sensitivity or True Positive Rate, measures the ratio of true positive detections to the total number of actual nodules in the dataset. The F1-score is the harmonic mean of Precision and Recall, providing a single metric that balances both metrics and gives an overall indication of the model's performance. In the final step, we evaluated the trained model's accuracy on the testing data, which the model has never seen during training or cross-validation. This evaluation allowed us to measure the model's performance on unseen data, providing insights into its ability to generalize to new lung CT images with lung nodules. By assessing Precision, Recall, and F1-score during cross-validation and testing, we obtained a comprehensive evaluation of our model's performance for tumor identification in lung CT images. These metrics are crucial for understanding the model's strengths and limitations and for comparing its performance to other state-of-the-art approaches. Additionally, we might have also visualized the model's predictions to gain insights into its behavior and potential areas for improvement.

3.4 Experimental Results

CT images play a crucial role in identifying tumors on lung. However, acquiring a large and diverse dataset of real LUNA16 images can be challenging due to various factors such as patient privacy, cost, and limited availability. To address this issue and enhance the training data for identify tumors on lung CT images, YOLO v5x, a state-of-the-art object detection, is utilized to identify tumors from available lung CT images. The process starts by collecting a set of CT images with corresponding lung tumor annotations. These CT images are then used as the input for our model, a network specifically designed for medical image synthesis. Our model consists of a detection network and a classification network, which are trained to Identify tumors on lung CT images. Once our model is trained, it starts identify tumors from the input CT images. The Identified images aim to be distinguishable from real CT images and should capture the essential position required for accurate lung tumor identification.

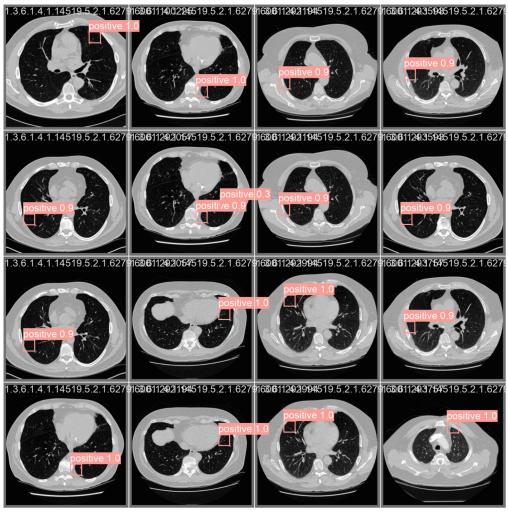

Fig. 4. Experiment to identify tumors on lung CT images

The results shown in Fig. 4 provide valuable insights into the performance of the deep learning model after 100 epochs of training for tumor identification in lung CT images. The identified images exhibit numerous areas with correctly identified tumors. This indicates that the deep learning model is successfully detecting and localizing tumors in the lung CT scans, which is the primary objective of our model. The ability to identify tumors accurately is crucial for early detection and subsequent medical interventions. The explanation for the presence of correctly identified tumors in the source CT images can be attributed to the nature of the training data. During the training process, the model was exposed to a large dataset of CT images that were annotated with lung tumor labels. This annotated data allowed the model to learn the visual patterns and features associated with tumors, enabling it to recognize similar patterns in unseen CT scans. The experimental results indicate that the deep learning model achieves convergence relatively quickly, typically within 100 epochs. Convergence refers to the point in training where the model stabilizes and the learning process reaches a state of equilibrium. In this state, the identified images closely resemble real CT images in terms of quality and appearance, showcasing the model's ability to generalize well to unseen data. As the training progresses, the identified images gradually improve in quality and appearance. This is a positive sign as it indicates that the model is continuously refining its understanding of tumor characteristics and is capable of making more accurate predictions. The ability to generalize to new and unseen CT images is crucial for the model's practical application in real-world medical settings. The results demonstrate that the deep learning model is effective in tumor detection and identification on lung CT images. The model's ability to quickly converge and produce high-quality results speaks to its efficiency and reliability as part of the CAD system. These findings are promising and contribute to the advancements in medical image analysis, particularly in the early detection and treatment of lung tumors, which can lead to improved patient outcomes. However, further evaluation

and validation on larger and diverse datasets may be necessary to solidify the model's performance and generalization capabilities.

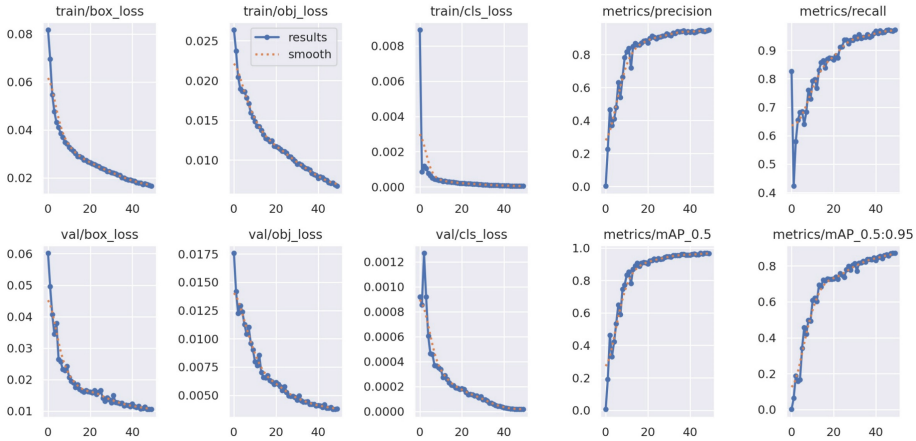

Fig. 5. Progress of train with loss function in identifying tumors on lung CT images with LUNA16 dataset.

The graph in Fig. 5 presents the loss function values of the validation set and the training set during the training process. It shows that the loss function of the validation set converges quite well without significant divergence or overfitting compared to the loss function of the training set. This observation is crucial as it indicates that the model generalizes well to data, and the performance on the validation set is not compromised due to overfitting on the training data. The success of the model in generalizing can be explained by considering several key factors in the data and the training process. The LUNA16 database used for training and validation is characterized by a substantial amount of lung CT images, providing a rich and diverse dataset for the model to learn from. Having a large database helps the model to capture the complexity and variety of lung CT images, improving its ability to generalize to new. The images in the LUNA16 database are in good shape, meaning they are well-annotated, have clear lung boundaries, and exhibit minimal artifacts or distortions. High-quality images contribute to more reliable training and allow the model to learn accurate representations of the lung structures. The synthetic identified images generated by our model are of pretty good quality, as mentioned earlier. These identified images serve as additional data for training and enrich the dataset, enhancing the diversity and increasing the effective size of the dataset. The good quality of these identified images ensures that they contribute valuable information during training, allowing the model to learn relevant features for identifying tumors on lung CT images. By using the identified images generated by our model to enrich the dataset, the overall quantity of available training data is increased. More data means the model has access to a broader set of examples, which aids in building a robust and generalized representation of lung tumor annotation patterns. The combination of a large and high-quality database, along with the introduction of synthetic lung CT images for data enrichment, results in a more effective level of learning. The model can leverage the

rich information. It is worth noting that while the current approach has shown promise, additional analysis and evaluation on different datasets or in a cross-validation setting could further validate the effectiveness of the model and its generalization capabilities. Nonetheless, the results presented in Fig. 5 suggest that the proposed method is on the right track in achieving accurate lung tumors identification without compromising generalization performance. The utilization of GIoU (Generalized Intersection over Union) loss function during the training of YOLO v5x plays a crucial role in achieving excellent performance in detecting nodules on lung CT images. GIoU loss is an improved version of the traditional Intersection over Union (IoU) loss, which takes into account the overlap between predicted and ground-truth bounding boxes. By utilizing GIoU loss, the model can better capture the spatial relationships between the predicted bounding boxes and the actual nodules, leading to improved accuracy. To provide a comprehensive assessment of the superiority of the proposed model, we conducted a rigorous comparative analysis with the previously recognized advanced lung nodule detection method, as presented in Table 1. This comparison is intended to introduce the advantages and advancements achieved by our model in terms of efficiency, accuracy, and computational complexity.

Table 1. Compare with various methods

Method	Precision	Recall
Swetha Subramanian [8]	0.893	0.712
Ours	0.947	0.975

In addition, the faster inference time in our model allows for rapid and real-time tumor identification, which is important in time-sensitive medical situations where decisions Prompt identification and prompt patient care are of the utmost importance. The enhanced speed of our model makes it more suitable for use in large-scale lung cancer screening programs where a significant number of CT scans are efficiently processed is essential for early tumor detection and improving public health outcomes.

4 Conclusions

In our research paper, the primary objective is to develop an efficient and accurate model for detecting nodules and distinguishing between nodular and non-nodular regions in lung CT images. To achieve this, we leverage the YOLO (You Only Look Once) v5x model, which is a state-of-the-art object detection algorithm known for its speed and accuracy. Our model is trained using the LUNA16 dataset, which is a well-established and widely used dataset for lung nodule detection and evaluation. The model's performance is assessed using standard evaluation metrics, including precision, recall, and F1-score. The experimental results demonstrate the effectiveness of our proposed model. The evaluation metrics on the LUNA16 dataset show a precision of 94,7%, and recall of 97,5%. These high values indicate that our model is capable of accurately identifying the locations of nodules while maintaining a good balance between precision (low

false positives) and recall (low false negatives). The improved efficiency of our model is attributed to the streamlined architecture of our model, which allows for real-time inference without compromising accuracy. The model's lightweight design and efficient use of resources make it a practical and viable solution for lung nodule detection tasks, especially in scenarios where real-time or near-real-time processing is required. Our research demonstrates the effectiveness and efficiency of our proposed model with GIoU loss for lung nodule detection and classification. The high precision, recall, and F1-score obtained on the LUNA16 dataset showcase the model's capability to accurately locate and differentiate between nodular and non-nodular regions in lung CT images. Additionally, the comparison with the previous state-of-the-art model confirms the superiority of our approach in terms of efficiency, making it a promising contribution to the field of lung nodule detection in medical imaging.

References

1. YOLO 5 Homepage. https://github.com/ultralytics/yolov5. Accessed 19 Apr 2023
2. LUNA16 Homepage. https://luna16.grand-challenge.org/ Accessed 09 Apr 2023
3. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., Savarese, S.: Generalized intersection over union: a metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June, pp. 658–666 (2019).
4. Lung Cancer Statistic Homepage. https://www.wcrf.org/cancer-trends/lung-cancer-statistics. Accessed 09 Apr 2023
5. Bhandary, A., Ananth Prabhu, G., Basthikodi, M., Chaitra, K M.: Early diagnosis of lung cancer using computer aided detection via lung segmentation approach. Int. J. Eng. Trends and Technol. (2021)
6. Bhattacharjee, A., Majumder, S.: Automated computer-aided lung cancer detection system. In: Advances in Communication, Devices and Networking. Lecture Notes in Electrical Engineering. pp. 425–433 (2019)
7. Zhang, D., Hu, J., Li, F., Ding, X., Sangaiah, A.K., Sheng, V.S.: Small Object Detection via Precise Region-Based Fully Convolutional Networks. Comput. Mater. Continua. (2021)
8. Swetha Subramanian Homepage. https://git.io/Jf9Og. Accessed 19 Apr 2023
9. Xie, H., Yang, D., Sun, N., Chen, Z., Zhang, Y.: Automated pulmonary nodule detection in CT images using deep convolutional neural networks. Comput. Sci. Pattern Recognit. (2019)
10. Ren, S., et al.: Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. (2017)