YOLOv8-Based Framework for Accurate Lung CT Nodule Images Detection

Mengzhe Wang[a]
Institute of Artificial Intelligence, Xidian University, Xinglong Street, Xian, China

Keywords: Machine Learning, Lung Nodule, Image Recognition.

Abstract: Traditional Lung Computed Tomography (CT) Nodule Recognition Primarily Relies on Visual Inspection by doctors. However, recent advancements in image recognition models have significantly increased the feasibility of utilizing emerging image recognition models' powerful capabilities to provide doctors with a new auxiliary means of identifying lung nodules. This study aims to leverage the YOLOv8, a more advanced image processing model, to process the LUNA-16 dataset of lung CT nodule images. Through continuous optimization, the goal is to achieve a relatively ideal recognition accuracy. This paper anticipates evaluating the recognition results from the perspectives of accuracy, recall, and other metrics. By iteratively searching for the local optimal solution of model parameters, the model will be continuously improved. Through final optimization, the study aims to achieve a roughly twofold increase in recognition capability compared to the initial stage, while significantly reducing the false negative rate.

1 INTRODUCTION

Pulmonary nodules, small abnormal lesions detected in lung imaging, serve as pivotal indicators for the presence of lung cancer, highlighting the critical need for early and accurate detection to facilitate timely intervention and enhance patient outcomes. Traditional methods of nodule detection rely heavily on manual interpretation of imaging studies, a process susceptible to variability and oversight, thus driving the exploration of advanced Artificial Intelligence (AI) solutions for automated detection due to their exceptional performance in many tasks in the last decade (Li, 2024; Qiu, 2020; Qiu, 2024; Sun, 2020; Wang, 2024; Wu, 2024; Zhou, 2023).

The You Only Look Once (YOLO)v8 model (Github, 2024), renowned for its robust object detection capabilities, emerges as a promising tool in this pursuit, offering the potential to streamline the detection process and improve diagnostic accuracy. For instance, identifying fractures in X-ray images (Jum 2023), detecting fire alarms (Talaat, 2023).

Existing methodologies face challenges, including the diverse characteristics of pulmonary nodules and the scalability across large and heterogeneous datasets. While previous studies have made significant strides in nodule detection, many are constrained by their reliance on specific datasets or imaging modalities, limiting their generalizability and practical utility in real-world clinical settings.

Against this backdrop, this research seeks to bridge these gaps by leveraging the YOLOv8 model and the comprehensive LUNA16 dataset to develop a robust, scalable, and adaptable solution for pulmonary nodule detection. The significance of this endeavour lies in its potential to transform lung cancer screening practices by introducing automated detection algorithms that reduce variability, enhance efficiency, and improve diagnostic accuracy. By building upon previous research and harnessing the wealth of publicly available datasets, this paper developed a standardized and reproducible framework for nodule detection that can be seamlessly integrated into diverse clinical workflows. By enabling early detection and intervention for lung cancer, the research ultimately strives to improve patient outcomes and contribute to the global fight against this devastating disease.

[a] https://orcid.org/0009-0007-9794-8840

Wang, M.
YOLOv8-Based Framework for Accurate Lung CT Nodule Images Detection.
DOI: 10.5220/0012972700004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 775-780
ISBN: 978-989-758-713-9

2 METHOD

2.1 Dataset Preparation

The LUNA16 dataset (LUNA16, 2016) comprises 1186 lung nodules in grayscale, each sized at 330x330 pixels. These images showcase various types of nodules, serving as essential resources for medical imaging research and development.

Despite its comprehensiveness, integrating the LUNA16 dataset with the YOLOv8 model, renowned for its object detection prowess, presents a notable challenge due to format disparities. To surmount this obstacle, a critical preprocessing step becomes imperative. This entails transforming the dataset into a format compatible with YOLOv8, such as the VOC format.

Navigating through each CT image in the dataset involves meticulous iteration through individual nodules, each potentially indicative of a distinct medical anomaly. Accurately calculating precise coordinates for each nodule necessitates the conversion from world coordinates to image coordinates, ensuring precise localization within the images. Subsequently, XML tags are generated for each nodule, encapsulating crucial information such as bounding box coordinates and corresponding class labels.

Through this detailed preprocessing regimen, the LUNA-16 dataset is adeptly tailored to meet the requirements of the YOLOv8 model. By providing a standardized format aligning with the model's architecture, researchers and medical practitioners can effectively harness the capabilities of YOLOv8 to discern and analyse lung nodules within CT images with heightened accuracy and efficiency, thus propelling advancements in medical imaging and diagnosis.

2.2 YOLOv8-Based Lung Nodule Detection

You Only Look Once (YOLO) stands as a widely-utilized model in the realm of object detection and image segmentation [5]. Since its inception in 2015, YOLO has garnered widespread acclaim due to its exceptional speed and accuracy.

Currently, with the advent of YOLOv8, the latest iteration provided by Ultralytics, the capabilities of YOLO have been further expanded. YOLOv8 is not limited to mere object detection [6]; it extends its functionality to encompass a diverse array of computer vision tasks, including segmentation, pose estimation, tracking, and classification. This versatility is instrumental in empowering users across various applications and fields, enabling them to harness the full potential of YOLOv8 in their respective domains.

At its core, YOLOv8 employs a grid-based methodology to scrutinize input images. These images are meticulously dissected into a structured grid layout, with each grid cell serving as a focal point for analysis. Within these cells, the model adeptly predicts bounding boxes encapsulating potential objects, while concurrently assigning probabilities to different object classes. This unified approach, facilitated by a neural network architecture, seamlessly consolidates these predictions, culminating in rapid and precise object detection.

The inherent agility and accuracy of YOLOv8 renders it exceptionally well-suited for real-time applications across diverse domains. Whether in surveillance, autonomous vehicles, medical imaging, or any other field reliant on computer vision, YOLOv8 stands as a cornerstone, driving innovation and facilitating progress with its unparalleled capabilities.

The YOLOv8 model employs deep learning networks to train on lung CT images, learning features of nodules and performing object detection. Prior to training, a substantial annotated dataset is required, encompassing various types of nodules and cases. The network architecture includes convolutional, pooling, and fully connected layers. During training, model parameters are adjusted guided by a loss function to optimize consistency between predicted and true labels. In the prediction phase, new CT images are inputted into the model to obtain bounding boxes with nodule positions and class information. Post-processing steps such as non-maximum suppression are often conducted to enhance accuracy. In summary, YOLOv8 utilizes deep learning techniques to achieve automated detection of nodules in lung CT images.

2.3 Implementation Details

The GPU used for model training is the NVIDIA GeForce RTX 4090 Laptop, with 16GB of VRAM. The loss function used in the YOLOv8 model mainly combines several parts, including the loss function for object detection and the loss function for image segmentation. The loss function for object detection usually includes position loss, confidence loss, and class loss. The position loss measures the difference between the predicted bounding box and the true bounding box, the confidence loss measures the model's confidence in the presence of the target, and the class loss measures the accuracy of the model's prediction of the target category. The loss function for image segmentation is used to measure the accuracy of the model's segmentation of each pixel, usually including pixel classification loss and pixel position loss. The combination of these loss functions helps the model continuously optimize during training to improve the accuracy and performance of object detection and image segmentation.

As for the optimizer, The Adam optimizer is a widely-used algorithm in the field of machine learning, particularly for training models. It functions by iteratively updating the parameters of a model based on the gradients of the loss function. What sets Adam apart is its ability to adaptively adjust the learning rate during the training process. This adaptability is achieved by combining the advantages of two other optimization algorithms, AdaGrad and RMSProp. By dynamically adjusting the learning rate, Adam ensures that larger updates are made for infrequently occurring parameters and smaller updates for frequently occurring parameters, leading to more efficient and effective optimization. Its versatility and effectiveness make Adam a preferred choice for optimizing models across various tasks and datasets in the machine learning community.

3 RESULTS AND DISCUSSION

In this experiment, different ways were employed to see how well the YOLOv8 model works shown in Figure 1-Figure 13. First, confusion matrix is used to see how accurate the model is at predicting different things and where it makes mistakes. Then, this research used F1 curve. It's a way to see how good the model is overall, taking into account both how often it's right and how often it misses things. After that, the discussion shifted to two more concepts known as the P curve and the R curve, utilized to illustrate the frequency of the model's accuracy and its capability to capture all correct instances under various conditions. The conversation then moved to an analysis of the PR curve, which serves to demonstrate how the model manages the trade-off between accuracy and comprehensiveness, aiming to determine the optimal settings for the model's performance. Initially, the research employed the YOLOv8 model in its unaltered state on the LUNA-16 dataset. However, upon examining the outcomes, it was evident that the model struggled to effectively identify nodules.

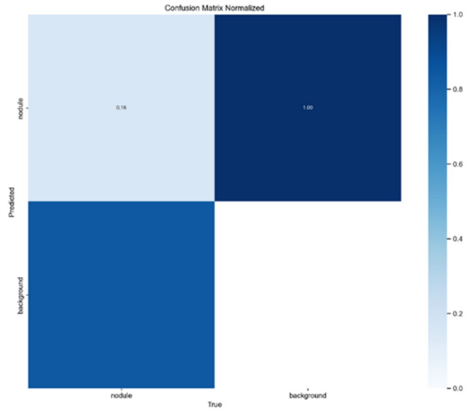

Figure 1: Confusion matrix normalized before adjusting the model (Photo/Picture credit: Original).

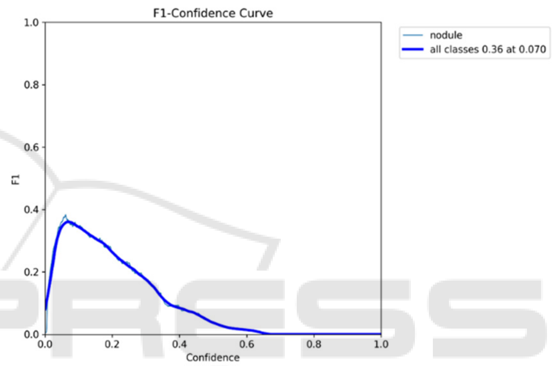

Figure 2: F1-confidence curve before adjusting the model (Photo/Picture credit: Original).

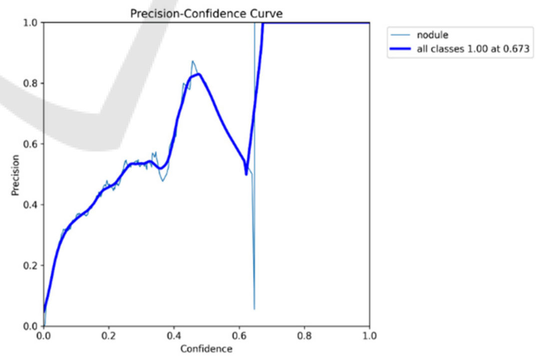

Figure 3: Precision-confidence curve before adjusting the model (Photo/Picture credit: Original).

Among the sixteen training rounds conducted, it's evident that the model succeeded in identifying nodules in only four instances. While the false positive rate remained relatively low, the false negative rate was notably high. This discrepancy in performance can be attributed to the model's reliance on default parameters, which may not adequately suit

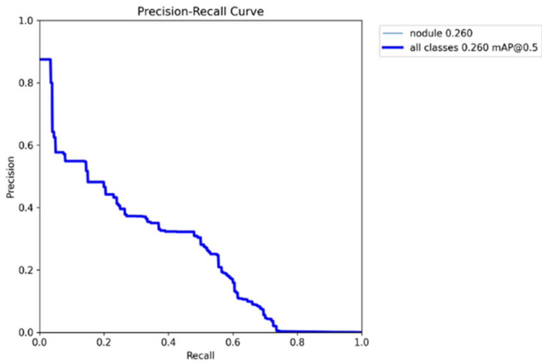

Figure 4: Precision-recall curve before adjusting the model (Photo/Picture credit: Original).

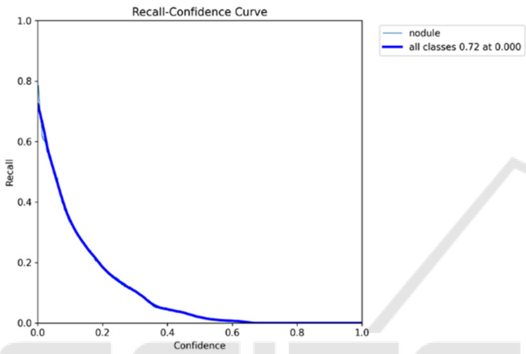

Figure 5: Recall-confidence curve before adjusting the model (Photo/Picture credit: Original).

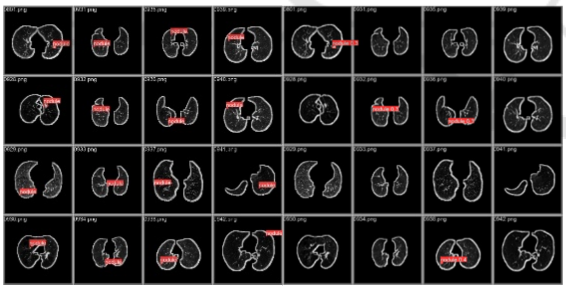

Figure 6: Detecting results before adjusting the model (Photo/Picture credit: Original).

the characteristics of the dataset. Therefore, the following subsequent efforts will be focused on conducting multiple iterations to explore optimal solutions for each parameter. Subsequently, armed with these refined parameters, the model will be adjusted accordingly and proceed with re-running it in order to achieve improved performance and accuracy.

Figure 9 provided depicts the local optimal solution attained after five hundred iterations. The next step involves translating the acquired parameters into modifications within the default.yaml file. Subsequently, this experiment will integrate the best model obtained throughout the iterative process, identified as best.pt, into the processing pipeline for a fresh round of dataset handling.

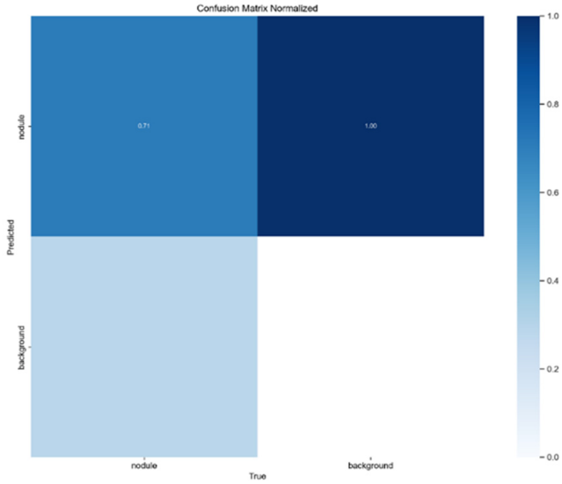

Figure 7: Confusion matrix normalized after adjusting the model (Picture credit: Original).

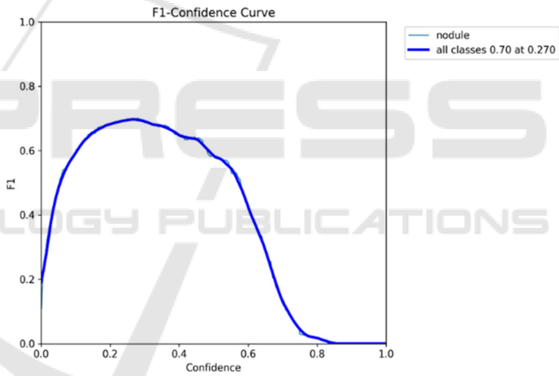

Figure 8: F1-confidence curve after adjusting the model (Picture credit: Original).

The results obtained from the new run clearly indicate a significant improvement in the model's accuracy, with the precision soaring from 31% to 71%. This marks a notable and substantial enhancement. At the same time, among the sixteen sets of validation images, there was only one instance of missed detection, indicating a significant improvement compared to the initial results. However, there were four occurrences of false alarms, suggesting a potential issue with model overfitting.

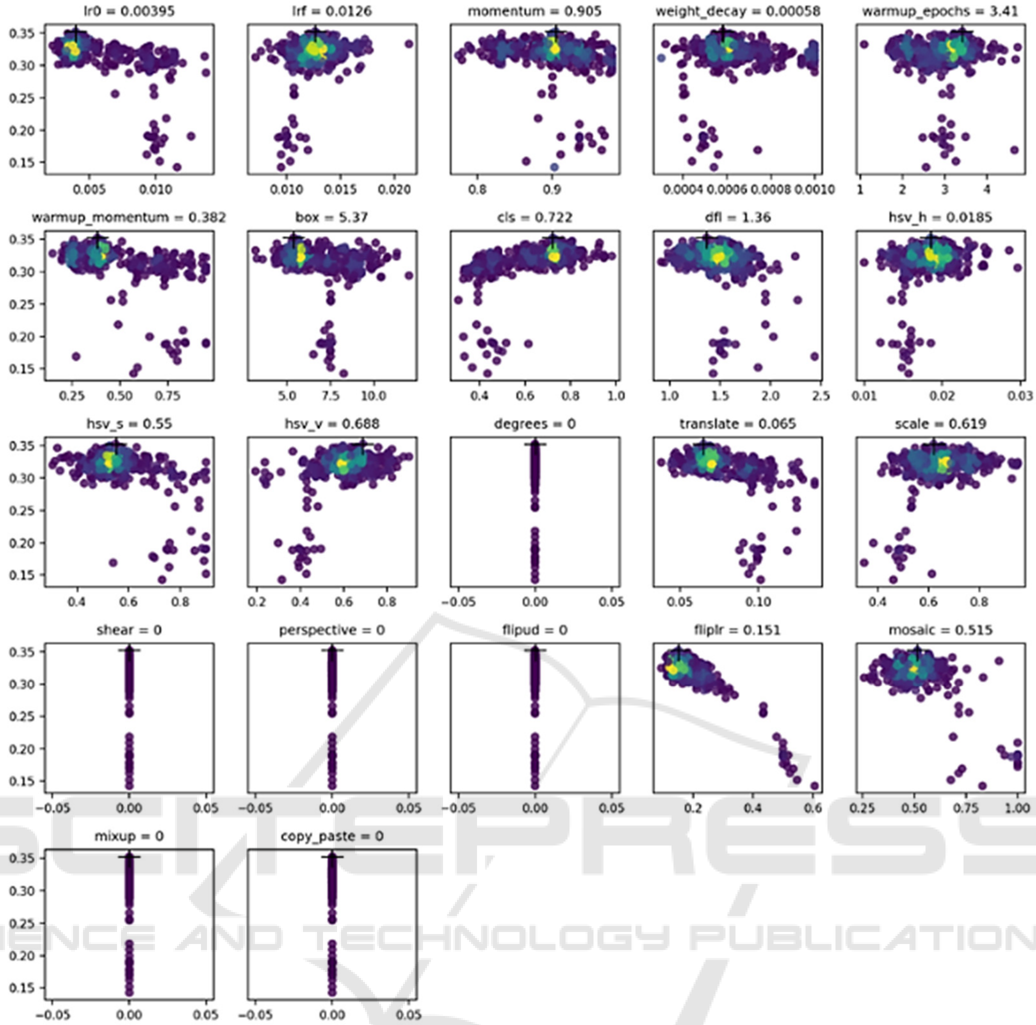

Figure 9: Local optimal solution (Picture credit: Original).

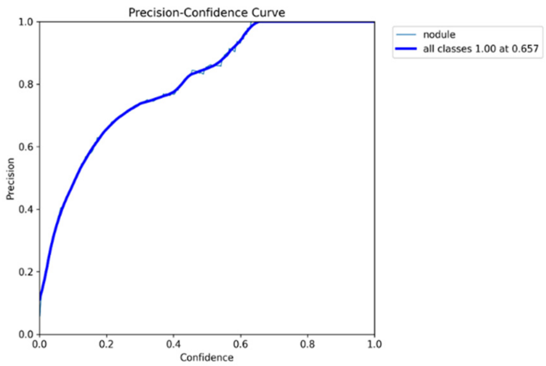

Figure 10: Precision-confidence curve after adjusting the model (Picture credit: Original).

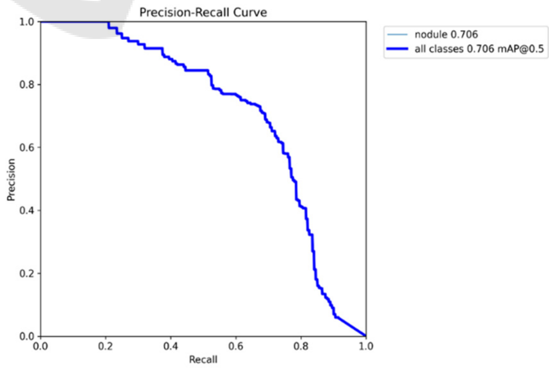

Figure 11: precision-recall curve after adjusting the model (Picture credit: Original).

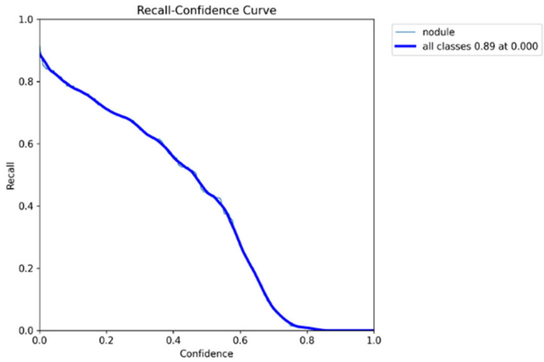

Figure 12: Recall-confidence curve after adjusting the model (Picture credit: Original).

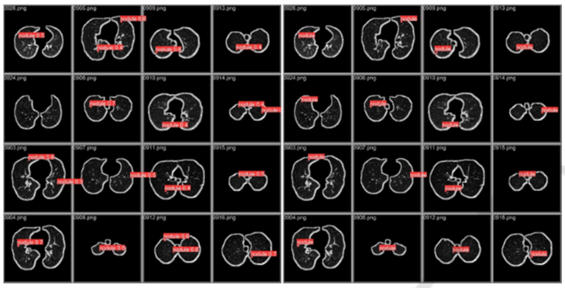

Figure 13: Detecting results after adjusting the model (Picture credit: Original).

4 CONCLUSIONS

This study utilized the YOLOv8 model for the detection of nodules within lung CT images sourced from the LUNA-16 dataset. Upon initial assessment, the model exhibited suboptimal performance in accurately identifying these nodules. However, through iterative parameter tuning during subsequent phases of experimentation, noticeable enhancements in the model's detection capabilities were observed, leading to a more consistent and stable performance trajectory. Although the study succeeded in identifying nodules within lung CT scans, the encountered challenges pertaining to limited accuracy pose obstacles to the practical application of the model within clinical settings. Additionally, the intricate nature of the algorithm contributes to its high level of complexity, signalling significant potential for optimization aimed at improving efficiency. While commendable progress has been made in the domain of lung nodule recognition through the utilization of the YOLOv8 model, the study underscores the imperative for further refinement and optimization endeavours. Addressing the constraints associated with accuracy and complexity will be pivotal in fully unlocking the model's capabilities for deployment in medical imaging and diagnostic applications, thereby facilitating advancements in the field.

REFERENCES

Github, Yolov8, https://github.com/ultralytics/ultralytics, 2024.

Ju, R. Y., & Cai, W. 2023. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm. Scientific Reports, 13(1), 20077.

Li, M., He, J., Jiang, G., & Wang, H. 2024. DDN-SLAM: Real-time Dense Dynamic Neural Implicit SLAM with Joint Semantic Encoding. arXiv preprint arXiv:2401.01545.

LUNA16, Lung Nodule Analysis 2016, https://luna16.grand-challenge.org/, 2016

Qiu, Y., Hui, Y., Zhao, P., Cai, C. H., Dai, B., Dou, J., ... & Yu, J. 2024. A novel image expression-driven modeling strategy for coke quality prediction in the smart cokemaking process. Energy, 130866.

Qiu, Y., et al 2020. Improved denoising autoencoder for maritime image denoising and semantic segmentation of USV. China Communications, 17(3), 46-57.

Sun, G., et al. 2020. Revised reinforcement learning based on anchor graph hashing for autonomous cell activation in cloud-RANs. Future Generation Computer Systems, 104, 60-73.

Talaat, F. M., & ZainEldin, H. 2023. An improved fire detection approach based on YOLO-v8 for smart cities. Neural Computing and Applications, 35(28), 20939-20954.

Wang, H., Zhou, Y., Perez, E., & Roemer, F. 2024. Jointly Learning Selection Matrices For Transmitters, Receivers And Fourier Coefficients In Multichannel Imaging. arXiv preprint arXiv:2402.19023.

Wu, Y., Jin, Z., Shi, C., Liang, P., & Zhan, T. 2024. Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis. arXiv preprint arXiv:2403.08217.

Zhou, Y., Osman, A., Willms, M., Kunz, A., Philipp, S., Blatt, J., & Eul, S. 2023. Semantic Wireframe Detection. publica.fraunhofer.de.