

Research on Automatic Elimination of Laptop Computer in Security CT Images Based on Projection Algorithm and YOLOv7-Seg

Fei Wang¹, Baosheng Liu¹, Yijun Tang², Lei Zhao¹

¹School of Computer Science and Technology, Shandong University of Technology, Zibo, China

²Shanghai Wuying Technology Co., Ltd., Shanghai, China

Email: herefhuar@gmail.com

How to cite this paper: Wang, F., Liu, B.S., Tang, Y.J. and Zhao, L. (2023) Research on Automatic Elimination of Laptop Computer in Security CT Images Based on Projection Algorithm and YOLOv7-Seg. *Journal of Computer and Communications*, 11, 1-17. <https://doi.org/10.4236/jcc.2023.119001>

Received: August 9, 2023

Accepted: September 17, 2023

Published: September 20, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In civil aviation security screening, laptops, with their intricate structural composition, provide the potential for criminals to conceal dangerous items. Presently, the security process necessitates passengers to individually present their laptops for inspection. The paper introduced a method for laptop removal. By combining projection algorithms with the YOLOv7-Seg model, a laptop's three views were generated through projection, and instance segmentation of these views was achieved using YOLOv7-Seg. The resulting 2D masks from instance segmentation at different angles were employed to reconstruct a 3D mask through angle restoration. Ultimately, the intersection of this 3D mask with the original 3D data enabled the successful extraction of the laptop's 3D information. Experimental results demonstrated that the fusion of projection and instance segmentation facilitated the automatic removal of laptops from CT data. Moreover, higher instance segmentation model accuracy leads to more precise removal outcomes. By implementing the laptop removal functionality, the civil aviation security screening process becomes more efficient and convenient. Passengers will no longer be required to individually handle their laptops, effectively enhancing the efficiency and accuracy of security screening.

Keywords

Instance Segmentation, Projection, CT Image, 3D Segmentation, Real-Time Detection

1. Introduction

In civil aviation security checks, laptop computers may be used by criminals to conceal dangerous items, posing challenges and risks to security screening. Traditional security methods require passengers to remove laptops from their luggage and undergo separate screening, increasing the time and workload of security checks, causing inconvenience and delays.

With the development of industrial technology and science, the demands on civil aviation security checks have grown. The need to detect and remove laptops in real-time has become a priority. However, the large volume of three-dimensional (3D) data, the high equipment requirements for model training, the complexity of labeling 3D data compared to 2D data, and the overall cost, make instance segmentation in 3D data challenging, as it is difficult to achieve both high accuracy and high speed, which is essential for real-time segmentation in security equipment.

Therefore, this paper proposes a method that combines projection and 2D instance segmentation for real-time laptop removal. In the projection part, a positive projection method is employed. Considering that the laptop's shape is similar to a rectangle and that laptops are more likely to be placed horizontally in packaging and luggage during security checks, positive projection increases the likelihood of obtaining a more standardized laptop projection shape (here, "more standardized" refers to a projection shape that closely resembles the laptop's front, side, and bottom).

This novel method aims to address the challenges posed by real-time laptop removal in security checks, utilizing the advantages of projection and 2D instance segmentation techniques.

In the 2D instance segmentation part, this study adopts YOLOv7 [1] (You Only Look Once v7) as the neural network model. In the field of object detection, single-stage models excel in speed over two-stage models. Among these single-stage models, YOLO [2] (You Only Look Once) has garnered significant attention for its accurate recognition and rapid processing speed.

Building upon the foundation of YOLOv5 (You Only Look Once v5), YOLOv7 incorporates a deeper network architecture, comprising additional convolutional layers and residual blocks. This augmentation enhances the model's representation capacity and detection accuracy. The integration of various data augmentation techniques, such as random cropping and rotation, diversifies the training data, thereby bolstering model robustness. By employing the swish activation function, the model's non-linear expressive power is heightened, contributing to further improvements in detection accuracy.

Experimental results demonstrate that the fusion of positive projection and YOLOv7-seg (You Only Look Once v7 Segment) yields satisfactory outcomes, as evidenced by the data collected from Shanghai Wuying Technology Co., Ltd. This approach effectively meets the real-time detection and removal requirements for laptops, rendering it viable for integration into civil aviation security screening equipment.

2. Literature Review

2.1. Traditional Security CT Image Segmentation Algorithm

Based on public resources, there is limited literature available on CT image segmentation for aviation security. Zhang Xian *et al.* [3] once segmented items with higher atomic numbers by establishing appropriate thresholds based on images of package materials. However, threshold segmentation has great limitations, for example, when the package contains objects close to the threshold of the object to be divided, threshold segmentation cannot work. And there is another problem with threshold segmentation, when the object to be divided is relatively pure, the threshold division may be effective, but if the division of more complex objects, such as laptops, will lead to incomplete segmentation, because the composition of notebook computers is complex and diverse, the material of the components is varied, there are parts with low thresholds and parts with high thresholds, if the threshold segmentation is processed, the parts with low thresholds will be divided. Zhang Yanzhu *et al.* [4] segmented various objects within packages through region growing with selected seed points based on pseudo-color images from security checks. However, this method also has limitations, regional growth selection seed point segmentation, for non-overlapping objects segmentation effect is better, but the data in the luggage is difficult to appear non-overlapping objects, if two similar threshold objects in the luggage partially overlap, it will be difficult to segment the complete data. This method has poor segmentation effect in real CT data and is not suitable for CT data segmentation.

Bai Cong *et al.* [5] employed dual grayscale transformations on security images to segment overlapping regions within complex high atomic number objects. However, this segmentation method is not suitable for laptop segmentation, as the internal components of laptops are complex and diverse, containing both high atomic number components and low atomic number components. The dual grayscale transformation can only segment the high atomic number portion, unable to separate the low atomic number portion of laptops. This method is better suited for objects with relatively pure compositions and is not suitable for segmenting objects like laptops. Gu Zhu *et al.* [6] focused on contour emphasis and distinct color information in test images. They employed a polygonal approximation based on the HSI color space to simplify image edges and performed segmentation based on geometric features. However, this approach is not applicable to grayscale images. The experimental subject of this study is grayscale images, and the HSI color space polygonal approximation method cannot be utilized to capture edges in grayscale images. Therefore, the aforementioned method is not suitable for segmenting laptop data in this experiment.

2.2. Two-Dimensional Instance Segmentation Algorithm Based on YOLO

Zhang Zehua [7] improved the YOLACT (You Only Look At Coefficients) algorithm to achieve pedestrian multi-object tracking combined with instance seg-

mentation, accurately segmenting target edges for finer tracking handling.

Chen Jianxiong [8] combined YOLOv2 (You Only Look Once v2) with Mask R-CNN to achieve instance segmentation for identifying loose fastening components on medium-to-low-speed maglev contact tracks.

He Jinqiang *et al.* [9] combined the YOLOv5 model with the graph-based segmentation Grabcut algorithm in a two-stage image recognition and segmentation process. This approach automatically locates and segments insulators with high efficiency and accuracy in complex backgrounds without requiring segmentation labeling or manual interaction.

Yu Bo *et al.* [10] optimized the YOLO detection and segmentation network model, introduced the K-means++ clustering algorithm to find multi-scale anchor box sizes, and employed a local detection position adaptive threshold segmentation method for pixel-level instance segmentation of detected objects. This work achieves fast and effective detection of pedestrians and generates instance masks in far-infrared images.

Yang Kuihe *et al.* [11] used MobileNetv3 as the feature extraction network in YOLOSeg, integrated PANet to fuse features of different scales, utilized dilated convolutional pooling pyramid for increased receptive field in the semantic segmentation branch, and obtained image segmentation results through bilinear interpolation. They proposed a YOLOSeg algorithm that jointly trains object detection and semantic segmentation, catering to the needs of multiple environmental information perception in the autonomous driving field, including vehicles, pedestrians, and lane markings.

Petr Hurtik1CA1 *et al.* [12] introduced Poly-YOLO, a new version of YOLO that offers improved speed and more precise detection, along with instance segmentation capabilities. Poly-YOLO is built upon the foundational concepts of YOLOv3, addressing two of its weaknesses: the need for a large number of rewritten labels and an inefficient distribution of anchor points. By leveraging features from a lightweight SE-Darknet-53 backbone using a hypercolumn technique and employing stairstep upsampling, Poly-YOLO generates a single-scale output with high resolution. Compared to YOLOv3, Poly-YOLO achieves a 40% relative improvement in mean average precision while utilizing only 60% of its trainable parameters. Additionally, Poly-YOLO lite is introduced, boasting fewer parameters and lower output resolution. Despite its reduced size, Poly-YOLO lite maintains the same precision as YOLOv3 and offers a threefold reduction in size and a twofold increase in speed, making it suitable for embedded devices. Notably, Poly-YOLO performs instance segmentation by identifying size-independent polygons on a polar grid, predicting polygon vertices along with their associated confidence levels, resulting in polygons with varying numbers of vertices.

Since civil aviation still requires laptops to be taken out separately for security checks, the technology of laptop removal has not been widely applied to current civil aviation security checks. Therefore, there is almost no publicly available research data online. This study draws on literature related to traditional security

data segmentation and YOLO instance segmentation to validate the infeasibility of traditional security data segmentation methods and the feasibility of YOLO instance segmentation. Through validation, a new approach is proposed in this study, which combines projection and instance segmentation to achieve security data segmentation.

By reviewing the literature, it is known that there is not much information available on traditional security data segmentation, and most of the publicly available information is quite outdated. Traditional security data segmentation methods that can be found are not suitable for laptop segmentation, as detailed in the previous section. Laptops have complex structures and specific characteristics, and the environments they are placed in are also diverse. The data used in this experiment are grayscale data, which makes it difficult to satisfy all these conditions. The methods found in the public literature cannot simultaneously satisfy these conditions, hence the infeasibility of traditional methods.

Therefore, this study draws on cases of YOLO used for instance segmentation. In the literature that can be accessed, the majority of YOLO instance segmentation is focused on two-dimensional instances. The approach of this study is to achieve three-dimensional data segmentation through two-dimensional instance segmentation. Some scholars have improved existing YOLO algorithms to achieve better results, while others have combined YOLO with other algorithms and made various improvements for different industrial scenarios. This demonstrates the high adaptability and accuracy of the YOLO algorithm. For example, in the aforementioned papers, YOLO is used for precise segmentation and tracking of pedestrians, instance segmentation of loose fastening components on medium-to-low-speed maglev contact tracks, accurate and efficient location and segmentation of insulators in complex backgrounds, pixel-level instance segmentation of pedestrians, and meeting the needs of environmental information perception in autonomous driving, including vehicles, pedestrians, and lane markings. YOLO algorithm shows high accuracy and adaptability. These examples showcase the wide applicability, technical maturity, and accuracy of the YOLO algorithm, making it a powerful tool that can meet the requirements of different scenarios by improving YOLO or combining it with different algorithms. Therefore, this study adopts YOLO for the instance segmentation part of the laptop removal algorithm. Through validation, YOLO is confirmed to be suitable for the instance segmentation part of the proposed laptop removal algorithm. However, whether YOLO is the most suitable instance segmentation algorithm for the proposed laptop removal algorithm remains to be further verified in the future.

3. Method

The innovative aspect of our designed 3D segmentation algorithm lies in its utilization of projection to transform 3D objects into 2D data. Through 2D instance segmentation, masks are obtained, and subsequently, 3D data is synthesized to achieve 3D instance segmentation.

The main steps are as follows:

- 1) Employing a positive projection algorithm to project the XYZ dimensions of the package data containing laptops, thereby obtaining three views of this 3D data.
- 2) Applying a pre-trained YOLOv7-seg instance segmentation model to the three views, extracting masks through instance segmentation.
- 3) Generating a 3D mask from the masks of the three views, and intersecting it with the original data to obtain segmented data.

The overall algorithmic process is illustrated in **Figure 1**.

3.1. Obtaining Three Views through Projection

As a valuable portable device, laptop computers are typically examined during civil aviation security checks with their orientation primarily horizontal or similar to horizontal positioning within luggage. Rarely, they are placed vertically or at other unusual angles.

Considering the aforementioned circumstances, this paper adopts the technique of positive projection to visualize laptop computers. Through positive projection, it is possible to generate a horizontal view parallel to the X-direction, a horizontal view parallel to the Y-direction, and a horizontal view parallel to the Z-direction from a security CT scan data, as depicted in **Figure 2**. This projection technique emulates the common placement posture of laptops during real security checks. Such projection allows for a better capture of the laptop's form and features, providing more accurate input for subsequent instance segmentation. Furthermore, positive projection preserves the laptop's geometric shape and dimensions, preventing information loss and distortion, thus enhancing the stability and reliability of the entire removal process.

The principle of the positive projection algorithm employed in this study is as follows:

Taking the three-dimensional dataset T as shown in **Figure 3** as an example, with dimensions of length n , width l , and height m . When projecting this three-dimensional dataset T along the y -axis from top to bottom, a two-dimensional dataset S is obtained. The length of this two-dimensional dataset S is n , and its width is l .

In this two-dimensional dataset S , each pixel value represents the cumulative sum of pixel values intersected along the y -direction of the three-dimensional dataset T . Put simply, when traversing all the pixels of the three-dimensional dataset T along the y -direction, their values are accumulated to obtain the final projection value, which becomes the value of each pixel in the two-dimensional

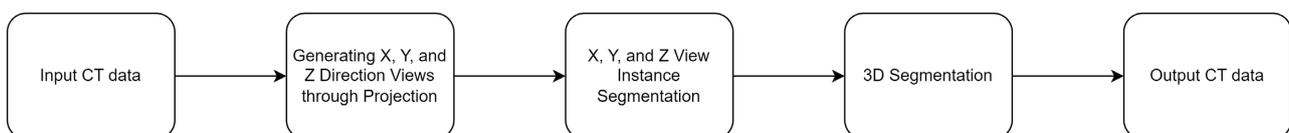


Figure 1. Algorithm overall flowchart.

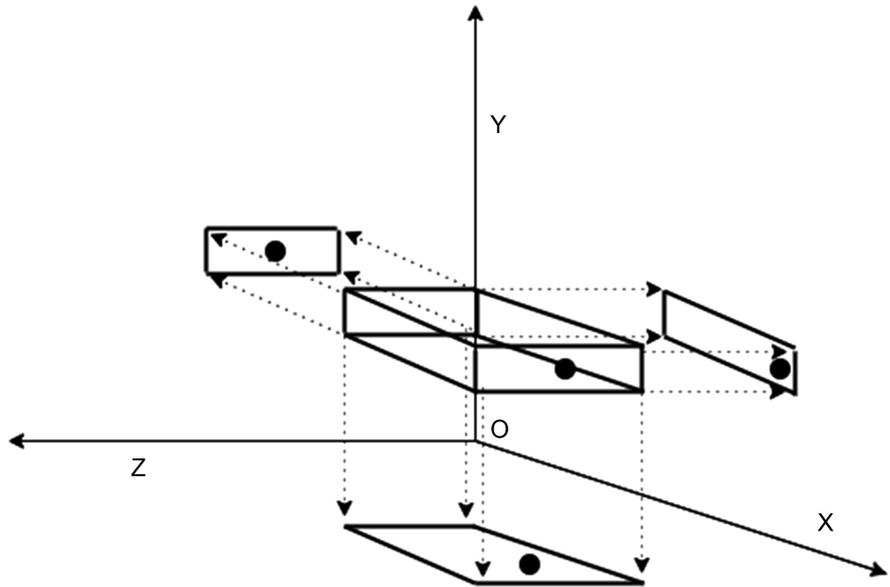


Figure 2. Principle of three-view projection.

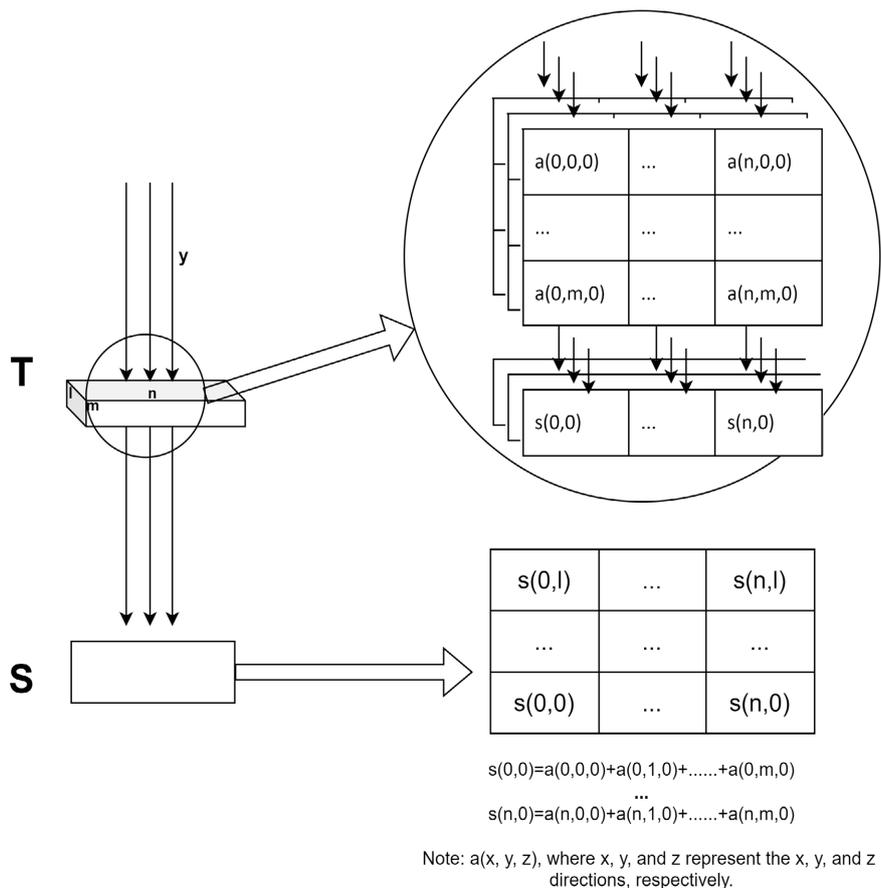


Figure 3. Projection principle diagram.

data.

Assuming the top-left corner vertex coordinate of the three-dimensional da-

taset T is $a(0,0,0)$, a two-dimensional dataset is generated along the y-axis with the top-left corner vertex coordinate being $s(0,0)$. Taking the top-left corner pixel $s(0,0)$ in the two-dimensional dataset as an example, its computation process is as follows:

The value of $s(0,0)$ is obtained by summing the values of all pixels in the three-dimensional dataset T that intersect along the y-direction, starting from $a(0,0,0)$. In essence, the process involves traversing pixels along the y-direction in the three-dimensional dataset T, beginning at $a(0,0,0)$, and accumulating their values until the boundary of the y-direction is reached.

This accumulation process yields the value of $s(0,0)$, which represents the projection value of the top-left corner pixel in the two-dimensional dataset. Mathematically, it is expressed as:

$$S(0,0) = \text{sum}(a(0,0,0) + a(0,1,0) + a(0,2,0) + \dots + a(0,m,0)) \quad (1)$$

Using this projection approach, information along the y-direction is extracted from the three-dimensional dataset T and projected onto a two-dimensional plane. This can be mathematically represented as:

$$S(x,z) = \sum_{y=0}^m a(x_i, y, z_j) \quad (2)$$

where i represents the x-coordinate and j represents the z-coordinate.

Similarly, projecting along the z-axis from front to back, the expression is:

$$S(x,y) = \sum_{z=0}^l a(x_i, y_j, z) \quad (3)$$

where i represents the x-coordinate and j represents the y-coordinate.

Similarly, projecting along the x-axis from front to back, the expression is:

$$S(y,z) = \sum_{x=0}^n a(x, y_i, z_j) \quad (4)$$

where i represents the y-coordinate and j represents the z-coordinate.

When dealing with larger security CT three-dimensional data, due to the possibility of encountering significant cumulative projection values, it becomes necessary to employ higher-bit data formats for storage. However, mindful of transportation and memory limitations, an effective strategy of normalizing the data has been adopted, constraining the projected pixel values within the range of 0 to 255.

Normalization helps maintain the data's relative relationships and inherent features, while also reducing the bit count, resulting in reduced storage space and memory demands. This approach simplifies data complexity and improves both the execution and storage efficiency of the algorithm.

The normalization method used in this context is referred to as *Min-Max* normalization [13], and it is represented by the following formula:

$$p(x, y) = \frac{p(x, y) - \text{Min}[p(x, y)]}{\text{Max}[p(x, y)] - \text{Min}[p(x, y)]} \quad (5)$$

Taking the real data example of 00012256. raw, with dimensions of $600 \times 400 \times 333$, the obtained three views are shown in **Figure 4**.

3.2. Obtaining Instance Segmentation Masks with YOLOv7-Seg

A pre-trained YOLOv7-seg model was employed for performing instance segmentation on the projected images. The images were organized into sets of three, with each set corresponding to a single data entry, as illustrated in **Table 1**. Instance segmentation was employed to extract masks and produce output.

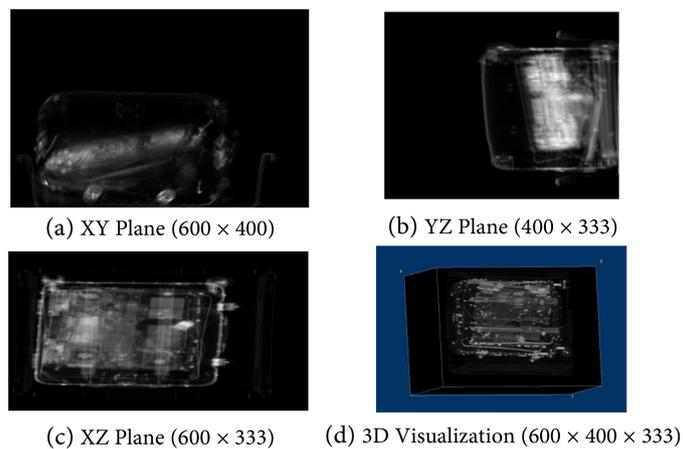


Figure 4. Corresponding three views. (a) Shows the projection along the z-axis, forming the XY plane. (b) Presents the projection along the x-axis, constituting the YZ plane. (c) Represents the projection along the y-axis, creating the XZ plane. (d) Depicts the original three-dimensional data visualized using the software ImageJ in Max Projection mode.

Table 1. Table of corresponding masks.

	Input Image	Instance Segmentation	Mask Extraction
XY Plane			
XY Plane			
YZ Plane			

3.3. Generating Three-Dimensional Masks for Obtaining Three-Dimensional Laptop Data

The Masks obtained from instance segmentation along the three angles are reconstructed along the original projection directions to generate a three-dimensional mask, as depicted in **Figure 5**.

Let x , y , and z denote variables representing the pixel quantities in three dimensions, i , j , and k respectively. The three views correspond to $xy(i,j)$, $yz(j,k)$, and $xz(i,k)$.

Assuming that the Laptop Computer region is assigned a value of 0 within the mask, while non-Laptop Computer regions have a value of 1, the three-dimensional mask must satisfy the condition:

$$xy(i, j) * yz(j, k) * xz(i, k) = 1 \tag{6}$$

This condition ensures the acquisition of the three-dimensional Laptop Computer mask.

By intersecting the three-dimensional mask with the original data, the segmented laptop data can be obtained. As depicted in the images on the right side of **Table 2**, the top section illustrates the three-dimensional mask, while the bottom section portrays the laptop computer after segmentation.

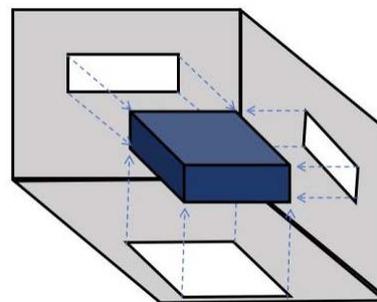
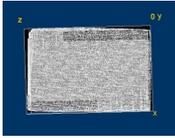


Figure 5. 3D mask principle diagram.

Table 2. Table of laptop computer segmentation.

Input Image	Output Image
 XY Plane	 three-dimensional mask
 XZ Plane	
 YZ Plane	
	 Laptop Computer

4. Experiment and Results Analysis

As illustrated in the above **Figure 6** and **Figure 7**, this approach, combining projection and instance segmentation, successfully achieves the laptop removal effect. In **Figure 6**, representing the original data, the laptop computer is visible within the red-boxed area, indicating it has not been removed. In **Figure 7**, which displays the original data with the laptop computer removed, it is noticeable that the red-boxed area no longer contains the laptop. Moreover, regions previously obscured by the laptop computer are now visible.

4.1. Experimental Environment

Operating on Windows 10 with 16 GB of RAM, the experimental environment includes an Intel Sliver 4210 CPU and an NVIDIA GeForce RTX 2080 Ti GPU. The software stack comprises PyTorch version 1.8, CUDA version 11.1, Visual Studio Community 2019, and OpenCV 4.8.0.

4.2. Experimental Data

In this study, a dataset of security checkpoint CT images containing laptop computers was collected. These image data were sourced from Shanghai Wuying

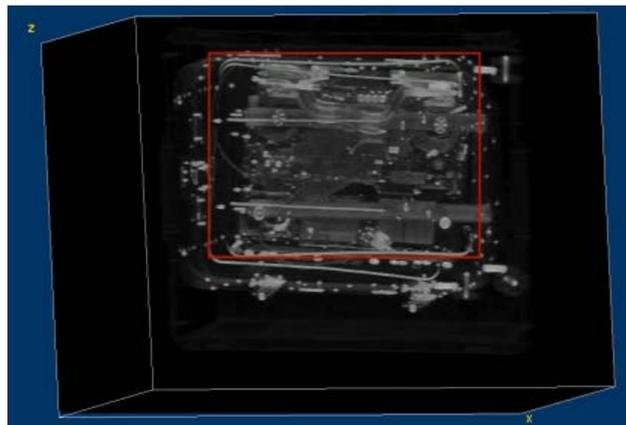


Figure 6. Original data.

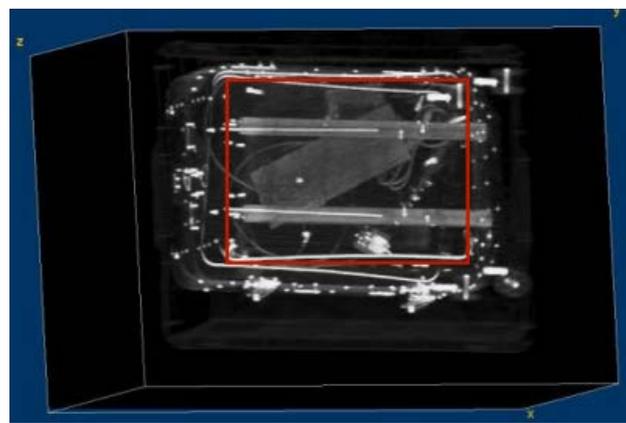


Figure 7. Data after laptop removal.

Technology Co., Ltd., encompassing CT data from 911 distinct parcels, each containing a laptop computer. The dataset covers various laptop brands, models, and sizes, resulting in a total of 2733 images of three different views. For model training and evaluation, 80% of the dataset was allocated for training, while the remaining 20% was used for testing.

Partial presentation of the actual dataset is shown in **Table 3**.

4.3. Experimental Parameter Setting

The network architecture used for model training is YOLOv7, with image scaling size of (640×640) , $\text{conf_thres} = 0.25$, and $\text{iou_thres} = 0.45$.

Table 3. Laptop model and brand.

Laptop Model	Brand
ThinkPad E450	Lenovo
17R5	Alienware
HP Omen	HP—Hewlett-Packard
redmibook pro14	Xiaomi
P15s	Lenovo
matebook14	Huawei
s1-01	Mechrevo
Edge E420	Lenovo
Pavilion	HP—Hewlett-Packard
	Apple
	HP—Hewlett-Packard
	Huawei
2010	Apple
Xiaoxin Pro16	Lenovo
Inspiron15	Dell
XiaoxinAir15	Lenovo
s2-01	Mechrevo
LegionY9000P	Lenovo
G Series	Dell
Zhaoyang E4	Lenovo
Xiaomi	Xiaomi
vostro	Dell
XiaoxinAir14	Lenovo
Honor HunterK16D	Hasee
G15	Dell
FL5900U	HP—Hewlett-Packard

4.4. Experimental Evaluation Metrics

Evaluation metrics are divided into two parts here: 1) Model evaluation metrics 2) Evaluation metrics for segmented laptop computers.

1) Model Evaluation Metrics

The evaluation metrics for this experiment's model involve Precision, Recall, and F1-score, aiming to assess the accuracy and effectiveness of the proposed method. Precision represents the ratio of correctly removed laptop computers to all removed laptop computers, Recall represents the ratio of correctly removed laptop computers to the actual number of existing laptop computers, and F1-score is the harmonic mean of Precision and Recall. TP, FP, TN, FN needs to be introduced here. TP means that the prediction is positive and the result is also positive. FP means that the prediction is positive and the result is negative. TN means that the prediction is negative and the result is negative. FN means that the prediction is negative and the result is positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{F1-score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} = 2 * \frac{\text{TP}}{\text{TP} + \text{FP} + \text{TP} + \text{FN}} \quad (9)$$

Figure 8 represents the Precision-Confidence Curve, **Figure 9** illustrates the Precision-Recall Curve, **Figure 10** displays the Recall-Confidence Curve, and **Figure 11** depicts the F1-Confidence Curve.

As shown in the graphs, when the confidence level is greater than 0.2, the precision approaches 1. This means that as the confidence level increases, the probability of correctly predicting positive samples in the test set also increases.

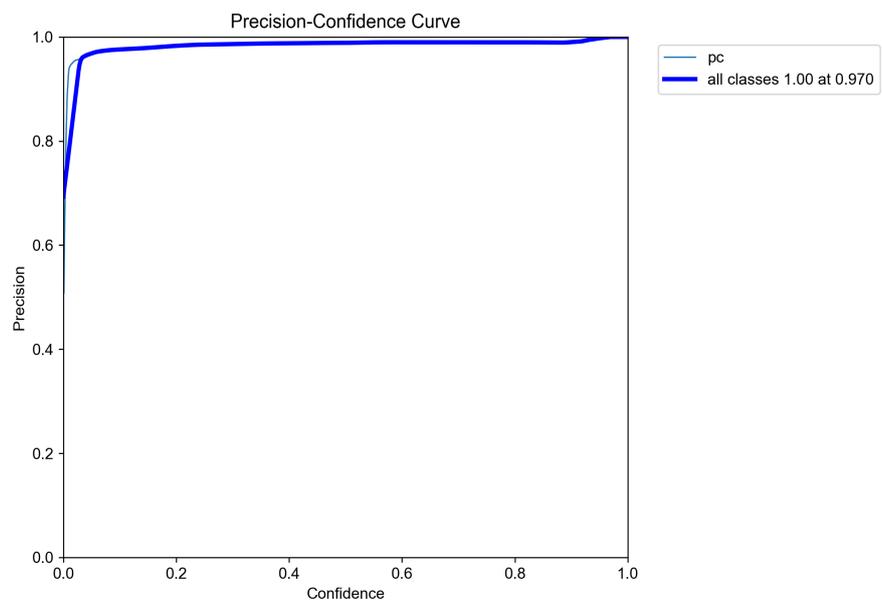


Figure 8. Precision-Confidence curve.

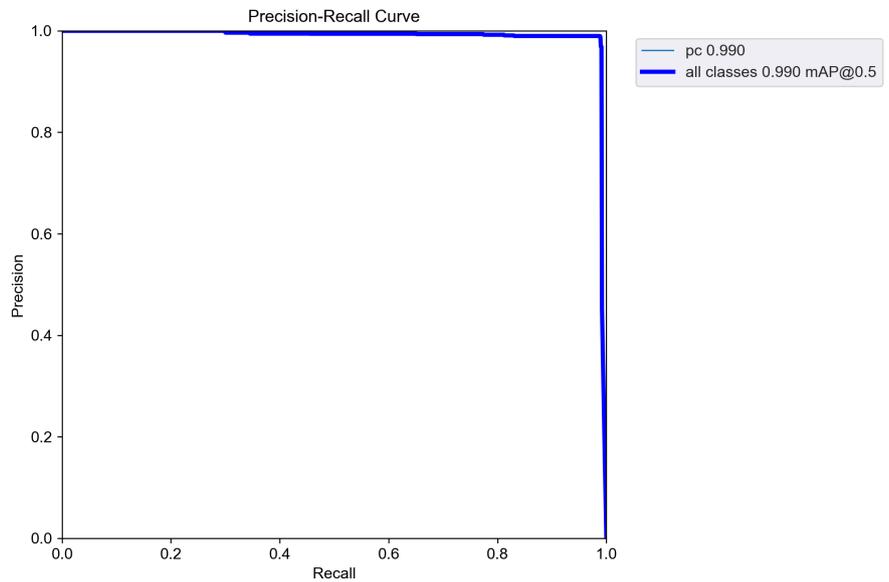


Figure 9. Precision-Recall curve.

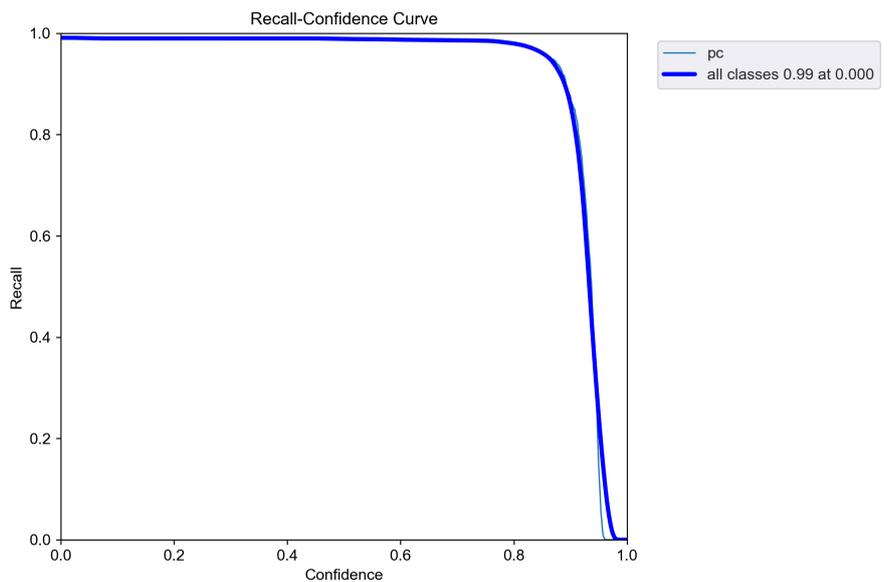


Figure 10. Recall-Confidence curve.

When the confidence level is less than 0.8, the recall approaches 1. In other words, as the confidence level decreases, the probability of correctly predicting all true positive samples in the test set also increases. To comprehensively measure precision and recall, the F1 score is introduced to balance these two metrics. When the confidence level ranges from 0.2 to 0.8, the recall rate can be adjusted. A higher F1 score indicates better model performance.

2) Evaluation metrics for segmented laptop computers primarily rely on visual assessment, categorized into four levels. A-level represents laptop computers that are completely segmented with accurate contours. B-level indicates laptop computers that are segmented but include fewer parts of other objects. C-level

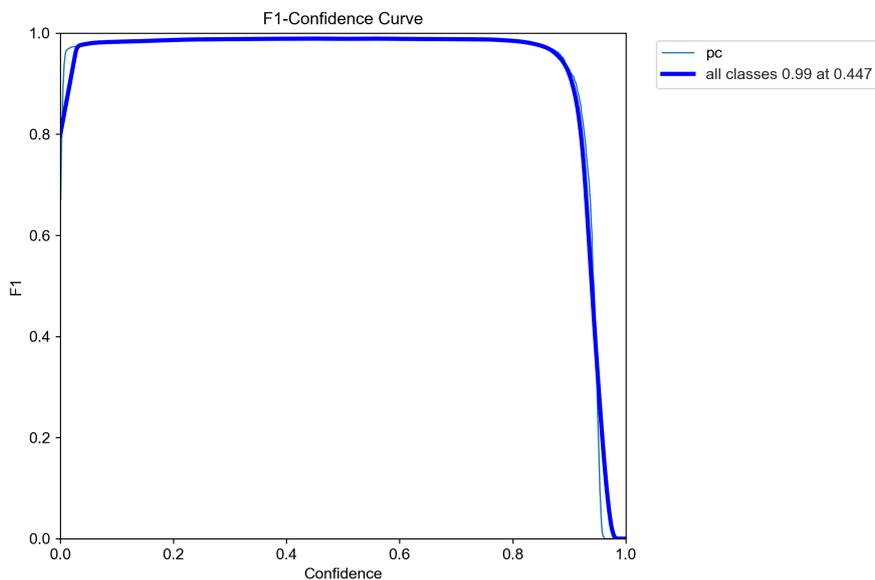


Figure 11. F1-Confidence curve.

Table 4. Segmentation results.

0012251(A)	0012252(A)	0012253(A)	0012254(A)	0012255(A)
0012256(A)	0012257(A)	0012258(A)	0012259(A)	0012260(A)
0012261(A)	0012262(A)	0012263(C)	0012264(A)	0012265(E)
0012266(A)	0012267(A)	0012268(A)	0012269(A)	0012270(A)
0012271(A)				

indicates slight incompleteness at the edges of the segmented laptop computers without significant impact on the overall result. D-level suggests that the seg-

mented laptop computers contains a significant portion of other objects. E-level indicates incomplete segmentation of the laptop computers. ABC cases are considered passing, while DE cases are considered failing.

Finally, an additional set of 21 parcel image data provided by Shanghai Wuying Technology Co., Ltd. was used to validate the proposed method. The segmentation results of this dataset are shown in **Table 4**. In **Table 4**, it can be seen that there are 19 data that meet level A, 0 data that meet level B, 1 data that meet level C, 0 data that meet level D, 1 data that meet level E.

The passing rate is calculated as $(19 + 1)/21 = 0.95\%$, which meets the requirement.

5. Conclusion

Through validation, it has been determined that the approach of integrating projection with the YOLOv7 instance segmentation model can indeed achieve the segmentation of laptop computers, and the accuracy aligns with the requirements of the security inspection system for laptop removal functionality. However, there are still some issues present. For instance, in the dataset of 21 samples, laptop segmentation failures were observed in data with the ID 0012263, where a laptop corner was missing, and in data with the ID 0012265, where laptop segmentation was incomplete. These instances of segmentation failure still have a probability of occurrence. Further optimization is necessary in the future to enhance the maturity of this approach.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/CVPR52729.2023.00721>
- [2] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [3] Zhang, X., Yang, L.R. and Zheng, Z.Y. (2014) The Segmentation Study of High Atomic Number Materials for Dual-Energy X-Ray Baggage Image. *Nuclear Electronics & Detection Technology*, **1**, 20-23.
- [4] Zhang, Y.Z., Song, X.Z. and Wang, Y.M. (2014) X-Ray Security Image Segmentation Algorithm Based on Material. *Equipment Manufacturing Technology*, **4**, 33-35.
- [5] Bai, C., Bai, Z.T., Cha, Y.L. and Zhang, S. (2022) Material Value Transform Segmentation of Overlapped Materials in Dual Energy X-Ray Image. *Nuclear Electronics & Detection Technology*, **10**, 963-968.
- [6] Gu, Z., Wu, W., Gao, Y.F., Wei, D. and Zhou, G.Z. (2009) The Segmentation Method for Testing Images of X-Ray Security Inspection Systems. *Journal of Nonde-*

- structive Evaluation*, **3**, 169-172, 176.
- [7] Zhang, Z.H. (2021) Research and Implementation of Pedestrian Multiple Object Tracking Algorithm Combined with Instance Segmentation. Master Degree Thesis, South China University of Technology.
<https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202301&filename=1021896001.nh>
- [8] Chen, J.X. (2019) Fastener Looseness Recognition of Medium-Low Speed Maglev Contact Rail Based on Instance Segmentation. Master Degree Thesis, Southwest Jiaotong University.
<https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202001&filename=1019951987.nh>
- [9] He, J.Q., Li, R.H., Li, H., Liao, Y.L., Gong, B., Hao, Y.P., Liang, W., Wu, J.R. and Wen, Y. (2023) Visible Light Image Automatic Recognition and Segmentation Method for Overhead Power Line Insulators Based on Yolov5 and Grabcut. *Southern Power System Technology*, **6**, 128-135.
- [10] Yu, B., Ma, S.H., Li, H.Y., Li, C.G. and An, J.B. (2020) Real-Time Pedestrian Detection for Far-Infrared Vehicle Images and Adaptive Instance Segmentation. *Laser & Optoelectronics Progress*, **2**, 293-303.
- [11] Yang, K.H. and Zhang, Y. (2023) Traffic Scene Object Detection and Segmentation Algorithm Based on YOLOv5. *Changjiang Information & Communications*, **4**, 48-50.
- [12] Hurtik, P., Molek, V., Hula, J., Vajgl, M., Vlasanek, P. and Nejezchleba, T. (2022) Poly-YOLO: Higher Speed, More Precise Detection and Instance Segmentation for YOLOv3. *Neural Computing and Applications*, **34**, 8275-8290.
<https://doi.org/10.1007/s00521-021-05978-9>
- [13] Patro, S.G. and Sahu, K.K. (2015) Normalization: A Preprocessing Stage. *International Advanced Research Journal in Science, Engineering and Technology*, **2**.
<https://doi.org/10.17148/IARJSET.2015.2305>