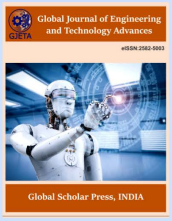
Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/



(RESEARCH ARTICLE)

Check for updates

Threat object detection and analysis for explosive ordnance disposal robot

Reagan Leoncio Galvez [1,*] and Elmer Pamisa Dadios [2]

[1] *Electronics Engineering Department, Bulacan State University, Malolos, Philippines.*
[2] *Manufacturing Engineering and Management Department, De La Salle University, Manila, Philippines.*

Global Journal of Engineering and Technology Advances, 2022, 11(01), 078–087

Publication history: Received on 20 March 2022; revised on 24 April 2022; accepted on 26 April 2022

Article DOI: https://doi.org/10.30574/gjeta.2022.11.1.0074

Abstract

Explosive Ordnance Disposal (EOD) robots are useful in military applications like the safe disposal of explosives. However, many of these robots do not have the capability to identify threat objects using their onboard vision system due to data unavailability for training an improvised explosive device (IED) detector. As a solution, this study used image processing and object detection algorithms to detect and analyze threat objects inside the baggage. A threat object detector was developed and composed of two separate modules such as baggage detection and IED detection and analysis modules. The experiments showed that baggage detection achieved 22.82% mean average precision (mAP) using Single Shot Detector (SSD) in the Microsoft Common Objects in Context (COCO) dataset, while IED detection achieved 77.59% mAP using Faster R-CNN in the X-ray dataset. The threat objects from the X-ray image were also analyzed using image processing techniques to get the dimension of the object and the distance from a reference object. Also, the baggage detection module was successfully deployed in Jetson TX2, which runs at a frame rate of 12 frames per second (FPS).

Keywords: Baggage Detection; Computer Vision; Explosive Ordnance Disposal; Image Processing; Threat Objects

1. Introduction

Explosive Ordnance Disposal (EOD) robot, also known as a bomb disposal robot, is a robot that is used to disarm or dispose of improvised explosive devices (IEDs) and other hazardous materials by controlling it at a safe distance. It is crucial in military applications because instead of bringing a human to complete a dangerous task, robots can be assigned to do this without risking any life. However, several bomb disposal robots such as [1, 2, 3, 4] do not have the capability to analyze and detect threat objects. It always depends on the expert to decide whether a given object poses a threat or not. Currently, researchers have already developed a mobile robot that has object recognition capabilities [5, 6]. This capability can be useful in some other tasks; however, there are few studies that are concerned with the detection of threat objects for EOD robots. Most of them focus on the specific part of the robot, like the robot arm [7, 8, 9], instead of its vision system, which is also a vital part of the robot design. The vision system serves as the guide of the robot operator in navigating the environment and in detecting unknown objects.

Threats can be identified using computer vision [10], unintended radiated emission (URE) [11], ground-penetrating radar [12], [13], and autoencoder [14]. However, identifying threats using computer vision is challenging due to the lack of available data needed to implement the task. In this study, a threat object detector is proposed that can be used to detect potential threats and aid human experts in examining unknown objects. The authors concentrated on the vision system of the robot because this part plays an important role in explosive detection and disposal.

* Corresponding author: Reagan Leoncio Galvez
Bulacan State University, Electronics Engineering Department, Malolos, Philippines.

2. Material and methods

The threat object detector in this study consists of two separate modules, such as the baggage detection module and the IED detection and analysis module, as shown in the threat object detector framework in Figure 1. The inputs are the live camera feed from the Jetson TX2 and the X-ray image from the X-ray machine. Two different object detection architectures were used in these modules, such as SSD and Faster R-CNN. The outputs are predictions about the classes of baggage and IED components.

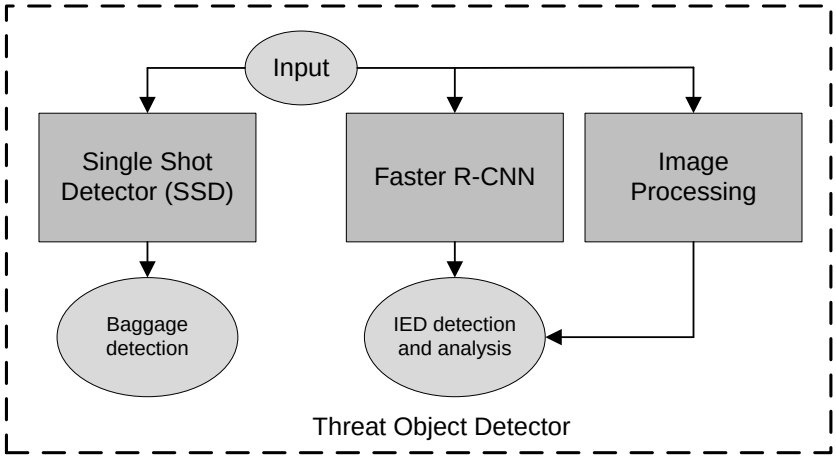


Figure 1 Threat object detector framework

2.1. Single Shot Detector (SSD)

SSD was selected for the detection of baggage (e.g., backpack, handbag, suitcase). This method is particularly useful in the real-time detection of objects due to its fast inference speed. SSD was introduced by [15] to address the slow inference problem of previous object detectors such as R-CNN [16], Fast R-CNN [17], and Faster R-CNN [18]. In SSD, both localization and detection are performed in a single forward pass of the network. This is similar to YOLO by [19], which scans the image only once during inference. The advantage of SSD is that it eliminates bounding box proposals and subsampling of pixels [15]. Figure 2 shows the SSD architecture. SSD initially used VGG-16 [20] as a base network, but in this paper, MobileNet by [21] was used to extract low-level features from the input image. It is followed by several convolutional layers that decrease in size to allow detection at multiple scales. Each convolutional layer is connected to the output fully connected layer, followed by a non-maximum suppression to prune the excess bounding box and obtain final detections.

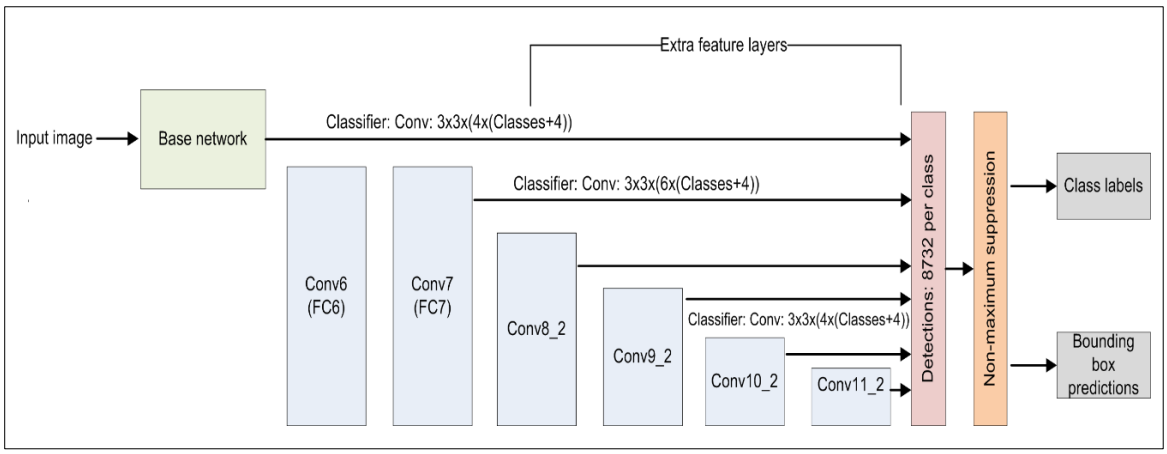


Figure 2 SSD architecture

2.2. Faster R-CNN

In this study, Faster R-CNN was chosen for IED detection because of its state-of-the-art performance in object detection tasks. Figure 3 shows Faster R-CNN architecture. ResNet-101 by [22] was used as a base network to extract features

from the input image. Faster R-CNN consists of two primary modules, namely Region Proposal Network (RPN) and Region of Interest (ROI) pooling. RPN accepts anchor boxes (bounding boxes with different scales and aspect ratios) and determines their objectness. The output of the RPN is bounding box proposals. On the other hand, the ROI pooling module accepts all the proposals from the RPN and extracts its ROI features from the feature map. The features are then resized before sending them to the fully connected (FC) layers. The final output is class labels and bounding box predictions.

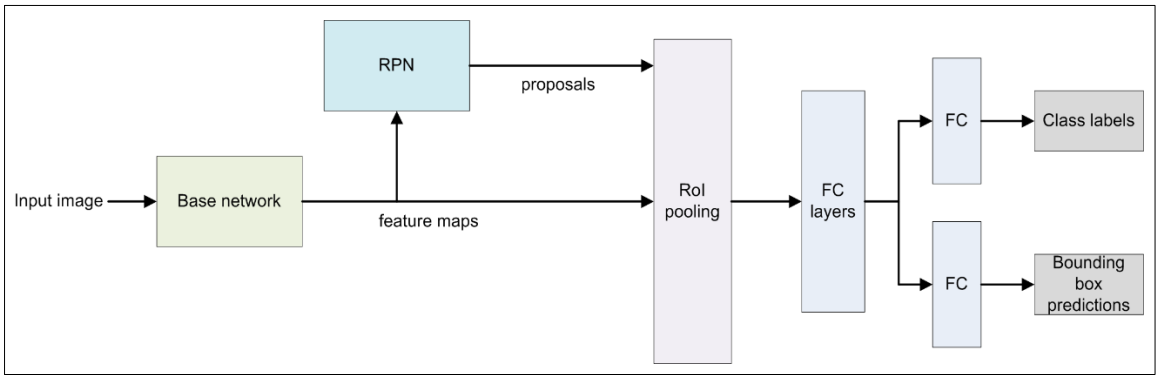


Figure 3 Faster R-CNN architecture

2.3. Baggage Detection Dataset

The images used in the study came from the MS COCO dataset (2014 train-val). Three classes were chosen such as a backpack, handbag, and suitcase. The dataset was collected individually using a Python script by looking for each category ID and downloading the image that matches the desired ID, e.g., ['backpack']. Then, annotations were downloaded using another script. It contains extensible markup language (XML) files of each image. After this, the XML files were converted to a comma-separated value (CSV) format to extract the ground truth bounding boxes and finally converted to TFrecord file format for efficient data storage. The baggage dataset was divided into train and test data. Training data consists of 8,461 images (90%), while testing data consists of 960 images (10%). The total number of annotations (ground truth bounding boxes) was 21,466 for both training and testing. Figure 4 shows the sample dataset with labeled ground truth bounding boxes.

Figure 4 Baggage dataset from MS COCO [23]

2.4. IED Detection and Analysis Dataset

IEDXray dataset by [24], [25] was used in IED detection and analysis. Figure 5 shows the sample dataset with labeled ground truth bounding boxes. These are composed of X-ray images of IED replicas scanned from an X-ray machine. Because these are replicas, the explosive material is not present in the IED. The study only focuses on three IED components such as battery, mortar, and wires.

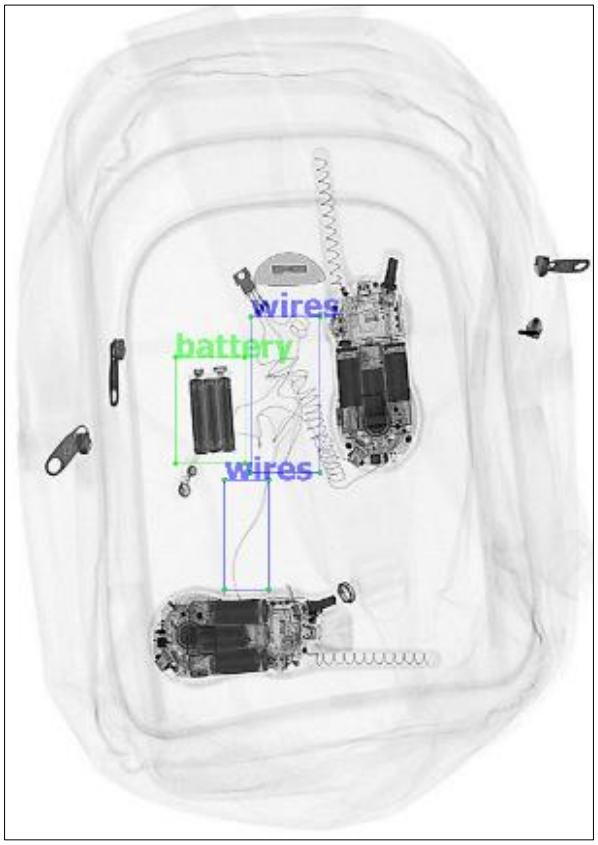


Figure 5 IED Xray dataset

The data was annotated manually using an image labeling tool. The annotations were also converted from XML file to TFrecord format, as discussed previously in the baggage dataset. Training data consists of 1,209 images (90%), while testing data consists of 134 images (10%). The total number of annotations for this dataset was 3,939 for both training and testing. The summary of the datasets used in the study is shown in Table 1.

Table 1 Dataset summary

	Baggage			**IED Xray**		
	annotations					
	backpack	handbag	suitcase	battery	mortar	wires
Training	6200	8778	4251	1159	529	1872
Testing	728	946	563	158	29	192
	images					
Training	8461			1209		
Testing	960			134		

2.5. Training and Evaluation

The baggage detector was trained using RMSprop with exponential decay learning rate. The values of the hyperparameter are the following: initial learning rate = 0.001, decay steps = 5000, decay factor = 0.8, momentum = 0.9, decay = 0.9, epsilon = 1, batch size = 10, number of steps = 100,000. Data augmentation was also applied such as horizontal flip, and random crop. In contrast, the IED detector was trained using stochastic gradient descent (SGD) with momentum. The values of the hyperparameter are the following: learning rate = 0.0003, momentum = 0.9, batch size = 1.

Both detectors were evaluated using mean average precision (mAP) following the PASCAL VOC metric by [26] at 0.5 intersection over union (IOU) threshold defined using Eq. (1). Where X_G is the ground truth bounding box, X_P is the predicted bounding box, and A is the area.

$$IOU = \frac{A(X_G \cap X_P)}{A(X_G \cup X_P)} \tag{1}$$

2.6. IED Analysis

IED analysis consists of measuring the size and the distance between objects. The contour in an image was used to analyze the X-ray image [27]. Contours are curves joining all the consecutive points that have the same color and intensity. It is useful for tasks like shape approximation and analysis [28]. Image binarization was applied using a Canny edge detector in order to find the contour. Then, morphological transformations were applied, specifically erosion and dilation. Erosion erodes away the boundaries of the foreground object, therefore shrinking the foreground and removing small white noise in the image. On the other hand, dilation is the opposite of erosion. It increases the size of the foreground object. Dilation followed by erosion was applied to close gaps in between the object edges.

The first step in measuring the size of the object is to sort the contours in the image from left to right. This ensures that the left-most contour corresponds to the reference object. Then, each contour was examined. If the contour size is small, this is discarded to avoid noises in the image. After examining the contours, the rotated bounding box is drawn around the object. Then, the midpoints M in Eq. (2) and Euclidean distances d between midpoints in Eq. (3) are computed. Finally, the pixel per metric P in Eq. (4) is used to obtain the object size. Where w_o is the width of the object in pixels, while w_r is the known width of the reference object in millimeters (mm). The same idea was used to measure the distance from a reference object to another object, except that the Euclidean distance between the reference object and the object location was calculated instead of the Euclidean distance between midpoints previously [29].

$$M = \frac{a_1 + a_2}{2}, \frac{b_1 + b_2}{2} \tag{2}$$

$$d(a,b) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2} \tag{3}$$

$$P = \frac{w_o}{w_r} \tag{4}$$

3. Results and discussion

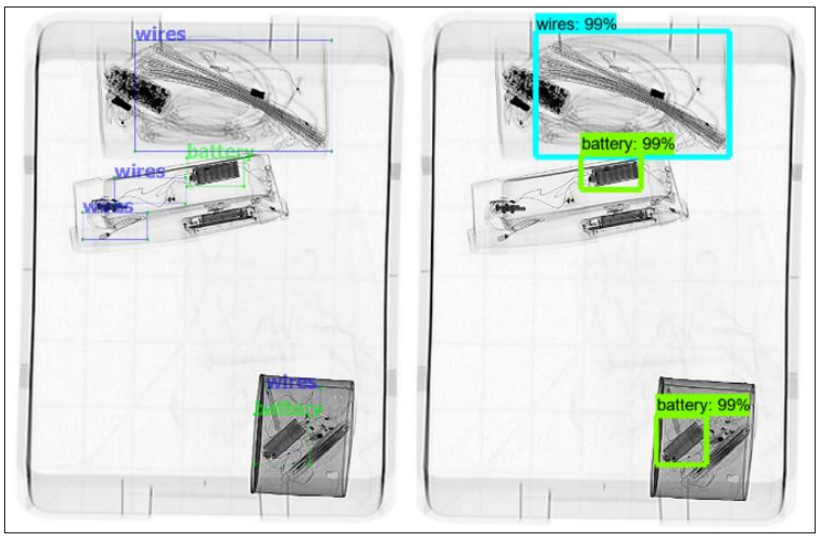
3.1. Baggage Detection

Figure 6 shows the detection output using SSD in the Jetson TX2 development kit. The 5MP fixed focus MIPI CSI camera of the Jetson TX2 was used to test the performance of the baggage detector. As can be seen in the figure, the two backpacks and the suitcase were detected successfully. However, false positives are still present, especially if the object or the camera is moving. Regarding the evaluation performance, SSD achieves 22.82% mAP on the test data and runs on the average frame rate of 12 FPS on Jetson TX2.

Figure 6 Sample detection on Jetson TX2 development kit using SSD

3.2. IED Detection and Analysis

The comparison between the ground truth bounding box and predicted bounding box using Faster R-CNN in the detection of IED components is shown in Figure 7. As can be seen from the figure, the model successfully detected the threat objects (2 batteries and 1 wire) but failed to detect the other objects due to occlusion. The model achieved 77.59% mAP in all categories using 900×1536 resolution of the training images. The Faster R-CNN took 208.96 milliseconds (ms) per image to evaluate the test data. The test data can be evaluated faster by decreasing the resolution of the input image. However, there is a tradeoff in the mAP.



(a) (b)

Figure 7 Sample IED detection using Faster R-CNN: (a) ground truth bounding box; (b) predicted bounding box

Figure 8 shows the sample image analysis of the X-ray dataset. The right side of the figure shows the resulting edges of the image using a Canny edge detector. The left side of the figure shows the measurements of the object. Object measurements such as area = 1,130.13 mm, contour number = 23, length = 47.22 mm, and width = 23.94 mm are displayed in the top left corner of the image. The contour number represents the contour ID of the object, which is unique in every given image. There were 31 contours found in the image, but only 5 objects were measured, while small objects were discarded. In the image, the measured object is a battery.

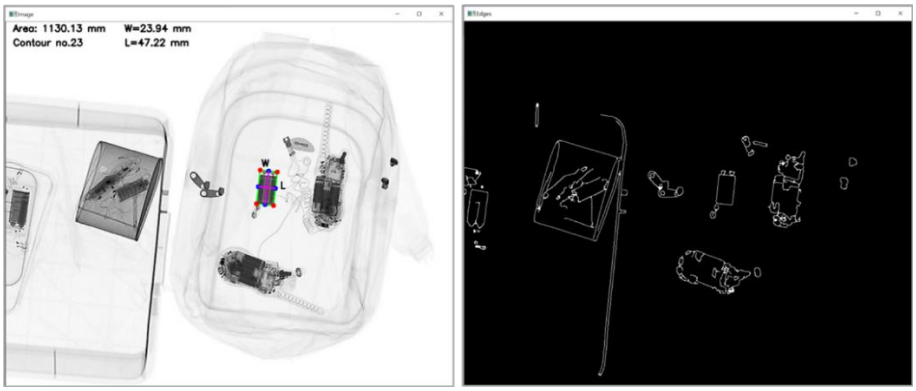


Figure 8 Analysis of IED components (battery)

Figure 9 shows another sample image analysis in a cellphone object showing its area = 4,941.71 mm, contour number = 29, length =107.32 mm, and width = 46.05 mm. This information from the X-ray image is useful in estimating the possible blast radius of the IED and identifying the components used.

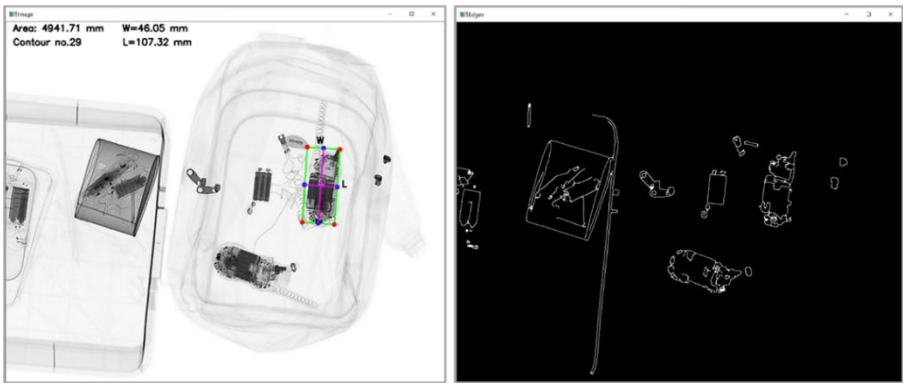


Figure 9 Analysis of IED components (cellphone)

In the measurement of the distance between a reference object to another object, the sample image analysis of the X-ray dataset is shown in Figure 10. The right image shows the edges detected on the X-ray image, while the left image shows the object measurements in millimeters (mm). There was a total of 31 contours found in the image, but only 4 objects were analyzed. The reference object is a battery, which is the left-most object in the image. The distance between the respective corners and centroids of the object was measured using the Euclidean distance defined previously in Section 2. The centroid to centroid distance between the battery and cellphone measures approximately 430.33 mm (the value of E in the figure). While the corner to corner distance is A = 422.62 mm, B = 444.25 mm, C = 439.44 mm, and D = 418.96 mm.

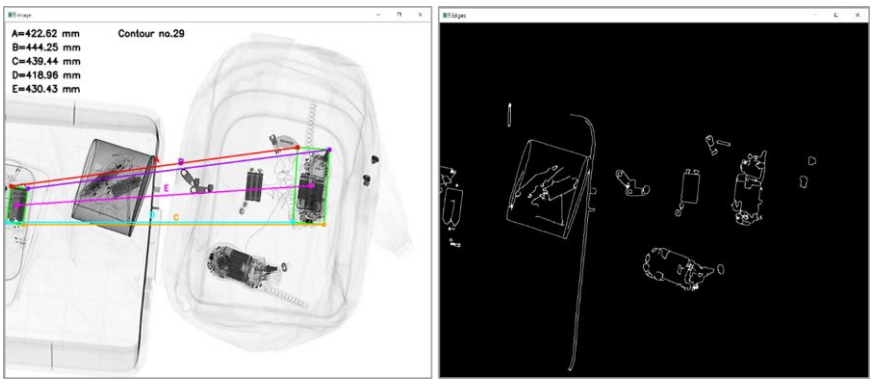


Figure 10 Distance measurement between battery and cellphone

Figure 11 shows another measurement using a different X-ray dataset. The right image shows the edges detected on the X-ray image, while the left image shows the object measurements in millimeters (mm). A total of 21 contours were found in the image, but only 2 objects were analyzed. The reference object is a cellphone, which is again the left-most object in the image. The centroid to centroid distance between the cellphone and battery is approximately 25.96 mm (the value of E in the figure). On the other hand, the corner to corner distance is A = 23.44 mm, B = 25 mm, C = 29.70 mm, and D = 28.36 mm. This information is important to know the exact location of the objects with reference to other objects so that the operator of the EOD robot can easily disable the IED circuitry.

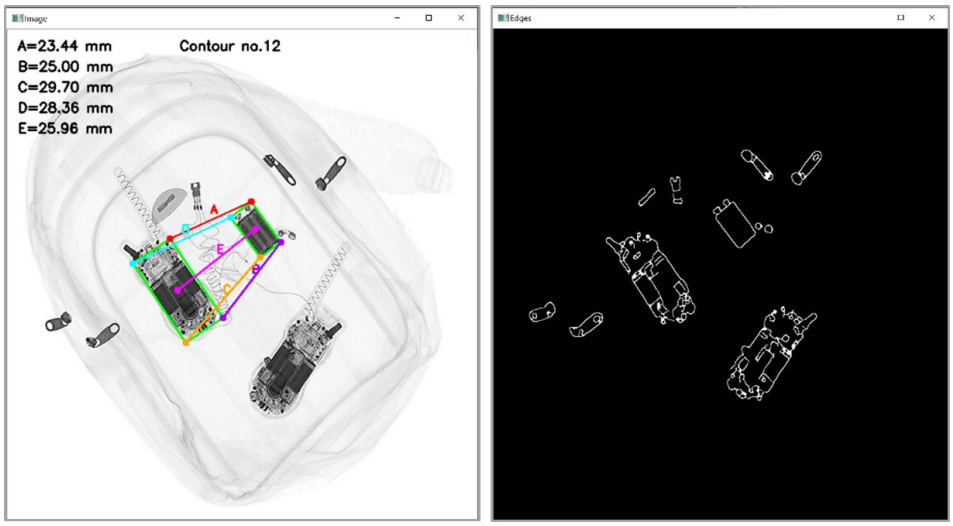


Figure 11 Distance measurement between cellphone and battery

4. Conclusion

The aim of the research was to develop a threat object detector for the EOD robot. This study has shown the possibility of using object detection algorithms and image processing to identify and analyze threat objects. One of the algorithms, i.e., the SSD, was successfully deployed in Jetson TX2 that can detect three types of baggage. In addition, the Faster R-CNN was able to identify IED components in an X-ray image correctly. These results will help experts in decision-making whether an unknown object poses a threat or not. The major limitation of this study is that the threat object detector can only be used effectively if the image is not heavily occluded. Therefore, future research should focus on minimizing the effect of occlusion in identifying threat objects and improving the detection performance of the threat object detector. Additional training data is recommended to improve the mAP results reported in this study.

Compliance with ethical standards

Acknowledgments

This research was supported by the Department of Science and Technology (DOST). We would also like to show our appreciation to the Explosive Ordnance Disposal and Canine Group (EOD-K9) of the Philippine National Police (PNP) for providing the data needed to finish this study.

Disclosure of conflict of interest

The authors affirm that the study was done without any financial or commercial relationships that could be interpreted as a potential conflict of interest.

References

[1] Su Y, Wang T, Zhang K, Yao C, Wang Z. Adaptive nonlinear control algorithm for a self-balancing robot. IEEE Access [Internet]. 2020; 8:3751–60. Available from: https://ieeexplore.ieee.org/document/8945362/

[2] Sabarivani A, Anbarasi A, Vijayaiyyappan A, Sindhuja S. Wireless synchronization of robotic arm with human movements using arduino for bomb disposal. In: 2018 International Conference on Smart Systems and Inventive

Technology (ICSSIT) [Internet]. IEEE; 2018; p. 229–34. Available from: https://ieeexplore.ieee.org/document/8748690/

[3] Ibarra A, Condor A, Martinez P, Tipan E. Control reengineering used for rehabilitation of Andros Remotec bomb disposal robot. In: 2020 IEEE ANDESCON [Internet]. IEEE; 2020; p. 1–6. Available from: https://ieeexplore.ieee.org/document/9272161/

[4] Ismail RM, Muthukumaraswamy S, Sasikala A. Military support and rescue robot. In: 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) [Internet]. IEEE; 2020; p. 156–62. Available from: https://ieeexplore.ieee.org/document/9121041/

[5] Islamgozhayev T, Kalimoldayev M, Eleusinov A, Mazhitov S, Mamyrbayev O. First results in the development of a mobile robot with trajectory planning and object recognition capabilities. Open Eng. 2016 Nov 2; 6(1):347–52.

[6] Yamauchi B, Moseley M, Brookshire J. LABRADOR: a learning autonomous behavior-based robot for adaptive detection and object retrieval. In: Proc SPIE 8662, Intelligent Robots and Computer Vision XXX: Algorithms and Techniques [Internet]. 2013; p. 1–8. Available from: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2011834

[7] Deng S, Cai H, Li K, Cheng Y, Ni Y, Wang Y. The design and analysis of a light explosive ordnance disposal manipulator. In: 2018 2nd International Conference on Robotics and Automation Sciences (ICRAS) [Internet]. IEEE; 2018; p. 1–5. Available from: https://ieeexplore.ieee.org/document/8443234/

[8] Islam AJ, Alam SS, Ahammad KT, Nadim FK, Barua B. Design, kinematic and performance evaluation of a dual arm bomb disposal robot. In: 2017 3rd International Conference on Electrical Information and Communication Technology (EICT) [Internet]. IEEE; 2017; p. 1–6. Available from: http://ieeexplore.ieee.org/document/8275193/

[9] Oke AO, Afolabi A. Development of a robotic arm for dangerous object disposal. In: 2014 6th International Conference on Computer Science and Information Technology (CSIT) [Internet]. IEEE; 2014; p. 153–60. Available from: http://ieeexplore.ieee.org/document/6805994/

[10] Griffin LD, Caldwell M, Andrews JTA, Bohler H. "Unexpected item in the bagging area": anomaly detection in X-Ray security images. IEEE Trans Inf Forensics Secur 2019 Jun; 14(6):1539–53.

[11] Friedel JE, Holzer TH, Sarkani S. Development, optimization, and validation of unintended radiated emissions processing system for threat identification. IEEE Trans Syst Man, Cybern Syst. 2020 Jun; 50(6):2208–19.

[12] Malof JM, Reichman D, Karem A, Frigui H, Ho KC, Wilson JN, et al. A large-scale multi-institutional evaluation of advanced discrimination algorithms for buried threat detection in ground penetrating radar. IEEE Trans Geosci Remote Sens. 2019 Sep; 57(9):6929–45.

[13] Moalla M, Frigui H, Karem A, Bouzid A. Application of convolutional and recurrent neural networks for buried threat detection using ground penetrating radar data. IEEE Trans Geosci Remote Sens. 2020 Oct; 58(10):7022–34.

[14] Bestagini P, Lombardi F, Lualdi M, Picetti F, Tubaro S. Landmine detection using autoencoders on multipolarization GPR volumetric data. IEEE Trans Geosci Remote Sens. 2021 Jan; 59(1):182–95.

[15] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. SSD: single shot multibox detector. In: European conference on computer vision [Internet]. 2016; p. 21–37. Available from: http://link.springer.com/10.1007/978-3-319-46448-0_2

[16] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition [Internet]. IEEE; 2014; p. 580–7. Available from: http://ieeexplore.ieee.org/document/6909475/

[17] Girshick R. Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision (ICCV) [Internet]. IEEE; 2015; p. 1440–8. Available from: http://ieeexplore.ieee.org/document/7410526/

[18] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017 Jun 1; 39(6):1137–49.

[19] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; p. 779–88.

[20] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv Prepr arXiv14091556 [csCV] [Internet]. 2014 Sep 4; Available from: http://arxiv.org/abs/1409.1556

[21] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv Prepr arXiv170404861 [csCV] [Internet]. 2017 Apr 16; Available from: http://arxiv.org/abs/1704.04861

[22] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; p. 770–8.

[23] Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: common objects in context. In: European conference on computer vision. 2014; p. 740–55.

[24] Galvez RL, Dadios EP, Bandala AA, Vicerra RRP. Object detection in X-ray images using transfer learning with data augmentation. Int J Adv Sci Eng Inf Technol. 2019; 9(6):2147–53.

[25] Galvez RL, Dadios EP, Bandala AA, Vicerra RRP. Threat object classification in X-ray images using transfer learning. In: 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication, and Control, Environment and Management (HNICEM) [Internet]. IEEE; 2018; p. 1–5. Available from: https://ieeexplore.ieee.org/document/8666344/

[26] Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A. The pascal visual object classes (voc) challenge. Int J Comput Vis. 2010; 88(2):303–38.

[27] Rosebrock A. Measuring size of objects in an image with OpenCV [Internet]. Pyimagesearch. 2016 [cited 2019 Dec 23]. Available from: https://www.pyimagesearch.com/2016/03/28/measuring-size-of-objects-in-an-image-with-opencv/

[28] Rosebrock A. Practical Python and OpenCV + Case Studies: An Introductory, Example Driven Guide to Image Processing and Computer Vision. 3rd ed. PyImageSearch; 2016.

[29] Rosebrock A. Measuring distance between objects in an image with OpenCV [Internet]. Pyimagesearch. 2016 [cited 2019 Dec 23]. Available from: https://www.pyimagesearch.com/2016/04/04/measuring-distance-between-objects-in-an-image-with-opencv/