International Journal of Application on Sciences, Technology and Engineering
(IJASTE)
Volume 1, Issue 3, 2023 ISSN:2987-2499

Distance and Accuracy in Object Detection Based on YOLOv8 Computer Vision Algorithm

Vinsensius Reinard[1], Yulius Kristianto[2], Meirista Wulandari[3*]

[1] Electrical Engineering Department, Universitas Tarumanagara, Jakarta, Indonesia
Email: vinsensius.525200005@stu.untar.ac.id
[2] Electrical Engineering Department, Universitas Tarumanagara, Jakarta, Indonesia
Email: yulius.525200008@stu.untar.ac.id
[3] Electrical Engineering Department, Universitas Tarumanagara, Jakarta, Indonesia
Email: meiristaw@ft.untar.ac.id

Submitted: 05-04-2023, Revised: 06-03-2023y, Accepted: 08-03-2023

ABSTRACT

Artificial intelligence is on the rise and has undergone massive growth in the industry, especially in computer vision. The emergence of computer vision from autonomous cars, robotics, surveillance, and many more has led to challenge Artificial intelligence's confidence accuracy in detecting an object. Many artificial intelligence algorithms are used by the industry, one of them is You Only Look Once version 8 (YOLOv8). YOLOv8 is a deep-learning model for object detection. YOLOv8, which is developed by Joseph Redmon and Ali Farhadi is a powerful method to detect an object in real time because YOLOv8 has the capability of processing high-resolution images at high speeds. The research discusses about accuracy YOLOv8 in detecting object with a certain distance from far to very close distance. The dataset is used to train the model of YOLOv8. The dataset is collected by taking photos of an object with constant lighting but different distances. This research aims to obtain the most effective distance that the YOLOv8 computer vision algorithm model could detect. The hipothessis is there is connection between distance and detection accuracy of YOLOv8. If the distance increases, the detection accuracy decreases. However, if the object is close the detection accuracy increases. So based on the results, a conclussion could be concluded that a YOLOv8 model would have the highest accuracy at a certain distance.

Keywords: *computer vision, detection, object, YOLOv8, accuracy*

1. INTRODUCTION

Artificial intelligence (AI) is expanding very rapidly. Advancement on the last two decade has been shaping the way we live. Application of Artificial intelligence is broad, one of them being object detection in computer vision. The capability of detection has challenges AI's algorithm to detect more accurately than ever before [1] [2] [3]. There are many algorithms developed by teams of people with multitude of purpose. Computer vision algorithms were aimed to detect, classify, and segmented objects based on the dataset [4]. You Only Look Once (YOLO) algorithm, later develop into You Only Look Once version 8 (YOLOv8) was designed to detect and classify object. YOLOv8 is design for application in high speed and fast processing which is ideal for application for autonomous vehicle deployment was proposed to only use camera for decision-making

Ultralytics released YOLOv8 in January 2023. As of writing this paper, the authors, Joseph Redmon and Ali Farhadi, have not provided a technical explanation of how YOLOv8 works [5]. Nevertheless, the authors stated that YOLOv8 builds on the YOLO algorithm.

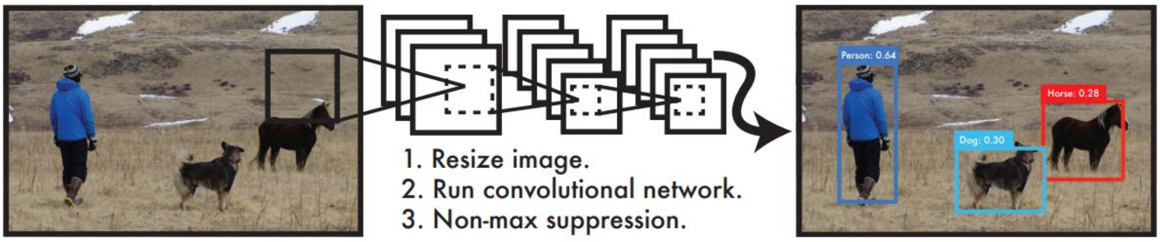

Figure 1 YOLO Diagram
Source: [6]

Figure 1 describes YOLO image processing. The processing is simple and straight forward. The YOLO system resizes the input of image to a 448 × 448. The resized image then processed by a single convolutional network. The resulting detections was evaluated by the model's accuracy. [6]

Related Works

The YOLO is an algorithm that is designed to classify an object in a specific detection time for real-time application while keep maintaining the accuracy. Since 2015, YOLO algorithm has been developed to many versions while principle has been same for each version but the efficiency of detection speed has been improved consistently.

The original version of YOLO or YOLOv1 had 448 × 448 × 3 as size of the input image and applied GoogLeNet to calculate the value of feature maps. At the end of the maps there are two fully connected layers that produce the output, and direct regression was employed to reduce computation and to minimize processing time. YOLOv2 as an enhance version of YOLO, which had been improved to detect small objects. It used DarkNet-19. YOLOv3 which is the further improved version of YOLOv2 adopted Feature Pyramid Network (FPN) to obtain an advanced feature map in order to detect object. It used DarkNet-53. YOLOv4 which the advancement of YOLOv3, applied Cross Stage Partial (CSP) and DarkNet-53 to create CSPDarkNet-53, The CSPDarkNet-53 facilitated richer gradient fusion information, in order to expand processing field Spatial Pyramid Pooling (SPP). YOLOv5 the next version of YOLOv4 used Spatial Pyramid Pooling - Fast (SPPF) method instead of SPP, also introduced a focus module to improve performance. YOLOv6 is the next generation after YOLOv5, which in terms of accuracy and speed, the single-path structure in YOLOv6-N outperforms the multi-branch structure. YOLOv7 as an improved version of YOLOv6, is able to accurately plot model scaling ideas to reduce computational costs and also increase inference speed. Based on the object map pyramid experiment, to jointly predict YOLOv8 uses multiple pyramids to the object detection results, and is able to directly connect additional heads to the pyramids in the middle layer to obtain the results [7] [8]

These advancements have made the YOLO series algorithms a widely used and popular object detection model for real-world applicationsClick or tap here to enter text.. As a little comparison YOLOv5 and the recent version of YOLO that is YOLOv8 simulated on the original dataset, the performance was compared between them. The comparison result showed those architectures used weighting per class techniques, which are comparable to the ones utilized in prior research, to inform the problem of class imbalance [9]. All of this kind of works in accuracy has been researched in due to massive growth in computer vision object detection in the industry. Research ranging from robotics object detection for road safety Click or tap here to enter text., and traffic management Click or tap here to enter text.

YOLOv8 as an advancement method in computer vision technology, has been deployed in the transportation system. It has made traffic video analytics became a crucial component. Automated traffic management systems applied computer vision techniques. The state-of-the-art YOLOv8 is employed to reduce the tediousness of manual monitoring of video stream from security cameras. The crucial step in the traffic management systems is object detection. Some research has been conducted in traffic scenes to identify objects. The techniques are grouped into subcategories, and the benefits and drawbacks of each approach are examined [10]. Additionally, this also discusses the main obstacles, restrictions, and prospective solutions, as well as future directions. YOLO is utilized to identify cars in a variety of conditions for traffic detection and provides output accordingly.

2. RESEARCH METHOD

The block diagram of the research method is displayed in the figure 2. The block diagram includes input, pre-process, process and output. The input is images of objects at distance a known distance X. Pre-process converts the images from Red Green Blue colour scheme to grayscale. Algorithm used to transform image from RGB to grayscale is included in opencv 4.7.0 [11].

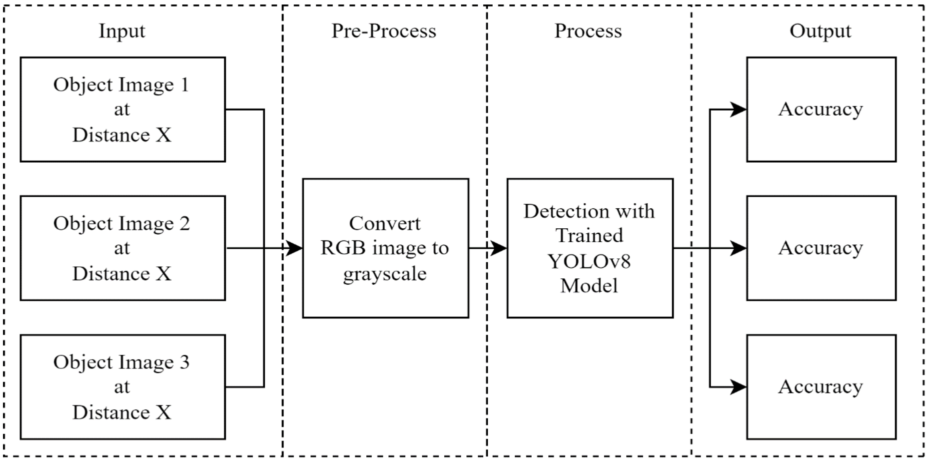

Figure 2 Block Diagram of the Research Method
Source: own photo

The distance used in the research was from 5m to 20cm with -20cm interval. The image was taken and then processed to grayscale image. The greyscale images are used because the model is intended to process high resolution images at high speeds. Greyscale images has the information of RGB images but required less processing power and time for the model to process [12].

Then grayscale images were processed using the YOLOv8 model. The YOLOv8 model was trained with a custom dataset. The model used in this research was trained with 45 epochs. The YOLOv8 model then output object detected and accuracy. The YOLO architecture is based on the Convolutional Neural Network (CNN). CNN does not have significant improvement after training more than 45 epochs.

Traning Datasets and Sample Objects

The objects used in this research were cylinder (figure 3), ball (figure 5), and box (figure 7). Different shapes are used to compare accuracy based on the object. The transformation for cylinder, ball, and box are shown in figure 4, 6, and 8 respectively.

The first object used in the research was a solid cylinder, characterized by its height and diameter measurements. The cylinder's dimensions were 30cm in height and 11.3cm in diameter. The original colour of the cylinder, prior to any transformations, was yellow.


Figure 3 cylinder object before transformation
Source: own photo

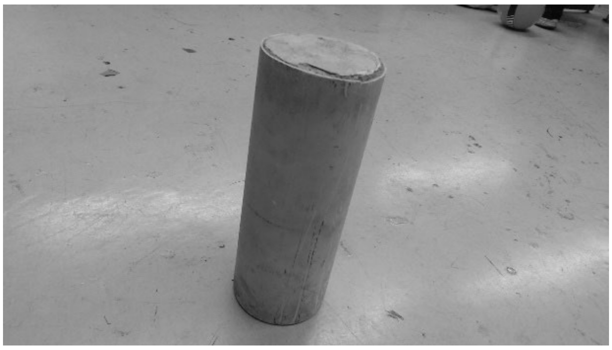
Figure 4 cylinder object after transformation
Source: own photo

The second object of the study was a spherical ball with a diameter of 21cm. The ball's colour was predominantly orange with a black stripe running through the center. Inside the black stripe, there were yellow stripes.

International Journal of Application on Sciences, Technology and Engineering (IJASTE)
Volume 1, Issue 3, 2023 ISSN:2987-2499

Figure 5 ball object before transformation
Source: own photo

Figure 6 ball object after transformation
Source: own photo

Object 3 in this study was a box, which was characterized by three key attributes: length, width, and height. The box had a length of 20.5cm, a height of 7cm, and a width of 7.3cm. The colour of the box was solely black, with no other colours or pattern.

Figure 7 box object before transformation
Source: own photo

Figure 8 box object after transformation
Source: own photo

The data for the study was collected by capturing images of each object 100 times, with consistent lighting conditions maintained across all three objects. This was done in a controlled environment, where the lighting remained constant throughout the process. The background used in the dataset was different for each photo, while the angles used for capturing the images were random. [12] [13]

3. RESULTS AND DISCUSSIONS

The research has yielded significant insights into the essential elements of object accuracy and distance. We have collected and analyzed relevant data that highlighted the interconnection between these factors and our model's overall performance. To provide a better understanding of our study, we will discuss the data gathered, the primary results of our findings, and the potential impact of our research on future advancements in the field.

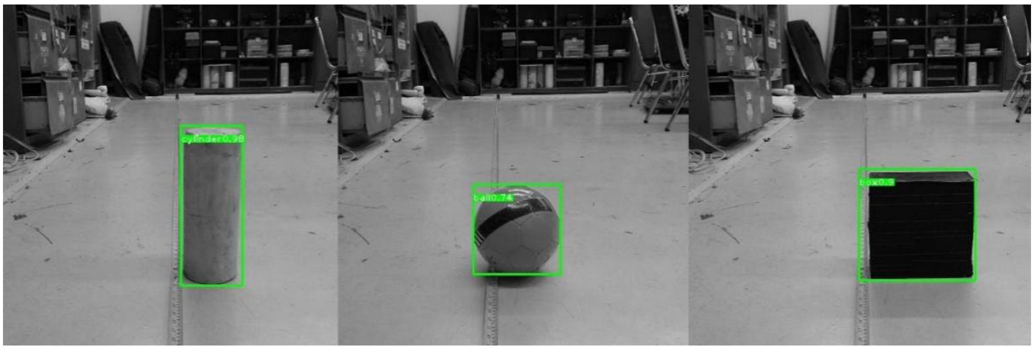

Figure 9 sample of detected objects
Source: own photo

The image provided in Figure 9 offers an example of the outcomes produced by the model. The model has detected various objects and has highlighted them with green boxes while also indicating their classification. The confidence was measured 25 times on each distance. The accuracy was calculated by taking the average of the measured confidence within the same distance. Additionally, Classification is presented alongside the respective label. If the object is not detected the model will not highlight the objects with green boxes.

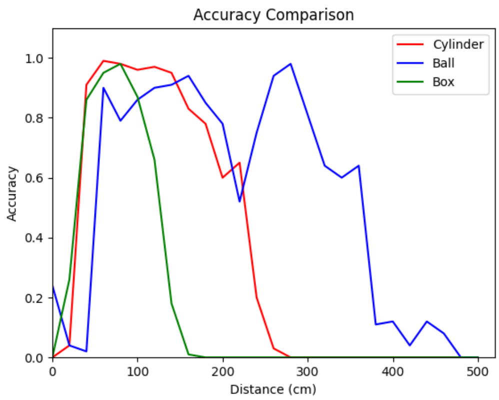

Figure 10 Accuracy vs Distance graph Object 1,2, and 3
Source: own photo

Figure 10 shows a graph indicating the accuracy of object detection at various distances. The x-axis represents the distance in centimeters, while the y-axis represents the accuracy of the detected object, ranging from 0 (0%) to 1 (100%). The higher the accuracy the better. The data is presented in a line format, with three different lines representing the accuracy of the cylinder, ball, and box objects, respectively. The lines are color-coded as red, blue, and green respectively for easy identification.

The data analysis provided valuable insights into the object detection performance and revealed that the optimal range for accurate detection across the three objects fell between 60cm to 120cm.

Further, the cylinder object exhibited the highest accuracy rate, with an average of 94% between 40cm to 160cm. Interestingly, the accuracy rate for distances greater than 140cm was observed to drop below 80%, indicating that the cylinder object detection was more challenging for longer distances. Notably, the trend for the cylinder object was that smaller distances gave better detection, whereas higher distances led to lower accuracy.

In contrast, the ball object had an accuracy rate that peaked at three specific points, namely, 60cm, 160cm, and 280cm. These results suggest that an optimal range exists for ball detection, with the highest accuracy rate observed from 60cm to 300cm, at an average of 84%, with a slight drop in accuracy at distance 200cm to 240cm.

The box object exhibited a relatively narrow distance of optimal range for detection between 40cm to 100cm compared to the other two objects. Making box accuracy less versatile, with an accuracy rate averaging at 91%. Thus, the findings suggest that the model could be optimized to detect box objects more accurately at longer distances to improve its versatility.

Table 1 result on accuracy based on distance
Source: own table

No	Distance (cm)	Accuracy (%)		
		Cylinder	Ball	Box
1	10	0	24	0
2	20	4	4	26
3	40	91	2	86
4	60	99	90	95
5	80	98	79	98
6	100	96	86	87
7	120	97	90	66
8	140	95	91	18
9	160	83	94	1
10	180	78	85	0
11	200	60	78	0
12	220	65	52	0
13	240	20	75	0
14	260	3	94	0
15	280	0	98	0
16	300	0	81	0
17	320	0	64	0
18	340	0	60	0
19	360	0	64	0
20	380	0	11	0
21	400	0	12	0
22	420	0	4	0
23	440	0	12	0
24	460	0	8	0
25	480	0	0	0
26	500	0	0	0

The data listed in Table 1 illustrates the precision of detecting three distinct objects (cylinder, ball, and box) at various distances from the observer. The cylinder object has a stable high level of accuracy ranging between 91-99%, with the greatest accuracy observed at a distance of 40-140 cm. The ball object's accuracy peaks at distance 60 cm, 160cm, and 280 cm, with an accuracy ranging between 90-98%. Similarly, the box object also has high precision rates varying from 95-98%, but it has a narrowest optimal range 40-100 cm. Overall, the data implies that object detection accuracy decreases as the distance between the object and the observer increases.

4. CONCLUSIONS AND SUGGESTIONS

Image processing method YOLOv8 could be implemented to detect an object in front of an observer with a certain distance. There were some variances of shapes in this object detection, such as cylinder, ball, and box. The results show that shape of the object and distance affects accuracy value. The data shows a trend that as the distance increase the accuracy tends to decrease, and vice versa. Therefore, distance and accuracy are inversely proportional.
In summary, the data shows that object detection has an optimum distance range. The cylinder object has the highest accuracy rate up to 99% at distance 60cm. The ball object performed the best accuracy at three distinct points, 60cm, 160cm, and 280cm, with an overall accuracy rate ranging from 90-98%. Meanwhile, the box object had a detection range of 40cm to 100cm, with an average accuracy rate of 91%.

Our findings have significant implications in some fields such as robotics, security systems, and autonomous vehicles, where object detection is an essential element of the process. Understanding the optimal detection range and the most effective algorithm for a particular object can improve the efficiency and accuracy of these systems.

Suggestion from the research for further implementation as follows, Firstly, adding more variance to the dataset is essential to achieve more robust object detection models. Diverse dataset can help the algorithm learn and generalize better. Therefore, it is crucial to ensure that the dataset used in the experiment comprises a wide range of objects with different shapes, sizes, textures, and colours. This can help to increase the algorithm's accuracy and reliability in real-world scenarios.

Secondly, conducting experiments in a general environment can provide a more realistic evaluation of the object detection system's performance. This experiment was conducted in a controlled environment, which can limit the algorithm's ability to perform in real-world scenarios. Therefore, it is crucial to test the algorithm in various environments with different lighting conditions, backgrounds, and weather conditions. This can help to ensure that the algorithm performs well in practical scenarios, which is essential for applications such as autonomous driving or security systems.

ACKNOWLEDGEMENT

The authors of this research paper express their appreciation for the support rendered by the Mechatronics Laboratory at the Faculty of Engineering, Universitas Tarumanagara, during the course of this study. The laboratory's aid was crucial in facilitating the collection and analysis of data, and the resources they provided were fundamental in establishing the groundwork for our research. We value their assistance and acknowledge their contributions to this study.

REFERENCES

[1] Alnujaidi, K., Alhabib, G., & Alodhieb, A. (2023). *Spot-the-Camel: Computer Vision for Safer Roads*. https://arxiv.org/abs/2304.00757v1

[2] Humayun, M., Ashfaq, F., Jhanjhi, N., & Alsadun, M. (2022). Traffic management: Multi-scale vehicle detection in varying weather conditions using yolov4 and spatial pyramid pooling network. *Electronics*, *11*(17), 2748. https://www.mdpi.com/2079-9292/11/17/2748

[3] Schwarting, W., … J. A.-M.-, Autonomous, and, & 2018, undefined. (2018). Planning and decision-making for autonomous vehicles. *Annualreviews.Org*, *1*, 187–210. https://doi.org/10.1146/annurev-control-060117

[4] Khan, A. A., Laghari, A. A., & Awan, S. A. (2021). Machine Learning in Computer Vision: A Review. *EAI Endorsed Transactions on Scalable Information Systems*, *8*(32), e4–e4. https://doi.org/10.4108/EAI.21-4-2021.169418

[5] Terven, J. R., & Cordova-Esparaza, D. M. (2023). *A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond*. https://arxiv.org/abs/2304.00501v1

[6] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

[7] Ju, R., & Cai, W. (2021). *arXiv : 2304 . 05071v1 [cs . CV] 11 Apr 2023 Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm*.

[8] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*.

[9] Ruiz-Ponce, P., Ortiz-Perez, D., Garcia-Rodriguez, J., & Kiefer, B. (2023). POSEIDON: A Data Augmentation Tool for Small Object Detection Datasets in Maritime Environments. *Sensors*, *23*(7). https://doi.org/10.3390/s23073691

[10] Ghahremannezhad, H., Shi, H., & Liu, C. (2023). Object Detection in Traffic Videos: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 1–20. https://doi.org/10.1109/TITS.2023.3258683

[11] Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

[12] Kanan, C., & Cottrell, G. W. (2012). Color-to-Grayscale: Does the Method Matter in Image Recognition? *PLOS ONE*, *7*(1), e29740. https://doi.org/10.1371/JOURNAL.PONE.0029740

[13] Triantafillou, E., Larochelle, H., Zemel, R., & Dumoulin, V. (2021). Learning a universal template for few-shot dataset generalization. *International Conference on Machine Learning*, 10424–10433. http://proceedings.mlr.press/v139/triantafillou21a.html