JIKO (Jurnal Informatika dan Komputer) Accredited KEMENDIKBUD RISTEK, No.105/E/KPT/2022
Vol. 7, No. 2, August 2024, pp. 123-130 p-ISSN: 2614-8897
DOI: 10.33387/jiko.v7i2.8535 e-ISSN: 2656-1948

DETECTION OF THE SIZE OF PLASTIC MINERAL WATER BOTTLE WASTE USING THE YOLOV5 METHOD

Dony Dwi Karyanto*[1], Jamaludin Indra[2], Adi Rizky Pratama[3], Tatang Rohana[4]

[1,2,3,4] Informatics engineering, faculty of computer science, Buana Perjuangan University Karawang
Email: : [1]if20.donykaryanto@mhs.ubpkarawang.ac.id, [2]jamaludin.Indra@ubpkarawang.ac.id,
[3]adi.Rizky@ubpkarawang.ac.id, [4]tatang.Rohana@ubpkarawang.ac.id

(Received: 29 July 2024, Revised: 5 Auguts 2024, Accepted: 18 August 2024)

Abstract

The use of plastic bottles for various needs is increasingly massive, especially in consumption needs such as mineral water bottles. The use of plastic bottles is used to reduce costs and be effective in maintaining the quality of mineral water, but its impact can affect natural conditions if not managed properly. Plastic bottle waste if left buried in the ground will have difficulty expanding, which can cause environmental pollution. Therefore, we can take advantage of technology to sort plastic bottle waste using a camera based on the size of plastic bottles. Differentiating the size of bottles aims to distinguish the economic value when exchanged at the waste bank. This technology utilizes object detection and recognition functions such as the YOLO (You Only Look Once) method. YOLO is a detection method that is a development of the CNN (Convolutional Neural Network) algorithm. By using YOLOv5, we can detect objects in the form of plastic bottle waste of various different sizes. To maximize object detection according to size, data annotation is done by creating a Bounding Box on each dataset according to its size. The test was carried out with several different distance configurations including 40cm, 80cm and 1m. Detection results using YOLOv5 produce up to 84% accuracy in real-time.
.

Keywords: *Plastic Bottle Waste, YOLOv5, CNN, Data Annotation, Real-Time*

**Corresponding Author: Dony Dwi Karyanto*

1. INTRODUCTION

At this time, almost all people around the world are very dependent on plastic, one of which is the people of Indonesia. Based on data from the 2023 National Waste Management Information System (SIPSN), every year Indonesia can produce 6.1 million tons of plastic waste. One of the most widely produced waste is plastic bottle waste. In the current era of digitalization and technology, the use of automation technology can be a promising solution. One of the innovation steps is to develop an automatic tool to detect plastic bottle waste of mineral water using a camera. This technology utilizes the camera's automatic object detection and recognition functions to enable highly accurate identification of plastic bottles. This system is expected to simplify the automatic collection and sorting of mineral water plastic bottle waste, reduce human burden, and improve the efficiency of waste sorting, especially mineral water plastic bottle waste.

In this context, technology is an attractive development option. The camera not only enables visual identification of objects, but also facilitates accurate data collection in real-time[1]. This mineral water plastic bottle waste automatic detection device is said to be an innovative solution that combines visual sensor technology and artificial intelligence to efficiently detect, classify, and record data about plastic bottles [2]. In addition, the concept of sustainability and carbon footprint reduction is an important aspect in the development of this system. By using an automated approach, we hope to optimize the process of collecting mineral water plastic bottle waste without significantly increasing CO2 emissions. In this context, this introduction outlines the problem, presents a key vision of the proposed system and aims to develop this technology in support of global efforts to address the plastic waste crisis[3]. By utilizing smart technology, it is hoped that this tool can make a meaningful contribution in creating a cleaner and more sustainable future. This type of plastic bottle waste has economic value that can be used for profitable economic activities. In addition to being able to reduce the pollution generated by the use of waste sorters, plastic bottles of mineral water can be used for economic activities by recycling waste into more useful goods and have a fairly high selling value. Use framework YOLO, which is based on the CNN algorithm[4], is used in this study to make it easier to detect plastic bottle waste and can distinguish the size of plastic bottles that have different economic values.

There are several studies that have existed before regarding the use of YOLO as a method in the detection process, which is a research reference as well as a comparison of updates, namely the first research conducted by Mohammed Sajidi and Nimali T Medagedara with title A New Paradigm for Waste Classification Based on YOLOv5 in 2021 which resulted in a test accuracy value of all types of detected objects resulting in a value of 61%[5]. The second research conducted by Lydia Palupi, Eko Ihsanto and Fifto Nugroho with the title Analysis of Validation and Evaluation of the Ginger Variant Object Detection Model Using the Yolov5 Algorithm in 2023 which produced a mAP value that was still relatively small at only 0.5 due to the use of less varied data and less % [6]. Furthermore, the 2023 research entitled Detection of Chili Plant Diseases Using the YOLOv5 Algorithm with Variations in Data Sharing by Setiana Riva produced a good mAP value of around 0.959 due to good data sharing and varied and many datasets[7].

From the above research, the use of YOLOv5 can be said to be quite good if it has a large enough and varied dataset in order to produce good accuracy values. For this reason, it is necessary to make an experiment to detect plastic bottle waste, namely using a camera using the YOLOv5 method. The use of the YOLO Framework makes it easier for system developers to implement it directly and produces a fairly high accuracy value in detecting objects[8]. This experiment is expected to be able to reduce the waste of plastic bottles of mineral water that are thrown in indiscriminately and be able to sort bottle waste so that it is easy for him to recycle it to become a more economical product. The use of the YOLO framework is expected to make it easier to detect objects to sort plastic mineral water bottle waste which has various sizes that have different economic values.

2. RESEARCH METHOD

This research begins with the data acquisition stage. Data acquisition was carried out directly taking pictures with plastic bottles of mineral water with various sizes from 150ml, 600ml and 1.5L using a smartphone camera with a 50MP resolution lens. Data collection needs to pay attention to the distance of each image in each class. The method used is an object detection based on the concept of artificial intelligence, where computers will be trained to be able to recognize objects of plastic and mineral bottles of various sizes. Built using basic algorithms Convolutional Neural Network (CNN) with the YOLOv5 model [9]. The following is the flow of the research stages carried out:

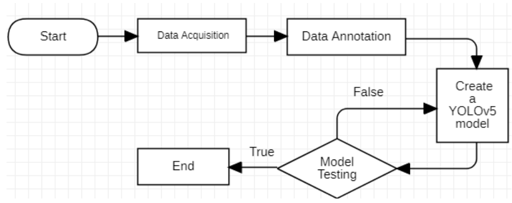

Figure 1**.** Flowchart of Research Stages

2.1. Data Acquisition

The stages of the data acquisition process are carried out by taking pictures directly using a camera *smartphone* Vivo X80 has 500 data divided into 3 sizes, including 150ml, 600ml and 1500ml. Data acquisition is carried out by paying attention to the distance taken every 2cm from the object, so that the system can distinguish the size of plastic bottles without being affected by the distance at the time of testing. At this stage, it is carried out to prepare a dataset which is divided into 3 parts, namely, training/train data, validation data, and test/testing data. The dataset used amounted to 500 photos with 70% data sharing for training data, 20% for validation data and 10% for test data. Here's an example of a dataset used:

Figure 2. Example Dataset

2.2 Data Annotation

Data annotation is a process that is carried out by labeling data by creating a bounding box on each image data using tools Label Image on the Roboflow website. Each labeled photo will be pinned with a class that adjusts to the size of the plastic bottle of mineral water.

Figure 3. Labeling Dataset

The distribution of datasets (split data) is carried out on the Roboflow website dataset that has been resizing and labelling, then the data is divided into 3 types of data, namely 70% train data, 20% valid data, and 10% test data. The data split process serves to facilitate the training process and to reduce the occurrence of overfitting and underfitting in the dataset. Figure 3 shows the split data done on the website.

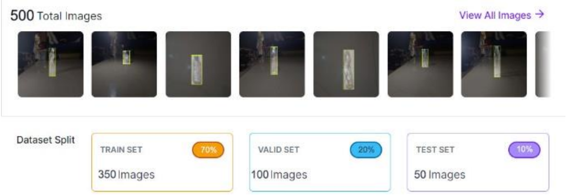

Figure 4. Data Sharing

2.3 Creating a YOLOv5 Model

The process of converting annotated datasets to YOLOv5 is carried out through the Google Collaboratory kernel by installing several main dependencies, namely YOLOv5 and *framework PyTorch* and the data training process is carried out with GPU accelerator hardware[10]. The initial process of YOLOv5 performs the cloning process of YOLOv5 Ultralytics process and installs *framework pytorch* and install *framework* roboflow in order to retrieve data on the robflow website. Next, import the data that has been labeled to be used as a dataset in training using the kernel *Google Collaboratory*. Then do the training process on the imported dataset by adding other parameters such as epoch, batch-size and img-size according to what you want. After the process is complete *Google Collaboratory* will create a model that can be used in testing to detect plastic bottles of mineral water.

2.3.1 How YOLO Works

YOLO will divide the image in a grid of SxS size, if there is an object in the S region, the region will detect. Each region will predict *bounding box* and class maps on each grid[11]. So that each region will have a bounding box and a confidence value. Every *bounding box* Contains 5 prediction values, namely x, y, w, h, conf. The values x, y represent the center of the box against the boundary boundary of the region. The w, h values represent the width and height values of the entire image. While conf represents the value of *confidence* between the ground truth box and the prediction box. At the time of testing, YOLO calculates the probability of the conditional class with the prediction of the box *confidence* with the equation:

$$Pr(Class\ i\ |Object)*\ Pr(Object)*IOU = Pr(Class\ i)*IOU$$

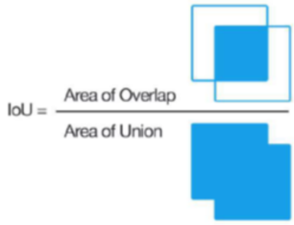

Figure 5. IOU[12]

Pr(*Object*) is the value of objectivity, Object if there is no object, it will have a value of zero and a value of one if there is an object. IOU is the value of the *Intersection Over Union ratio* between the ground truth box and the prediction box. Truth is the area of the object's certainty while pred is the area of the prediction box.

The IoU compares the ground truth bounding box and the prediction bounding box located on an object in the image image. The resulting IoU value will serve as the *confidence* to the process stage *Non Max Supression*[13].

Process *Non Max Supression* functions to eliminate the bounding box that *overlap* / overlap with an object in an image. *Non Max Supression* will utilize two values, namely IoU and *Threshold*. Value *Threshold* is used as a threshold against the IOU value, if the IOU value is less than the *Threshold* so *Bounding Box* will be eliminated[14].

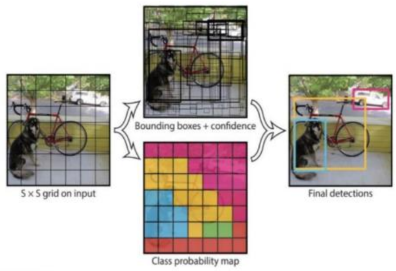

Figure 6. The concept of how YOLO works [15]

2.3.2 YOLOv5

YOLO has three main frameworks, namely *backbone, neck* and *head.*

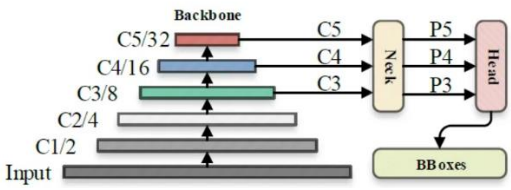

Figure 7. YOLOv5 Core Framework[16]

The backbone of YOLOv5 uses CSPDarknet53 which combines and shapes imagefeatures and integrates gradient change into the feature map so that it can improve accuracy, reduce inference speed, and reduce model weight size by reducing the number of params.

The neck of YOLOv5 uses *Path Aggregation Network* (PANet) to increase the flow of information. This panel is a development of *Feature Pyramid Network* (FPN) in YOLOv3 which includes the *bottom up* and *Top Down*. Part *Head* YOLOv5 produces 3 outputs from different features in order to perform multi-process *prediction*, so that it can help in improving the process of predicting objects both large and small in size as efficiently as possible[17].

2.4 Model Testing

This process is carried out to assess the quality of the model produced. Model testing is done by testing the system using the same objects as the dataset used. The data used testing data that had been previously shared by 10% of the dataset used.

This test can be said to be appropriate if it has test results that are equivalent to the model produced or have better performance when compared to the model used in real conditions.

2.4.1 Confusion Matrik

Evaluation metrics are also used to obtain precision, recall and f1-score values from previously trained data and match them with new data. The main evaluation metric used is Mean Average Precision (mAP)[18]. This metric is the one that is popularly used in calculating the accuracy of a detection object. The mAP value is obtained through average precision, where Average Precision (AP) [19] It is obtained from the following equation:

		Actual Value	
		Positive	**Negative**
Prediction Value	**Positive**	TP	FP
	Negative	FN	TN

Figure 8. Confusion Matrik

Here's the description:

- True Positive - The prediction label is equal to the class and the basic truth label is equal to the class.
- True Negative - The predicted label is its class while the underlying truth label is not its class.
- False Positive - The predicted label is its class and the underlying truth is not its class.
- False Negative - The predicted label is not a class and the base truth label is a class.

1. Accuracy
   It is a parameter used to measure how close a value is to the actual value[20].

$$\text{Accuracy} = \frac{TP+TN}{TP + TN + FP + FN}$$

2. Precision
   It is a parameter used to determine how reliable and consistent the model's measurements are if performed repeatedly[21].

$$\text{Precision} = \frac{TP}{TP + FP}$$

3. Recall
   It is a parameter used to know the model's ability to find all relevant objects as detectable positive values.

$$\text{Recall} = \frac{TP}{TP + FN}$$

4. F1-score
   It is a machine learning evaluation metric which is the harmonic average of the presicion and recall[22].

$$\text{F1 - Score} = 2x\frac{\text{Precision x Recall}}{\text{Precision+Recall}} = \frac{2TP}{2TP + FP + FN}$$

5. mAP
   It is the average of the precision values that are commonly used to measure the performance of detection models[23].

$$\text{mAP} = \frac{1}{n}\sum_{k=1}^{k=n} AP^{k}$$

3. RESULT AND DISCUSSION

3.1 Training Model

The schema used in the training model is predetermined by combining each other with each other to find a more optimal result. There are several training model schemes that are carried out which can be seen in the following table 1.

Table 1. YOLOv5 Model Training Scheme

Case	Img	Epoch	Batch
1	640	100	8
2	640	100	16
3	640	100	32
4	640	200	8
5	640	200	16
6	640	200	32
7	416	100	8
8	416	100	16
9	416	100	32
10	416	200	8
11	416	200	16
12	416	200	32

Table 1 shows 12 configuration training models that will be carried out in this study. The case or configuration of the YOOv5 training model will produce a training model that includes mAP, *F1-score, Recall (R) and Precission* (P) values, along with the length of time needed in the training process (time) as shown in Table 2.

Table 2. YOLOv5 Model Training Results

Weight : Yolov5m.pt								
Case	**Img**	**Epoch**	**Batch**	**P**	**R**	**F1**	**mAP**	**Time(s)**
1	640	100	8	0,993	1	0,877	0,994	1802
2	640	100	16	0,997	1	0,876	0,995	1537
3	640	100	32	0,996	1	0,878	0,995	1400
4	640	200	8	0,998	1	0,726	0,995	3438
5	640	200	16	0,998	1	0,868	0,995	3078
6	640	200	32	0,998	1	0,891	0,995	2872
7	416	100	8	0,992	1	0,875	0,994	1195
8	416	100	16	0,998	1	0,852	0,995	882
9	416	100	32	0,997	1	0,882	0,995	788
10	416	200	8	0,998	1	0,876	0,995	2120
11	416	200	16	0,998	1	0,868	0,995	1782
12	416	200	32	0,998	1	0,905	0,995	1580

From some of the research configurations carried out, where the results of the model training have been presented in table 2, it can be seen that the use of a larger image size (img) does not have a significant effect on the results obtained but sufficiently affects the time required in the training process such as in *case* 1 and *case* 7 where the *iamge* size (img) *case* 7 smaller results in a 33% faster time when compared to *case* 1 which has the same epoch and batch values. The resulting results are not too far apart so that they do not affect the results of the model evaluation later. For the use of large epoch values, it does not always guarantee a higher accuracy value, such as in *case* 3 which has the same value as *case* 6. Based on the results of the above study, it is known to produce an accuracy value range between 0.994 to 0.995 with the highest and fastest overall value obtained by *case* 12. The following is a graph of the results of the training results for *case* 12.

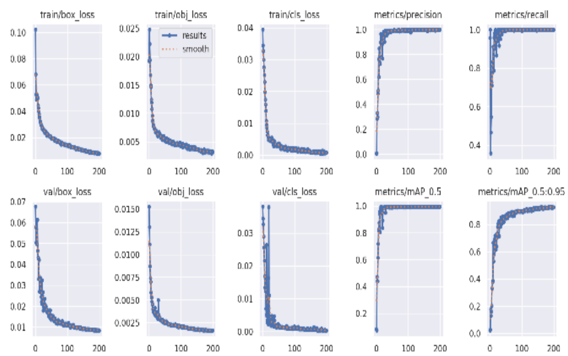

Figure 9. Result from Case 12

3.2 Model Evaluation

To validate the training results with several configurations that have been carried out, it is necessary to evaluate the model. The evaluation will use a real-time method that is carried out with sufficient lighting and has several configuration distances between the camera and different objects ranging from 40cm and 80cm and 1m. The following are the results of the model evaluation in *real time* with a 40cm distance configuration.

Figure 10. Evaluation of 150ml bottle models

For the evaluation results, the size of a 150ml bottle with a distance of 40cm can be detected well according to the actual size.

Figure 11. Evaluation of 600ml bottle models

For the evaluation results of the 600ml size with a distance of 40cm, there is a discrepancy between the detection results and the actual size where the image above detects the size of 1500ml which should be dectting the size of 600ml.

Figure 12. Evaluation of 1500ml bottle models

The results of the evaluation above for a bottle size of 1500ml with a distance of 40cm are in accordance with the actual size where for the pink bounding box has a size of 150ml, orange has a size of 600ml and red has a size of 1500ml. The following are the results of the model evaluation in *real time* with a distance configuration of 80cm.

Figure 13. Evaluation of 150ml bottle models

The results of the evaluation of the size of the 150ml bottle with a distance of 80cm can be detected well according to the actual size.

Figure 14. Evaluation of 600ml bottle models

For the 80cm distance configuration, the bottle can be detected well according to the actual size, which is 600ml.

Figure 15. Evaluation of 1500ml bottle models

The evaluation results for the 80cm distance configuration have quite good results in detecting objects in *real time* where all objects can be detected according to their respective sizes. In addition to the image above, there is an additional evaluation by adding 3 objects in the same image. The following are the results of the evaluation of the image below.

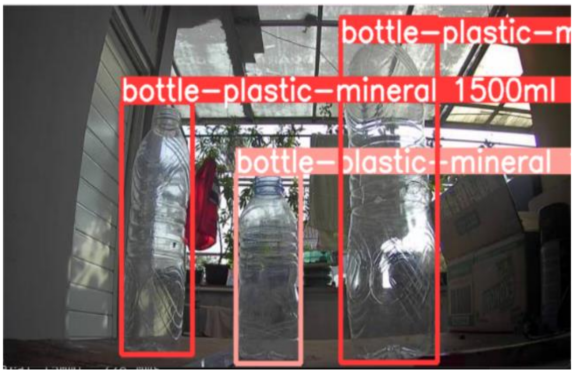

Figure 16. Evaluation of 3 objects 40cm distance

In the evaluation above, it is according to the situation when 1 object per image in the configuration is 40cm distance. Where a 600ml bottle was detected with a size of 1500ml while other objects were in accordance with the actual situation.

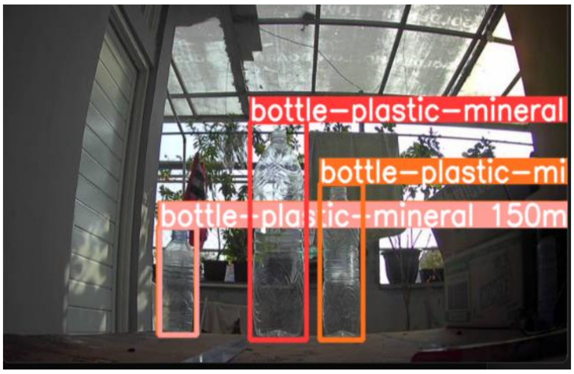

Figure 17. Evaluation of 3 objects 80cm distance

For the evaluation of image 15, it is in accordance with the situation when 1 object per image is configured at a distance of 80cm. Where all objects have been detected according to the actual situation.

The results of the model evaluation above can be summarized in the Table 3 below.

Table 3. Results of Model evaluation

No	Distance	Result	True	False
1	40cm	*"bottle-plastic-mineral 150ml"*	1	0
2	40cm	*"bottle-plastic-mineral 1500ml"*	0	1
3	40cm	*"bottle-plastic-mineral 1500ml"*	1	0
4	80cm	*"bottle-plastic-mineral 150ml"*	1	0
5	80cm	*"bottle-plastic-mineral 600ml"*	1	0
.				
.				
45	80cm	*"bottle-plastic-mineral 150ml" "bottle-plastic-mineral 600ml" "bottle-plastic-mineral 1500ml"*	3	0
46	40cm	*"bottle-plastic-mineral 150ml" "bottle-plastic-mineral 1500ml" "bottle-plastic-mineral 1500ml"*	2	1
47	80cm	*"bottle-plastic-mineral 150ml" "bottle-plastic-mineral 600ml" "bottle-plastic-mineral 1500ml"*	3	0
48	1m	*"bottle-plastic-mineral 150ml"*	1	0
49	1m	*"bottle-plastic-mineral 600ml"*	1	0
50	1m	*"bottle-plastic-mineral 1500ml"*	1	0

Based on the table above, it can be known that the use of a distance that is too close to the object can affect the results of object detection in accordance with the evaluation results in the table above, where a distance of 40cm affects the detection results on a 600ml bottle object that reads 1500ml, while for a distance configuration of 80cm, it can be detected well for all objects. The results of the above tests can be summarized in the following Table 4.

Table 4. Test Results

Confusion Matrik				
		Prediction		
		150ml	600ml	1500ml
Actual	150ml	18	0	2
	600ml	0	12	7
	1500ml	0	0	19

From the table above, it can be concluded that the resulting accuracy level has a lower value than the model used, which is 84%. This can happen due to several factors such as the distance between the camera and the object and poor lighting that can affect the detection results. The value of the results of the above tests can be summarized in the table below.

Table 5. Evaluation Result Value

	Result			
	150ml	600ml	1500ml	Total
Accuracy				0,8448276
Precision	0,9	0,632	1	0,8438596
Recall	1	1	0,67857	0,8928571
F1 Score				0,7313433
mAp				0,8448276

4. CONCLUSION

The implementation of *the You Look Only Once* (YOLO) method in detecting the size of plastic bottle waste objects is quite effective in detecting the size of plastic bottle waste. In this study, a yolo variant was used which has a medium weight size, namely yolov5m with different *batch*, *epoch* and *img* parameter values. So that various kinds of training are carried out which will be followed by the best training to be tested. *Case* 12 was chosen because it had the best overall score with the fastest training time. The *case* 12 model is tested in *real time* with different distance configurations, namely 40cm, 80cm and 1 meter.

From the above study, a lower accuracy level value was obtained compared to the model accuracy value, which was 84%. This can be influenced by various factors such as the distance between the object and the camera that is too close and the lack of datasets used for a fairly close distance and less bright lighting. So it can be seen that great accuracy does not guarantee a high level of truth.

5. REFERENCE

[1] O. E. Karlina and D. Indarti, “Pengenalan Objek Makanan Cepat Saji Pada Video Dan Real Time Webcam Menggunakan Metode You Look Only Once (Yolo),” *J. Ilm. Inform. Komput.*, vol. 24, no. 3, pp. 199–208, 2019, doi: 10.35760/ik.2019.v24i3.2362.

[2] L. C. Prita1, “ALAT PEMILAH SAMPAH ORGANIK ANORGANIK DAN LOGAM SECARA OTOMATIS MENGGUNAKAN SENSOR PROXIMITY,” vol. 2, no. 10, p. 6, 2021.

[3] M. Faisal *et al.*, “Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation,” *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/3639222.

[4] A. A. Rahmawati, A. Muhaimin, and D. A. Prasetya, “Classification of Javanese Nglegena Script Using Complexvalued Neural Network,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 7, no. 1, pp. 30–35, 2024, doi: 10.33387/jiko.v7i1.7808.

[5] M. Sajid and N. T. Medagedara, “A New Paradigm for Waste Classification Based on YOLOv5,” vol. 8, no. 4, pp. 9–17, 2021.

[6] L. Palupi, E. Ihsanto, and F. Nugroho, “Analisis Validasi dan Evaluasi Model Deteksi Objek Varian Jahe Menggunakan Algoritma Yolov5,” *Univ. Mercu Buana, Jakarta Jl. Menteng Raya No*, vol. 5, no. 1, 2023, doi: 10.47065/josh.v5i1.4380.

[7] L. Setiana Riva, “Deteksi Penyakit Tanaman Cabai Menggunakan Algoritma YOLOv5 Dengan Variasi Pembagian Data,” *J. Pengemb. IT*, vol. 8, no. 3, pp. 248–254, 2023.

[8] B. Putra, G. Pamungkas, B. Nugroho, and F. Anggraeny, “Deteksi dan Menghitung Manusia

Menggunakan YOLO-CNN," *J. Inform. dan Sist. Inf.*, vol. 02, no. 1, pp. 67–76, 2021.

[9] J. Zophie and H. Himawan Triharminto, "9. Implemetasi Algoritma You Only Look Once (YOLO) menggunakan Web Camera untuk Mendeteksi Objek Statis dan Dinamis," *TNI Angkatan Udar.*, vol. 1, no. 1, 2023, doi: 10.62828/jpb.v1i1.50.

[10] K. A. Baihaqi and C. Zonyfar, "Deteksi Lahan Pertanian Yang Terdampak Hama Tikus Menggunakan Yolo v5," *Syntax J. Inform.*, vol. 11, no. 02, pp. 1–9, 2022.

[11] E. R. Justitian, I. Y. Purbasari, and F. T. Anggraeny, "Perbandingan Akurasi Deteksi Kelelahan pada Pengendara Menggunakan YOLOv3-Tiny YOLOv4-Tiny," *J. Inform. dan Sist. Inf.*, vol. 3, no. 1, pp. 21–30, 2022, doi: 10.33005/jifosi.v3i1.440.

[12] M. Dio Riza Pratama, B. Priyatna, S. S. Hilabi, and A. L. Hananto, "Deteksi Objek Kecelakaan Pada Kendaraan Roda Empat Menggunakan Algoritma YOLOv5," *J. Ilm. Sist. Informas*, vol. 12, no. 2, pp. 15–26, 2022.

[13] D. Wahiddin, "Klasifikasi Kadar Hidrasi Tubuh Berdasarkan Warna Urine dengan Metode Ekstraksi Fitur Citra dan Euclidean Distance," *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 5, no. 1, pp. 16–20, 2020, doi: 10.36805/technoxplore.v5i1.887.

[14] D. G. Arwindo, E. Y. Puspaningrum, and Y. V. Via, "Identifikasi Penggunaan Masker Menggunakan Algoritma CNN YOLOv3-Tiny," *Pros. Semin. Nas. Inform. Bela Negara*, vol. 1, pp. 153–159, 2020, doi: 10.33005/santika.v1i0.41.

[15] J. Redmon and A. F. , Santosh Divvala, Ross Girshick, "You Only Look Once: Unified, Real-Time Object Detection Joseph," *ACM Int. Conf. Proceeding Ser.*, 2018, doi: 10.1145/3243394.3243692.

[16] H. Liu, F. Sun, J. Gu, and L. Deng, "SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode," *Sensors*, vol. 22, no. 15, pp. 1–14, 2022, doi: 10.3390/s22155817.

[17] J. Indra, "Penerapan Artificial Neural Network Untuk Klasifikasi Fertilitas Telur Itik Menggunakan Raspberry Pi," *Buana Ilmu*, vol. 3, no. 1, 2018, doi: 10.36805/bi.v3i1.460.

[18] P. Henderson and V. Ferrari, "End-to-end training of object class detectors for mean average precision," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10115 LNCS, pp. 198–213, 2017, doi: 10.1007/978-3-319-54193-8_13.

[19] K. Oksuz, B. C. Cam, E. Akbas, and S. Kalkan, "Localization recall precision (LRP): A new performance metric for object detection," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 521–537, 2018, doi: 10.1007/978-3-030-01234-2_31.

[20] D. Iskandar Mulyana and M. A. Rofik, "Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5," *J. Pendidik. Tambusai*, vol. 6, no. 3, pp. 13971–13982, 2022, doi: 10.31004/jptam.v6i3.4825.

[21] L. Suroiyah, Y. Rahmawati, and R. Dijaya, "Facemask Detection Using Yolo V5," *J. Tek. Inform.*, vol. 4, no. 6, pp. 1277–1286, 2023, doi: 10.52436/1.jutif.2023.4.6.1043.

[22] P. Mustamo, "Object detection in sports: TensorFlow Object Detection API case study," no. January, pp. 1–43, 2018.

[23] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," *Int. Conf. Syst. Signals, Image Process.*, vol. 2020-July, pp. 237–242, 2020, doi: 10.1109/IWSSIP48289.2020.9145130.