A detection method for floating debris in waterways using YOLOv9

Jan Edgar E. Tupas *, Miguel T. Soniel, Christian Henry Miguel E. Caruz,

Darryll B. Babar, John Paul Q. Tomas

School of Information Technology, Mapúa University, Makati, Philippines

* Corresponding author.

E-mail address: jeetupas@mymail.mapua.edu.ph

ABSTRACT

The Philippines is home to vast networks of waterways which serve as a vital resource for communities. However, waterways are subject to pollutions such as the accumulation of debris. The Pasig River releases an approximate 32 to 64 thousand tons of plastic every year and is one of the most polluting rivers in the world. This study explores the development of a YOLOv9 model for detecting floating debris in waterways, focusing on the classification of various debris types such as garbage, leaves, branches, and aquatic plants. The study involved creating a dataset from images of floating debris on the Pasig River, training the model, and evaluating its performance. The YOLOv9 GELAN C variant was utilized, and hyperparameter tuning was conducted to optimize the performance of the model. The model attained promising results for detecting floating debris with overall scores of 86.9%, 90.2%, and 78.8% for mAP@50, precision, and recall respectively. The model excelled in classifying aquatic plants but faced challenges with garbage classification. Future recommendations include addressing dataset imbalance, and adding variances when acquiring images, such as different camera angles, to improve the robustness of the model. Additionally, the researchers suggest exploring other data preprocessing techniques and hyperparameter optimization beyond just batch size, as these could potentially further enhance the performance of the model.

Keywords: YOLOv9, Object detection, Dataset, Floating debris

1. Introduction

Waterways have a vital role in supporting our society, such as in transportation, supplying drinking water, and facilitating our agriculture. Philippines is home to a vast and interconnected networks of lakes, rivers and coastal areas that serve as an important resource for the communities. According to a study of Guerrero [1], the country is provided with over 400,000 hectares of inland waters which include lakes, reservoirs and especially, waterways. With the recent advancements in our society such as urbanizations of most areas and the continuous improper waste management of the people, these valuable waterways are at risk to pollutions and accumulations of debris in the water bodies. It is estimated that 0.28-0.75 million metric tons of plastic waste was released by the country into the world's ocean and Pasig River is considered as one of the most polluting rivers where it releases estimated 32 to 64 thousand tons of plastic every year [2]. This raises a concern regarding monitoring and management of debris in waterways in the Philippines.

Monitoring and detection of water surface debris is essential to the maintenance of the health and sustainability of our environment. Water surface debris such as plastic wastes, plants, and other materials do not only pose a risk to the environment, but also impacts the quality of the water and the life within it [7][9]. By using an object detection model, it can classify different types of water surface debris which is beneficial in terms of monitoring floating debris in waterways.

In this study, an object detection model for detecting floating debris in waterways was developed by the researchers. Specific floating debris types encompassing garbage, leaves, branches, and aquatic plants were classified and detected. The YOLOv9 object detection algorithm was used in the study as a review of object detection studies utilizing YOLOv9 provided effective results and using the algorithm for the detection of floating debris is yet to be conducted.

The primary objective of the study was to develop a YOLOv9 model to effectively detect and classify debris floating on a water surface. The specific objectives of the study are the following: (1) To develop a YOLOv9c model for detecting floating debris and its performance; (2) To determine the hyperparameter values that enhance the performance of the YOLOv9 model using hyperparameter tuning; (3) To assess the performance of the YOLOv9 model in classifying the various floating debris types.

2. Review of related literature

2.1. Floating debris in waterways

The study of Rocamora et al. [3] shows that only a small part of floating debris on the water is caused by agricultural activity, while most are anthropogenic, and this kind of waste has significant consequences in terms of environment and tourism. According to a study by Agamuthu et al. [4], floating debris, specifically plastics, may cause permanent damage to the environment as these wastes carry chemicals and contaminants. The study also states that this water debris mistakenly appears to be food for marine organisms, which they often consume. Consumption of seafood may be a risk due to plastic exposure. Floating debris negatively impacts the water quality and surface water aesthetics [7][9]. The current system of monitoring floating debris on water depends on human resources and needs to be more efficient in managing the issue in real time [5]. Thus, navigating emerging technologies that allow real-time object detection can be a response to the concern.

2.2. Floating debris detection using YOLO algorithm

The YOLO algorithm is an approach used to detect objects in real-time [6]. Multiple studies have utilized YOLO to detect floating debris on water. The study of Qiao et al. [7] utilized coordinate attention (CA) module on the backbone network of YOLOv5 and replaced its Path Aggregation Network module with Bidirectional feature pyramid network (BiFPN) module, which improves the overall detection capability. The proposed CA-YOLOv5 with

BiFPN outperformed YOLOv5s, CA-YOLOv5, and Bi-YOLOv5 (YOLOv5 with BiFPN) in terms of Recall and AP, which achieved 96.5% and 95.8%, respectively. In the study of Zailan et al. [8], an improved YOLOv4 is proposed where a spatial pyramid pooling (SPP) block is applied. The performance of the SPP-YOLO model has significantly increased in terms of precision, F1 score, and computation time of training. The study of Lin et al. [9] proposed an improved YOLOv5s where a feature map attention (FMA) layer is integrated. The result of FMA-YOLOv5s model outperformed YOLOv5s by 2.18% in terms of mean Average Precision (mAP). The studies present the effectiveness of YOLO for object detection, specifically floating debris detection.

3. Methodology

3.1.Data collection

This section discusses the methods and materials used in the study. A dataset of floating debris images was created by the researchers. The debris instances were classified according to type through annotation. The images were subject to preprocessing methods to enhance data quality. Data augmentation was implemented to artificially increase the size of the dataset. The dataset was split and used for model training. The study focuses on the YOLOv9 object detection algorithm. Specifically, the YOLOv9 GELAN C variant was trained with the dataset. Hyperparameter tuning focusing on the batch size was conducted. The model with the best performance was identified through validation and used for testing and inference. Figure 1 shows the conceptual framework of the study.

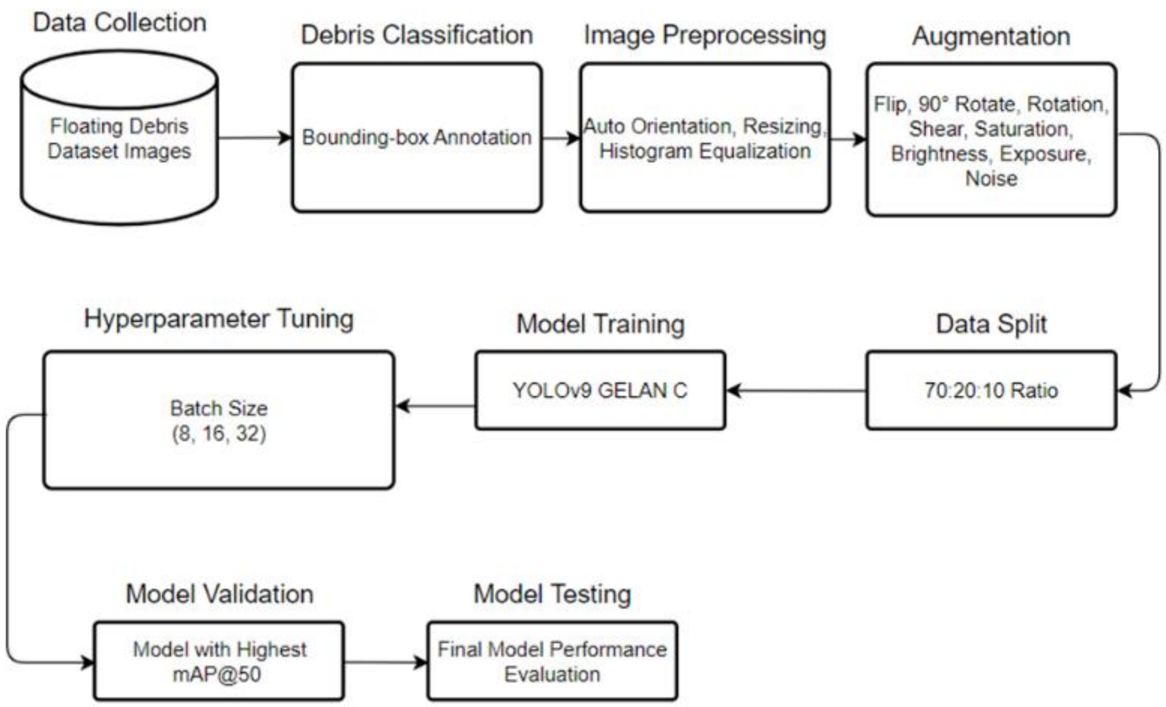

Fig. 1. Conceptual Framework of the Study

A localized dataset of debris floating on a water surface was produced by the researchers. The researchers initially recorded a 19-minute video of floating debris on the water surface of Pasig River. To capture the video, the researchers utilized a 24.1 megapixel CMOS M50 Mk II camera. The video was captured in 1080p Full HD resolution. The recording process was conducted at a controlled environment. The recording was capture at daytime to ensure that the floating debris were clear and visible. The camera was set up at a height of 162 cm from the ground, distanced 10 feet from the water, and positioned at an angle of 45 degrees in recording the 19-minute video.

3.2.Dataset

With the recorded 19-minute video of debris floating on the water surface of the Pasig River, the researchers created the dataset used for the study. A total of 1144 images was extracted from the video. The dataset classes were garbage, aquatic plants, branches, and leaves, as these were the notable floating debris recognized from the footage. Table 1 shows the total instances of each floating debris class, attaining a total of 10432 floating debris instances.

Table 1

Total Instances of Each Class

Debris Type	Total Instance
Garbage	7284
Leaves	2830
Branches	169
Aquatic Plants	149
Total	10432

3.3.Data processing

Data preprocessing methods were implemented on the dataset to enhance its quality through the online platform Roboflow. Using the determined classes, the researchers annotated each image. The dataset was split in a 70:20:10 ration for training, validation, and testing. The preprocessing methods applied to the images are shown in Figure 2. The dataset was significantly imbalanced with garbage and leaves had significantly more instances than branches and aquatic plants. Data augmentation was performed on the dataset to potentially alleviate the effects of the dataset imbalance as shown in Figure 3. Through augmentation, the dataset was further increased artificially, resulting in 2746 images.

Auto-Orient: Applied

Resize: Stretch to 640x640

Auto-Adjust Contrast: Using Histogram Equalization

Fig. 2. Data Preprocessing Methods Applied to the Images

Outputs per training example: 3

Flip: Horizontal, Vertical

90° Rotate: Clockwise, Counter-Clockwise, Upside Down

Rotation: Between -15° and +15°

Shear: ±10° Horizontal, ±10° Vertical

Saturation: Between -25% and +25%

Brightness: Between -15% and +15%

Exposure: Between -10% and +10%

Noise: Up to 0.1% of pixels

Fig. 3. Augmentations

3.4.Model training and validation

The researchers used the YOLOv9 object detection algorithm for the detection of floating debris in waterways. Specifically, the researchers evaluated the small, medium, and large variants of YOLOv9. The online platform Google Colab and a L4 GPU was used for model development. A YOLOv9 model pre-trained on the COCO dataset was used to attain better model performance and avoid training from scratch. The YOLOv9 model was further developed using the floating debris dataset. To evaluate and improve the performance of the model, hyperparameter tuning was conducted, focusing on batch size and optimizer. For batch size, a value of 8 and 16 were evaluated. For optimizer, SGD and Adam were assessed. A weight decay of 0.001, initial learning rate of 0.0005, and epoch value of 100 were kept constant during training. To visualize and keep track of the model during training, Tensorboard was used. Upon training completion, the models were further assessed using the validation set to determine the best performing model.

3.5.Data collection

The most effective and best performing model identified through validation is selected for testing. The model was tested using the test set to assess how well the model generalizes with unseen data. This process aids in determining the efficacy of the model for the object detection task as it simulates the deployment of the model on real-world data. To evaluate model performance, the metrics used in the study were mean average precision at IoU@0.5 (mAP@50), Precision (P), and Recall (R). To further assess the classification performance of the model, the confusion matrix was used.

4. Materials and methods

This section discusses the results obtained from the developed YOLOv9 model to detect floating debris in waterways. The performance of the model after training and validation, hyperparameter tuning, and the final model testing, as well as the classification performance of the model are discussed.

4.1.Training, hyperparameter tuning, and validation

The researchers developed a pre-trained YOLOv9 GELAN C model for detecting debris floating on a water surface. The training and validation sets from the dataset of floating debris images were used for training and validating the model. To further explore and enhance the performance of the model, hyperparameter tuning was conducted that centered on the batch size. To simplify, model 1, model 2, and model 3 used batch size values of 8, 16, and 32 respectively. Other parameters were kept constant throughout the study, where learning rate was set to 0.01, epoch was set to 50, and optimizer was set to Stochastic Gradient Descent (SGD). The loss graphs of models 1, 2, and 3 during training are shown in figures 4, 5, and 6 respectively. The validation results of the models are shown in table 2.

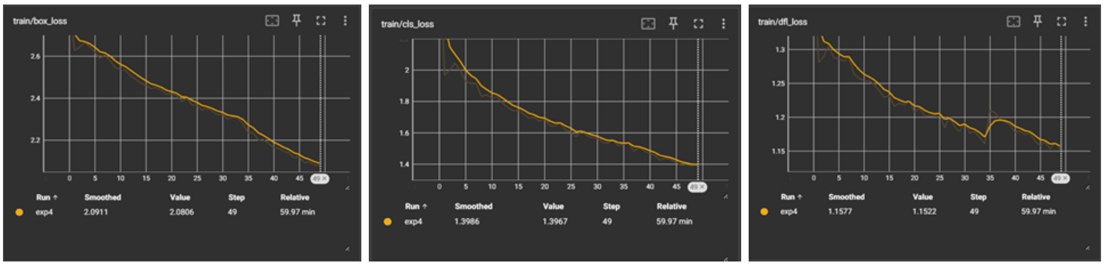

Fig. 4. Loss graphs of Model 1

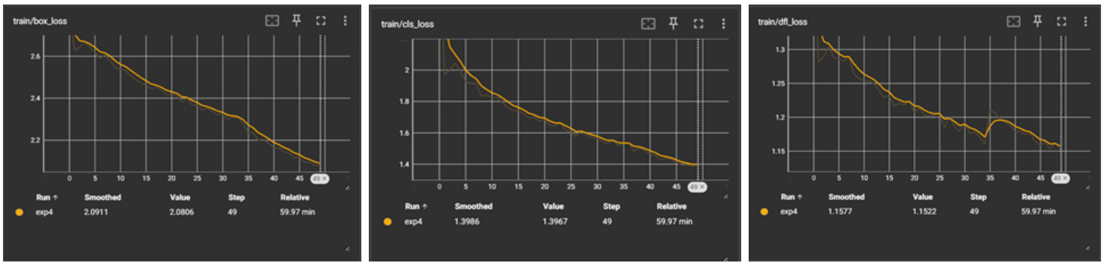

Fig. 5. Loss graphs of Model 2

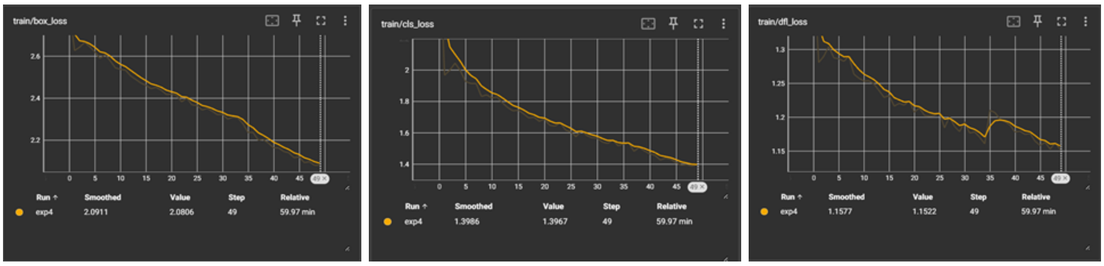

Fig. 6. Loss graphs of Model 3

In training the model for the object detection task, the optimizer attempts to minimize the loss function to attain better performance. The box loss and classification Cross Entropy Loss (CLS) loss of all three models show a smooth descent, indicating that the performance of the model during training was improving. Additionally, the Distribution Focal Loss (DFL) of all models were also descending, but manifested an inconsistency between the 35th and 40th epoch, indicating a limited struggle during training. It can be observed that all losses continued to descend, signifying that increasing the epoch during model training could potentially further enhance the performance of the model.

Table 2

Model Validation Results

Model 1	Batch Size (8)		
Classes	mAP@50	Precision	Recall
All	0.865	0.833	0.839
Aquatic Plants	0.937	0.806	0.924
Branch	0.995	0.972	1
Garbage	0.684	0.740	0.618
Leaf	0.843	0.815	0.816
Model 2	**Batch Size (16)**		
Classes	mAP@50	Precision	Recall
All	0.866	0.861	0.836
Aquatic Plants	0.925	0.799	0.963
Branch	0.995	0.976	1
Garbage	0.678	0.785	0.6
Leaf	0.867	0.884	0.779
Model 3	**Batch Size (32)**		
Classes	mAP@50	Precision	Recall
All	0.87	0.84	0.814
Aquatic Plants	0.934	0.784	0.889
Branch	0.993	0.892	1
Garbage	0.676	0.818	0.562
Leaf	0.878	0.866	0.804

The performance of the model for detecting and classifying floating debris is determined through the validation results. Batch size values of 8,16, and 32 were used while values for parameters such as epoch, learning rate, and optimizer were kept constant. In observing the mAP@50 scores obtained from the study, all models achieved close performances with a value of 0.865, 0.866, and 0.87 for model 1, model 2, and model 3, respectively. Performance for all models were strong and similar. Further evaluating the precision and recall values show variances in the detection and classification task. In terms of precision, model 2 attained the highest precision with 0.861 or 86.1%, followed by model 3 with 0.84 or 84% and lastly model 1 with 0.833 or 83.3%. Meanwhile, the recall values obtained also showed small variance. Model 1 achieved the highest recall with 0.839 or 83.9%, followed by model 2 with 0.836 or 83.6%, and lastly model 3 with 0.814 or 81.4%. In observing the performance of the model in classifying each class, the mAP@50 values attained were high indicating good performance. However, classifying debris as garbage obtained an average mAP@50 for all models, with 0.684 or 68.4% from model 1, 0.678 or 67.8% from model 2 and 0.676 or 67.6% from model 3. With the obtained results, the developed model performed well in detecting floating debris in waterways. In evaluating the models based on overall mAP@50, precision, and recall metrics, the researchers determined that model 2 that used a batch size of 16 offered the best overall performance.

4.2.Model classification performance

Table 3

Model 1 Confusion Matrix

	Class	True				
		Aquatic Plant	Branch	Garbage	Leaf	Background
Predicted	Aquatic Plant	0.96	0	0	0	0.01
	Branch	0	1	0	0	0
	Garbage	0	0	0.66	0.04	0.78
	Leaf	0	0	0.02	0.83	0.21
	Background	0.04	0	0.32	0.13	0

Table 4

Model 2 Confusion Matrix

	Class	True				
		Aquatic Plant	Branch	Garbage	Leaf	Background
Predicted	Aquatic Plant	0.89	0	0	0	0.02
	Branch	0	1	0	0	0.01
	Garbage	0.07	0	0.68	0.04	0.79
	Leaf	0	0	0.02	0.83	0.19
	Background	0.04	0	0.31	0.13	0

Table 5

Model 3 Confusion Matrix

	Class	True				
		Aquatic Plant	Branch	Garbage	Leaf	Background
Predicted	Aquatic Plant	0.93	0	0	0	0.02
	Branch	0	1	0	0	0.01
	Garbage	0.04	0	0.63	0.04	0.71
	Leaf	0	0	0.02	0.83	0.25
	Background	0.04	0	0.35	0.12	0

The confusion matrices of each model using different batch size values are shown in tables 3, 4, and 5. The debris classes of aquatic plant, branch, and leaf attained scores above 0.80 r

80% which indicate the effective capacity of the models to correctly classify floating debris of these types. All instances of branches were correctly classified in each model with a score of 1.00 or 100%. Majority of the instances of aquatic plants were correctly classified by the models with scores of 0.96 or 96%, 0.93 or 93% and 0.89 or 89% from models 1, 3 and 2 respectively. Meanwhile, leaves manifested lower scores then branches and aquatic plants with 0.83 or 83% for all models. Garbage obtained the lowest scores for all models with 0.68 or 68%, 0.66 or 66%, and 0.63 or 63% from models 2, 1, and 3 respectively. It is important to note the imbalance of the curated dataset with leaves and branches as the minority class while garbage had the most instances. The low scores for classifying garbage can be attributed to the high number of instances and the over-all visibility of the garbage instances on the images. Observing the dataset images, garbage instances were small as compared to the larger instances for branches and aquatic plants. The dark colors of numerous garbage instances allowed it to blend with the water, further limiting visibility. Additionally, the models attained high background scores for garbage indicating that a significant portion of garbage instances were overlooked and not classified at all.

4.3.Model testing and inference

The results from the validation phase of the study indicate that model 2 that used a batch size value of 16 provided the best performance for detecting floating debris. The test set was used to test the model and determine the capacity and performance of the model in detecting floating debris with data is has not yet seen. This simulates model deployment with real-world data and provides insights on the generalization capabilities of the model. Table 6 shows the test results of the YOLOv9 GELAN C model. Figure 7 shows an output of the developed model.

Table 6
Test Results

Class	Images	Instances	mAP@50	Precision	Recall
All	114	1041	0.869	0.902	0.788
Aquatic Plant	114	12	0.995	0.992	1.00
Branch	114	16	0.958	0.893	0.875
Garbage	114	747	0.651	0.841	0.525
Leaf	114	266	0.873	0.881	0.752

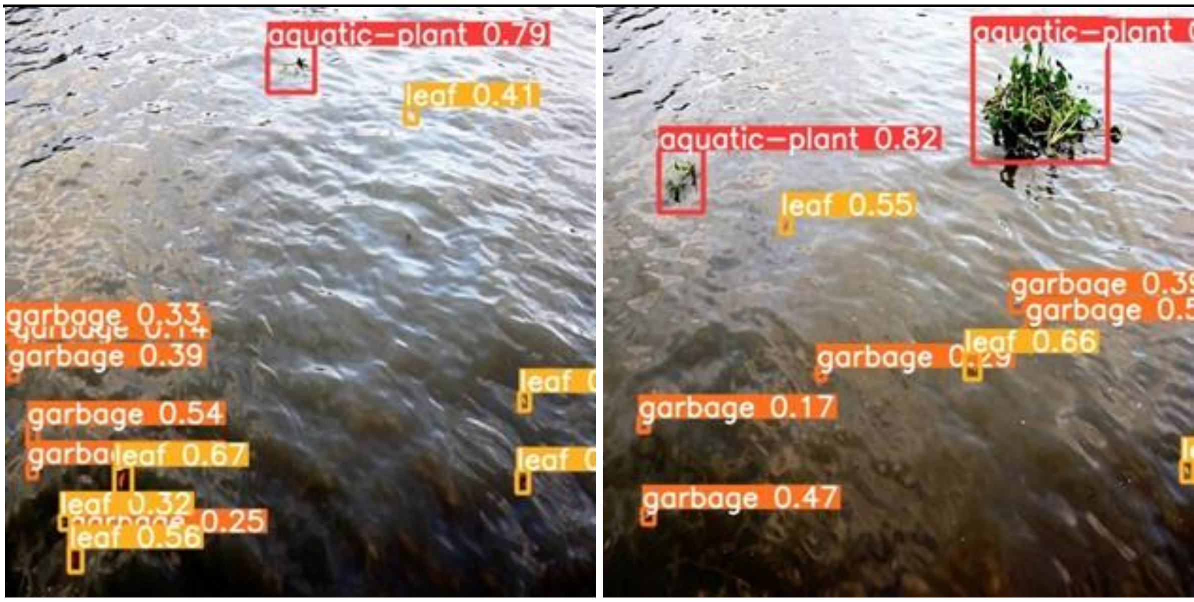

Fig. 7. Output of the Developed Model

The test results indicate that the developed model performed with good efficacy. The model obtained an overall mAP@50 of 0.869 or 86.9%, precision of 0.902 or 90.2% and recall of

0.788 or 78.8%. The scores obtained by the model suggest that the model is effective in detecting floating debris instances in waterways, as well as classifying the detected floating debris. In classifying the floating debris, the model performed best in classifying aquatic plants with a mAP@50 of 0.995 or 99.5%, precision of 0.992 or 99.2%, and recall of 1 or 100%. The model was also able to effectively classify branch and leaf debris types with limited error. However, the model manifested lower performance scores in classifying garbage instances following a mAP@50 of 0.651 or 65.1%, precision of 0.841 or 84.1%, and recall of 0.525 or 52.5%. It is important to note that the dataset is imbalanced. As shown, garbage had the most instances and is significantly larger than leaf, aquatic plant, and branch debris types. Ensuring that the dataset is balanced where each class is equally represented could provide more insights into model performance and may potentially enhance it. Overall, the model attained good performance but may still be further enhanced.

5. Conclusion and recommendation

5.1.Conclusion

The study focused on the development of a YOLOv9 model for the detection of floating debris in waterways. The researchers produced their own dataset of floating debris images and used it to train, validate, and test the model. However, the researchers acknowledge the limitations of the dataset, primarily due to the imbalance of instances per class. Data preprocessing and augmentation methods were employed to enhance the quality of the dataset. Hyperparameter tuning was employed during training and validation where variances in batch size was the focus. Comparing the results from the validation phase, model 2 with a batch size of 16 was determined as the best model. The model was tested using images it has not yet seen to simulate model deployment with real-world data. The model obtained scores of 86.9%, 90.2%, and 78.8% for mAP@50, precision, and recall respectively. The classification performance of the model was also assessed. The results indicate that the model is effective in classifying aquatic plant, branch, and leaf debris types manifested by the high metric scores attained. However, the model struggled in correctly classifying garbage instances. This can be attributed to how the garbage instances tend to blend with the water and the small size of the instances limiting visibility. Improvements can still be made to enhance the performance of the model. In conclusion, the developed model manifested good efficacy for the detection of floating debris in waterways indicating that it can be effectively deployed for the task.

5.2.Recommendation

In conducting the study, the authors acknowledge the limitations that persisted. Despite the good results attained from the model, there is still potential to further enhance its performance for detecting floating debris. The researchers acknowledge the limitations of the created dataset. Although the dataset consisted of a significant number of images, it was imbalanced where the number of instances per class were not equal. The instances of garbage in the dataset significantly outnumbered the remaining classes. To address this, the researchers recommend creating a dataset of floating debris where classes are equally represented. This ensures that there is no minority class in the dataset. Following a controlled approach in data acquisition, the researchers recommend acquiring images through various lighting and camera angles and positions. This can further improve the robustness of the model as it is trained with images of different variances. The researchers also recommend using other data preprocessing techniques to enhance the quality of the images. Additionally, the researchers recommend exploring other hyperparameters during the hyperparameter optimization process since the study only focused on batch size. Variances in learning rate, epoch, and optimizer can potentially further enhance model performance. Lastly, the researchers recommend using other deep learning algorithms for the detection of floating debris in waterways and assess its performance with the model developed in this study.

Acknowledgement(s)

The researchers would like to express their sincere gratitude to their advisor, Prof. John Paul Tomas, for his invaluable guidance and support throughout this study. His supervision and mentorship were essential in the successful completion of this research.

References

[1] Rafael D. Guerrero III. 2022. Commercially Caught Freshwater Fishes in the Philippines: Status, Issues, and Recommendations. Transactions NAST PHL, 44, 15 pages. http://dx.doi.org/10.57043/transnastphl.2022.2568

[2] Laurenz Luigi B. Cruz and Takenori Shimozono. 2021. Transport of floating litter within Manila Bay, Philippines. Marine Pollution Bulletin, 163, 111944. https://doi.org/10.1016/j.marpolbul.2020.111944

[3] Carmen Rocamora, Herminia Puerto, Ricardo Abadía, Margarita Brugarolas, Laura Martínez-Carrasco, and José Cordero. 2021. Floating Debris in the Low Segura River Basin (Spain): Avoiding Litter through the Irrigation Network. Water 2021, 13(8), 1074. DOI: https://doi.org/10.3390/w13081074

[4] P Agamuthu, S Mehran, and A Norkhairah. 2019. Marine debris: A review of impacts and global initiatives. Waste Management & Research, 37(10), 987-1002. https://doi.org/10.1177/0734242X19845041

[5] Zhanjun Jiang, Baijing Wu, Long Ma, Huawei Zhang and Jing Lian. 2023. APM-YOLOv7 for Small-Target Water-Floating Garbage Detection Based on Multi-Scale Feature Adaptive Weighted Fusion. Sensors 2024, 24(1), 50. https://doi.org/10.3390/s24010050

[6] Juan R. Terven and Diana M. Cordova-Esparza. 2024. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. arXiv:2304.00501v6 [cs.CV]. Retrieved from https://arxiv.org/html/2304.00501v6

[7] Guangchao Qiao, Mingxiang Yang, and Hao Wang. 2022. A Detection Approach for Floating Debris Using Ground Images Based on Deep Learning. Remote Sens. 2022, 14(17), 4161. https://doi.org/10.3390/rs14174161

[8] Nur Athirah Zailan, Anis Salwa Mohd Khairuddin, Uswah Khairuddin, and Akira Taguchi. 2021. YOLO-based Network Fusion for Riverine Floating Debris Monitoring System. In Proceedings of the 2021 International Conference on Electrical, Communication and Computer Engineering (ICECCE), 2021, Kuala Lumpur, Malaysia, pp. 1-5. https://doi.org/10.1109/ICECCE52056.2021.9514096

[9] Feng Lin, Tian Hou, Qiannan Jin, and Aiju You. 2021. Improved YOLO Based Detection Algorithm for Floating Debris in Waterway. Entropy 2021, 23(9), 1111. https://doi.org/10.3390/e23091111