Comprehensive Analysis of On-Site Riparian Waste Pollution: A Case Study on the Hyakken River Basin

Shijun PAN[1], Keisuke YOSHIDA[2], Takashi KOJIMA[3]

[1]Member of JSCE, Doctoral Student, Graduate School of Environmental and Life Science, Okayama University
(Tsushima-naka 3-1-1, Kita-ku, Okayama 700-8530, Japan)
E-mail: p4b36znn@s.okayama-u.ac.jp (Corresponding Author)
[2]Member of JSCE, Associate Professor, Graduate School of Environmental and Life Science, Okayama University
[3]Senior Researcher, TOKEN C. E. E. Consultants Co., Ltd

Environmental degradation due to waste pollution (i.e., abandon artificial objects in the natural environment) is a growing concern that demands comprehensive understanding and effective solutions. This paper discussed the intricate issue of waste pollution, focusing on the Hyakken River Basin. This study presents of the waste pollution under long-term sunshine and/or water flow brush-up in the area, emphasizing the transformative effects of these environmental factors on the waste pollution. To substantiate this findings, the authors utilized smartphones to capture high-resolution images. To train these images in a YOLOv8 model, the datasets were meticulously annotated using roboflow, incorporating advanced data augmentation techniques. Subsequent testing involved the application of the model to smartphone videos from distinct sections of the field. Engaging actively in the cleanup process alongside dedicated volunteers, the authors successfully cleared the riparian area, providing a firsthand account of the positive impact of collaborative efforts. The on-site waste pollution after long-term in the natural environment is not similar to commercial goods (i.e., color and shape). Several models (i.e., pre-trained and custom-trained) have been tested with the Hyakken River Basin Wild Dataset. Derived from the results, custom-trained YOLOv8n-seg model has better results. Comparing with the SAM result without specified label, custom-trained YOLOv8n-seg model also has potential of improving. Looking ahead, the authors envision the utilization of this dataset as a valuable tool for supporting both volunteers and government staff in waste amount analysis. By contributing to environmental protection initiatives, this research aspires to pave the way for informed decision-making and sustainable solutions in the ongoing battle against waste pollution.

Key Words: *Instance Segmentation, Riparian Waste Detection, River Monitoring, Smartphone*

1. INTRODUCTION

These years, the riparian environment around the Hyakken River, Japan has become much worse than before with a persistent environmental challenge: waste pollution. And at the same time, the waste pollution problem is not just in Japan, also existed around the worldwide[1,2,3]. Even there are several datasets[1,2,3] that have been published online, the dataset related to long-term waste pollution in the riparian environment is also very rare.

Based on the mentioned practical problems, the authors trained the AI models, and applied the models to detect the waste pollution[4] in the environmental management. But the targets in the previous research[4] were settled on-site manually, that have some different features comparing with the real waste pollution in the natural environment.

This study embarks on an exploration of the dynamics of waste pollution in this region, aiming to offer a comprehensive understanding of its multifaceted nature. The utilization of smartphones allowed the authors to capture high-resolution images, forming the foundation for an image-based dataset. During the process of the on-site smartphone data collection, the drone-derived images around the same area have also been collected. Because of the time limitation, in this research, the authors just focused on the dataset derived from the on-site smartphone-based images.

Owing to the limitation of the on-site data accumulations, the authors considered that the
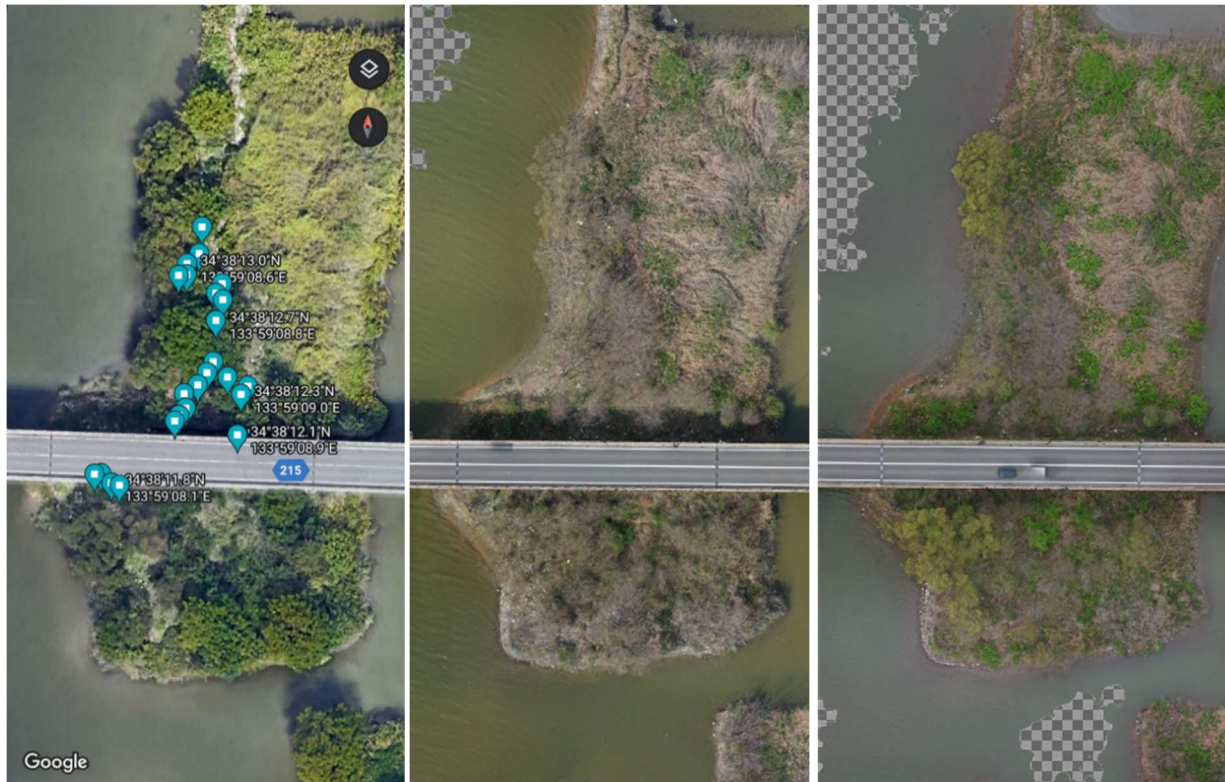
Fig.1 Bird's eye view images of the study site (left-side from Google map, collected at 20th, Sep, 2022; middle- and right- sides from DJI Mavic2, collected at 17th and 29th, Mar, 2022 before and after the waste pollution collection by volunteers, respectively) and positions of the data-collection points derived from the GPS data in each on-site images.

application of the data augmentation technology to increase the data amount is necessary. Thus, leveraging roboflow[5] (i.e., an open source online website for data processing) for dataset annotation and incorporating data augmentation techniques to ensure an examination of the waste landscape.

Considering of applying the state-of-the-art open source models on the dataset examination, the implementation of the YOLOv8[6] model (i.e., a cutting-edge, state-of-the-art model for computer vision tasks) and SAM[7] (i.e., segment anything model, performs pixel-level classification by categorizing pixels into various classes. These classified pixels correspond to different objects or regions within an image.) were considered, further validated through rigorous testing with smartphone videos from distinct sections of the field.

Beyond the realm of data analysis, this involvement in on-the-ground cleanup efforts alongside dedicated volunteers provided a tangible connection to the impact of waste pollution on the riparian area. This firsthand experience serves as a poignant backdrop against which the gravity of the waste pollution issue is vividly portrayed, challenging preconceived notions that the condition of waste on-site is similar to waste samples for the previous research.

As the authors present the findings, the authors aspire to contribute not only to academic discourse but also to the practical tools available for environmental stewardship. Envisaging the application of this dataset as a resource for volunteers and government staff involved in waste amount analysis, the authors seek to catalyze a paradigm shift in approaching waste pollution. In doing so, this research emerges as a pivotal endeavor in the ongoing quest for informed decision-making and sustainable environmental solutions.

2. STUDY SITE AND METHODS

(1) Study site

Fig.1 showed the bird's eye view of the Hyakken River basin in Okayama City, Japan. The Hyakken River is an artificial river that was constructed in the 17th century as a floodway of the Asahi River to prevent flooding of Okayama Castle. The river has a unique structure that combines a low stone embankment (arate) and a drainage gutter gate at its mouth, which allows water to overflow and drain into the sea during floods. The river also serves as an irrigation canal for the surrounding farmland. In the **Fig.1**, the authors also showed the positions of the data-collection points (i.e., 24 pins in the map), that are located in the middle of the tidal zone with
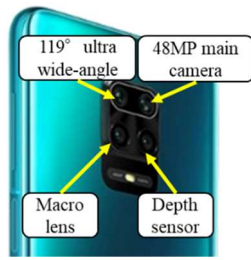
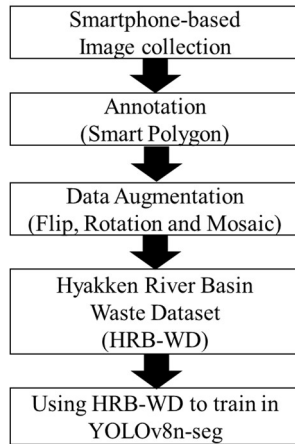Fig.2 Details of the cameras in Redmi 9s (MP: mega pixel).

Fig.3 Work flow of collecting the HRB-WD.

approximate 4.2 km distance from the river mouth.

(2) Device

In this research, data collection was derived from an Android phone called Redmi 9s, that is a low-cost smartphone from Xiaomi that comes with four cameras as shown in **Fig.2** (i.e., 119° ultra wide-angle, 48MP main camera, macro lens and depth sensor). The images were taken by the main camera, that is a 48MP high-resolution sensor with a bright f/1.79 lens. The size of the image is 4000 pixels × 3000 pixels, width and height, individually.

(3) Model

YOLOv8 is a model developed by Ultralytics, YOLOv8 is faster and more accurate than previous versions of YOLO in this company, and can handle various computer vision tasks (i.e., Object Detection and Instance Segmentation) with a single model. In this model, there are several sizes (i.e., n, s, m, l and x). YOLOv8n is the smallest version, optimized for lower speed but with lower detection performance. YOLOv8x achieves remarkable accuracy while handling a wide range of objects. Considering of how accurate the YOLOv8 model can extract the waste pollution from the complicated backgrounds, rather than only bounding-box from Object Detection, the masks aided with bounding-boxes from Instance Segmentation are more suitable for this research.

YOLOv8 model has a pre-trained version, that has

been trained on various datasets (e.g. the COCO dataset, one of the most widely used benchmarks for object detection and segmentation). In this research, the authors also trained YOLOv8 model using the dataset derived from the on-site images to compare with the results derived from the pre-trained version.

As a comparing reference of the mentioned two models (i.e., custom- and pre-trained YOLOv8 model), the segmented mask results derived from SAM also have been in consideration. The reason of choosing this model is derived from its extensive training dataset comprising 11 million images and 1.1 billion masks and zero-shot performance across a variety of segmentation tasks, even SAM cannot identify the specific classes using "Everything mode". If SAM also cannot extract the objects from the background, these objects are comparatively difficult to detect only using the image.

(4) Workflow

Fig.3 introduces a workflow designed for a waste dataset collection and application, that can support the other researchers in the data collection as a reference manual. There are simply four steps included in this workflow: Data collection, Data annotation, Data augmentation and Data combination, individually.

1. Data Collection from Smartphone Devices: The primary stage is data acquisition. Leveraging the 48MP main camera in the smartphone, the authors gathered data (i.e., images) directly from a near distance to the objects using this portable device.

2. Annotation via Roboflow's Smart Polygon Function: The next phase, the authors transited into the realm of annotation, a step in elucidating the collected data. Roboflow, an annotation generation platform, with its cutting-edge Smart Polygon function (i.e., powered by the segment anything model from Meta AI), emerges as the tool of choice for this task. As performed in **Fig.4**, there are four steps for annotating the objects, and the annotations files can be outputted automatically.

3. Data Augmentation for Image Preprocessing: Embarking on the quest for enhanced dataset quality, the authors applied data augmentation techniques on the preprocessed images. This stage aims to diversify and fortify the dataset, making it resilient to various real-world scenarios as shown in **Fig.5**. Considering of using un-color-change data augmentation methods only, **Fig.6** displayed the order of how to increase the dataset in this research.

4. Compilation of HRB-WD (Hyakken River Basin Waste Dataset): The researchers culminate in the creation of the Hyakken River Basin Waste Dataset (HRB-WD), that include the corresponding images, annotations.
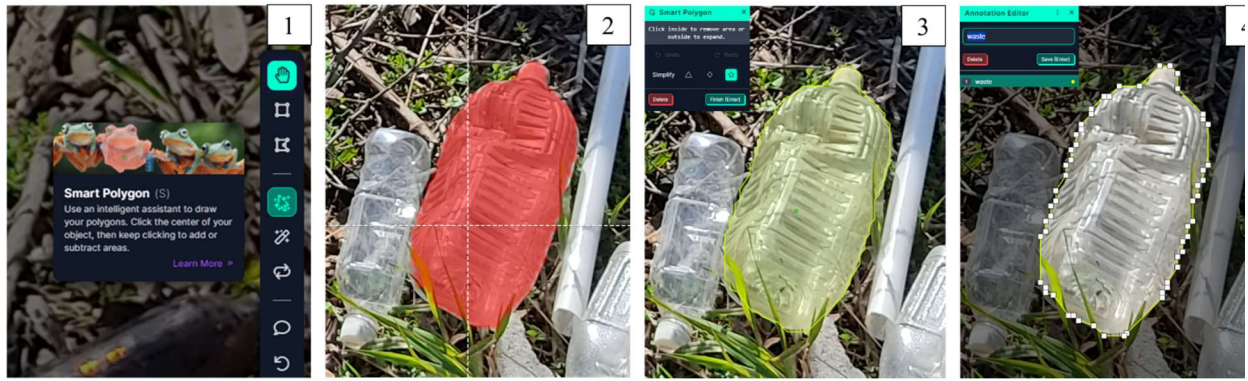
Fig.4 Process of annotating the object with roboflow: 1. Selecting the toolbox; 2. Clicking on the object; 3. Generating the masks; 4. Positions derived from the mask.

Fig.5 On-site individual waste pollution samples, that included but not limited in plastic bags, PET bottles that have been the objects for the other datasets.
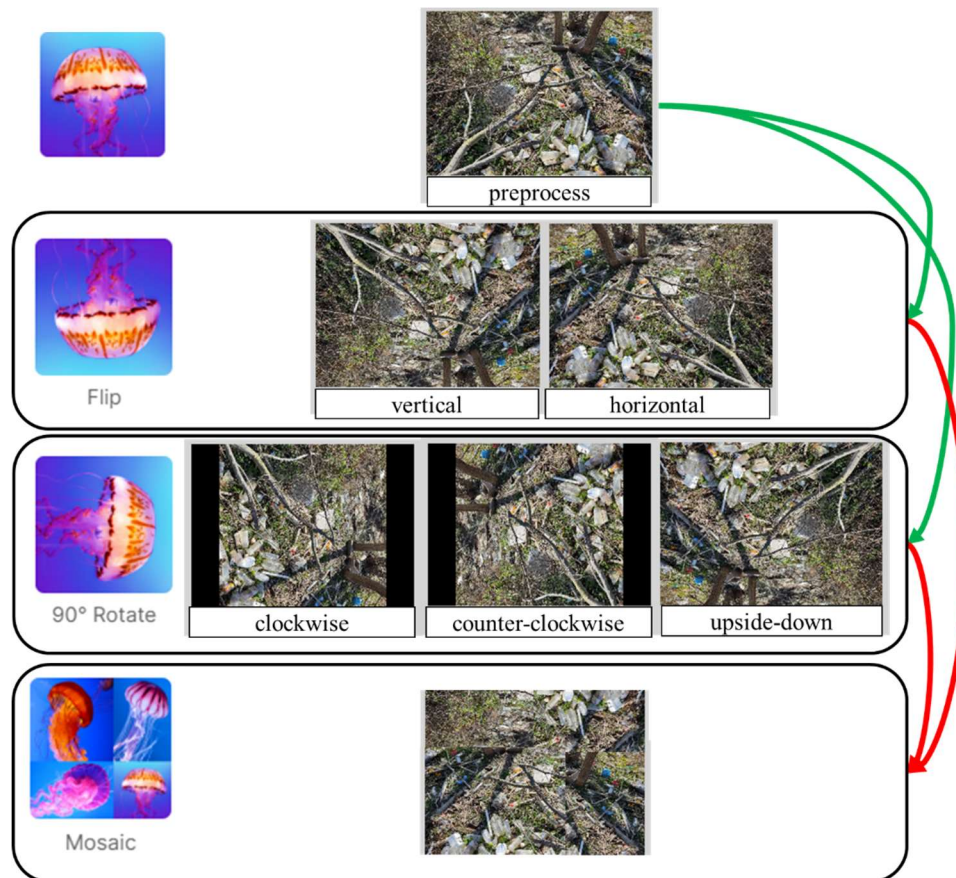

Fig.6 Data augmentation methods used in this research (Flip, 90° Rotate, Mosaic). First step is to transfer the preprocess image to Flip and 90° Rotate ones (i.e., the green line), the second step is to re-combine the transferred ones in the mosaic way.

Table 1 Model-related parameter setting.

Parameters	Configuration (YOLOv8n-seg)
Operating system (Version)	Linux (Ubuntu 20.04.4 LTS)
GPU (VRAM)	NVIDIA GeForce RTX 3090 Ti (24 GB)
CUDA	11.7
Driver Version	515.65.01
Imgsz (pixel)	1024
Initial learning rate	0.01
Final learning rate	0.1
Optimizer	Adam
Momentum	0.937
Batch size	10
Epochs	1000
Patience	50

5. Training the YOLOv8 model using HRB-WD: As the parameters shown in **Table 1**, the authors trained the YOLOv8n-seg model for the on-site practical waste segmentation. Shown in the **Fig.7**, the model has stopped at epochs 251 with a smooth going-down val/box_loss and val/seg_loss curves.

(5) Objects of this work

The work aims to validate the efficacy of the pre-trained models (YOLOv8n-seg, YOLOv8x-seg and SAM) and custom-trained YOLOv8n-seg model through on-site testing using smartphone videos from distinct field sections, providing real-world validation on the waste pollution after long-term sunshining and water brush-up.

In this research, the authors did not separate the waste pollution into different species, just considered as one class. The reason of using just one label is based on the randomness of the on-site waste pollution species, and the similar color of the objects. In another word, the authors are seeking for something that are not made by natural environment with the specified shapes (i.e., bottle shape).

Additionally, the research seeks to contribute to environmental protection initiatives and emphasize the imperative for sustainable solutions. The generated dataset is envisioned as a valuable resource supporting both volunteers and government staff in waste amount analysis, thereby facilitating informed decision-making in the future, especially in the case of analyzing the amount of the workload for calculating the volunteers' numbers estimation. In this research, the authors prefer to upload the dataset as the benchmark for the practical application.

3. RESULTS AND DISCUSSION

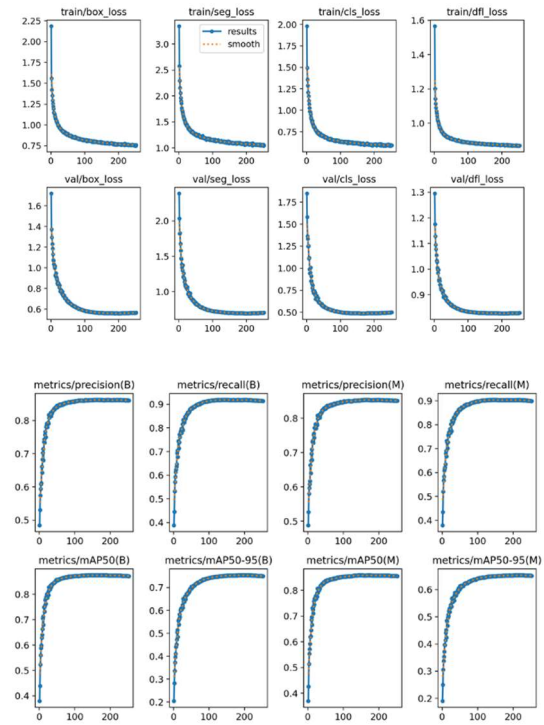**(1) Result derived from YOLOv8n-seg (HRB-WD)**

Fig.7 Results of the trained YOLOv8n-seg derived from HRB-WD.

The results showcased in **Fig.8 (a)** that the trained model has detected 24 waste instances. And all the waste instances have been detected over 0.6 confidence. When the waste pollution overlapped with each other, especially the colors of the objects are close, the instance segmentations among the objects were difficult.

(2) Result derived from YOLOv8n/x-seg (pre-trained)

The results in **Fig.8 (b)** performed that the pretrained model has not successfully detected any bottle instance. And in the **Fig.8 (c)**, 16 bottle instances (total approximate 40 bottle instances in this image, detection ratio is almost 40%) have been detected over 0.6 confidence. Comparing with the Result-1, even the YOLOv8x-seg (pre-trained) has more parameters and trained with the COCO dataset, the YOLOv8n-seg (HRB-WD) performed much better.

(3) Result derived from SAM

Fig.8 (d) has showed the segmentation of all the factors in the image, and almost all the waste pollutions have been separated from the background as individual instances or partly. Based on this result, YOLOv8 model also has the potential of improving the accuracy in extracting the waste pollution from the background to a much better level with more data.

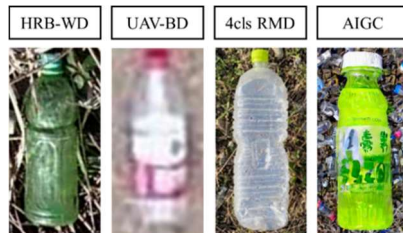**Fig.8** Comparison of waste-based segmentation from different models.

Fig.9 Comparison of PET-bottle-objects used in different dataset

(4) Discussion

From the mentioned results, YOLOv8n-seg (HRB-WD) has performed the advantages comparing with the pre-trained YOLOv8n-seg, even the YOLOv8x-seg (pre-trained) cannot also overcome the result derived from YOLOv8n-seg (HRB-WD). **Fig.8 (d)** displayed the well-segmented result without labels. In the future, the combination of YOLOv8 and SAM are in consideration for the waste pollution detection.

Contrary to the instances settled in the other benchmark dataset[4] (i.e., UAV-BD and 4cls RMD) or AIGC[8] as performed in the **Fig.9**, this research underscores the practical riparian waste pollutions are in a totally different situation (i.e., similar color to the background and natural environment).

This study also emphasizes the potential of the generated dataset as a valuable tool for waste amount analysis, supporting both volunteers and government staff in environmental protection initiatives[9]. Looking ahead, this research aspires to contribute to informed decision-making and sustainable solutions in the ongoing battle against waste pollution.

4. CONCLUSION

This study presented a novel dataset for waste pollution segmentation in riparian area using a smartphone equipped with a 48MP digital camera for data collection, the YOLOv8n model for data analysis, and collected dataset was uploaded for the benchmark. There are also some limitations in waste instances segmentation, and the dataset also need to be improved in the future.

REFERENCES

1) Hong, J., Fulton, M., & Sattar, J. Trashcan: A semantically-segmented dataset towards visual detection of marine debris, *arXiv preprint arXiv:2007.08097*, 2020.
2) http://tacodataset.org/
3) https://github.com/garythung/trashnet
4) Pan, S., Yoshida, K., Boney, A. S., & Nishiyama, S. The Application of Drone-Assisted Deep Learning Technology in Riverbank Garbage Detection. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, 78(2), I_133-I_138, 2022.
5) https://roboflow.com/
6) https://github.com/ultralytics/ultralytics
7) Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P. & Girshick, R. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023.
8) Shijun, P. A. N., YOSHIDA, K., & KOJIMA, T. Application of the Prompt Engineering-assisted Generative AI for the Drone-based Riparian Waste Detection. *Intelligence, Informatics and Infrastructure*, 4(2), 50-59, 2023.
9) Shovon, S. M., Akash, F. A., Rahman, M. A., Rahman, W., Chakraborty, P., Monir, M. U., Karim, K. M. R., Habib, M. A. & Khan, M. F. H. The potential for sustainable waste management and energy recovery in Bangladesh: A review. *Sustainable Energy Technologies and Assessments*, 64, 103705, 2024.