Balancing Accuracy and Efficiency for River Plastic Monitoring

IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Proceedings

Saddi, Khim Cathleen; Miglino, Domenico; Isgro, Francesco; Van Emmerik, Tim H.M.; Manfreda, Salvatore

https://doi.org/10.1109/IGARSS53475.2024.10641940

IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10641940

BALANCING ACCURACY AND EFFICIENCY FOR RIVER PLASTIC MONITORING

Khim Cathleen Saddi [1,2,3], Domenico Miglino [1], Francesco Isgro [4], Tim H. M. van Emmerik [5], and Salvatore Manfreda [1]

1 Dipartimento di Ingegneria Civile, Edile e Ambientale (DICEA), Università degli Studi di Napoli Federico II, Napoli, Italy;
2 Istituto Scuola Superiore Pavia (IUSS Pavia), Pavia, Italy;
3 Department of Civil Engineering and Architecture, Ateneo de Naga University, Naga, Philippines;
4 Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione (DIETI), Università degli Studi di Napoli Federico II, Napoli, Italy;
5 Hydrology and Environmental Hydraulics Group, Wageningen University, Wageningen, The Netherlands

ABSTRACT

The availability and use of drones, fixed cameras, and smartphones allows the advancement in close range optical remote sensing due to its operability and viability for large-scale implementation. Specifically, RGB cameras are particularly advantageous for low-cost close-range sensing, offering higher spatial resolution at maximum two orders of magnitude compared to satellite sensors. This characteristic proves invaluable in environmental monitoring applications, such as tracking plastic movement in riverine environments. However, one of the pressing issues in object detection is to find a balance between accuracy and computational cost even with single stage object detection algorithms such as You Only Look Once (YOLO). Considering the environmental variabilities such as weather and different lighting conditions, here we attempt to demonstrate uncomplicated data augmentation techniques for both train and test images with primary focus on lighting adjustment. The goal is to determine which approach will increase performance with minimum cost.

Index Terms— river plastic detection, optical remote sensing plastic, YOLO, improving object detection, image processing techniques

1. INTRODUCTION

Global quantification of plastic pollution abundance and fluxes is hindered by a lack of field observations. It is undeniable that it is crucial to have extensive data on plastic release and transport to advance the understanding of plastic fates and optimize management and mitigation efforts.

The unbalanced plastic release versus those found in the environment, especially in marine and coastal areas, led to refocus the spotlight to riverine environments–which are also now seen as potential plastic reservoirs [1]. Many diverse attempts to advance plastic monitoring were undertaken, mostly for the marine environment, to develop plastic monitoring methodologies which would harmonize and widen data collection, including the use of Uncrewed Aerial Systems (UAS), sonar, hyper and multispectral spectral imaging, optical sensing, and capacitance proximity sensors [2,3]. However, most of these technologies are expensive or in its early development, and not yet suitable for large-scale implementation.

Recent advances in computer vision allow detection of objects in the environment. This approach for plastic transport monitoring, also including other hydrological data and different camera systems, was highlighted to have a promising potential [4]. In fact, plastic object detection has been demonstrated effective in plastic recycling facilities using Near-Infrared Spectroscopy [5], and recently optical sensing coupled with computer vision [6]. However, in these sites and other enclosed environments, lighting conditions are often regulated, and other environmental variabilities do not exist. In the case of river plastic monitoring, water background, lighting conditions, and varying geometry during transport poses the biggest challenge.

The YOLO (You Only Look Once) family of single detection algorithms is one of the most popular object detection algorithms in recent years due to its fast inference time [7]. Its implementation is simple and readily applicable for user-defined object detection needs, hence its growing popularity for different monitoring systems including traffic [8], and UAS [9] detection. Unlike river plastics, these objects do not geometrically transform rapidly during transport, but are also often influenced by weather variability.

In addressing adverse weather conditions for YOLO detection, techniques like adding a dark channel defogging algorithm [10], or as far as using different image-adaptive hyperparameters such as Defog, White Balance, Gamma, Contrast, Tone and Sharpen [11] may help improve object detection for cars and pedestrians. There are also more

complicated layers being added to the YOLO network architecture like context loss [12], positive sample selection [13], and transformer module [14]. These approaches change the model architecture, which adds more complexity, and most probably increases inference time. Data augmentation, also another approach to increase the dataset size, on the other hand also provides an opportunity for increasing performance with additional inference cost and minimal training cost [15]. These entails image processing techniques such as translation, rotation, or shearing [16] and additional instance processing cost, which could be labor intensive for large datasets.

To increase performance, high interest in recent research focuses more on developing different transformations of available base algorithms rather than developing methods to improve the maximum learning curve of datasets. While data augmentation has been implemented in previous works, little is known on how to develop a good dataset. The application of dark channel has provided improvements on images with weather disturbance in previous works when the image is augmented prior to detection (test image augmentation). This is beneficial for seasonal variabilities in the image, but not for long-term variability in lighting conditions and effect of the water spectral properties in river plastics. This case might require full inclusion of these special image characteristics in the training dataset (train image augmentation).

We decided to incorporate both and focus on understanding the potential gains of this proposed dataset manipulation before and after the training process. Utilizing both YOLOv7 and YOLOv8, this paper demonstrates the potential value addition of simple data augmentation using image processing techniques to improve river plastic detection without suffering from increased labor and computational cost–which the YOLO family has always been.

2. MATERIALS AND METHODS

We explored two approaches to improve YOLO detection performance: 1) train image augmentation, and 2) test image augmentation.

Train image augmentation includes dark channel adjustment (gamma=0.4) to all images of the dataset, doubling its size, without complex annotation processing since the bounding boxes are still the same with the raw images. Essentially, adjusting the gamma value of an image results in a lighter or darker version, in this case, a darker version was created. For the test image augmentation, apart from the dark (gamma=0.4), bright (gamma=1.7), and contrast (0.2 0.3 0; 0.6 0.7 1) adjustments were added to the input image prior to feeding into the YOLO detector, to widen the assessment about these lighting adjustments.

A publicly available river plastic dataset was used as reference taken from Jakarta, Indonesia [17]. From this, there were three dataset versions utilized in this paper: 1) J-excerpt from the Jakarta dataset, 2) J+Dark- J and dark version of J, and 3) J50+Dark50- half of the previous version (see Table 1). Both YOLOv7 and YOLOv8 were trained for all three datasets.

Table 1. Dataset Description

Name	Images (T/V)	Annotations (T/V)
J	237/27	9,582/644
J+Dark	474/54	19,164/1,288
J50+Dark50	238/28	1712/576

**T-train, V-Validation*

There were three classes in the training: 0-plastic, 1-plastic bottle, 2-plastic bag. Plastic bottles and plastic bags were two of the most common plastic litter found in oceans in the region of Indonesia (South Asia) [18]. We decided to limit the classes to three, so all the other types of plastic fell under the class 0 for brevity.

Image processing was performed in Matlab while the algorithm training and detection was performed in python (Google Colab Notebook). Both YOLOv7 and YOLOv8 were trained using 100 epochs to make the performance results comparable. The test image used was from another river (Sarno, Italy) with a different camera angle from the train dataset to set the analysis in transfer learning levels.

3. RESULTS

Training results showed that YOLOv8 outperforms YOLOv7 in terms of mAP@50-95 performance. The highest mAP@50-95 was observed for J+Darkv8 (+17% improvement from base reference Jv8) and lowest performance for J50+Dark50v7 (-61% decrease from base reference Jv7, base reference has the highest mAP@50-95 for all YOLOv7 models) (see Figure 1).

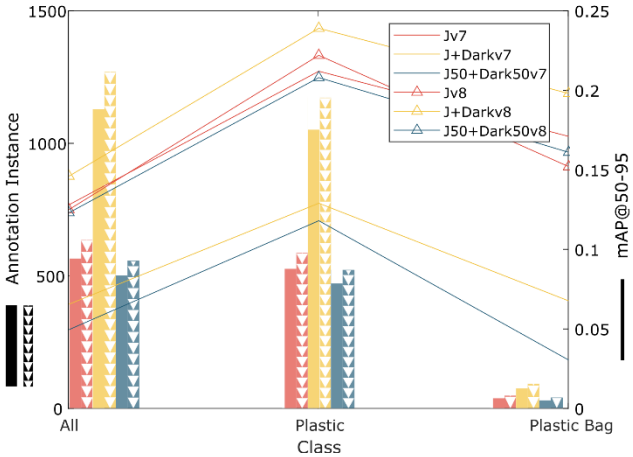

Figure 1. Instance and mAP for YOLOv7 and YOLOv8

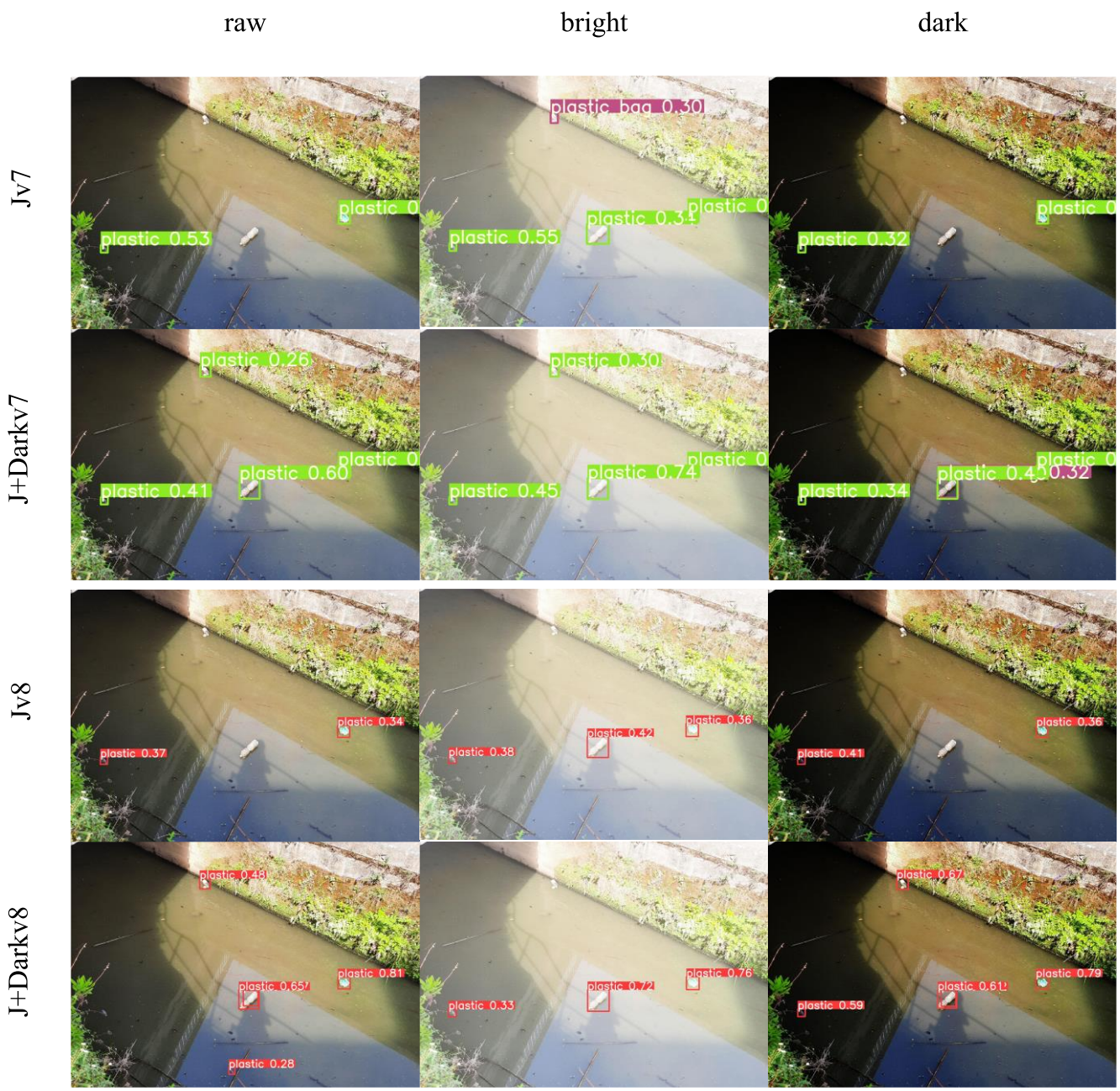

Figure 2. Detection using YOLOv7 and YOLOv8 models. The algorithms were tested on an image taken in Sarno River, Italy, which is different from the dataset site (Jakarta, Indonesia) to demonstrate transferability of the algorithm, and avoid biased analysis which would lead to overfitting. All models performed poorly when tested to the augmented image with adjusted contrast settings (samples not shown here).

Generally, an increased mAP@50-95 performance for all YOLOv8 models can be observed with the incorporation of dark images in the dataset independent of the annotation instance (class count in the validation set). It is also interesting that most of the model performance gains (J+Darkv7 and J+Darkv8) from incorporating dark images was negated later when the dataset was halved (J50+Dark50v7 and J50+Dark50v8), but not so far off from the raw dataset (Jv7 and Jv8).

The models were tested with a river image independent from the Jakarta dataset using raw and augmented versions. For this specific image, the use of an augmented version which is brighter resulted in higher inference scores for most of YOLOv7 models while there is a minimal increase observed for the YOLOv8 models. In fact, there were additionally detected plastics in JDarkv7

and JDarkv8 which their base counterparts Jv7 and Jv8 were not able to detect (see Figure 2).

Approximately 38% of the inference scores are <0.50 for YOLOv8 models, while only 18% for YOLOv7 models. False positives were also produced mainly due to insufficient class annotation instances. Jv7 and J+Darkv7 detected a plastic bottle as a plastic bag and plastic, respectively, while J+Darkv8 model tagged it as plastic. This is not surprising since the plastic bottle class only comprises a minimal annotation instance in the dataset (annotations=30).

4. DISCUSSION

Taking into consideration the pressing need to advance river plastic monitoring, without compromising accuracy, it is critical to balance time and resources to address policy gaps [19] coupled with effective monitoring technologies.

In object detection, high accuracy often entails cost, both computational and labor. In this case, the +17% mAP@50-95 increase of J+Darkv8 cost only an additional 24 mins (approximately doubling) in the training time from the base Jv8. Minimal labor was also added for the data processing and dataset preparation (maximum of 30 mins per dataset, execution only). Since there was no change implemented in the YOLO network architecture, which is often the case in the literature, the challenge of deployment cost for hardware may not threaten adoption of this approach in the future. The uncomplicated train and test data augmentation proves to have its pros and cons and though it is still early to conclude that we should use darker over brighter augmented images, it is highly encouraging to use a simple threshold to automate brightness adjustment and explore object specific augmentations in the future.

Interestingly, despite the general low performance of the YOLO family in terms of detecting small objects, our current models were able to detect a 50 x 70 px plastic bottle.

It is also important to highlight that setting the training parameters (e.g. batch and epoch) as constant is beneficial to compare algorithm performance, this might also lead to limiting specific algorithms. YOLOv8 detected a relatively lower increase in the learning curve within 100 epochs and automatically stopped the training, but it's not the case with YOLOv7, which could still continuously run for higher epochs (Jv7 with epoch=300 outperform Jv8 with epoch =100).

5. CONCLUSION

This work demonstrated the effectiveness of increasing river plastic detection using simple data augmentation methods, especially the dark channel in the training set, without suffering from increased computational cost or changing the architecture of the object detector. Moving forward, it is essential that the community continue to build simpler systems which can be easily integrated with those that are already working to advance to the next level of science in close range optical remote sensing of river plastic.

Related to this work is an ongoing project, RiverWatch, which aims to develop optical monitoring systems for Sarno River using unsupervised computer vision and citizen science.

11. REFERENCES

[1] T. van Emmerik, Y. Mellink, R. Hauk, K. Waldschläger and L. Schreyers, "Rivers as Plastic Reservoirs". *Front. Water*, vol. 3, 786936, 2022.

[2] M. Farré, "Remote and in situ devices for the assessment of marine contaminants of emerging concern and plastic debris detection", *Current Opinion in Environmental Science & Health*, vol. 18, pp. 79-94, 2020.

[3] M. Sandra, L. I. Devriese, A. M. Booth, B. De Witte, G. Everaert, J. Gago, F. Galgani, K. Langedock, A. Lusher, T. Maes et al., "A systematic review of state-of-the-art technologies for monitoring plastic seafloor litter", *Journal of Ocean Engineering and Science,* 2023, in press.

[4] S. Manfreda, D. Miglino, K.C. Saddi, S. Jomaa, A. Etner, M. Perks, S. Peña-Haro, T. Bogaard, T.H.M. van Emmerik, S. Mariani et al., "Advancing river monitoring using image-based techniques: challenges and opportunities", *Hydrological Sciences Journal*, pp. 1-21, 2024.

[5] J. Beigbeder, D. Perrin, J.F. Mascaro, and J.M. Lopez-Cuesta, "Study of the physico-chemical properties of recycled polymers from waste electrical and electronic equipment (WEEE) sorted by high resolution near infrared devices", *Resources, Conservation and Recycling,* vol. 78, pp. 105-114, 2013.

[6] J. Choi, B. Lim, and Y. Yoo, "Advancing Plastic Waste Classification and Recycling Efficiency: Integrating Image Sensors and Deep Learning Algorithms", *Applied Sciences,* vol. 13 (18), 10224, 2023.

[7] T. Diwan, G. Anirudh and J.V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications", *Multimed Tools Appl,* vol. 82, pp. 9243–9275.

[8] A. Ćorović, V. Ilić, S. Đurić, M. Marijan, and B. Pavković, "The real-time detection of traffic participants using YOLO algorithm", In *2018 26th Telecommunications Forum (TELFOR)*. IEEE, 2018, pp. 1-4.

[9] X. Zhai, Z. Huang, T. Li, H. Liu, H., and S. Wang, "YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection", *Electronics*, vol. 12(17), 3664, 2023.

[10] X. Liu and Y. Lin, "YOLO-GW: Quickly and Accurately Detecting Pedestrians in a Foggy Traffic Environment", *Sensors*, vol. 23, 5539, 2023.

[11] W. Liu, G. Ren, R. Yu, S. Guo, J. Zhu, and L. Zhang, "Image-Adaptive YOLO for Object Detection in Adverse

Weather Conditions", In *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Press, 2020, vol. 36 (2), pp. 1792-1800.

[12] G. Oreski, "YOLO* C—Adding context improves YOLO performance", *Neurocomputing*, vol. 555, 126655, 2023.

[13] C. Wang, Q. Wang, Y. Qian, Y. Hu, Y. Xue, and H. Wang, "DP-YOLO: Effective Improvement Based on YOLO Detector", *Applied Sciences*, vol. 13(21), 11676, 2023.

[14] Z. Guo, C. Wang, G. Yang, Z. Huang, and G. Li, "Msft-yolo: Improved yolov5 based on transformer for detecting defects of steel surface", *Sensors*, vol. 22(9), 3467, 2022.

[15] B. Zoph, E.D. Cubuk, G. Ghiasi, T.Y. Lin, J. Shlens and Q.V. Le, "Learning data augmentation strategies for object detection", In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020*, *Proceedings, Part XXVII 16*. Springer International Publishing, 2020, pp. 566-583.

[16] E.D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q.V. Le, "AutoAugment: learning augmentation policies from data", arXiv preprint arXiv:1805.09501, 2018.

[17] C. van Lieshout, K. van Oeveren, T.H.M. van Emmerik, and E. Postma, "Automated river plastic monitoring using deep learning and cameras", *Earth and space science*, vol. 7 (8), e2019EA000960.

[18] C. Morales-Caselles, J. Viejo, E. Martí, D. González-Fernández, H. Pragnell-Raasch, J.I. González-Gordillo, E. Montero, G. M. Arroyo, G. Hanke, V. S. Salvo et al. An inshore–offshore sorting system revealed from global classification of ocean litter, *Nature Sustainability*, vol. 4(6), pp. 484-493, 2021.

[19] A. L. Lusher and S. Primpke, "Finding the Balance between Research and Monitoring: When Are Methods Good Enough to Understand Plastic Pollution?", *Environmental Science & Technology*, vol. 57 (15), pp. 6033-6039, 2023.