Volume 4 Number 2 (2026)
July-December 2026
Page: 614-624

Journal of Creative Power and Ambition (JCPA)
https://edujavare.com/index.php/jcpa
E-ISSN: 3031-4054
DOI: 10.70610/jcpa.1666

Improving YOLOv8n Robustness via Deformation-Aware Augmentation and Class Balancing for Plastic Resin Code Classification

Nyoman Sarasuartha Mahajaya[1], I Gusti Made Ngurah Desnanjaya[2]
[12] Institut Bisnis dan Teknologi Indonesia, Denpasar, Indonesia
* Correspondence e-mail; nyoman_sarasuartha@instiki.ac.id, ngurah.desnanjaya@instiki.ac.id

Article history Submitted: 2026/04/01; Revised: 2026/05/11; Accepted: 2026/07/17

Abstract

Computer vision-based plastic resin code classification is an important component in supporting automated plastic waste sorting systems. However, deep learning models often experience performance degradation when plastic objects undergo physical deformation, such as dents, folds, or damage. In addition, class imbalance may cause models to be biased toward majority classes and perform poorly on minority classes. This study aims to improve the robustness of YOLOv8n for plastic resin code classification through a combination of deformation-aware augmentation and class balance. The WaDaBa dataset was used, consisting of 4,000 plastic images categorized into five classes, namely PET, PE-HD, PP, PS, and Others, with four physical deformation levels: intact, mild, moderate, and severe. The experiment was conducted using four scenarios: YOLOv8n baseline, YOLOv8n with class balance, YOLOv8n with deformation-aware augmentation, and the proposed YOLOv8n combining both strategies. The evaluation metrics include Accuracy, Macro-F1, Weighted-F1, per-class recall, accuracy across deformation levels, latency, and FPS. The experimental results show that the proposed YOLOv8n achieved the best overall performance, with an Accuracy of 76.82%, Macro-F1 of 0.7026, Weighted-F1 of 0.7613, latency of 4.87 ms per image, and throughput of 205.2 FPS. Compared with the baseline, the proposed method improved Accuracy by 4.37 points and Macro-F1 by 0.0715. Furthermore, accuracy under severe deformation increased from 70.59% to 73.53%. These findings indicate that the combination of deformation-aware augmentation and class balance can improve YOLOv8n robustness without sacrificing inference efficiency.

Keywords YOLOv8n, plastic, resin, deformation, class balance, robustness

INTRODUCTION

Plastic waste is a growing environmental challenge, driven by the increasing consumption of plastic products. The success of the recycling process depends heavily on the system's ability to accurately sort plastic types, as each resin type has different material characteristics and processing requirements. Resin codes such as

PET, PE-HD, PP, PS, and Others need to be accurately recognized to ensure effective, fast, and consistent sorting. In practice, manual identification still has limitations due to the labor-intensive nature of the process, the relatively long time required, and the potential for error due to visual similarities between plastic types.

Developments in computer vision and deep learning have opened up opportunities for developing automated plastic waste sorting systems. Recent studies have shown that deep learning approaches can be used to improve the effectiveness of waste classification and detection, both in general classification scenarios, plastic waste, and real-time sorting systems [1]–[5]. Convolutional Neural Network-based models and the You Only Look Once (YOLO) family are widely used due to their ability to balance accuracy and inference speed, making them potentially applicable to both automated sorting systems and edge devices.

However, the application of deep learning to plastic resin code classification still faces two major challenges. First, plastic objects in real environments often experience physical deformation, such as dents, folds, deformation, or damage due to use, collection, and transportation. Physical deformation can alter the object's visual pattern, thus reducing the model's ability to consistently recognize resin classes. Second, the class distribution in plastic waste datasets tends to be imbalanced. Some classes, such as PET, can have a significantly larger number of images than minority classes, such as Others. This condition can cause the model to more easily recognize the majority class, but less optimally recognize the minority class.

The WaDaBa dataset is a relevant dataset for plastic resin code classification research because it provides metadata containing information about the resin type and the degree of physical deformation of the object. Research related to WaDaBa has shown that this dataset can be used as a benchmark for deep learning-based plastic waste classification and detection [6], [7]. The presence of deformation information makes WaDaBa suitable for evaluating model robustness, not just overall accuracy.

Previous research, cross-generational robustness evaluation of YOLO, showed that YOLOv8n outperformed YOLOv11n and YOLOv12n in plastic resin code classification. However, these results also indicate that model performance can still decline under severe deformation conditions, and that certain classes are more difficult to classify. Therefore, further research is needed to improve YOLOv8n's robustness to physical deformation and class imbalance.

Based on these issues, this study proposes an approach to improve YOLOv8n's robustness through a combination of deformation-aware augmentation and class balance. Deformation-aware augmentation is used to enrich the visual diversity of training data through transformations that mimic the deformation conditions of plastic objects, while class balance is used to mitigate the influence of data imbalance between classes. This study evaluates four experimental scenarios: baseline YOLOv8n, YOLOv8n with class balance, YOLOv8n with deformation-aware augmentation, and the proposed YOLOv8n, which combines both strategies.

The main contributions of this study are as follows. First, this study develops a deformation-aware augmentation approach to improve YOLOv8n's robustness to variations in the shape of plastic objects. Second, this study applies a class balance strategy to reduce model bias toward the majority class. Third, this study systematically compares the four experimental scenarios to assess the impact of each strategy. Fourth, this study evaluates the model performance multidimensionally using Accuracy, Macro-F1, Weighted-F1, recall per class, accuracy per deformation level, latency, and FPS..

METHODS

This study employed a quantitative experimental approach to evaluate the effects of **deformation-aware augmentation** and **class balancing** on the performance of the YOLOv8n classification model for plastic resin code recognition. The research workflow consisted of dataset acquisition, metadata parsing, resin class mapping, dataset splitting, experimental scenario development, model training, performance evaluation, and result analysis. The WaDaBa dataset was used as the primary data source because it provides image-level metadata embedded in file names, including object identity, resin code, color, lighting condition, deformation level, dirt, cap, ring, and object position. Since deformation information is explicitly available, the dataset is particularly suitable for evaluating model robustness under varying physical object conditions. The metadata followed the format *ObjectID + ResinCode + Color + Light + Deformation + Dirt + Cap + Ring + Position*. Resin codes were mapped into five classification categories: PET, PE-HD, PP, PS, and Others, where the Others category combined PVC, PE-LD, Other, and unreadable objects. After metadata parsing, a total of 4,000 valid images representing 100 unique objects were successfully processed. The resulting class distribution was highly imbalanced, with PET accounting for 2,200 images, whereas the Others class contained only 40 images, indicating the need for class balancing to reduce prediction bias toward majority classes. In contrast, the deformation levels were evenly distributed across four categories—intact, light, moderate, and severe—with 1,000 images each, enabling a fair evaluation of model robustness across different degrees of physical deformation. The dataset was divided into training, validation, and testing subsets using a 70:15:15 ratio at the object level to prevent data leakage, ensuring that images of the same object did not appear in multiple subsets. Because the Others category contained only one unique object, its split was performed at the image level and acknowledged as a limitation of the study. The final split consisted of 2,748 training images, 566 validation images, and 686 testing images, with the same testing set used across all experimental scenarios to ensure fair performance comparisons.

Four experimental scenarios were designed to investigate the contributions of data balancing and deformation-aware augmentation: (S0) the baseline YOLOv8n model trained on the original dataset, (S1) YOLOv8n with class balancing, (S2) YOLOv8n with deformation-aware augmentation, and (S3) the proposed method combining both strategies. The baseline scenario served as the reference model, whereas the proposed approach represented the complete framework evaluated in this study. Deformation-aware augmentation enriched the training data by simulating realistic variations in plastic waste through rotations, perspective transformations, shear, translation, brightness and contrast adjustments, slight blur, and other visual modifications while preserving the original class labels, mathematically represented as ($X' = A(X)$), where (X) is the original image, (A) denotes the augmentation function, and (X') is the augmented image. Meanwhile, class balancing increased the number of minority-class samples until they approached the size of the majority class, thereby providing a more balanced learning opportunity for all resin categories. As a result, the proposed method produced the largest and most balanced training dataset, containing 17,100 training images with 14,352 augmented samples equally distributed across all five classes. All models were trained using the YOLOv8n-cls architecture under identical settings, including an image size of 224 × 224 pixels, batch size of 128, 30 training epochs, AdamW optimizer, random seed 42, and Tesla T4 GPU within the Kaggle Notebook environment. Maintaining identical training configurations ensured that performance differences originated solely from the proposed data preparation strategies rather than hyperparameter variations. Model performance was evaluated comprehensively using Accuracy, Macro-F1 Score, Weighted-F1 Score, per-class recall, deformation-level accuracy, inference latency, and Frames Per Second (FPS), allowing the assessment to measure not only overall classification accuracy but also class balance, robustness against physical deformation, and suitability for real-time deployment.

FINDINGS AND DISCUSSION

The results of the evaluation across the four experimental scenarios are presented in

Table 1. Overall Performance Comparison

Scenario	**Accuracy (%)**	**Macro-F1**	**Weighted-F1**	**Parameters (M)**	**Model Size (MB)**	**Latency (ms/image)**	**FPS**
YOLOv8n Baseline	72.45	0.6311	0.7298	1.44	2.97	4.91	203.8
YOLOv8n + Class Balance	75.95	0.6466	0.7465	1.44	2.97	4.96	201.5
YOLOv8n + Deformation Augmentation	76.38	0.6482	0.7495	1.44	2.97	4.78	209.4
Proposed YOLOv8n	76.82	0.7026	0.7613	1.44	2.97	4.87	205.2

Table 1 shows that the proposed YOLOv8n achieved the best overall performance among all experimental scenarios. Accuracy increased from **72.45%** in the baseline model to **76.82%** in the proposed method. Likewise, the Macro-F1 score improved from **0.6311** to **0.7026**. The improvement in Macro-F1 indicates that the proposed method not only increased the overall number of correct predictions but also enhanced the balance of performance across different resin classes. The class balance scenario improved Accuracy to **75.95%** and Macro-F1 to **0.6466**, demonstrating that balancing the training data positively contributed to model performance. Meanwhile, the deformation augmentation scenario achieved an Accuracy of **76.38%** and a Macro-F1 score of **0.6482**, indicating that deformation-based augmentation enabled the model to better recognize objects with more diverse visual characteristics.

From an efficiency perspective, the proposed YOLOv8n maintained a low inference latency of **4.87 ms per image**, corresponding to **205.2 FPS**. These results demonstrate that the performance improvements were achieved without sacrificing inference speed. Therefore, the proposed method remains suitable for deployment in real-time applications or edge computing devices.

Per-Class Recall Analysis

Per-class recall was used to evaluate the model's ability to recognize each plastic resin category. The per-class recall results are presented in **Table 8**.

Table 2. Per-Class Recall

Scenario	Others	PE-HD	PET	PP	PS
YOLOv8n Baseline	100.00	50.83	95.56	50.00	32.50
YOLOv8n + Class Balance	100.00	45.00	98.61	67.50	31.25
YOLOv8n + Deformation Augmentation	100.00	44.17	98.33	73.33	28.75
Proposed YOLOv8n	100.00	57.50	99.44	40.00	57.50

Based on Table 8, all experimental scenarios successfully recognized the **Others** class with a recall of **100%**. However, it should be noted that the **Others** category contains only a small number of testing samples; therefore, the perfect recall for this class should be interpreted cautiously. For the **PET** class, the proposed YOLOv8n achieved the highest recall of **99.44%**, indicating excellent recognition of the majority class. The proposed YOLOv8n also produced the best recall values for the **PE-HD** and **PS** classes, each reaching **57.50%**. The improvement in the PS class is particularly significant because the baseline model achieved only **32.50%** recall. These findings indicate that the combination of deformation-aware augmentation and class balancing effectively improved the recognition of classes that were previously difficult to classify.

However, different results were observed for the **PP** class. The highest PP recall was achieved by the deformation augmentation scenario at **73.33%**, whereas the proposed YOLOv8n obtained a PP recall of only **40.00%**. This finding indicates that although the proposed method achieved the best overall performance, the improvements were not distributed evenly across all classes. The PP class remains challenging because it likely shares similar visual characteristics with the PE-HD and PS classes. This observation highlights that the class balancing and augmentation strategies require further refinement to prevent performance degradation in the PP class when both approaches are combined within the proposed framework.

Robustness Analysis Based on Deformation Level

Performance evaluation across different deformation levels was conducted to assess the model's ability to maintain classification performance when objects undergo physical shape changes. The evaluation results are presented in **Table 9**.

Table 3. Accuracy Across Deformation Levels

Scenario	Intact	Light	Moderate	Severe	Robustness Gap (Intact–Severe)
YOLOv8n Baseline	71.26	75.44	72.51	70.59	0.67
YOLOv8n + Class Balance	73.56	77.19	81.29	71.76	1.80
YOLOv8n + Deformation Augmentation	73.56	78.95	79.53	73.53	0.03
Proposed YOLOv8n	74.14	77.78	81.87	73.53	0.61

Table 9 shows that the proposed YOLOv8n achieved the highest accuracy under the **intact deformation condition** at **74.14%** and under the **moderate deformation condition** at **81.87%**. Under **severe deformation**, both the proposed YOLOv8n and the deformation augmentation scenario achieved the same accuracy of **73.53%**, outperforming the baseline model, which achieved **70.59%**.

The improvement in accuracy under severe deformation demonstrates that deformation-aware augmentation enhances the model's robustness against physical shape variations. The deformation augmentation scenario achieved the smallest **Robustness Gap** of **0.03**, indicating excellent stability between intact and severely deformed conditions. Meanwhile, the proposed YOLOv8n achieved a Robustness Gap of **0.61**, which remains considerably better than the class balance scenario with a gap of **1.80**.

These findings indicate that deformation-aware augmentation plays an important role in improving model robustness. When combined with class balancing, the model achieves the best overall performance while maintaining high classification accuracy even under severe deformation conditions.

Inference Efficiency Analysis

In addition to classification performance, inference efficiency is an important consideration because automated plastic waste sorting systems require fast classification. Based on the experimental results, all scenarios achieved an inference latency below **5 ms per image** and throughput exceeding **200 FPS**. These results demonstrate that YOLOv8n remains computationally efficient even when the training dataset is expanded through augmentation and class balancing.

The deformation augmentation scenario achieved the fastest inference latency of **4.78 ms per image**, corresponding to **209.4 FPS**. The proposed YOLOv8n achieved a latency of **4.87 ms per image** with a throughput of **205.2 FPS**. Compared with the baseline model, which achieved **4.91 ms** latency and **203.8 FPS**, the proposed method remained highly competitive in terms of inference speed.

With a model size of only **2.97 MB** and approximately **1.44 million parameters**, YOLOv8n remains a lightweight architecture suitable for edge-device applications. Therefore, the robustness improvements achieved by the proposed method do not result from increasing model complexity but rather from improved training data preparation strategies.

Confusion Matrix Analysis

Confusion matrix analysis was performed to examine classification error patterns among resin classes. In this study, the confusion matrix is recommended to be presented for the two primary scenarios: the **baseline model** and the **proposed method**. The baseline serves as the initial reference model, whereas the proposed method represents the framework introduced in this study. The confusion matrices are necessary to identify which classes are most frequently misclassified. Based on the per-class recall results, the **PP** class remains the most challenging category for the proposed method. Therefore, the confusion matrix can further strengthen the discussion by illustrating the potential visual similarity between the **PP**, **PE-HD**, and **PS** classes, which likely contributes to the observed classification errors.

General Discussion

In general, the results show that the combination of deformation-aware augmentation and class balance can improve the performance of YOLOv8n in classifying plastic resin codes. The main improvements are seen in Accuracy, Macro-F1, Weighted-F1, and accuracy under severe deformation. Compared to the baseline, the proposed YOLOv8n improved Accuracy by 4.37 points, from 72.45% to 76.82%. The Macro-F1 increased from 0.6311 to 0.7026, indicating that the model is more balanced in recognizing various classes. These results demonstrate that improving robustness does not necessarily require changing the model architecture to a larger or more complex one. Appropriate data processing strategies can provide significant performance improvements for lightweight models like YOLOv8n. This is crucial for automated waste sorting applications, as these systems require models that are not only accurate but also fast and efficient.

Although the proposed method provided the best overall performance, the recall results per class indicate that the PP class remains a challenge. The deformation augmentation scenario actually produced the highest PP recall at 73.33%, while the proposed method produced a PP recall of 40.00%. This indicates that combining class balance and augmentation does not always provide uniform improvements across classes. There may be a pattern in the augmentation or synthetic sample distribution that shifts the model's decision boundary and affects PP class recognition. Another limitation of this study is the limited number of unique objects in the Others class, requiring dataset division for that class at the image level. This condition may affect the interpretation of the Others class recall results. Therefore, future research should consider adding external datasets or collecting new data, particularly for minority classes, to improve model evaluation.

CONCLUSION

This research developed an approach to improve the robustness of YOLOv8n through deformation-aware augmentation and class balance for plastic resin code classification. Experiments were conducted using the WaDaBa dataset, which consists of 4,000 plastic images with five classes and four levels of physical deformation. Four experimental scenarios were tested: baseline YOLOv8n, YOLOv8n with class balance, YOLOv8n with deformation augmentation, and the proposed YOLOv8n, which combines both strategies. The test results showed that the proposed YOLOv8n achieved the best overall performance with an accuracy of 76.82%, a macro-F1 of 0.7026, a weighted-F1 of 0.7613, a latency of 4.87 ms per image, and a throughput of 205.2 FPS. Compared to the baseline, the proposed method improved accuracy by 4.37 points and macro-F1 by 0.0715. Furthermore, accuracy for severe deformation increased from 70.59% in the baseline to 73.53% in the proposed method.

The results also show that deformation-aware augmentation contributes significantly to improving model robustness, particularly for severe deformations. Meanwhile, class balance helps improve overall performance by reducing bias toward the majority class. However, the recall of the PP class in the proposed method is still suboptimal, making this class a challenge for further research. Overall, the combination of deformation-aware augmentation and class balance has been shown to improve the robustness of YOLOv8n without sacrificing inference efficiency. Future research can focus on developing more selective augmentation strategies for the PP class, implementing class weight or focal loss-based loss functions, and validating using external datasets or plastic data from real environments..

REFERENCES

Ahmad, G., et al. (2025). Intelligent waste sorting for urban sustainability using deep learning. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-08461-w

Abdu, H., & Noor, M. H. M. (2022). A survey on waste detection and classification using deep learning. IEEE Access, 10, 128151–128165. https://doi.org/10.1109/ACCESS.2022.3226682

Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. Journal of Imaging, 9(2), Article 46. https://doi.org/10.3390/jimaging9020046

Dipo, M. H., et al. (2025). Real-time waste detection and classification using YOLOv12-based deep learning model. Digital, 5(2), Article 19. https://doi.org/10.3390/digital5020019

Islam, M. M., Hasan, S. M., Hossain, M. R., Uddin, M. P., & Mamun, M. A. (2025). Towards sustainable solutions: Effective waste classification framework via enhanced deep convolutional neural networks. PLOS ONE, 20(6). https://doi.org/10.1371/journal.pone.0324294

Jin, S., Yang, Z., Królczyk, G., Liu, X., Gardoni, P., & Li, Z. (2023). Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. Waste Management, 162, 123–130. https://doi.org/10.1016/j.wasman.2023.02.014

Khanam, R., & Hussain, M. (2024). YOLOv11: An overview of the key architectural enhancements (arXiv:2410.17725). arXiv. https://arxiv.org/abs/2410.17725

Kunwar, S., Owabumoye, B. R., & Alade, A. S. (2024). Plastic waste classification using deep learning: Insights from the WaDaBa dataset (arXiv:2412.20232). arXiv. https://arxiv.org/abs/2412.20232

Kunwar, S., Owabumoye, B. R., & Alade, A. S. (2025). Plastic waste detection using deep learning: Insights from the WaDaBa dataset. Industrial and Domestic Waste Management, 5(1), 1–11. https://doi.org/10.53623/idwm.v5i1.580

Majchrowska, S., et al. (2022). Deep learning-based waste detection in natural and urban environments. Waste Management, 138, 274–284. https://doi.org/10.1016/j.wasman.2021.12.001

Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. Array, 16, 100258. https://doi.org/10.1016/j.array.2022.100258

Nahiduzzaman, M., et al. (2025). An automated waste classification system using deep learning techniques: Toward efficient waste recycling and environmental sustainability. Knowledge-Based Systems.

Nasien, D. (2025). Automated waste classification using YOLOv11: A deep learning approach for sustainable waste management. Journal of Advanced Biomedical and Technology, 6(1). https://doi.org/10.35145/jabt.v6i1.205

Qibthiyah, M., Mutiara, A. B., & Refianti, R. (2023). Resin-based classification of plastic bottles waste using CNN architecture model. In Proceedings of the International Conference on Computer Science and Its Application in Agriculture.

Ranjbar, I., Ventikos, Y., & Arashpour, M. (2025). Deep learning-based construction and demolition plastic waste classification by resin type using RGB images. Resources, Conservation and Recycling, 212. https://doi.org/10.1016/j.resconrec.2024.107937

Sayem, F. R., et al. (2025). Enhancing waste sorting and recycling efficiency: Robust deep learning-based approach for classification and detection. Neural Computing and Applications, 37(6), 4567–4583. https://doi.org/10.1007/s00521-024-10855-2

Singh, J., Beeche, C., Shi, Z., & Beale, O. (2023). Batch-balanced focal loss: A hybrid solution to class imbalance in deep learning. Journal of Medical Imaging, 10(5). https://doi.org/10.1117/1.JMI.10.5.051809

Tian, Y., Ye, Q., & Doermann, D. (2025). YOLOv12: Attention-centric real-time object detectors (arXiv:2502.12524). arXiv. https://arxiv.org/abs/2502.12524

Xu, Y., & Lyu, C. (2024). Class-balanced regularization for long-tailed recognition. Neural Processing Letters, 56, Article 158. https://doi.org/10.1007/s11063-024-11624-x

Yaseen, M. (2024). What is YOLOv8: An in-depth exploration of the internal features of the next-generation object detector (arXiv:2408.15857). arXiv. https://arxiv.org/abs/2408.15857