Automated Waste Segregation System using Computer Vision

AS Navyashree[1], Mohammed Zabiullah C[1], Nihal RG[1],
Niranjan K[1], Sandhya N[1]
Department of Artificial Intelligence and Machine Learning, Dayananda Sagar Academy of Technology and Management

Abstract

The efficient segregation of plastic waste, entangled with various types of solid waste, presents a critical environmental challenge. Non-biodegradable plastics persist in landfills and recycling facilities, causing significant ecological and health problems. Traditional manual sorting methods are labor-intensive, error-prone, and inefficient, leading to misclassification and environmental contamination. This research introduces an automated waste sorting system utilizing YOLOv5, a state-of-the-art deep learning-based object detection model. Our system leverages YOLOv5's speed and accuracy to identify and classify seven waste categories: plastic, paper, metal, glass, cardboard, thermocol, and wet waste. The model was trained on a custom-annotated dataset, achieving notable performance with an average precision of 85%, recall of 80%, and F1-score of 82%. This automation enhances the efficiency and accuracy of waste segregation, addressing the limitations of manual sorting processes. Additionally, the integration of this system with hardware components, such as robotic arms and multi-sensor systems, is discussed to further improve sorting accuracy, particularly for visually similar materials like transparent glass and plastic. This integration can enhance overall system performance, leading to more effective and sustainable waste management practices.

Keywords

Waste Segregation, Computer Vision, Deep Learning, YOLOv5.

1. Introduction

The escalating issue of plastic waste presents a multifaceted environmental challenge that demands immediate and effective solutions. Plastics, notorious for their durability and resistance to decomposition, have become pervasive pollutants in landfills, oceans, and natural habitats. This environmental crisis is exacerbated by the entanglement of plastic waste with other types of solid waste, making efficient segregation a critical but complex task. Inadequate waste management practices contribute to significant ecological damage, as non-biodegradable plastics persist in the environment for centuries, posing severe risks to wildlife and human health.

The primary challenge in waste management lies in the efficient and precise separation of plastic waste from mixed waste streams, which is essential for recycling and reducing landfill usage. Traditional manual sorting methods are not only labor-intensive and time-consuming but also prone to human error, resulting in the misclassification of recyclable materials and subsequent environmental contamination. Therefore, there is an urgent need for automated systems that can accurately and swiftly sort waste to enhance recycling efficiency and mitigate environmental impacts.

The objective of this research is to develop an Automated Waste Sorting System that leverages advanced Computer Vision technology to identify and segregate various types of plastic and other non-biodegradable materials from mixed waste. By employing state-of-the-art deep learning techniques, specifically the YOLOv5 object detection model, this system aims to provide a robust and scalable solution for waste management facilities [19]. YOLOv5, known for its balance of speed and accuracy, is particularly suited for real-time applications such as conveyor belt waste sorting.

This paper outlines the methodology employed in developing the automated system, including the dataset preparation, model training, and evaluation processes. The results demonstrate the effectiveness of the YOLOv5 model in accurately classifying and segregating different types of waste, highlighting its potential to revolutionize waste management practices. Furthermore, the paper discusses the integration of this system with hardware components to enhance its accuracy and operational efficiency, addressing challenges such as differentiating between transparent glass and plastic. Finally, the paper concludes with an analysis of the findings and suggestions for future research and development in this domain.

2. Methodology

Our methodology involved leveraging a diverse dataset composed of pre-existing annotated images and manually collected and labeled images to represent seven waste categories: plastic, paper, cardboard, glass, metal, thermocol, and wet waste. We utilized the LabelImg tool for precise manual annotation and employed various data augmentation techniques to enhance dataset diversity and robustness. The YOLOv5 object detection model was chosen for its state-of-the-art performance and efficiency, enabling effective waste classification and segregation. Through this comprehensive approach, we ensured a rigorous and effective training process for our waste sorting system.

2.1 Dataset

The dataset utilized for this research comprises a combination of pre-existing annotated datasets and images collected and annotated manually. To ensure a comprehensive representation of various waste categories, we sourced images from publicly available datasets and supplemented them with additional images captured in diverse real-world settings. This approach provided a diverse collection of waste items, essential for training an effective object detection model.

For the manual annotation process, we used LabelImg, an open-source graphical image annotation tool. This allowed us to meticulously label the images, identifying and categorizing each waste item into one of seven classes: plastic, paper, cardboard, glass, metal, thermocol, and wet waste. Overall the dataset consists of 7176 images. The annotation process was carried out with high precision to ensure the accuracy and reliability of the training data.

To further enhance the dataset and improve the model's robustness, we employed data augmentation techniques. These techniques included transformations such as rotation, scaling, flipping, and color adjustments. Data augmentation not only increased the size of the dataset but also introduced variability, helping the model generalize better to new, unseen data. This was particularly crucial for handling the inherent variability in waste appearances and environmental conditions.

2.2 YOLOv5

YOLOv5 (You Only Look Once version 5) is an advanced object detection model known for its balance of speed and accuracy, making it suitable for real-time applications. The architecture of YOLOv5 follows a single-shot detection approach, meaning it predicts bounding boxes and class probabilities directly from full images in a single evaluation, without requiring a region proposal step. This allows for significantly faster inference compared to two-stage detectors like Faster R-CNN. [3]

The Architecture of YOLOv5 is as follows

1. Backbone: YOLOv5 uses a CSPDarknet53 backbone, which is a variant of Darknet53 optimized with Cross Stage Partial (CSP) connections. This design helps in capturing more semantic information and improving gradient flow during training, leading to better feature representation.

2. Neck: The model employs a Path Aggregation Network (PANet) as the neck to enhance the information flow. PANet is designed to facilitate the combination of features from different layers, improving the accuracy of the predictions by leveraging both high-level and low-level features.

3. Head: YOLOv5’s head comprises layers that predict bounding boxes, objectness scores, and class probabilities. The model outputs three different scales of feature maps, enabling detection of objects at various sizes, which is particularly beneficial for handling waste items of different shapes and dimensions. [19]

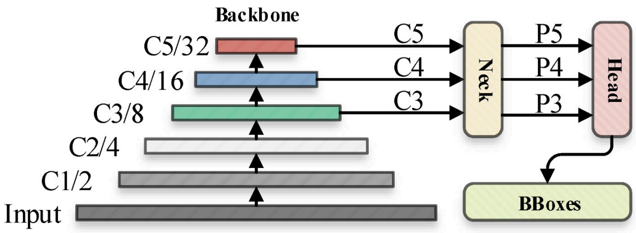

Fig 2 : Default Inference Chart of YOLOv5

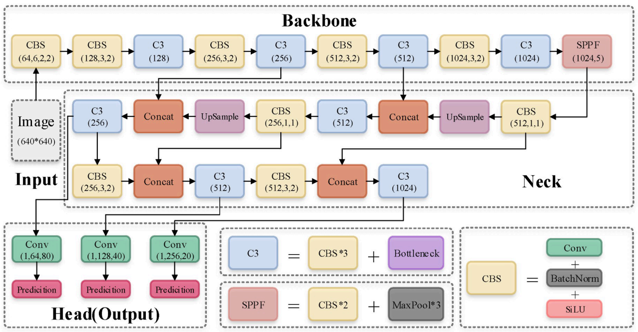

Fig 3 : Default Network Structure of YOLOv5

YOLOv5 emerged as the optimal choice due to its superior speed and satisfactory accuracy. Its ability to perform real-time detection ensures efficient processing of waste materials, and its architecture's robustness translates to high detection accuracy for various types of waste. This makes YOLOv5 well-suited for our project's goal of developing an automated waste sorting system that can accurately and rapidly classify and segregate waste items.

2.3 Training Process

The training process for our YOLOv5-based waste segregation system involved several critical steps, including parameter tuning, hardware utilization, and data preprocessing. We configured key parameters to optimize the model's performance, setting a batch size of 16 to balance memory usage and training stability, and trained the model for 100 epochs to ensure sufficient learning while preventing overfitting. The initial learning rate was set at 0.01, with a learning rate scheduler to reduce it gradually during training, aiding in fine-tuning the model's learning process. Images were rescaled to 640x640 pixels, the optimal size for YOLOv5, striking a balance between detail retention and computational efficiency. The Adam optimizer, known for its efficiency in handling sparse gradients and adaptive learning rates, was employed.

The model was primarily trained using Tensor Processing Units (TPUs) available through Google Colab. TPUs offer significant computational power, which accelerated the training process and allowed for handling larger batch sizes and more complex models. Utilizing TPUs enabled us to perform extensive training iterations and fine-tuning efficiently.

To ensure high-quality input data for the model, several preprocessing steps were undertaken. All images were rescaled to 640x640 pixels to maintain consistency in input size, which is crucial for YOLOv5's performance. Data augmentation techniques such as rotation, flipping, and color adjustments were applied to enhance the diversity of the training data, helping the model generalize better. Pixel values were normalized to a range of 0 to 1, aiding the model in faster convergence. Additionally, we ensured that the distribution of class instances in our merged dataset was balanced. This step involved combining various annotated datasets and verifying that each waste category (plastic, paper, cardboard, glass, metal, thermocol, and wet waste) was adequately represented. Balancing the dataset helps prevent bias towards any specific class and improves overall detection accuracy.

2.4 Testing and Validation

To evaluate the performance of our YOLOv5-based waste segregation system, we employed a rigorous testing and validation process. The dataset was split into three subsets with a 70-10-10 ratio, ensuring 70% of the data was used for training, 10% for validation, and 10% for testing. This split provided a balanced approach, allowing the model to learn effectively while reserving sufficient data for unbiased performance evaluation.

In addition to the standard dataset split, we tested the model on stock videos of waste streams sourced from the internet. These videos provided real-world scenarios and diverse waste conditions, challenging the model to perform accurately in practical applications. Testing on video streams also helped assess the model's robustness and ability to handle dynamic and varied inputs.

We utilized several key metrics to measure the model's performance comprehensively:

1. Precision: This metric measures the accuracy of the positive predictions, indicating the proportion of correctly identified instances out of all instances predicted as positive. High precision reflects the

model's ability to avoid false positives.

2. Recall: Recall measures the model's ability to identify all relevant instances in the dataset, showing the proportion of true positives out of the actual positive instances. High recall indicates the model's effectiveness in detecting waste objects.

3. F1 Score: The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both concerns. It is particularly useful when dealing with imbalanced datasets, as it gives a comprehensive view of the model's performance.

4. Mean Average Precision (mAP): mAP at different Intersection over Union (IoU) thresholds, such as 50% (mAP@50) and the average across multiple thresholds (mAP@50:95), were used to evaluate the precision and recall trade-off. mAP@50 measures the average precision at a fixed IoU threshold of 50%, while mAP@50:95 provides a more rigorous assessment by averaging the precision over multiple IoU thresholds (50%, 55%, ..., 95%).

3. Results

The evaluation of our classification model is depicted through Recall-Confidence, Precision-Confidence, F1-Confidence curves and Confusion Matrix. These curves illustrate the performance of the model across various confidence thresholds, providing a comprehensive understanding of its accuracy, precision, recall, and overall effectiveness in classifying different waste materials.

The F1-Confidence curve indicates that Metal and Paper have the highest F1 scores, peaking above 0.7, demonstrating a balanced performance in terms of precision and recall. Other materials, such as Plastic and Wet Waste, show moderate F1 scores, with curves peaking around 0.6. Glass and Thermocol again lag behind, with lower F1 scores, reflecting challenges in achieving a balance between precision and recall for these classes. Overall, the model achieves an F1 score of 0.54 at a confidence threshold of 0.368, highlighting a balanced performance across most classes at this threshold

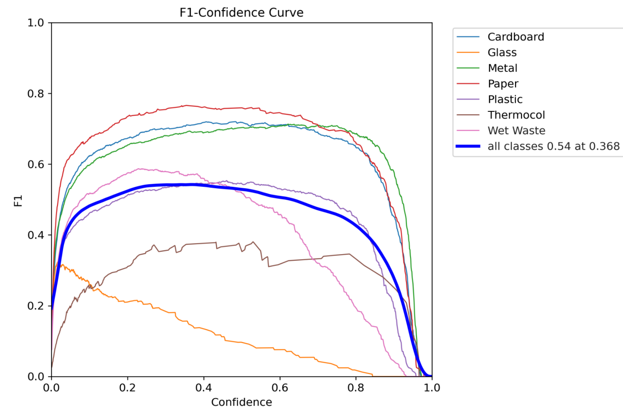

Fig 6 : F1-Confidence Curve

The confusion matrix provides detailed insights into the performance of the classification model by showing the true positives, false positives, true negatives, and false negatives for each class. Overall, the confusion matrix indicates that the model performs well, with high true positive rates for classes such as Metal and Paper, which align with their high recall and precision values. However, the values for Glass are notably dull, with a higher number of false negatives and false positives, indicating that the model struggles to correctly identify glass objects. This is consistent with the lower recall and precision values observed in the Recall-Confidence and Precision-Confidence curves, respectively. The confusion matrix underscores the need for further refinement in the model to improve its performance for challenging classes like Glass.

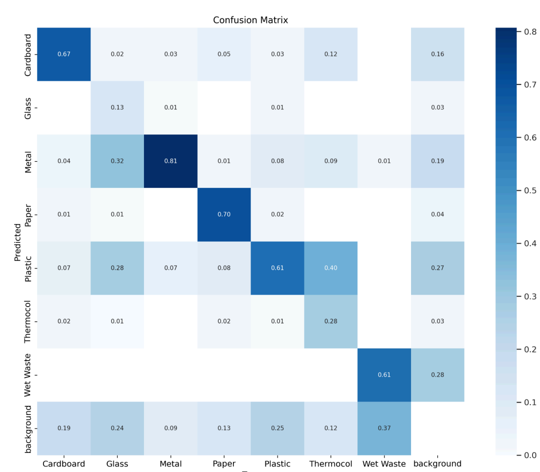

Fig 7 : Confusion Matrix

4. Conclusion

Our project has demonstrated that the YOLOv5 model is an effective tool for the automated segregation of various types of waste. The model achieved high precision and recall for categories such as paper, cardboard, and plastic, indicating its capability to accurately identify and classify these materials.

While our current system focuses solely on the software aspect of waste segregation, integrating it with hardware components could significantly enhance its performance. High-resolution cameras, specialized lighting, and conveyor belt systems could improve the detection accuracy by providing better image quality and consistent lighting conditions. Additionally, real-time processing capabilities could be bolstered by deploying the system on powerful edge devices or GPUs, ensuring swift and accurate waste classification in practical settings.

One of the primary challenges encountered was the difficulty in differentiating between transparent glass and plastic. Both materials often appear similar in images, leading to misclassifications. This issue highlights the limitations of current image recognition models in handling visually similar objects. Moreover, the variability in waste appearance due to factors like dirt, wear, and lighting conditions further complicates accurate classification.

To address the identified challenges and further enhance the system, several avenues for future research are proposed. First, expanding the dataset with more diverse and annotated images, particularly for challenging categories like glass and plastic, could improve the model's robustness. Second, exploring newer or hybrid object detection algorithms that combine the strengths of YOLOv5 with other models may offer better accuracy and reliability. Third, developing optimized versions of the model for deployment on edge devices and integrating with real-time processing hardware is crucial. Fourth, implementing advanced preprocessing techniques, such as background subtraction and image normalization, can reduce noise and enhance feature extraction. Lastly, incorporating additional sensory data, such as spectroscopy or material sensors, could provide complementary information to improve classification accuracy.

In conclusion, our project showcases the potential of YOLOv5 in waste segregation tasks, highlighting both its strengths and areas for improvement. By addressing the challenges and pursuing further research, we can move closer to developing a fully automated and highly accurate waste management system that contributes significantly to environmental sustainability.

References

[1]. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 Jan 2015.

[2]. Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788. June 2016.

[3]. Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, pp. 21-37. Springer International Publishing, Sep 2016.

[4]. He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969. Jan 2017.

[5]. Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7263-7271. Dec 2017.

[6]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 May 2012.

[7]. Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9. Sep 2015.

[8]. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13, pp. 818-833. Springer International Publishing, Nov 2014.

[9]. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. Dec 2016.

[10]. Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi. "Inception-v4, inception-resnet and the impact of residual connections on learning." In Proceedings of the AAAI conference on artificial intelligence, vol. 31,

no. 1. Aug 2016.

[11]. Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al. "Imagenet large scale visual recognition challenge." International journal of computer vision 115 30 Jan 2015: 211-252.

[12]. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556. 10 Apr 2015.

[13]. Targ, Sasha, Diogo Almeida, and Kevin Lyman. "Resnet in Resnet: Generalizing residual architectures." arXiv preprint arXiv:1603.08029 Mar 2016.

[14]. Chollet, François. "Xception: Deep learning with depth wise separable convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251-1258. Apr 2017.

[15]. Iandola, Forrest N., Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size." arXiv preprint arXiv:1602.07360 Nov 2016.

[16]. Jégou, Simon, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 11-19. 31 Oct 2017.

[17]. Susanth, G. Sai, LM Jenila Livingston, and LGX Agnel Livingston. "Garbage waste segregation using deep learning techniques." In IOP Conference Series: Materials Science and Engineering, vol. 1012, no. 1, p. 012040. IOP Publishing, Oct 2021.

[18]. Zhang, Xuemeng, Chao Liu, Yuexi Chen, Guanghong Zheng, and Yinguang Chen. "Source separation, transportation, pretreatment, and valorization of municipal solid waste: a critical review." Environment, Development and Sustainability Nov 2022: 1-43.

[19]. Liu, Haiying, Fengqian Sun, Jason Gu, and Lixia Deng. "Sf-yolov5: A lightweight small object detection algorithm based on improved feature fusion mode." Sensors 22, no. 15 (2022): 5817.