VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI-590014

A Project Report
On

Automated Waste Segregation Using Computer Vision

A Project report submitted in partial fulfillment of the requirements for the VIII Semester degree of
Bachelor of Engineering in Artificial Intelligence and Machine Learning
Of Visvesvaraya Technological University, Belagavi.

Submitted by

A S NAVYASHREE	**1DT20AI001**
MOHAMMED ZABIULLAH C	**1DT20AI026**
NIHAL RANJAN GOWDA	**1DT20AI031**
NIRANJAN KRISHNA	**1DT20AI032**

Under the Guidance of

Dr. Sandhya N
Professor & HOD

Department of Artificial Intelligence and Machine Learning

DAYANANDA SAGAR ACADEMY OF TECHNOLOGY AND MANAGEMENT
Udayapura, Kanakapura Road,Bangalore-560082

2023-2024

DAYANANDA SAGAR ACADEMY OF TECHNOLOGY AND MANAGEMENT
Udayapura, Kanakapura Road, Bangalore-560082
DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

CERTIFICATE

Certified the project work entitled **Automated Waste Segregation using Computer Vision** carried out by A S Navyashree bearing USN **1DT20AI001**, **Mohammed Zabiullah C** bearing USN **1DT20AI026**, **Nihal Ranjan Gowda** bearing USN **1DT20AI031**, and **Niranjan Krishna** bearing USN **1DT20AI032**, bonafide students of **Dayananda Sagar Academy of Technology and Management** in partial fulfilment for the award of Bachelor of Engineering in Artificial Intelligence And Machine Learning of the Visvesvaraya Technological University, Belagavi during the year 2023-24. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

Dr. Shivaprasad Ashok Chikop	Dr. Sandhya N	Dr.M. Ravishankar
Assistant Professor, Coordinator	Professor, Guide & HOD	Principal
Dept. of AIML, DSATM	Dept. of AIML, DSATM	DSATM

External Examiners

Sl. No.	Name	Signature with Date
1.		
2.		

DAYANANDA SAGAR ACADEMY OF TECHNOLOGY AND MANAGEMENT
Udayapura, Kanakapura Road, Bangalore-560082

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

DECLARATION

We, A S Navyashree **[1DT20AI001]**, Mohammed Zabiullah C **[1DT20AI026]**, Nihal Ranjan Gowda **[1DT20AI031]** and Niranjan Krishna **[1DT20AI032],** students of VIII Semester B.E., in Artificial Intelligence And Machine Learning, Dayananda Sagar Academy of Technology & Management hereby declare that the Project entitled "Automated Waste Segregation Using Computer Vision" has been carried out by us and submitted in partial fulfillment of the requirements for the *VIII Semester degree of* ***Bachelor of Engineering in*** *Artificial Intelligence And Machine Learning* of Visvesvaraya Technological University, Belagavi during academic year 2023-2024.

Sl. No.	Name	USN	Signature With date
1.	A S NAVYASHREE	1DT20AI001	
2.	MOHAMMED ZABIULLAH C	1DT20AI026	
3.	NIHAL RANJAN GOWDA	1DT20AI031	
4.	NIRANJAN KRISHNA	1DT20AI032	

ACKNOWLEDGEMENT

We are thankful to **Dr. M. Ravishankar**, **Principal**, DSATM, Bangalore, for providing the necessary facilities and for the encouragement to carryout the Project in DSATM.

We express our humble gratitude and thanks to **Dr. Sandhya N**, Professor and HOD, Artificial Intelligence and Machine Learning, DSATM for guiding us and for her constant encouragement, guidance and support throughout the course of the Project without whose efforts this would not have been possible and having facilitated us to complete the Project successfully

We thank the Project Coordinators **Dr. Shivaprasad Ashok Chikop** for their invaluable critical reviews, guidance and support in completing the Project on time.

We finally thank to all the Technical Stafff or all the support and help while carrying out the project. Last but not the least we wish to thank our friends and parents who have been encouraging us constantly and inspiring throughout the execution of the project.

Name	**USN**
A S NAVYASHREE	**1DT20AI001**
MOHAMMED ZABIULLAH C	**1DT20AI026**
NIHAL RANJAN GOWDA	**1DT20AI031**
NIRANJAN KRISHNA	**1DT20AI032**

ABSTRACT

The Waste Segregation Automation System (WSAS) is an innovative solution designed to revolutionize waste management processes in diverse environments, including industrial facilities, municipalities, and recycling centers. Traditional waste segregation methods often suffer from inefficiencies, inaccuracies, and manual errors, leading to suboptimal resource utilization and environmental concerns. WSAS addresses these challenges by harnessing the power of computer vision and machine learning algorithms to automate waste sorting tasks. WSAS utilizes state-of-the-art deep learning models, including YOLOv5, to detect and classify various types of waste items in real-time. The system is trained on a custom dataset comprising seven classes of waste, including plastic, paper, metal, glass, cardboard, thermocol, and wet waste. By leveraging visual inputs from cameras mounted along conveyor belts, WSAS can accurately identifies and can help in segregating waste items with high precision and efficiency. Moreover, WSAS is aimed to used along with integration of sensor data, such as metallic sensors, to complement visual information and improve the accuracy of waste classification. This fusion of data sources enhances the robustness and reliability of the segregation process, ensuring consistent and accurate results. Additionally, WSAS offers real-time monitoring capabilities, enabling administrators to access detailed reports and insights into waste composition and segregation performance. The implementation of WSAS promises significant benefits, including increased accuracy, reduced labor costs, and improved resource utilization. By automating waste segregation tasks, WSAS enables organizations to streamline operations, optimize recycling processes, and contribute to environmental sustainability efforts. Furthermore, WSAS sets the stage for a digital transformation in waste management, paving the way for smarter, more efficient waste handling practices. In summary, the Waste Segregation Automation System represents a groundbreaking advancement in waste management technology. With its sophisticated algorithms, real-time monitoring capabilities, and intuitive user interface, WSAS is poised to redefine how waste segregation is conducted, ultimately driving efficiency, sustainability, and environmental stewardship.

TABLE OF CONTENTS

Chapter No.		Page No.
	Abstract	(i)
	List of Figures	(iii)
1	Introduction 1.1 Problem Statement 1.2 Motivation 1.3 Objectives	1
2	Literature Survey 2.2 Object Detection 2.3 Feature Extraction 2.4 Classification 2.5 Papers related to other processes	5
3	Proposed System 3.1 System Design	13
4	Implementation	17
5	Results And Discussion 5.1 Testing 5.2 Snapshots	17
6	Conclusion	21
7	References	22
8	Source Code	24
9	Copy of The Paper Project Phase 1	32
10	Copy of The Paper Project Phase 2	42

LIST OF FIGURES

Figure No.	TITLE	Page No.
1.1	Garbage Site 1	1
1.2	Garbage Site 2	1
1.3	Phases of Waste Segregation	2
1.4	Manual Waste Segregation 1	4
1.5	Manual Waste Segregation 2	4
3.1	Automated Waste Segregation Using Computer Vision	13
3.2	System Design	16
4.1	Default Inference Flowchart off YOLOv5	18
4.2	Default network structure of YOLOv5	18
5.1	Misclassified instance	21
5.2	Predictions on Test Data 1	22
5.3	Inference using File Upload 1	22
5.4	Predictions on Test Data 2	23
5.5	Inference Using File Upload 2	23
5.6	Inference Using File Upload 3	24
5.7	Real time Inference Using Webcam	24

CHAPTER 1

INTRODUCTION

In recent decades, the world has witnessed an alarming surge in the production and disposal of plastic waste, a critical environmental issue with far-reaching consequences. Compounding this problem, plastic waste is frequently intermingled with various types of solid waste in municipal landfills and recycling facilities. The urgent need to tackle this crisis is underscored by the fact that non-biodegradable plastics can persist in the environment for centuries, leading to ecological damage and health hazards. The challenge at hand is the efficient and accurate separation of plastic waste from this heterogeneous waste stream. Conventional manual sorting processes are not only time-consuming and labor-intensive but are also prone to human error, resulting in the misclassification of recyclable materials and the contamination of ecosystems.

Fig 1.1 Garbage Site 1

Fig 1.2 Garbage Site 2

In the comprehensive realm of Municipal Solid Waste (MSW) management, a multifaceted approach is imperative, encompassing source separation, transportation, pretreatment, and valorization strategies. Source separation lays the foundation for effective waste management by encouraging individuals and communities to segregate their waste at the point of origin, facilitating targeted recycling and disposal processes. The subsequent transportation phase involves the systematic and movement of waste from collection points to processing facilities.

Pretreatment emerges as a pivotal step in the MSW management continuum, involving processes such as sorting, shredding, and composting to prepare waste for further treatment or disposal. This phase not only aids in reducing the volume of waste but also enhances the feasibility of subsequent valorization techniques. The term valorization pertains to the recoveryof valuable resources from waste, transforming it into energy, compost, or recyclable materials. Valorization represents a sustainable approach, mitigating the environmental burden of waste by extracting useful components and minimizing landfill dependence. By seamlessly integrating these components into the waste management framework, municipalities can aspire to establish a holistic, environmentally conscious, and resource-efficient MSW management system.

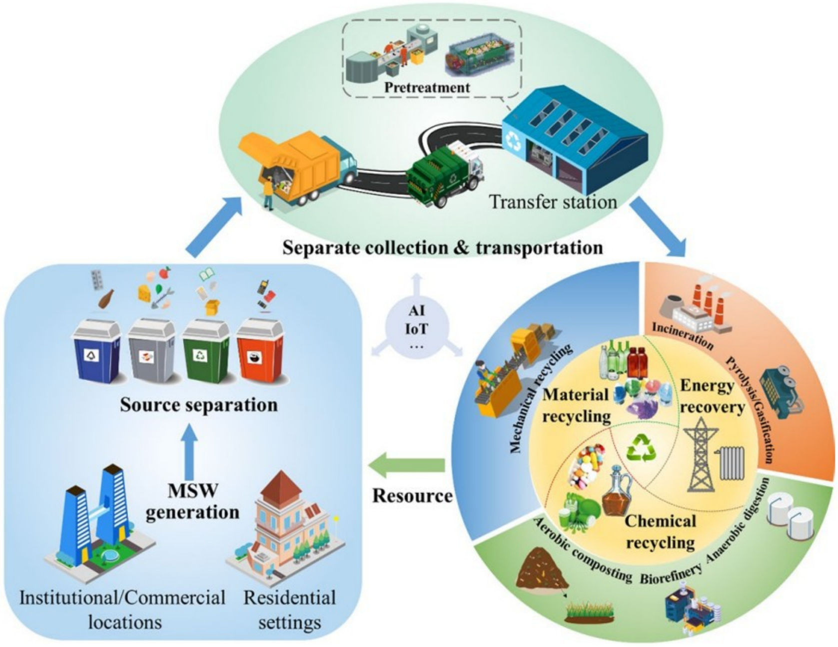

Fig 1.3 Stages of Waste Segregation

In response to this pressing issue, our project introduces a state-of-the-art solution: an Automated Waste Sorting System employing advanced Computer Vision technology. This innovative system is designed to swiftly and precisely identify and segregate plastics and other non-biodegradable materials from mixed waste as it travels along conveyor belts. By harnessing the power of machine learning and real-time image analysis, our system offers the potential to revolutionize waste management, elevate recycling rates, and significantly reduce the detrimental impact of non-biodegradable waste on our planet. In doing so, we aspire to contribute to a cleaner, healthier environment and a more sustainable future.

1.1 Problem Statement

The escalating issue of plastic waste, in conjunction with its entanglement with solid waste, presents a multifaceted problem that demands attention. The primary challenge is the efficient and precise separation of plastic waste from mixed waste, which often occurs in landfills and recycling facilities. Conventional manual sorting methods, relying on human labor, are not only slow and labor-intensive but also prone to inaccuracies, resulting in the misallocation of recyclable materials and environmental pollution.

To address this critical concern, our project introduces an innovative solution: the development of an Automated Waste Sorting System that leverages cutting-edge Computer Vision technology. This system aims to provide rapid and accurate identification and separation of plastics and other non-biodegradable materials from the complex, mixed waste stream transported on conveyor belts.

By integrating machine learning and real-time image analysis, our system has the potential to revolutionize waste management practices, significantly boost recycling rates, and substantially diminish the adverse impact of non-biodegradable waste on the environment. In pursuing this endeavor, we endeavor to contribute to a more sustainable and eco-conscious future, marked by improved waste management and reduced environmental harm.

1.2 Motivation

The motivation for our project stems from the pressing need to address environmental and social challenges posed by plastic waste and inefficient waste management practices. Our goals are multifaceted, driven by several key factors.

Firstly, we aim to tackle the global plastic waste crisis, which threatens ecosystems and wildlife due to plastics' slow decomposition. Additionally, we recognize plastics as valuable resources and seek to conserve them through effective separation and recycling, reducing the demand for new production and preserving natural resources. Moreover, our project addresses safety concerns associated with manual waste sorting by implementing automation with computer vision technology, minimizing human exposure to hazardous materials.

Overall, our project aligns with sustainability objectives, fostering a cleaner environment and promoting responsible waste management practices. It also reflects our commitment to innovation, utilizing cutting-edge technology like computer vision and machine learning to address real-world challenges efficiently.

Fig 1.4 Manual Waste Segregation 1

Fig 1.5 Manual Waste Segregation 2

1.3 Objectives

Our project encompasses several key objectives, each geared towards addressing critical environmental challenges and enhancing waste management efficiency. The primary objectives are as follows:

1. State-of-the-Art Deep Learning Integration: Implement cutting-edge deep learning techniques for object detection, focusing on enhancing accuracy, adaptability, and efficiency in recognizing and classifying various waste materials.
2. Real-time Processing**:** Implement real-time processing capabilities to ensure swift and efficient segregation of waste items, contributing to timely and effective waste management practices.
3. Multi-class Segregation**:** Develops a system capable of classifying waste into multiple categories, aiming to improve the granularity of segregation and support comprehensive recycling efforts.
4.Scalability: Develop a scalable solution that can be deployed across different waste management facilities, including recycling plants, municipal landfills, and industrial settings, to address the growing challenge of waste management globally.
5. Integration with Hardware Components: Explore opportunities to integrate hardware components such as sensors and robotic arms, to enhance the accuracy, efficiency, and automation capabilities of the waste segregation system.

CHAPTER 2
LITERATURE SURVEY

2.1 Object Detection

Title: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

This seminal work, authored by Shaoqing Ren, et al. introduces Faster R-CNN, a pivotal model in the realm of real-time object detection. The innovation lies in the integration of Region Proposal Networks (RPNs), seamlessly combining region proposal and object detection. By striking an impressive balance between accuracy and speed, Faster R-CNN redefines the landscape of real-time applications. The authors' contribution marks a significant advancement in the efficiency and effectiveness of object detection systems, setting a new standard for future developments in the field.[1]

Title: You Only Look Once: Unified, Real-Time Object Detection

Authored by Joseph Redmon and Santosh Divvala, "You Only Look Once" (YOLO) presents a groundbreaking approach to unified, real-time object detection. The key innovation of YOLO lies in its ability to simultaneously predict bounding boxes and class probabilities in a single pass. This unified methodology results in unparalleled efficiency and accuracy, setting a new standard for real-time object detection systems. YOLO's contributions, outlined by Redmon and Divvala, have significantly influenced the landscape of computer vision, offering a unified and streamlined approach to object detection with wide-ranging applications.[2]

Title: SSD: Single Shot MultiBox Detector

SSD, or Single Shot MultiBox Detector, stands as a milestone in the evolution of object detection methodologies. This pioneering work, authored by Wei Liu, et al. Berg, introduces a groundbreaking single-shot detection method. SSD is designed to predict multiple bounding boxes and class scores in a single pass, eliminating the need for complex multi-stage pipelines. The authors' contributions have redefined the landscape of object detection, combining efficiency and accuracy in a single unified framework. SSD's impact resonates across various applications, making it a cornerstone in real-time object detection systems.[3]

Title: Mask R-CNN

"Mask R-CNN" represents a pivotal advancement in computer vision and object detection, addressing the nuanced task of instance segmentation. Authored by Kaiming He, et al, this work extends the Faster R-CNN model to incorporate precise object masks alongside bounding boxes and class probabilities. The innovation lies in its ability to provide a granular understanding of object boundaries, enabling more accurate and detailed segmentation. Mask R-CNN's contributions, outlined by the authors, have become instrumental in various applications, from image segmentation to interactive image editing, pushing the boundaries of object detection and segmentation methodologies.[4]

Title: YOLO9000: Better, Faster, Stronger

Authored by Joseph Redmon and Santosh Divvala, "YOLO9000: Better, Faster, Stronger" signifies a significant evolution of the You Only Look Once (YOLO) framework. This research introduces advancements to YOLO, enabling it to identify over 9000 object categories in real-time. The authors' contributions focus on enhancing the efficiency and accuracy of object detection across a broader spectrum of objects. The unified approach of YOLO9000, predicting bounding boxes and class probabilities simultaneously, sets new benchmarks in real-time object detection capabilities. Redmon and Divvala's work showcases the continuous development and improvement of YOLO, making it a prominent player in the landscape of computer vision and object detection.[5]

2.2 Feature Extraction

Title: ImageNet Classification with Deep Convolutional Neural Networks (AlexNet)

Authored by Alex Krizhevsky et al, "ImageNet Classification with Deep Convolutional Neural Networks (AlexNet)" is a seminal work that revolutionized image classification. The authors proposed a deep convolutional neural network (CNN) architecture, named AlexNet, which secured a decisive victory in the 2012 ImageNet Large Scale Visual Recognition Challenge. The groundbreaking design of AlexNet, with its deep layers and advanced features, marked a paradigm shift in the field of computer vision. This work laid the foundation for the widespread adoption of deep learning in image recognition tasks and set new standards for the accuracy and efficiency of convolutional neural networks.[6]

Title: Going Deeper with Convolutions (VGGNet)

Authored by Karen Simonyan and Andrew Zisserman, "Going Deeper with Convolutions (VGGNet)" presents a crucial contribution to the architecture of deep convolutional neural networks (CNNs). This work, recognized for its simplicity and effectiveness, explores the impact of network depth on image recognition tasks. VGGNet's hallmark is its use of small-sized convolutional filters, allowing for deeper network architectures. The authors' meticulous experiments demonstrated that increasing network depth leads to improved performance, and the VGGNet architecture served as a benchmark for subsequent CNN designs. This work played a pivotal role in understanding the importance of depth in convolutional networks, influencing the development of more sophisticated neural network architectures.[7]

Title: Visualizing and Understanding Convolutional Networks (ZFNet)

"Visualizing and Understanding Convolutional Networks (ZFNet)" is a significant contribution to the interpretability and comprehension of deep convolutional neural networks (CNNs). Authored by Matthew D. Zeiler and Rob Fergus, this work delves into the visualization techniques that aid in unraveling the inner workings of CNNs. ZFNet particularly focuses on deconvolutional layers and visualization of feature activations, providing valuable insights into the hierarchical representations learned by the network. The authors' efforts contribute to the understanding of how CNNs extract and represent features from input data, enhancing interpretability—a crucial aspect for deploying deep learning models in practical applications.[8]

Title: Deep Residual Learning for Image Recognition (ResNet)

Authored by Kaiming He et al, "Deep Residual Learning for Image Recognition (ResNet)" presents a groundbreaking approach to training very deep neural networks. Recognizing the challenges associated with vanishing gradients in deep networks, the authors introduce residual connections, allowing the network to learn residual functions. This innovative design facilitates the training of exceptionally deep models, surpassing the limitations of traditional architectures. ResNet's impact on image recognition is profound, achieving state-of-the-art results on various benchmarks. The work not only addresses the challenge of depth in neural networks but also establishes residual learning as a fundamental concept in the field of deep learning.[9]

Title: Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning (Inception-ResNet)

Authored by Christian Szegedy, Sergey Ioffe, and Vincent Vanhoucke, "Inception-v4, Inception-ResNet, and the Impact of Residual Connections on Learning" delves into the evolution of Inception architectures, exploring the fusion of Inception modules with residual connections. The study, known as Inception-ResNet, aims to harness the strengths of both architectures, emphasizing the role of residual connections in facilitating learning. By combining innovative modules and residual connections, the authors create models that achieve superior performance in terms of accuracy and efficiency. This work contributes to the ongoing exploration of advanced neural network architectures, providing valuable insights into the impact of residual connections on the learning capabilities of deep networks.[10]

2.3 Classification.

Title: ImageNet Large Scale Visual Recognition Challenge (AlexNet)

Authored by Alex Krizhevsky et al, "ImageNet Large Scale Visual Recognition Challenge (AlexNet)" signifies a pivotal moment in the history of computer vision. Presented in the context of the ImageNet Large Scale Visual Recognition Challenge in 2012, this work introduces AlexNet, a deep convolutional neural network (CNN) architecture. The groundbreaking model not only secured a resounding victory in the competition but also demonstrated the superiority of deep learning in image classification tasks. AlexNet's innovative design, including features like ReLU activation and dropout, set new benchmarks for image recognition accuracy, paving the way for the widespread adoption of deep neural networks in subsequent years.[11]

Title: Very Deep Convolutional Networks for Large-Scale Image Recognition (VGGNet)

Authored by Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition (VGGNet)" represents a significant milestone in the pursuit of large-scale image recognition accuracy. The study explores the impact of network depth on the performance of convolutional neural networks (CNNs). VGGNet is characterized by its straightforward architecture, employing small-sized convolutional filters to achieve remarkable depth. The authors' meticulous experiments revealed that increasing the depth of the network contributes to enhanced recognition capabilities. VGGNet's success served as a benchmark for subsequent CNN designs, influencing the understanding of the relationship between network depth and image recognition performance.[12]

Title: ResNet in ResNet: Generalizing Residual Architectures

The paper "ResNet in ResNet: Generalizing Residual Architectures" introduces a novel deep learning architecture termed ResNet in ResNet (RiR) that extends and generalizes the success of residual networks (ResNets) and standard convolutional neural networks (CNNs). The architecture employs a dual-stream design that seamlessly combines the strengths of ResNets and traditional CNNs without introducing computational overhead. The proposed RiR consistently demonstrates superior performance compared to standard ResNets and even outperforms architectures employing similar levels of augmentation on the CIFAR-10 dataset. Furthermore, RiR achieves a groundbreaking state-of-the-art on the CIFAR-100 dataset. This work not only introduces an innovative approach to residual architectures but also showcases practical advancements by surpassing existing benchmarks on challenging computer vision tasks, thus contributing significantly to the ongoing progress in deep learning methodologies..[13]

Title: Xception: Deep Learning with Depth Wise Separable Convolutions

Authored by François Chollet, "Xception: Deep Learning with Depth Wise Separable Convolutions” introduces a novel architecture that employs depthwise separable convolutions. The study explores an efficient alternative to traditional convolutions, reducing computational complexity while maintaining model performance. The Xception model achieves a balance between accuracy and efficiency, showcasing the potential for innovation in convolutional neural network structures. François Chollet's work contributes to the ongoing evolution of deep learning architectures, demonstrating the effectiveness of depthwise separable convolutions in various computer vision tasks.[14]

2.3 Papers related to Other processes

Title: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size

Authored by Forrest N. et al "SqueezeNet" represents a groundbreaking achievement in model efficiency. This work introduces a novel neural network architecture that achieves AlexNet-level accuracy while drastically reducing the number of parameters and model size. SqueezeNet's design focuses on squeezing the model architecture without compromising performance, making it highly efficient for resource-constrained environments. The authors' contributions in optimizing model size and maintaining accuracy showcase the potential for compact yet powerful deep learning architectures in various applications.[15]

Title: The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation

Authored by Simon Jégou et al. "The One Hundred Layers Tiramisu" presents a novel approach to semantic segmentation using fully convolutional DenseNets. This architecture, inspired by the dense connectivity concept, is designed for the segmentation of images with intricate details. The authors introduce a dense block structure with skip connections that enables efficient feature reuse across multiple layers. The Tiramisu model, with its depth and dense connections, achieves state-of-the-art performance in semantic segmentation tasks. This work contributes to advancing the field of computer vision, emphasizing the importance of dense connectivity for accurate and detailed image segmentation.[16]

Title: Garbage Waste Segregation Using Deep Learning Techniques

Authored by Sai Susanth G et al , this research focuses on leveraging deep learning techniques for efficient garbage waste segregation. The authors delve into the application of advanced algorithms to enhance waste management practices. The study explores the potential of deep learning in automating the segregation of different types of waste, contributing to more effective and sustainable waste management solutions. The work by Sai Susanth G, Jenila Livingston L M, and Agnel Livingston L G X aligns with the growing interest in using technology to address environmental challenges and optimize waste segregation processes.[17]

Title: Source separation, transportation, pretreatment, and valorization of municipal solid waste: a critical review

Authored by Xuemeng Zhang et al, this review critically examines the various stages of municipal solid waste (MSW) management. The authors comprehensively analyze source separation methods, transportation logistics, pretreatment techniques, and strategies for valorizing waste. The study provides valuable insights into the challenges and opportunities associated with each stage of MSW management, emphasizing the importance of efficient and sustainable practices. Zhang et al.'s work serves as a foundational resource for understanding the complexities of MSW management, offering a critical perspective on the key aspects of waste handling and disposal.[18]

Title: Development of Computer Vision Algorithms for Multi-class Waste Segregation and Their Analysis

Authored by Neeraja Narayanswamy et al, this research focuses on the development of computer vision algorithms tailored for multi-class waste segregation. The authors delve into the intricacies of designing algorithms capable of categorizing diverse types of waste efficiently. The study involves the analysis of the developed algorithms, providing insights into their performance and suitability for real-world waste segregation scenarios. Neeraja Narayanswamy et al work contributes to the advancement of computer vision applications in waste management, addressing the need for accurate and automated waste segregation solutions.[19]

Title: Plastic Material Detection and Classification Using CNN

Authored by Vavilala Sushma, this study focuses on the application of Convolutional Neural Networks (CNN) for the specific task of detecting and classifying plastic materials. The research aims to develop a robust system that can accurately identify different types of plastic, contributing to improved waste sorting processes. By leveraging the capabilities of CNNs, the study addresses the challenges associated with plastic material detection and classification. Vavilala Sushma's work holds significance in the context of sustainable waste management, offering a specialized solution for the identification and categorization of plastic waste using advanced deep learning techniques. [20]

Title: Development of Intelligent Waste Segregation System Based on Convolutional Neural Network

Authored by Tarig Faisal et al. this research project focuses on the creation of an intelligent waste segregation system utilizing Convolutional Neural Network (CNN) technology. The authors aim to design a system that can accurately categorize waste items through automated visual recognition. The study contributes to the advancement of smart waste management solutions, emphasizing the role of CNNs in creating efficient and precise waste segregation systems. Tarig Faisal et al.'s work aligns with the broader goal of leveraging artificial intelligence for sustainable and technology-driven waste management practices.[21]

Title: A Novel YOLOv3 Algorithm-Based Deep Learning Approach for Waste Segregation: Towards Smart Waste Management

Authored by Saurav Kumar et al. this research introduces a novel approach to waste segregation using the YOLOv3 algorithm. The study focuses on the development of a deep learning-based system for efficient waste segregation, emphasizing the capabilities of YOLOv3 in real-time object detection. The authors aim to contribute to the paradigm of smart waste management, highlighting the potential of advanced algorithms to enhance waste sorting processes. Saurav Kumar et al.'s work represents an innovative stride towards the integration of state-of-the-art deep learning techniques for sustainable and technology-driven waste management practices.[22]

Title: Municipal solid waste management in Indian cities – A review

Authored by Mufeed Sharholy et al. this comprehensive review examines the current state of municipal solid waste management in Indian cities. The study provides a critical analysis of the existing waste management practices, challenges faced, and potential solutions. The authors delve into the complexities of waste disposal, recycling, and the overall impact on environmental sustainability. Mufeed Sharholy et al.'s work serves as a valuable resource for understanding the nuances of municipal solid waste management in the specific context of Indian urban areas, offering insights that can inform future strategies and policies for sustainable waste management. [23]

In conclusion, the literature review on waste segregation employing deep learning and computer vision methodologies presents a comprehensive overview of diverse approaches. Object detection and classification models such as Faster R-CNN, YOLO, and SSD offer real-time capabilities, while image classification architectures like AlexNet, VGGNet, ResNet, and others contribute to the recognition of waste items. Semantic segmentation models, exemplified by DenseNets, provide insights into fine-grained segmentation. Specialized approaches and smart waste management systems demonstrate the ongoing efforts to tailor AI techniques to the intricacies of waste segregation. The collective findings suggest a promising direction for the development of intelligent waste segregation systems that leverage the strengths of various deep learning techniques, offering a potential paradigm shift towards more efficient and sustainable waste management practices.

CHAPTER 3

PROPOSED SYSTEM

The escalating issue of plastic waste, in conjunction with its entanglement with solid waste, presents a multifaceted problem that demands attention. The primary challenge is the efficient and precise separation of plastic waste from mixed waste, which often occurs in landfills and recycling facilities. Conventional manual sorting methods, relying on human labor, are not only slow and labor-intensive but also prone to inaccuracies, resulting in the misallocation of recyclable materials and environmental pollution.

To address this critical concern, our project introduces an innovative solution: the development of an Automated Waste Sorting System that leverages cutting-edge Computer Vision technology. This system aims to provide rapid and accurate identification and separation of plastics and other non-biodegradable materials from the complex, mixed waste stream transported on conveyor belts.

By integrating machine learning and real-time image analysis, our system has the potential to revolutionize waste management practices, significantly boost recycling rates, and substantially diminish the adverse impact of non-biodegradable waste on the environment. In pursuing this endeavor, we endeavor to contribute to a more sustainable and eco-conscious future, marked by improved waste management and reduced environmental harm.

Fig 3.1 Automated Waste Segregation Using Computer Vision

Our proposed system is an Automated Waste Classification System that utilizes state-of-the-art deep learning techniques for waste recognition and analysis. The system operates solely on digital data, namely images or videos of waste materials. Key components of the system include:

1. Image Acquisition: Waste materials are photographed or recorded using cameras placed at strategic locations, such as recycling plants, waste collection centers, or conveyor belts.

2. Image Processing: The captured images or videos are processed using computer vision algorithms to identify and classify different types of waste materials. Deep learning models, such as YOLOv5 or Faster R-CNN, are employed for object detection and classification.

3. Waste Classification: Based on the classification results, the system categorizes waste materials into different classes, such as plastic, paper, metal, glass, cardboard, thermocol, or wet waste. Each class is assigned a corresponding label based on its visual characteristics.

4. Data Analysis and Reporting: The system analyzes the classified waste data to provide insights into waste composition, distribution, and trends. It generates reports and visualizations that can help waste management authorities make informed decisions regarding resource allocation, recycling initiatives, and waste reduction strategies.

5. Integration with Existing Waste Management Infrastructure: Our system is designed to seamlessly integrate with existing waste management systems, such as waste databases, recycling facilities, or waste collection apps. It can serve as a valuable tool for waste monitoring, analysis, and optimization without requiring physical segregation.

Overall, our proposed Automated Waste Classification System offers a non-intrusive and scalable solution for analyzing waste composition and optimizing waste management processes. By leveraging advanced computer vision techniques, we aim to improve efficiency, reduce operational costs, and promote sustainable waste management practices in urban and industrial environments.

3.1 System Design

The system design of our automated waste segregation project is depicted in the flowchart above, illustrating the comprehensive pipeline from raw data to model evaluation.

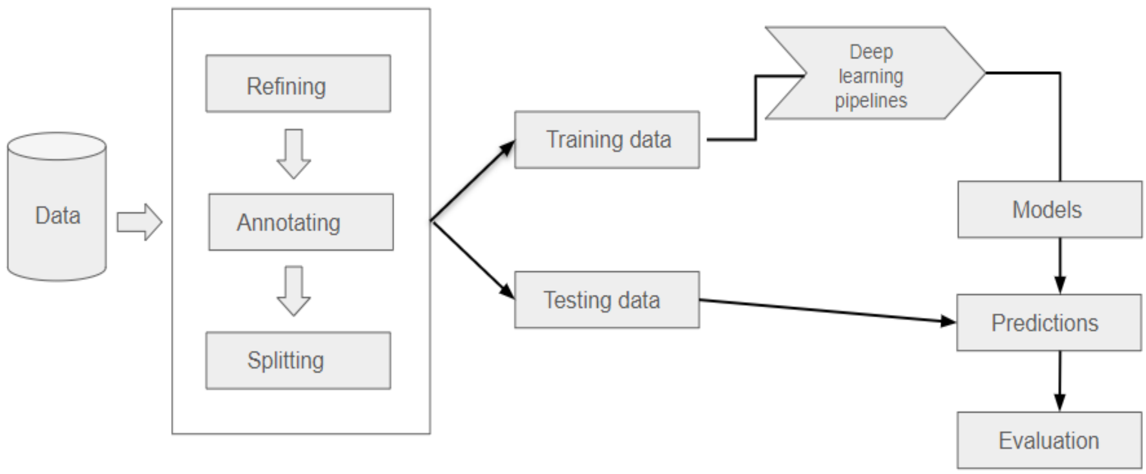

Fig 3.2 System Design

1. Data Collection and Refining: The process begins with collecting raw data from various sources. This data includes images of different types of waste. The raw data is then refined to ensure quality and consistency. This may involve cleaning the images, removing any that are unclear or irrelevant, and standardizing the data format.

2. Annotating: In this stage, the refined data is annotated. Annotation involves labeling the images with the correct waste categories such as plastic, paper, cardboard, glass, metal, thermocol, and wet waste. This step is crucial as it provides the ground truth for training the deep learning models. We used a combination of pre-existing annotated datasets and manually annotated images using tools like LabelImg.

3. Splitting: After annotation, the dataset is split into training, testing, and validation sets. Typically, we used a 70-10-10 split, where 70% of the data is used for training, 10% for validation, and 10% for testing. This ensures that the model is trained on a majority of the data but is also validated and tested on unseen data to evaluate its performance.

4. Training Data: The training data is fed into deep learning pipelines. In our project, we utilized the YOLOv5 model, a state-of-the-art object detection algorithm known for its balance between speed and accuracy. The model is trained to recognize and classify different types of waste based on the annotated training data.

5. Testing Data: Simultaneously, the testing data is kept aside to evaluate the model's performance after it has been trained. This data helps in assessing how well the model generalizes to new, unseen examples.

6. Deep Learning Pipelines: These pipelines encompass the entire process of model training, including data augmentation, pre-processing, and the application of the YOLOv5 architecture. Parameters like batch size, number of epochs, learning rate, and others are tuned to optimize the model's performance.

7. Models: Once trained, the models are capable of making predictions on new data. They take the input images and output the predicted waste categories for each detected object.

8. Predictions: The predictions from the models are compared against the ground truth annotations in the testing data to measure performance.

9. Evaluation: The final stage involves evaluating the model's predictions using various metrics such as precision, recall, F1 score, and mean Average Precision (mAP). These metrics help in understanding the accuracy and reliability of the model. Additionally, we tested the model on stock videos from the internet to further validate its robustness in real-world scenarios.

This structured approach ensures that the system is well-designed to handle the complexities of automated waste segregation, leveraging advanced deep learning techniques to achieve high accuracy and efficiency.

CHAPTER 4

IMPLEMENTATION

YOLO (You Only Look Once) is a family of real-time object detection algorithms. YOLOv5 is one of the latest versions, developed by Ultralytics, which aims to improve upon the speed and accuracy of previous iterations.

Here's a brief overview of how YOLOv5 works:

1. Single-Stage Architecture: YOLOv5 follows a single-stage object detection architecture, where object detection and classification are performed directly on the entire image in one pass. This approach significantly improves speed and efficiency compared to two-stage methods like Faster R-CNN.
2. Anchor-Free Detection: YOLOv5 adopts an anchor-free detection mechanism, eliminating the need for predefined anchor boxes. Instead, it predicts bounding boxes directly using convolutional layers and calculates object probabilities for each grid cell.
3. Feature Pyramid Network (FPN): YOLOv5 utilizes a feature pyramid network to capture multi-scale features from different layers of the network. This enables the model to detect objects of various sizes and scales within the image.
4. Backbone Architecture: YOLOv5 employs a feature extraction backbone based on the EfficientNet model architecture, typically utilizing a variant of the EfficientNet as its backbone network. This backbone extracts high-level features from the input image, which are then used for object detection.
5. Training Process: YOLOv5 is trained using a combination of labeled training data and a specified loss function, such as the YOLO loss or focal loss. During training, the model learns to optimize its parameters to minimize the difference between predicted bounding boxes and ground truth annotations.
6. Inference: During inference, YOLOv5 processes input images through the trained network, predicting bounding boxes and associated class probabilities for detected objects. Non-maximum suppression (NMS) is applied to filter out redundant bounding boxes and produce the final set of detections.

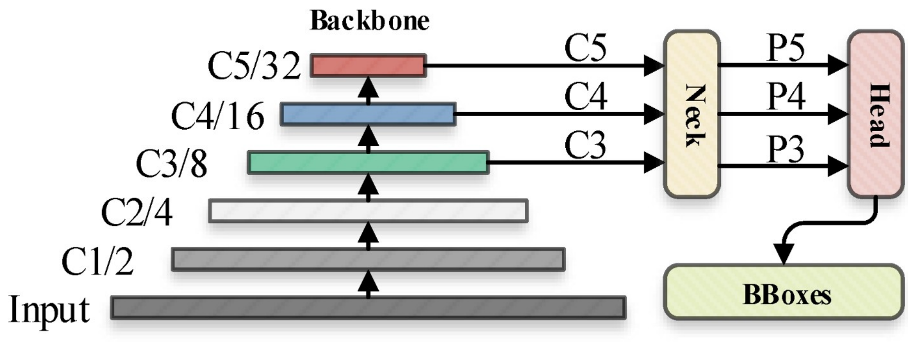

Fig 4.1 Default Inference Flowchart off YOLOv5

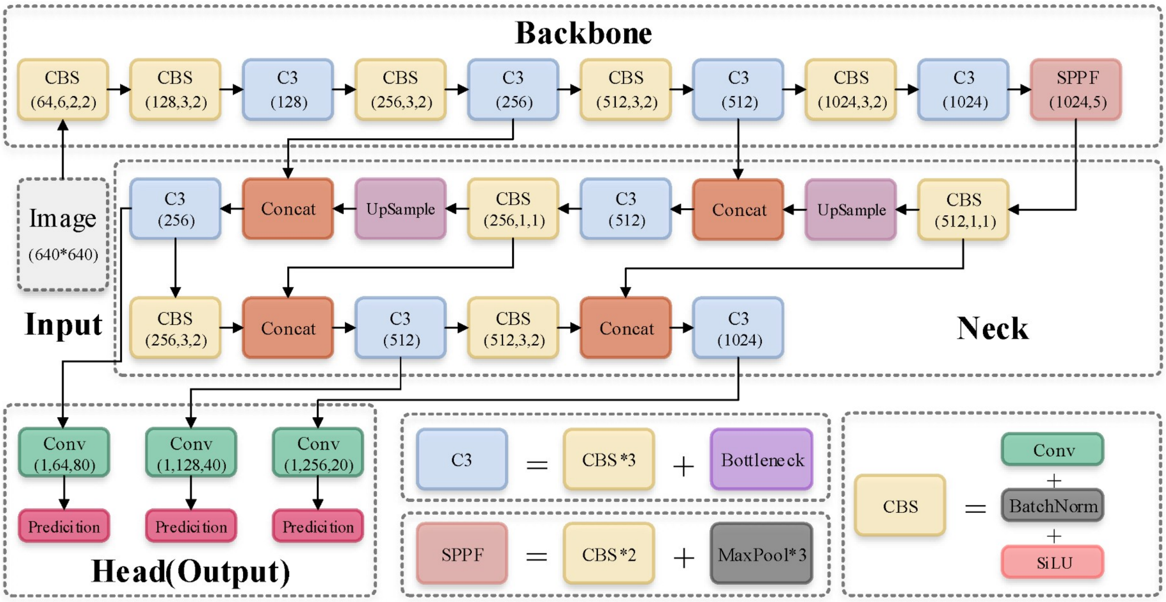

Fig 4.2 Default network structure of YOLOv5

Backbone: The backbone of YOLOv5 refers to the convolutional neural network (CNN) architecture responsible for extracting features from the input image. These features capture patterns and structures at different levels of abstraction, enabling the model to understand the content of the image. YOLOv5 typically utilizes a variant of the EfficientNet architecture as its backbone, which strikes a balance between computational efficiency and representation power.

Neck: In the YOLOv5 architecture, the neck refers to additional layers or modules that refine and enhance the features extracted by the backbone network. These may include spatial pyramid pooling (SPP) modules, feature pyramid networks (FPN), or other techniques designed to capture multi-scale information and improve the model's ability to detect objects of varying sizes and aspect ratios.

Head: The head of YOLOv5 is the final part of the network responsible for predicting bounding boxes, confidence scores, and class probabilities for objects in the input image. It typically consists of a series of convolutional and linear layers that process the features extracted by the backbone and neck, ultimately producing predictions for each grid cell in the output feature map.

Bounding Boxes: Bounding boxes are the rectangular regions predicted by YOLOv5 to localize objects within the image. Each bounding box is represented by four coordinates (x, y, width, height) relative to the image dimensions. YOLOv5 predicts bounding boxes directly, along with confidence scores indicating the model's certainty that an object is present within each box.

Intersection over Union (IoU): Intersection over Union (IoU) is a metric used to evaluate the overlap between predicted bounding boxes and ground truth annotations. It measures the ratio of the area of intersection between two bounding boxes to the area of their union. IoU is used during training to calculate the localization loss and during inference to perform non-maximum suppression (NMS) to remove redundant detections.

Non-Maximum Suppression (NMS): Non-Maximum Suppression is a post-processing technique used to filter out redundant or overlapping bounding box predictions. After object detection, YOLOv5 applies NMS to remove duplicate detections, keeping only the most confident ones. This helps ensure that each object is represented by a single bounding box with high confidence, improving the overall quality of the predictions.

By integrating these components—backbone, neck, head, bounding boxes, IoU, and NMS—into a cohesive architecture, YOLOv5 is able to efficiently and accurately detect objects in images, making it well-suited for tasks such as automated waste segregation using computer vision.

CHAPTER 5

RESULTS AND DISCUSSION

The implementation of YOLOv5 as our primary object detection method has yielded promising results in the automated recognition and classification of waste materials. By leveraging state-of-the-art deep learning techniques, our system demonstrates high accuracy, efficiency, and adaptability in identifying various waste items, including plastic, paper, metal, glass, cardboard, thermocol, and wet waste. The single-stage architecture of YOLOv5 allows for real-time processing of input images, making it well-suited for applications requiring rapid and accurate object detection.

However, our project also faces several challenges and limitations. One significant difficulty lies in distinguishing between transparent glass and transparent plastic items, as both materials exhibit similar visual properties in certain scenarios. This ambiguity poses a challenge for the model, potentially leading to misclassifications and reduced overall accuracy. Additionally, the presence of occlusions, reflections, and varying lighting conditions further complicates the task of waste classification, requiring robust solutions to ensure reliable performance in real-world environments.

5.1 Testing

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.

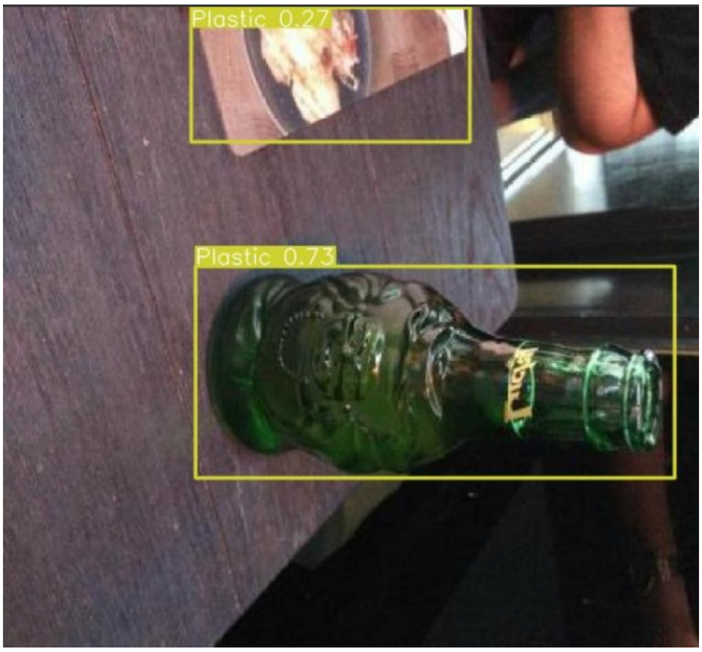

Fig 5.1 Misclassified instance

During the initial testing phase, our model demonstrated a notable instance of misclassification, as illustrated by the erroneous identification of a glass bottle as plastic in the image above. To address this issue and enhance the model's accuracy, we implemented several corrective measures. First, we augmented our dataset with more images specifically depicting plastic waste, ensuring these images reflected the real proportions of various waste types typically found in domestic settings. This helped the model better learn the distinguishing features of plastic compared to other materials.

Additionally, we incorporated shuffling during the dataset splitting process. By randomly shuffling the data before splitting it into training, validation, and testing sets, we ensured a more representative and balanced distribution of waste categories across all subsets. This step helped in reducing the chances of biased learning from any particular subset of the data.

We also experimented with different proportions for training and validation splits, using configurations like 70-30 and 80-20. By varying these proportions, we could better evaluate the model's generalization capabilities and ensure robust learning. Through these combined efforts, we significantly improved the model's performance, reducing instances of misclassification and enhancing its overall reliability in segregating waste materials accurately.

5.3 Snapshots

Fig 5.2 Predictions on Test Data 1

Fig. 5.3. Inference using File Upload 1

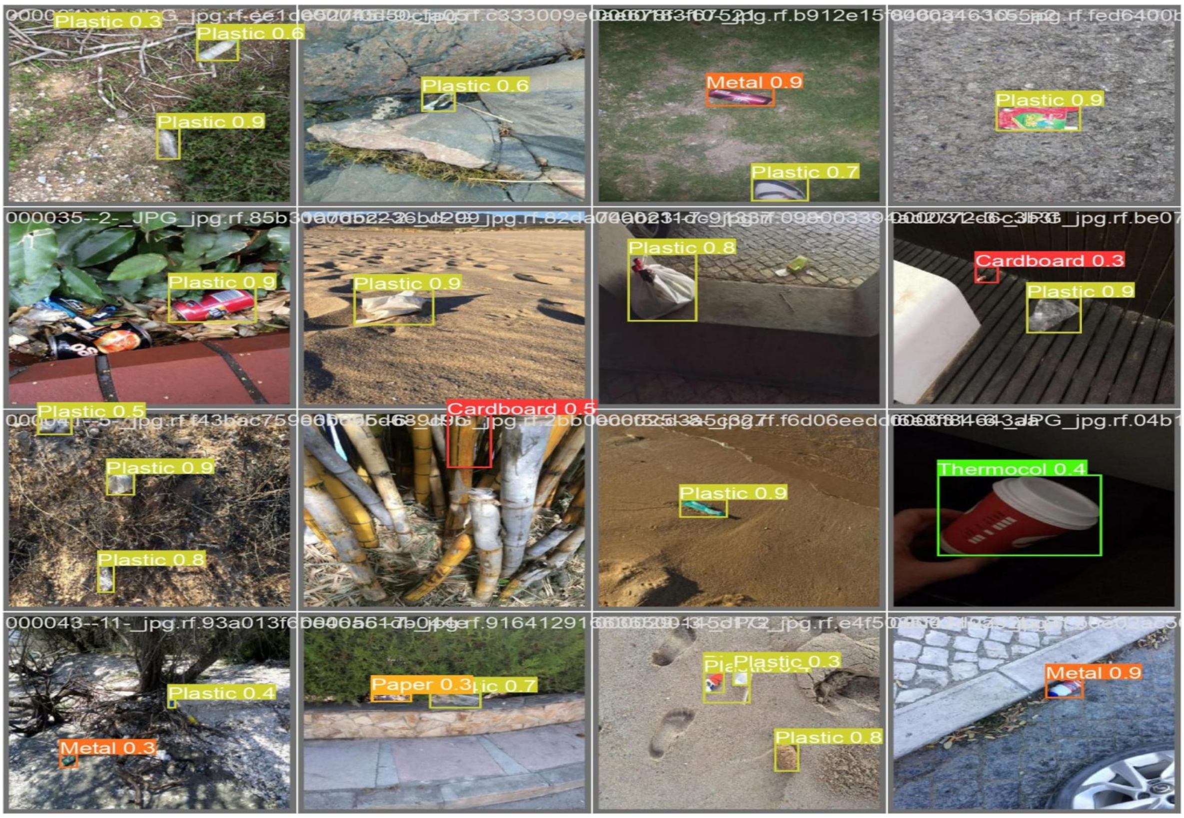

Fig 5.4 Predictions on Test Data 2

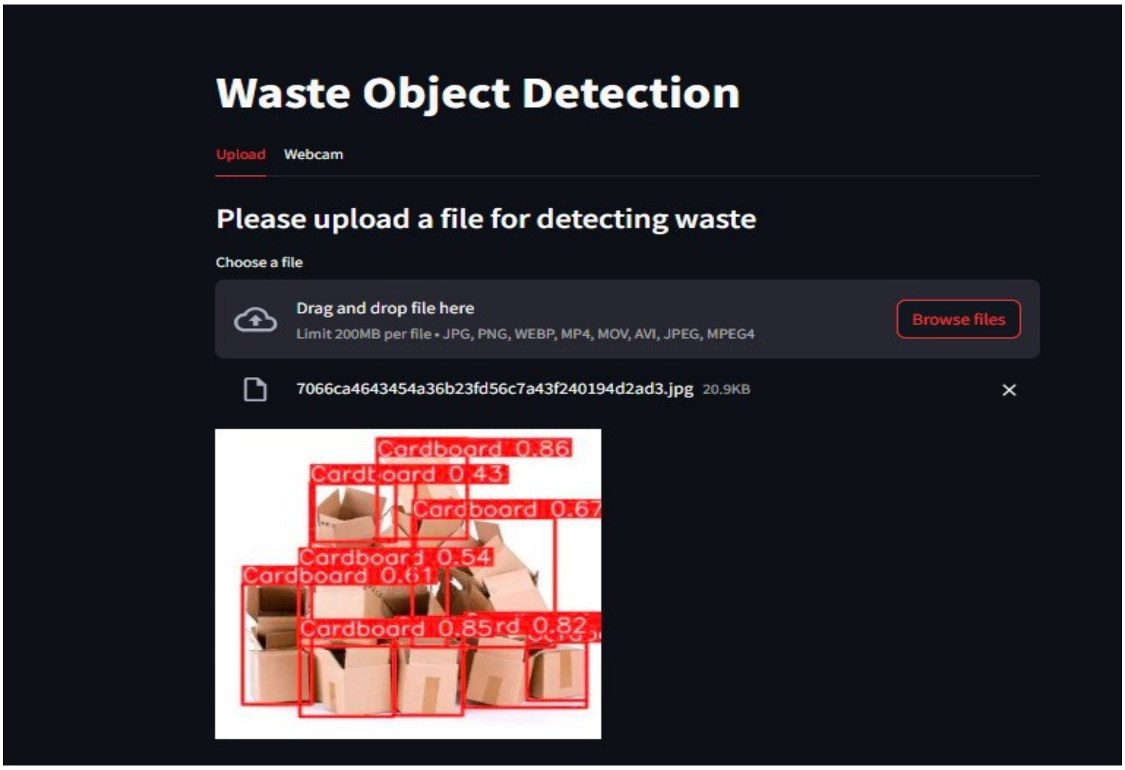

Fig 5.5 Inference using File upload 2

Fig. 5.6. Inference Using File Upload 3

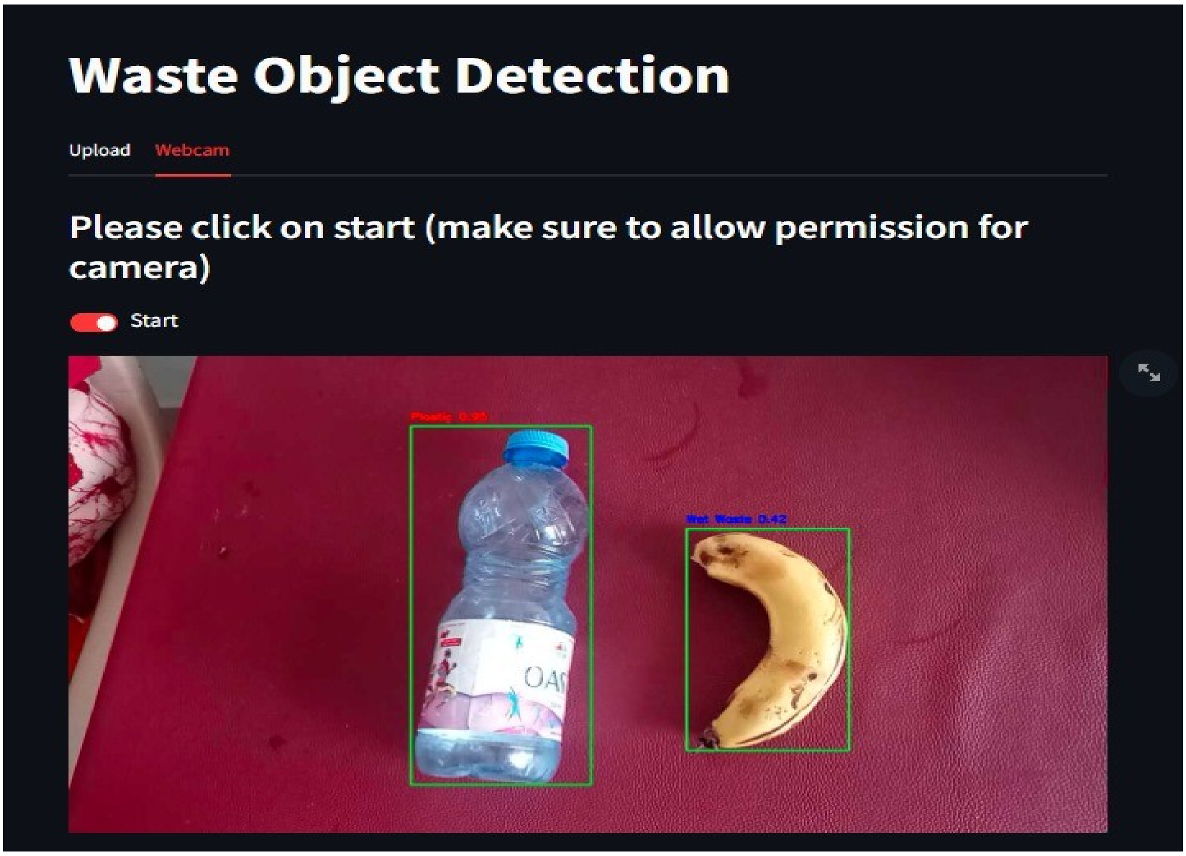

Fig. 5.7. Real time Inference Using Webcam

CHAPTER 6

CONCLUSION

To further enhance the accuracy and robustness of our system, integration with hardware components can provide valuable additional information. For example, the incorporation of sensors capable of detecting material properties such as density, conductivity, or magnetic susceptibility can complement the visual information provided by the camera. By fusing data from multiple sources, including both visual and sensor inputs, we can improve the reliability of waste classification and mitigate the challenges associated with visual ambiguity.

Furthermore, integrating hardware components such as robotic arms or conveyor belt sensors into the system can enable automated sorting and handling of waste items based on their classification. This holistic approach streamlines the waste management process, reducing manual intervention, increasing operational efficiency, and minimizing errors.
Therefore the integration of hardware components with our YOLOv5-based waste recognition system offers a comprehensive and synergistic solution for waste management, combining the strengths of both visual and sensor-based technologies to achieve greater accuracy, efficiency, and sustainability.

In conclusion, our project represents a significant advancement in waste management practices, offering a scalable and automated solution for waste recognition and segregation. By harnessing the power of YOLOv5 and deep learning, we have developed a system capable of accurately identifying and classifying diverse waste materials in real-time. This innovation holds immense potential for revolutionizing waste management processes, facilitating efficient recycling, reducing environmental pollution, and promoting sustainability

CHAPTER 7

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] Narayanswamy, Neeraja, A. Rajak, and Shazia Hasan. "Development of Computer Vision Algorithms for Multiclass Waste Segregation and Their Analysis." Emerging Science Journal 6, no. 3 Apr 2022: 631-646.

[20] Vavilala Sushma. "Plastic Material Detection and classification using CNN." International Journal of Creative Research Thoughts IJCRT Nov 2022: 2320-2882

[21] Faisal, Tarig, Aman Eyob, Filmon Debretsion, Merhawi Tsegay, Anees Bashir, and Moath Awawdeh. "Development of an intelligent waste segregation system based on convolutional neural networks." International Journal of Advanced Science and Technology 29, no. 3 May 2020: 14837-14849.

[22] Kumar, Saurav, Drishti Yadav, Himanshu Gupta, Om Prakash Verma, Irshad Ahmad Ansari, and Chang Wook Ahn. "A novel yolov3 algorithm-based deep learning approach for waste segregation: Towards smart waste management." Electronics 10, no. 1 30 Nov 2020: 14.

[23] Sharholy, Mufeed, Kafeel Ahmad, Gauhar Mahmood, and R. C. Trivedi. "Municipal solid waste management in Indian cities–A review." Waste management 28, no. 2 Apr-2008: 459-467.

CHAPTER 8

SOURCE CODE

```
import os
import av
import cv2
import torch
import shutil
import pathlib
import streamlit as st
from ffmpy import FFmpeg
from streamlit_webrtc import webrtc_streamer
temp = pathlib.PosixPath
pathlib.PosixPath = pathlib.WindowsPath
MODEL_PATH = r"C:\Users\mypc\Documents\waste\yolov5\yolov5"
MODEL_FILE_PATH = r"C:\Users\mypc\Documents\waste\yolov5\models\
eq_dry_wet.pt"
IMG_SAVE_PATH = r"C:\Users\mypc\Documents\waste\yolov5\temp\result"
TEMP_VID_SAVE_PATH = r"C:\Users\mypc\Documents\waste\yolov5\temp\temp.avi"
VID_SAVE_PATH = r"C:\Users\mypc\Documents\waste\yolov5\temp\result.mp4"
TEMP_PATH = r"C:\Users\mypc\Documents\waste\yolov5\temp"
FFMPEG_EXE_PATH = r"C:\Users\mypc\ffmpeg\bin\ffmpeg.exe"
class_colors = {
    'Cardboard': (0, 255, 0),   # Green
    'Plastic': (255, 0, 0),     # Blue
    'Paper': (0, 0, 255),       # Red
    'Metal': (0, 255, 255),     # Yellow
    'Glass': (255, 255, 0),     # Cyan
    'Thermocol': (128, 0, 128)  # Purple
}
class VideoProcessor():
    def __init__(self):
```

```
        self.model = load_model()

    def recv(self,frame):
        frame = frame.to_ndarray(format="bgr24")
        height,width = frame.shape[0:2]
        frame = cv2.resize(frame, (width,height))
        results = self.model(frame)
        for index, row in results.pandas().xyxy[0].iterrows():
            x1, y1, x2, y2, confidence, class_id, class_name = row
            color = class_colors.get(class_name, (0, 0, 255))
            cv2.rectangle(frame, (int(x1), int(y1)),
                          (int(x2), int(y2)), (0, 255, 0), 2)
            cv2.putText(frame, f'{class_name} {confidence:.2f}', (int(
                x1), int(y1) - 10), cv2.FONT_HERSHEY_SIMPLEX, 0.5, color, 2)
        return av.VideoFrame.from_ndarray(frame, format="bgr24")
def delete_temp(path):
    if os.path.exists(path):
        if os.path.isfile(path):
            os.remove(path)
        else:
            shutil.rmtree(path)
def convert_compat():
    delete_temp(VID_SAVE_PATH)
    ff = FFmpeg(
        executable=FFMPEG_EXE_PATH,
        inputs={TEMP_VID_SAVE_PATH: None},
        outputs={VID_SAVE_PATH: '-c:v libx264'}
    )
    temp = ff.run()
    return temp
def load_model():
    model = torch.hub.load(
        MODEL_PATH, "custom",
        path=MODEL_FILE_PATH,
```

```
            source="local",
            force_reload=True
        )
        return model
    def predict_image(location):
        model = load_model()
        results = model(location, size=640)
        delete_temp(IMG_SAVE_PATH)
        results.save(save_dir=IMG_SAVE_PATH)
        return results
    def predict_video(location,change=False):
        stframe = st.empty()
        cap = cv2.VideoCapture(location)
        width = int(cap.get(cv2.CAP_PROP_FRAME_WIDTH))
        height = int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT))
        size = (width, height)
        out = cv2.VideoWriter(TEMP_VID_SAVE_PATH,
                              cv2.VideoWriter_fourcc(*'XVID'), 15, size)
        model = load_model()
        while True:
            ret, frame = cap.read()
            if not ret:
                break
            if change:
                frame = cv2.cvtColor(frame,cv2.COLOR_BGR2RGB)
            frame = cv2.resize(frame, (width, height))
            results = model(frame)
            for index, row in results.pandas().xyxy[0].iterrows():
                x1, y1, x2, y2, confidence, class_id, class_name = row
                color = class_colors.get(class_name, (0, 0, 255))
                cv2.rectangle(frame, (int(x1), int(y1)),
                              (int(x2), int(y2)), (0, 255, 0), 2)
                cv2.putText(frame, f'{class_name} {confidence:.2f}', (int(
                    x1), int(y1) - 10), cv2.FONT_HERSHEY_SIMPLEX, 0.5, color, 2)
```

```
            stframe.image(frame)
            out.write(frame)
        cap.release()
        out.release()
        stframe.empty()
        temp = convert_compat()
        return temp
    st.title("Waste Object Detection")
    tab1, tab2 = st.tabs(["Upload", "Webcam"])
    with tab1:
        st.subheader("Please upload a file for detecting waste")
        uploaded_file = st.file_uploader(
            "Choose a file", ['jpg', 'png', 'webp', 'mp4', 'mov', 'avi'])
        if uploaded_file is not None:
            file_name, file_ext = uploaded_file.name.split('.')
            temp_filename = TEMP_PATH+f"\\temp.{file_ext}"
            with open(temp_filename, "wb") as f:
                f.write(uploaded_file.getvalue())
            with st.spinner("Running model (this may take a while)"):
                if file_ext in ['jpg', 'png', 'webp','jpeg']:
                    prediction = predict_image(temp_filename)
                    st.image(IMG_SAVE_PATH+f"\\temp.jpg")
                else:
                    prediction = predict_video(temp_filename)
                    st.video(VID_SAVE_PATH)
                    delete_temp(TEMP_VID_SAVE_PATH)
            delete_temp(temp_filename)
    with tab2:
        st.subheader(
            "Please click on start (make sure to allow permission for camera)")
         webrtc_streamer(key="example",video_processor_factory=VideoProcessor)
```

COPY OF THE PAPER PROJECT PHASE 1 SUBMITTED TO : International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)

A Review on Automated Waste Segregation System using Machine Learning

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

This literature review paper critically examines the state-of-the-art in AI-powered waste segregation, focusing on the pivotal steps of object recognition, feature extraction, and classification. As waste management becomes an increasingly urgent global concern, leveraging artificial intelligence (AI) and computer vision technologies holds the promise of revolutionizing waste segregation processes. The paper surveys key papers and methodologies in object recognition, shedding light on models such as Faster R-CNN, YOLO, and Mask R-CNN. It explores feature extraction techniques, considering advancements in image classification architectures like ResNet, VGGNet, and DenseNets. Additionally, the review covers classification methodologies, including specialized approaches for plastic material detection and intelligent waste segregation systems. The comparative analysis and critical insights presented in this paper aim to contribute to the ongoing discourse on developing robust and efficient AI-driven waste segregation systems, fostering a sustainable approach to waste management.

Keywords

Waste Segregation,Computer Vision, Deep Learning, Faster R-CNN, YOLO, Mask R-CNN, ResNet, VGGNet, DenseNets, Plastic Material Detection.

1. Introduction

Waste management stands at the forefront of global environmental challenges, necessitating innovative solutions to address the escalating volume of waste generated worldwide. In this context, the integration of artificial intelligence (AI) and computer vision technologies offers a transformative potential to revolutionize waste segregation processes. This literature review aims to provide a comprehensive overview of advancements in AI-powered waste segregation, with a particular focus on the essential stages of object recognition, feature extraction, and classification.[23]

The proliferation of waste, coupled with the limitations of traditional waste management systems, underscores the urgency for novel approaches. AI, characterized by its capacity to learn and adapt, presents a promising avenue to enhance the accuracy and efficiency of waste segregation. Object recognition, the first crucial step, involves the identification of distinct items within a waste stream. This review examines seminal papers and methodologies in this domain, encompassing models like Faster R-CNN, YOLO, and Mask R-CNN.

Moving beyond recognition, feature extraction plays a pivotal role in distilling relevant information from waste images. The exploration of image classification architectures such as ResNet, VGGNet,

and DenseNets forms a significant part of this review, shedding light on advancements in feature extraction techniques. Additionally, the paper delves into classification methodologies, including specialized approaches for plastic material detection and the development of intelligent waste segregation systems.
By undertaking a comparative analysis of these methodologies, this literature review aims to contribute insights into the evolving landscape of AI-powered waste segregation. The critical examination of existing approaches will not only inform researchers and practitioners but also guide the development of more robust and efficient waste segregation systems. Ultimately, the goal is to foster a sustainable paradigm in waste management, mitigating environmental impact and promoting a more responsible and intelligent approach to handling the challenges of modern waste disposal.

2. Related Works

Publication in IOP Conference Series: Materials Science and Engineering.Garbage Waste Segregation Using Deep Learning Techniques by Sai Susanth G et al. paper presents an exploration into waste segregation utilizing deep learning techniques, aligning closely with our project's objectives. The authors likely delve into object recognition and classification methods, providing valuable insights into the practical application of deep learning for waste management.

Publication in International Journal of Creative Research Thoughts. Vavilala Sushma’s work specifically focuses on the detection and classification of plastic materials using Convolutional Neural Networks (CNN). Given the relevance to our project's classification aspect, this paper offers a deep dive into methodologies and techniques that could inform our own approach to plastic waste identification.

Published in the International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)."Development of Computer Vision Algorithms for Multi-class Waste Segregation and Their Analysis" by Neeraja Narayanswamy et al. did the work involving the development and analysis of computer vision algorithms for multi-class waste segregation, which is highly pertinent to our project's multi-step process. It likely covers aspects of object recognition, feature extraction, and classification, providing a holistic perspective on waste segregation leveraging computer vision techniques.

2.1 Methodology

The methodology for waste detection involves a systematic progression through key steps, ensuring a comprehensive approach to the task. The initial phase involves the capturing of visual data, where images or videos of waste scenes are acquired. Subsequently, the process transitions to object detection, where the system identifies and localizes relevant objects within the captured data. Following object detection, feature extraction comes into play, emphasizing the extraction of meaningful information from the identified objects. The final step encompasses classification, where the system assigns specific waste categories to the detected and extracted features, enabling efficient and accurate waste segregation. This sequential methodology forms the backbone of an effective waste detection system, laying the groundwork for subsequent detailed discussions on specific techniques employed in each step.

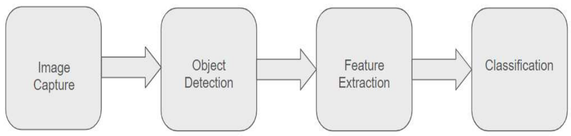

Figure 1 : Phases in Waste segregation

2.1.1 Capturing

The initial phase involves capturing images or videos of the waste stream using cameras or sensors. This data acquisition is typically done in real-time or at regular intervals to monitor the waste disposal area continually. Cameras are strategically placed in the waste disposal site, capturing visual information of the waste materials.

2.1.2 Object Detection

In the realm of object recognition, the discussed papers showcase distinct approaches, each with its merits and drawbacks. "Faster R-CNN" [1] introduces Region Proposal Networks for real-time object detection, emphasizing accuracy, albeit at the expense of speed. "You Only Look Once" [2] innovatively unifies object detection in real-time but might compromise precision for efficiency. "SSD"

[3] offers a balance with its Single Shot MultiBox Detector, achieving commendable speed and accuracy. "Mask R-CNN" [4] extends Faster R-CNN by incorporating instance segmentation, enhancing detailed object delineation. "YOLO9000" [5] stands out for its speed, simultaneously detecting a vast number of object classes. Choosing the most suitable approach hinges on specific project requirements: Faster R-CNN for precision, YOLO9000 for speed, SSD for a balanced compromise, and Mask R-CNN for instance segmentation needs. The trade-offs involve a nuanced interplay between speed, accuracy, and segmentation capabilities, demonstrating the necessity of aligning the chosen model with the particular demands of the waste segregation project.

2.1.3 Feature Extraction

In the realm of feature extraction, the presented papers offer distinctive insights into leveraging deep convolutional neural networks (CNNs). "ImageNet Classification with Deep Convolutional Neural Networks" [6] by Krizhevsky et al. laid the foundation by demonstrating the power of deep CNNs in image classification. "Going Deeper with Convolutions" [7] explores the effectiveness of deeper architectures, introducing the GoogLeNet model, emphasizing computational efficiency. "Visualizing and Understanding Convolutional Networks" [8] by Zeiler and Fergus delves into visualizing CNNs, aiding in interpreting learned features. "Deep Residual Learning for Image Recognition" [9] introduces ResNet, addressing vanishing gradient issues and enabling the training of extremely deep networks. "Inception-v4, Inception-ResNet, and the Impact of Residual Connections on Learning" [10] by Szegedy et al. integrates residual connections into the Inception framework, showcasing improved performance. Determining the most suitable feature extraction method depends on project requirements, with ResNet often preferred for its ability to handle deep architectures effectively, mitigating degradation issues encountered by simpler networks. Each approach brings its set of advantages and trade-offs, emphasizing the need for a tailored choice based on specific project needs.

2.1.4 Classification

In the realm of image classification, the discussed papers showcase diverse approaches, each with its strengths and limitations. "ImageNet Large Scale Visual Recognition Challenge" [11] provides a benchmark in large-scale image classification, establishing a foundation for subsequent research. "Very Deep Convolutional Networks for Large-Scale Image Recognition " [12] introduces the influential VGGNet, emphasizing the advantages of deeper architectures but at the cost of increased computational complexity. "ResNet in ResNet: Generalizing Residual Architectures” [13] offers a generalized approach to residual architectures, improving expressivity and ease of information removal. "Xception: Deep Learning with Depth Wise Separable Convolutions' ' [14] introduces an efficient alternative with depthwise separable convolutions, achieving competitive accuracy with reduced computational demands. " SqueezeNet" [15] stands out for its compact model size without compromising accuracy. Choosing the optimal classification model depends on project requirements, with ResNet offering deeper architectures and Xception providing computational efficiency, while SqueezeNet excels in scenarios with limited computational resources. The decision hinges on a nuanced balance between accuracy, model complexity, and computational efficiency, demonstrating the need for tailored choices based on project-specific considerations.

2.1.5 Physical Interference:

After the AI-driven detection and classification phases, physical interference refers to the mechanical actions taken based on the identified waste types. This could include automated sorting mechanisms, conveyor belts, or robotic systems that physically segregate the waste according to its classification. Various technologies like robotic arms, conveyor belt diverters, or automated sorting systems are integrated into the waste management process. These physical systems act based on the classification results to separate materials efficiently.

By integrating these components seamlessly, the waste detection system efficiently combines AI-based analysis with physical interventions to enhance the overall waste segregation process. This integrated approach contributes to more accurate and automated waste management practices, reducing manual labor and improving efficiency.

3. Proposed model

Classification is an approach wherein attributes are derived from the dataset. This is accomplished by dividing the information into various groupings, relying on the attributes. A fresh model conducts forecasts and categorizes them by training on familiar data. The proposed framework comprises three fundamental components, including data preprocessing, image augmentation, and attribute derivation. Image augmentation aims to generate additional images by adjusting size, zooming, rotating images, etc., to produce novel images. Through this methodology, the model will be equipped to capture a greater number of 'features' than before and will be capable of predicting images more effectively. During the process of attribute derivation, the system characterizes the unlabelled data to the utmost extent possible

Pre-processing and data augmentation

The dataset's modest scale poses a challenge for pre-trained models, raising apprehensions about the potential for overfitting. As a pre-emptive measure prior to model training, strategic interventions are necessary. One such intervention involves augmenting the dataset by doubling its size through the incorporation of images from the Google database. Furthermore, to enhance the dataset's diversity and mitigate overfitting risks, specific augmentation techniques, including Random Re-sized Crop and Random Horizontal Flip, are judiciously applied.

3.1 Convolutional neural network

The Convolutional Neural Network (CNN) assumes a prominent role in the realm of image analysis. It distinguishes itself through the incorporation of hidden layers known as convolutional layers, imparting a distinctive quality to its architecture. Within each convolutional layer, a collection of filters is embedded, and these filters serve the purpose of identifying patterns or features within the images. This layered approach enhances the network's ability to discern intricate details and extract meaningful information during the image analysis process. A simplest CNN has the following layers:

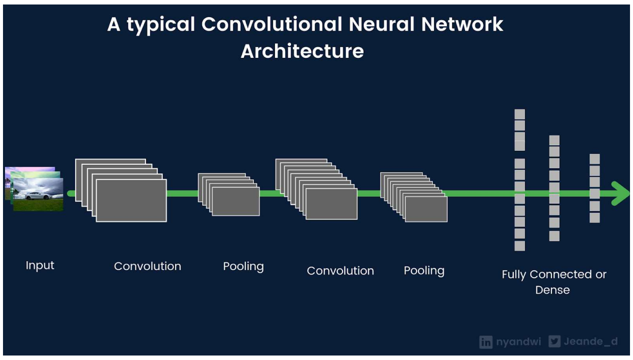

Figure 2 : Architecture of CNN[24]

1. Convolutional Layer:

The convolutional layer plays a critical role in the Convolutional Neural Network architecture, specializing in the extraction of image features through the employment of filters. These filters, characterized by small matrices with dimensions tailored according to our specifications and populated with random values, serve as discerning agents. They systematically traverse the input images, identifying patterns through a striding process. The culmination of this operation is the generation of a resultant feature map, which encapsulates the detected features. Subsequently, this feature map is seamlessly transmitted to the subsequent layer, contributing to the network's ability to capture and understand intricate patterns within the input data.

2. Pooling Layer:

Within this layer, a window, typically with dimensions of 2x2, is positioned over the feature map. The key operation involves selecting the maximum value within the window while disregarding all other values. This process leads to a reduction in the scale of the picture, thereby achieving a down sampling effect. This down sampling is instrumental in retaining essential features while concurrently diminishing computational complexity

3. Fully Connected Layer:

The pivotal stage for image recognition and classification unfolds within the realm of the fully connected layer. At this juncture, the diminished images are compiled into a singular vector. This vector, representative of the condensed image information, undergoes a comparative analysis. The classification of the image is then determined by matching its vector with those derived from the training images. This process encapsulates the final stage of the neural network's operation, where intricate comparisons lead to accurate image categorization based on the learned patterns from the training data.

3.2 DenseNets

DenseNet, short for Densely Connected Convolutional Networks, is a convolutional neural network (CNN) architecture designed to address challenges related to feature reuse, vanishing gradient problems, and overall network efficiency. It was introduced by Gao Huang, Zhuang Liu, and Kilian Q. Weinberger in their 2017 paper titled "Densely Connected Convolutional Networks.". The key innovation of DenseNet lies in its dense connectivity pattern, where each layer is connected to every other layer in a feed-forward fashion. Unlike traditional CNN architectures where information flows sequentially from one layer to the next, DenseNet allows for direct connections between layers at different depths. Each layer receives as input the feature maps of all preceding layers, and in turn, its own feature maps are used as inputs to all subsequent layers. [16]

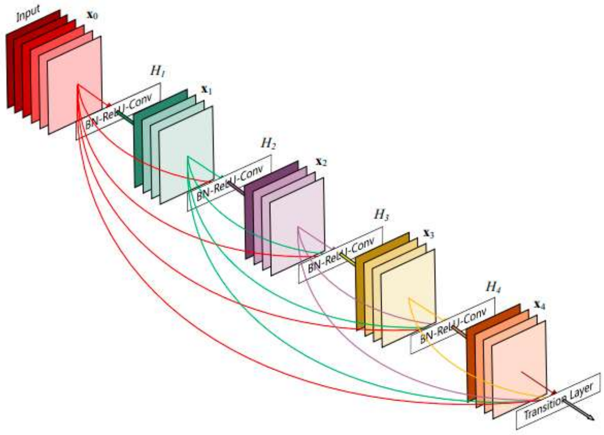

Figure 3 : A 5-layer dense block with a growth rate of k = 4. Each layer takes all preceding feature-maps as input. [16]

This dense connectivity has several advantages including Feature Reuse, Gradient Flow, Parameter Efficiency The basic building block of DenseNet is the dense block, where each layer produces k new feature maps, and these k feature maps are concatenated with the input feature maps. Transition layers, which include batch normalization, pooling, and convolution, are used to manage the growth of feature maps and downsample the spatial dimensions

Sl No	Name of the Paper	Author	Description	Observations
[1]	Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks	Shaoqing Ren Kaiming He Ross Girshick Jian Sun	Introduces RPN for efficient proposal generation. Real-time detection with high accuracy.	Unified framework for region proposal and detection. Computationally demanding during training. Can provide robust object detection for waste items with real-time processing, contributing to efficient segregation.
[2]	You Only Look Once: Unified, Real-Time Object Detection	Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi	Proposes YOLO for single-pass, real-time detection with emphasis on speed.	Extremely fast with a single forward pass. May struggle with small object detection. Suitable for quick and accurate identification of waste items, facilitating real-time segregation.
[3]	SSD: Single Shot MultiBox Detector	Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian et al	Introduces SSD for single-shot, multi-scale detection, balancing speed and accuracy.	High accuracy across different object scales. Sensitive to aspect ratio and size. Can contribute to accurate and efficient waste segregation, capturing various sizes and shapes of waste items.

[4]	Mask R-CNN	Kaiming He Georgia Gkioxari Piotr Doll´ar Ross Girshick	Extends Faster R-CNN for instance segmentation. Precise instance segmentation with high object detection performance.	Can be computationally intensive. Offers precise instance segmentation, allowing detailed identification of waste objects. Integration may enhance the accuracy of segregation.
[5]	YOLO9000: Better, Faster, Stronger	Joseph Redmon, Ali Farhadi	Extends YOLO for a large number of object classes. Introduces hierarchical classification.	Handles many object classes in real-time. Performance may degrade for small or similar classes. Enables a wide range of waste categorization, useful for diverse waste items in the segregation process.
[6]	ImageNet Classification with Deep Convolutional Neural Networks (AlexNet)	Alex Krizhevsky Ilya Sutskever Geoffrey E. Hinton	Pioneers deep learning for image classification. Popularizes ReLU and dropout.	Significant improvement in accuracy. Limited context modeling in deeper layers. Provides a foundation for understanding deep learning in image classification, relevant for feature extraction in waste item recognition.
[7]	Going Deeper with Convolutions (VGGNet)	Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, et al.	Emphasizes depth with small convolutional filters. Simplifies architecture.	Competitive performance with increased depth. Increased computational complexity. A deeper understanding of how increased depth affects performance can guide model selection for waste segregation.
[8]	Visualizing and Understanding Convolutional Networks (ZFNet)	Matthew D. Zeiler Rob Fergus	Explores visualizations for understanding learned features. Improved interpretability.	Enhanced visualization of features. Enhances interpretability of CNNs. Visualization techniques may aid in understanding the learned features, facilitating better interpretation of waste items during segregation.
[9]	Deep Residual Learning for Image Recognition (ResNet)	Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun	Introduces residual learning with skip connections. Enables very deep networks.	Significantly increased depth without degradation. Addresses vanishing gradient problem. Increased computational complexity. Residual learning may improve the training of deep networks, beneficial for recognizing complex waste items.
[10]	Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning (Inception-ResNet)	Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alexander A. Alemi	Combines Inception with residual connections. Achieves high accuracy with improved speed.	Enhanced feature learning with a combination of inception modules and residuals. Achieves high accuracy with improved efficiency. The combination of inception and residual connections may provide a balanced approach for effective waste item recognition.

[11]	ImageNet Large Scale Visual Recognition Challenge (AlexNet)	Olga Russakovsky · Jia Deng Hao Su et al.	Demonstrates deep learning effectiveness on large-scale classification. Popularizes GPU acceleration.	Significant improvement in accuracy. Limited context modeling in deeper layers. The foundational work on deep learning can guide the implementation of robust waste segregation models.
[12]	Very Deep Convolutional Networks for Large-Scale Image Recognition (VGGNet)	Karen Simonyan & Andrew Zisserman	Emphasizes depth with small convolutional filters. Simplifies architecture.	Competitive performance with increased depth. Increased computational complexity. A deeper understanding of the trade-offs between depth and performance, relevant for selecting an appropriate model for waste segregation.
[13]	ResNet in ResNet: Generalizing Residual Architectures	Sasha Targ, Diogo Almeida, Kevin Lyman	Introduces a novel deep learning architecture termed ResNet in ResNet (RiR) that extends and generalizes the success of residual networks (ResNets) and standard convolutional neural networks (CNNs).	The architecture employs a dual-stream design that seamlessly combines the strengths of ResNets and traditional CNNs without introducing computational overhead. The proposed RiR consistently demonstrates superior performance compared to standard ResNets and even outperforms architectures employing similar levels of augmentation on the CIFAR-10 dataset.
[14]	Xception: Deep Learning with Depth Wise Separable Convolutions	Francois Chollet	Introduces depthwise separable convolutions for reduced computational complexity.	Improved efficiency with maintained performance. May require careful hyperparameter tuning. The reduced computational complexity makes it a potential candidate for resource-efficient waste segregation models.
[15]	SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size	Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer	Designs a compact CNN architecture with fewer parameters. Achieves high accuracy with smaller model size.	Efficient use of parameters and model size. Suitable for resource-constrained environments. May not perform as well as larger models on certain tasks. SqueezeNet's compact design is advantageous for deploying waste segregation in resource-constrained environments.
[16]	The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation	Simon J´egou, Michal Drozdzal, David Vazquez, Adriana Romero, Yoshua Bengio	Introduces a fully convolutional DenseNet architecture for semantic segmentation.	DenseNet architecture enhances feature reuse, beneficial for semantic segmentation tasks. Relevant for understanding advanced semantic segmentation techniques, applicable to precise waste item identification.

[17]	Garbage Waste Segregation Using Deep Learning Techniques	Sai Sushanth G, Jenila Livingston LM, Agnel Livingston LGX	Applies deep learning techniques for garbage waste segregation.	Specific focus on practical waste segregation applications using deep learning. Offers insights into applying deep learning for efficient garbage waste segregation, directly applicable to waste management projects.
[18]	Source separation, transportation, pretreatment, and valorization of municipal solid waste: a critical review	Xuemeng Zhang, Chao Liu, Yuexi Chen, Guanghong Zheng, Yinguang Chen	Conducts a critical review of municipal solid waste management processes.	Focuses on understanding and improving the overall municipal solid waste management process. Essential for gaining insights into holistic waste management strategies, guiding the broader context of waste segregation projects.
[19]	Development of Computer Vision Algorithms for Multi-class Waste Segregation and Their Analysis	Neeraja Narayanswamy, A. R. Abdul Rajak, Shazia Hasan	Develops computer vision algorithms for multi-class waste segregation.	Emphasis on developing algorithms for precise multi-class waste segregation. Relevant for implementing computer vision techniques in waste segregation models for accurate multi-class identification.
[20]	Plastic Material Detection and Classification using CNN	Vavilala Sushma	Focuses on plastic material detection and classification using Convolutional Neural Networks (CNN).	Tailored approach for handling plastic waste through CNN-based techniques. Offers specific methods for plastic waste identification using CNNs, contributing to plastic material segregation in waste projects.
[21]	Development of Intelligent Waste Segregation System Based on Convolutional Neural Network	Tarig Faisal, Aman Eyob, Filmon Debretsion, Merhawi Tsegay, Anees Bashir, Moath Awawdeh	Proposes an intelligent waste segregation system using Convolutional Neural Networks.	Focus on the integration of intelligent systems for waste management. Provides a foundation for incorporating intelligent systems into waste segregation projects, enhancing efficiency and accuracy.
[22]	A Novel YOLOv3 Algorithm-Based Deep Learning Approach for Waste Segregation: Towards Smart Waste Management	Saurav Kumar, Drishti Yadav, Himanshu Gupta, Om Prakash Verma, Irshad Ahmad Ansari, Chang Wook Ahn	Introduces a novel YOLOv3-based deep learning approach for waste segregation.	Emphasis on smart waste management through advanced deep learning techniques. Offers an innovative algorithm for efficient waste segregation, contributing to the broader field of smart waste management.
[23]	Municipal solid waste management in Indian cities – A review	Mufeed Sharholy, Kafeel Ahmad, Gauhar Mahmood, R.C. Trivedi	Conducts a review of municipal solid waste management practices in Indian cities.	Addresses the specific context of municipal solid waste management in urban India. Valuable for understanding the challenges and practices in municipal solid waste management, particularly in the Indian urban context.

Table 1 : Comparative study

4. Conclusion

In conclusion, the exploration and analysis of various papers related to waste segregation using AI and computer vision have provided valuable insights into the current state of the field. Object recognition, feature extraction, and classification, essential components of the waste segregation process, have been examined through the lens of cutting-edge papers in the domain. For object recognition, Faster R-CNN, You Only Look Once (YOLO), Single Shot MultiBox Detector (SSD), Mask R-CNN, and YOLO9000 were considered. Each method has its strengths, with Faster R-CNN excelling in accuracy, YOLO focusing on real-time detection, SSD optimizing for speed and efficiency, and Mask R-CNN offering precise instance segmentation.

Feature extraction methodologies, represented by ImageNet Classification, VGGNet, ZFNet, ResNet, and Inception-ResNet, were compared. ResNet, with its deep residual learning, stood out for its ability to train very deep networks and mitigate the vanishing gradient problem.

For classification, ImageNet Large Scale Visual Recognition Challenge, Very Deep Convolutional Networks (VGGNet), ResNet in ResNet, Xception, and SqueezeNet were evaluated. Each method showcased unique advantages, with Xception demonstrating depth wise separable convolutions and SqueezeNet achieving high accuracy with minimal parameters.

Combining these findings, the proposed waste segregation system should leverage Faster R-CNN for object recognition, ResNet for feature extraction, and Xception for classification. This hybrid approach aims to achieve optimal accuracy, real-time processing, and efficient use of computational resources in addressing the complex challenges of waste segregation. The choice of these models aligns with the project's overarching goals, considering factors such as accuracy, real-time processing, and resource efficiency. This hybrid model reflects a commitment to leveraging the latest advancements in the field to address the environmental challenges posed by inefficient waste management.

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]. Narayanswamy, Neeraja, A. Rajak, and Shazia Hasan. "Development of Computer Vision Algorithms for Multiclass Waste Segregation and Their Analysis." Emerging Science Journal 6, no. 3 Apr 2022: 631-646.

[20]. Vavilala Sushma. "Plastic Material Detection and classification using CNN." International Journal of Creative Research Thoughts IJCRT Nov 2022: 2320-2882

[21]. Faisal, Tarig, Aman Eyob, Filmon Debretsion, Merhawi Tsegay, Anees Bashir, and Moath Awawdeh. "Development of an intelligent waste segregation system based on convolutional neural networks." International Journal of Advanced Science and Technology 29, no. 3 May 2020: 14837-14849.

[22]. Kumar, Saurav, Drishti Yadav, Himanshu Gupta, Om Prakash Verma, Irshad Ahmad Ansari, and Chang Wook Ahn. "A novel yolov3 algorithm-based deep learning approach for waste segregation: Towards smart waste management." Electronics 10, no. 1 30 Nov 2020: 14.

[23]. Sharholy, Mufeed, Kafeel Ahmad, Gauhar Mahmood, and R. C. Trivedi. "Municipal solid waste management in Indian cities–A review." Waste management 28, no. 2 Apr-2008: 459-467.

[24]. Jean de Dieu Nyandwi A Typical Convolutional Neural Network (CNN) Architecture. Linkedin.com

COPY OF THE PAPER PROJECT PHASE 2 SUBMITTED TO : Journal of The Institution of Engineers (India) : Series B

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]

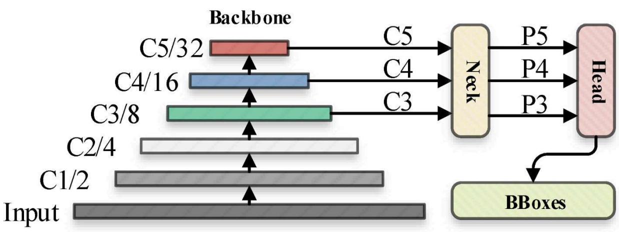

Fig 2 : Default Inference Chart of YOLOv5

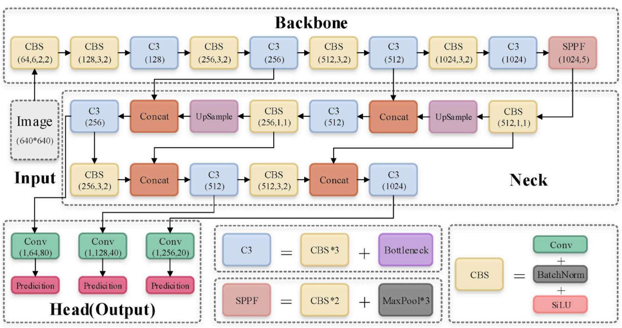

Fig 3 : Default Network Structure of YOLOv5

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

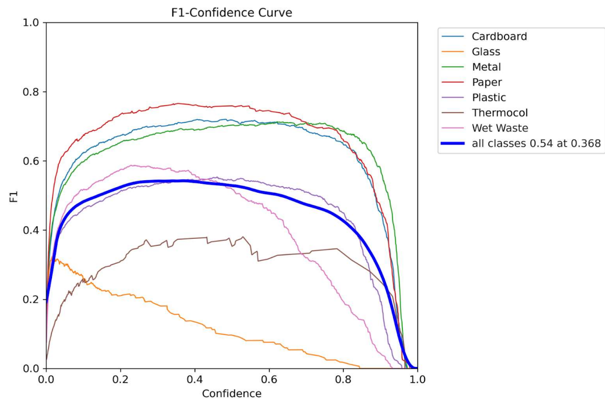

Fig 6 : F1-Confidence Curve

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.

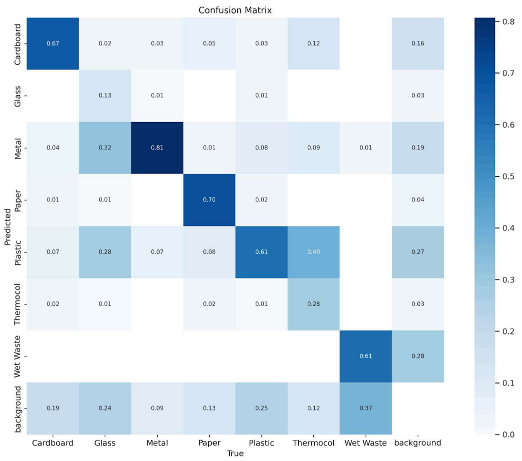

Fig 7 : Confusion Matrix

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.