Computer Vision Approaches for Waste Segregation: A Survey of Methods, Datasets, and Challenges

A M Sanjeev

M.Tech. Student, Department of Computer Science and Engineering, Gurukula Kangri Deemed to be University, Haridwar, India

Abstract
Municipal solid waste management has emerged as a critical environmental challenge due to rapid urbanization and increasing waste generation. Manual waste segregation is often inefficient, labour-intensive, and prone to inconsistency, leading to growing interest in automated approaches. In recent years, computer vision and machine learning techniques have been increasingly explored for image-based waste segregation, enabling the classification of waste materials using visual information.
This paper presents a survey of existing computer vision-based approaches for waste segregation reported in the literature. The review covers the evolution of image-based waste classification methods, including early feature-based techniques, classical machine learning models, and recent deep learning-based approaches. Publicly used datasets and commonly adopted evaluation practices in academic studies are also discussed at a high level.
In addition to summarizing existing methods, this survey highlights key challenges identified across prior studies, such as dataset limitations, environmental variability, class imbalance, and generalization issues in real-world scenarios. Based on the reviewed literature, open research gaps and future directions are outlined to guide researchers entering this domain. This survey focuses on synthesizing existing literature and identifying open challenges in image-based waste segregation.

Keywords: Waste Segregation, Computer Vision, Machine Learning, Deep Learning, Municipal Solid Waste, Image Based Classification

1. Introduction

1.1 Global Municipal Solid Waste Challenge

Municipal solid waste generation has increased significantly over the past few decades due to rapid urbanization, population growth, and changes in consumption patterns. Urban areas across the world are facing growing pressure to manage diverse waste streams in an environmentally sustainable and economically feasible manner. Inadequate waste segregation at the source often leads to inefficient recycling, increased landfill dependency, environmental pollution, and public health concerns. Developing countries face additional challenges due to limited infrastructure, high variability in waste composition, and resource constraints. As a result, effective waste segregation has become a critical component of sustainable municipal solid waste management strategies.

1.2 Limitations of Manual Waste Segregation

Traditionally, waste segregation has relied heavily on manual sorting processes carried out at households, collection points, or waste processing facilities. However, manual segregation is labor-intensive, time-consuming, and often inconsistent due to human fatigue, lack of awareness, and subjective judgment. The accuracy of manual sorting can vary significantly depending on working conditions, training levels, and the complexity of mixed waste streams. Furthermore, manual handling of waste poses occupational health risks and may not scale effectively with increasing waste volumes in urban environments. These limitations have motivated researchers and practitioners to explore automated and semi-automated approaches for improving waste segregation efficiency and reliability.

1.3 Emergence of Automation and Artificial Intelligence

Advancements in automation and artificial intelligence (AI) have opened new possibilities for addressing challenges in municipal solid waste management. Machine learning techniques enable systems to learn patterns from data and make predictions without explicit rule-based programming. In the context of waste management, AI-driven approaches have been investigated for tasks such as waste identification, classification, and monitoring. The increasing availability of digital imaging devices and computational resources has further facilitated the application of data-driven methods. As a result, research interest has grown in leveraging AI-based techniques to support automated waste segregation and enhance the effectiveness of existing waste management processes.

1.4 Role of Computer Vision in Waste Segregation

Computer vision has emerged as a key enabling technology for automated waste segregation by allowing machines to interpret and analyze visual information from images. Image-based waste classification leverages visual characteristics such as shape, color, texture, and material appearance to distinguish between different waste categories. The integration of machine learning techniques with computer vision has enabled data-driven models to learn discriminative features directly from waste images. In recent years, deep learning-based computer vision approaches, particularly convolutional neural networks, have demonstrated promising performance in academic studies under controlled conditions. Consequently, computer vision has gained significant attention as a potential tool for improving the accuracy and consistency of waste segregation processes compared to purely manual methods.

1.5 Need for Systematic Literature Synthesis

The growing body of research on image-based waste segregation has resulted in a diverse range of methodologies, datasets, and evaluation practices. While numerous studies reports encouraging results, variations in experimental settings, dataset characteristics, and methodological assumptions make direct comparison difficult. In addition, existing research often focuses on specific waste categories or controlled environments, which may limit the generalizability of reported findings. As a result, there is a need for a systematic synthesis of existing literature to organize current research, identify common trends, and critically examine reported challenges. A structured survey can help consolidate fragmented findings and provide clarity on the current state of computer vision-based waste segregation research.

1.6 Objective of the Survey

The primary objective of this survey is to systematically review existing computer vision-based approaches for waste segregation, classify methodologies adopted in prior studies, examine commonly used datasets, and highlight key challenges and open research gaps reported in the literature. By synthesizing existing research rather than proposing new models or systems, this survey aims to provide a consolidated reference for researchers seeking to understand current developments and limitations in image-based waste segregation.

2. Theoretical Foundations of Image-Based Waste Classification

This section outlines the theoretical concepts that underpin image-based waste classification approaches reported in existing literature. Understanding these foundations is essential for interpreting how computer vision and machine learning techniques are applied to waste segregation tasks. Rather than focusing on specific implementations, this section discusses general principles related to pattern recognition, supervised learning, and feature representation that form the basis of most image-based waste classification studies.

2.1 Pattern Recognition and Visual Classification

Pattern recognition is a fundamental concept in computer vision that involves identifying regularities and structures within visual data. In the context of image-based waste segregation, pattern recognition refers to the process of distinguishing different waste categories based on observable visual characteristics present in images. Waste materials such as plastic, paper, metal, and glass often exhibit distinct visual patterns, which can be leveraged for classification purposes.

Visual features play a crucial role in representing image content in a form suitable for computational analysis. Commonly used visual features in earlier studies include color information, texture descriptors, and shape-related attributes. These features are extracted from images to capture discriminative properties that help differentiate between waste types. For example, color distributions may assist in identifying certain materials, while texture patterns can provide cues about surface properties.

Object categorization is another key aspect of visual classification, where images are assigned to predefined categories based on learned patterns. In waste segregation, this typically involves assigning an input image to a waste class such as plastic, paper, or organic material. The effectiveness of object categorization depends on how well the extracted features represent the visual variability within each category while maintaining separation between different categories.

Inter-class similarity poses a significant challenge in waste classification tasks. Many waste materials share overlapping visual characteristics, leading to ambiguity during classification. For instance, certain plastics and metals may appear visually similar under specific lighting conditions. This overlap highlights the importance of robust feature representation and learning mechanisms capable of handling visual similarities and intra-class variations commonly observed in waste images.

2.2 Supervised Learning Paradigm

Most image-based waste classification approaches reported in the literature adopt a supervised learning paradigm. In supervised learning, models are trained using labelled datasets, where each input image is associated with a known waste category. The availability of labelled data enables learning algorithms to establish relationships between visual patterns and corresponding class labels.

The labelled data concept is central to supervised learning, as the quality and representativeness of labels directly influence model performance. In waste segregation studies, labelled datasets typically consist of images annotated according to waste material types. These labels serve as reference points during training, allowing models to learn decision boundaries that separate different waste categories based on visual cues.

Training and evaluation are conceptually distinct phases in the supervised learning process. During training, learning algorithms adjust their internal parameters to minimize classification errors on the training data. Evaluation is performed using separate data that were not seen during training, in order to assess how well the learned patterns generalize to new images. Although specific evaluation strategies vary across studies, the general objective is to measure classification consistency and robustness under varying conditions.

Supervised learning approaches are widely adopted due to their interpretability and effectiveness when sufficient labelled data are available. However, their performance is often influenced by factors such as dataset size, class balance, and visual diversity, which are frequently discussed as challenges in the context of waste image classification.

2.3 Feature Learning vs Representation Learning

Feature representation is a critical component of image-based classification systems, as it determines how visual information is encoded for learning algorithms. Earlier waste classification studies commonly relied on handcrafted features, which are manually designed descriptors intended to capture specific visual properties of images. Examples include color histograms, edge detectors, and texture descriptors. These features are typically selected based on domain knowledge and are used in combination with classical machine learning classifiers.

While handcrafted features can be effective in controlled settings, they often struggle to capture complex visual variations present in real-world waste images. Changes in lighting, background clutter, occlusion, and material deformation can significantly affect the reliability of manually designed features. As a result, the limitations of handcrafted feature extraction have been widely discussed in the literature.

Representation learning, particularly through deep learning techniques, has emerged as an alternative approach that enables models to automatically learn feature representations directly from data. Convolutional neural networks are commonly used for this purpose, as they are capable of learning hierarchical feature representations from raw image inputs. Lower layers typically capture simple visual patterns, while higher layers learn more abstract and task-specific representations.

The shift from handcrafted feature extraction to representation learning reflects a broader trend in computer vision research. In waste segregation studies, this transition has allowed models to adapt more flexibly to complex visual patterns without requiring explicit feature design. However, representation learning approaches also introduce new challenges, such as increased computational requirements and dependence on large, diverse datasets, which are further discussed in later sections of this survey.

Figure 1. Overview of Image-Based Waste Classification

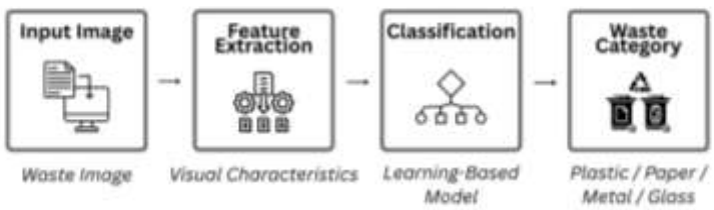

Figure 1 illustrates a high-level conceptual view of the image-based waste classification process commonly described in the literature. An input image is first represented through feature extraction, followed by a classification stage that assigns the image to a corresponding waste category. This conceptual illustration is intended to explain the general workflow at an abstract level and does not represent a specific system or implementation.

3. Evolution of Image-Based Waste Segregation Research

Research on image-based waste segregation has evolved significantly over the past decade, reflecting broader advances in computer vision and machine learning. Early studies primarily relied on handcrafted visual descriptors combined with traditional classifiers, followed by the adoption of classical machine learning models with engineered features. More recent work has increasingly shifted toward deep learning approaches, particularly convolutional neural networks, which enable end-to-end feature learning. This section outlines the major phases in the evolution of image-based waste segregation research and highlights the motivations underlying each transition.

3.1 Early Feature-Based Approaches

Initial research on image-based waste and material classification largely focused on handcrafted visual features designed to capture specific image characteristics. During this phase, features related to color, texture, and shape were commonly extracted from images and used to represent waste materials in a form suitable for classification. Examples of such descriptors include color moments, texture histograms, local binary patterns, co-occurrence matrices, and shape-based geometric features.

These handcrafted feature representations were motivated by domain knowledge and visual intuition regarding material appearance. In controlled experimental settings, such features demonstrated reasonable effectiveness in distinguishing between visually distinct material categories. As a result, early waste-related classification pipelines often followed a sequential process in which feature extraction was performed explicitly, followed by classification using a separate learning algorithm.

However, the reliance on manually designed features introduced several limitations. Handcrafted descriptors were often sensitive to variations in illumination, scale, viewpoint, background clutter, and image quality. Waste images captured in real-world environments frequently exhibit such variability, leading to reduced robustness and generalization performance. In addition, designing effective feature sets required substantial domain expertise and iterative tuning. These challenges highlighted the need for more adaptable approaches capable of handling the visual diversity inherent in waste images.

3.2 Classical Machine Learning Methods

Building upon early feature-based representations, subsequent studies increasingly adopted classical machine learning algorithms to improve classification performance. Between approximately 2014 and 2019, methods such as support vector machines, k-nearest neighbors, and random forest classifiers were widely applied to image-based waste segregation tasks. In these approaches, engineered visual features served as inputs to learning algorithms that aimed to model decision boundaries between different waste categories.

Classical machine learning methods offered advantages over purely rule-based systems by enabling data-driven classification and improved adaptability to complex feature spaces. Support vector machines were frequently employed due to their effectiveness in high-dimensional settings, while ensemble-based methods such as random forests provided robustness against noise and overfitting. k-nearest neighbor classifiers were also explored for their simplicity and interpretability, particularly in smaller datasets.

Despite achieving strong performance on moderate-scale and well-curated datasets, classical machine learning approaches remained dependent on feature engineering. As datasets grew in size and complexity, challenges related to scalability, computational cost, and feature selection became increasingly apparent. Training and tuning classical models for large and diverse datasets required careful parameter optimization, and inference efficiency could degrade as data volume increased. These limitations, combined with the growing availability of larger image datasets, motivated the exploration of feature learning approaches that could reduce reliance on manual feature design.

3.3 Deep Learning-Based Approaches

The emergence of deep learning marked a significant shift in image-based waste segregation research. From around 2018 onward, convolutional neural networks became the dominant approach for waste image classification, largely replacing traditional pipelines based on handcrafted features and classical classifiers. CNNs enable hierarchical feature learning directly from raw image data, allowing models to automatically discover discriminative visual patterns relevant to waste categorization.

Early CNN-based studies explored custom or shallow network architectures tailored to waste classification tasks. As research progressed, transfer learning emerged as a prevalent strategy, wherein CNNs pre-trained on large-scale image datasets were fine-tuned for waste classification. Transfer learning reduced the need for large task-specific datasets and improved convergence and performance, particularly in settings with limited labelled data.

Recent studies increasingly employ deeper architectures and ensemble-based strategies to enhance classification robustness under controlled conditions. While deep learning approaches have demonstrated clear performance gains compared to earlier methods in many reported studies, their success is often linked to curated datasets and constrained imaging environments. Nevertheless, the shift toward CNN-based feature learning reflects a broader trend in computer vision research and represents the current state of the art in image-based waste segregation.

Figure 2. Evolution of Image-Based Waste Classification Techniques

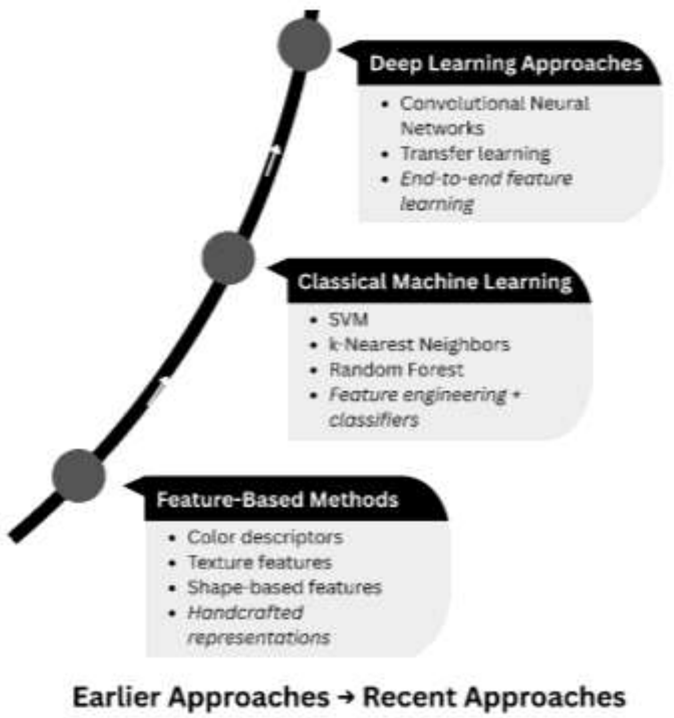

Figure 2 illustrates the conceptual evolution of image-based waste classification techniques over time. Early approaches relied on handcrafted color, texture, and shape features, followed by the adoption of classical machine learning classifiers using engineered features. More recent research has shifted toward deep learning-based methods, particularly convolutional neural networks, and transfer learning, enabling end-to-end feature learning and improved adaptability.

4. Literature Classification of Existing Approaches

The literature on image-based waste segregation can be broadly categorized based on the underlying learning paradigm and the degree of automation in feature extraction. Existing studies generally fall into three groups: classical machine learning-based approaches relying on engineered features, deep learning-based approaches employing convolutional neural networks, and hybrid or comparative studies that analyze trade-offs between different methodologies. This section organizes and synthesizes representative works within each category to highlight methodological trends, strengths, and limitations reported in prior research.

4.1 Classical Machine Learning-Based Studies

Early image-based waste segregation studies predominantly employed classical machine learning algorithms in combination with handcrafted or engineered visual features. In these approaches, feature extraction and classification were treated as separate stages. Commonly used features included color descriptors, texture measures such as local binary patterns and histogram-based representations, and keypoint-based descriptors such as SIFT. These features were then provided as inputs to classifiers including support vector machines, k-nearest neighbors, and random forest models.

A widely referenced benchmark in this category is the TrashNet dataset, which introduced a six-class waste categorization problem and enabled systematic evaluation of image-based approaches under controlled conditions. Many studies during this period either directly used TrashNet or constructed datasets with similar characteristics, typically consisting of single waste objects captured against clean backgrounds with limited variation in lighting and viewpoint. In addition to TrashNet, several works relied on small, domain-specific datasets tailored to particular waste categories, such as recyclable materials or plastic items.

Classical machine learning methods demonstrated strong performance on moderate-scale datasets, particularly when feature representations were carefully designed or when deep features extracted from pre-trained networks were combined with shallow classifiers such as SVMs or ensemble models. Among classical approaches, SVMs were frequently reported as the most robust and consistent classifier, while k-nearest neighbor methods were more sensitive to noise and feature dimensionality. Random forest and ensemble-based models provided competitive performance and improved robustness in some scenarios.

Despite these advantages, classical machine learning-based pipelines exhibited notable limitations. Their effectiveness was highly dependent on feature engineering, and performance often degraded under real-world conditions involving cluttered scenes, mixed waste, or varying illumination. Scalability also emerged as a concern as dataset size and class diversity increased. These limitations motivated the gradual transition toward end-to-end learning approaches that could reduce reliance on manual feature design.

4.2 Deep Learning-Based Studies

Recent research on image-based waste segregation has increasingly adopted deep learning techniques, particularly convolutional neural networks, as the primary modeling framework. Unlike classical pipelines, CNN-based approaches enable automatic feature learning directly from raw image data, allowing hierarchical representations to be optimized jointly with the classification objective.

A wide range of CNN architectures has been explored in the literature, including both custom-designed networks and widely used backbone models originally developed for large-scale image recognition tasks.

Commonly reported architectures include VGG, ResNet, DenseNet, MobileNet, Inception, and EfficientNet variants. In many studies, these architectures are not trained from scratch but instead fine-tuned using transfer learning from models pre-trained on large benchmark datasets.

Transfer learning has become a dominant strategy in waste image classification due to its ability to achieve strong performance with relatively limited labelled data. By leveraging pre-trained representations, researchers have reduced training time and improved convergence while mitigating overfitting on small datasets. In addition, lightweight CNN architectures have been increasingly favored in studies targeting real-time or resource-constrained deployment scenarios.

Although TrashNet remains a frequently used benchmark, recent works have also introduced larger and more diverse datasets to better reflect real-world conditions. Examples include datasets with increased class diversity, varied backgrounds, and multi-object scenes. Methodological trends in deep learning-based studies include extensive data augmentation, architectural optimization for efficiency, and the exploration of attention mechanisms and hybrid network designs to improve robustness. Overall, the literature indicates a clear shift toward deep learning as the prevailing approach for image-based waste segregation.

4.3 Hybrid and Comparative Studies

In addition to purely classical or deep learning-based approaches, several studies have conducted explicit comparisons between machine learning and deep learning methods or proposed hybrid pipelines that combine elements of both paradigms. Comparative studies typically evaluate classical classifiers such as SVMs and random forests alongside CNN-based models on the same image datasets, enabling direct assessment of trade-offs.

Across these comparisons, deep learning approaches are generally reported to provide superior robustness and adaptability, particularly for multi-class problems and images captured under variable environmental conditions. However, classical machine learning models remain competitive in constrained settings, especially when datasets are small, classes are visually distinct, or computational resources are limited. Hybrid pipelines that use CNNs as feature extractors followed by classical classifiers have also been explored as a compromise between performance and complexity.

Survey and review papers consistently emphasize that the choice between classical machine learning and deep learning depends on application context, data availability, and deployment constraints. While CNN-based models dominate current research, classical and hybrid approaches continue to be relevant for proof-of-concept systems, edge deployment, and scenarios where interpretability or low computational overhead is prioritized.

Table 1. Summary of Representative Studies on Image-Based Waste Segregation

Study	Method Type	Dataset Used	Key Observation
Thung & Yang	Classical ML (SVM)	TrashNet	Established benchmark for image-based recyclable waste sorting
Costa et al.	ML and DL comparison	TrashNet-like	Classical classifiers remain competitive on small datasets

Gruber et al.	Classical ML ensembles	Custom plastic	Engineered features effective for specialized waste types
Zhang et al.	Deep learning (CNN)	TrashNet	CNN-based models improve robustness under controlled settings
Poudel et al.	Deep learning (TL-CNN)	Waste images	Transfer learning reduces dependence on large datasets
Nahiduzzaman et al.	Hybrid ML + DL	Multi-source	Hybrid pipelines balance efficiency and classification depth
Vukicevic et al.	DL with preprocessing	Multiple sets	End-to-end pipelines adaptable across varied waste scenarios

5. Publicly Used Datasets in Literature

Publicly available image datasets play a crucial role in the development and evaluation of computer vision-based waste segregation methods. These datasets enable benchmarking, comparative analysis of algorithms, and reproducibility of research findings. However, the existing literature also highlights several limitations associated with commonly used datasets, particularly with respect to realism and generalization. This section summarizes widely used datasets reported in prior studies and discusses their key limitations as identified in review and survey papers.

5.1 Common Benchmark Datasets

TrashNet

TrashNet is one of the most widely used benchmark datasets in image-based waste classification research. Introduced as a six-class waste categorization dataset covering common recyclable and non-recyclable materials, it has been repeatedly employed in both classical machine learning and deep learning studies. Numerous papers use TrashNet to compare feature extraction techniques, classifier performance, and deep learning architectures under controlled experimental conditions.

The dataset's popularity stems from its public availability, simple class structure, and suitability for rapid experimentation. As a result, TrashNet has often served as a reference benchmark in survey and comparative studies. Several works further extend TrashNet by augmenting it with additional locally collected or web-sourced images, reinforcing its role as a foundational dataset rather than a comprehensive representation of real-world waste scenarios.

General-Purpose Image Datasets and Waste-Related Classes

In addition to waste-specific datasets, general-purpose image datasets are frequently used in waste classification research, particularly for transfer learning and background modeling. Large-scale datasets such as ImageNet are commonly employed for pretraining convolutional neural networks, enabling models to leverage generic visual representations before fine-tuning on waste images.

While datasets such as Open Images are not always explicitly named in waste classification studies, the broader practice of reusing large, non-waste image collections is well established. These datasets are often used to provide background or non-waste images, negative samples, or auxiliary training data to improve robustness. The literature also reports a growing trend of merging multiple public datasets into hybrid benchmarks to increase coverage and diversity, especially for detection-oriented tasks.

5.2 Limitations of Existing Datasets

Despite their widespread use, existing waste image datasets exhibit several limitations that affect the generalizability and real-world applicability of trained models. Survey and review studies consistently emphasize that many datasets do not adequately reflect the complexity of real-world waste environments.

One major limitation is the prevalence of controlled acquisition conditions. Many datasets are collected under uniform lighting and clean background settings, with waste objects placed in isolation. Such conditions simplify the classification task but fail to capture environmental variability encountered in practical deployments, including cluttered scenes, variable illumination, occlusion, and object deformation.

Another commonly reported issue is the dominance of single-object, single-label images. A large proportion of datasets focus on classifying one waste item per image, with limited support for multi-object scenes, instance-level annotations, or segmentation masks. This restricts the applicability of trained models to real-world scenarios where multiple waste items are often present simultaneously.

Limited category diversity and geographic coverage further constrain dataset usefulness. Many datasets contain a small number of waste categories, typically focusing on broad material types, despite the need for more fine-grained classification in practical waste management systems. In addition, the geographic origin of dataset images is often narrow or undocumented, raising concerns about regional bias and reduced cross-location generalization.

Finally, dataset imbalance, small scale, and lack of standardization are recurring challenges. Review papers highlight inconsistencies in labeling schemes, annotation formats, and category definitions across datasets, making cross-study comparison difficult. These factors collectively contribute to model bias and reduced robustness when systems trained on existing datasets are applied to uncontrolled environments.

Table 2. Overview of Commonly Used Waste Image Datasets

Dataset Name	Waste Types	Common Usage
TrashNet	Recyclable and non-recyclable materials	Benchmark classification and model comparison
Open Images*	Waste-related and background objects	Transfer learning and background modeling
ImageNet*	Generic object categories	Pretraining CNN backbones for waste classification
Hybrid sets	Mixed waste categories	Dataset augmentation and diversity enhancement

*General-purpose datasets are typically reused for pretraining or background classes rather than as dedicated waste datasets.

6. Comparative Discussion of Existing Approaches

A comparative analysis of existing image-based waste segregation approaches reveals distinct strengths and limitations across different methodological paradigms. Rather than identifying a universally superior technique, the literature highlights trade-offs related to robustness, scalability, generalization, and adaptability to urban variability. This section critically discusses these dimensions by comparing classical machine learning, deep learning, and hybrid approaches as reported in prior studies.

6.1 Robustness to Environmental Variability

Classical machine learning-based approaches and deep learning-based methods differ significantly in their robustness to environmental variability. Studies employing handcrafted color, texture, and shape features combined with classifiers such as support vector machines have demonstrated reliable performance under controlled conditions with uniform backgrounds and consistent lighting. However, multiple surveys note that these approaches are sensitive to changes in illumination, occlusion, and background clutter, leading to degraded performance in less controlled environments.

In contrast, deep learning-based approaches, particularly convolutional neural networks, are generally reported to be more robust to visual variability due to their ability to learn hierarchical feature representations directly from data. Comparative studies indicate that CNN-based models better handle variations in object appearance and background complexity than classical pipelines relying on fixed feature descriptors. Hybrid approaches, such as CNN feature extraction followed by shallow classifiers, are sometimes reported to improve robustness compared to purely handcrafted pipelines, while maintaining simpler decision models. Overall, the literature suggests that deep learning approaches offer improved robustness, although their effectiveness remains influenced by dataset quality and diversity.

6.2 Scalability and Computational Considerations

Scalability is a key factor distinguishing classical machine learning and deep learning approaches in waste image classification. Classical methods, including SVMs and random forest models, are often described as computationally efficient when operating on compact feature representations and moderate dataset sizes. Survey papers report that such models can be trained and deployed with limited computational resources, making them attractive for early-stage systems or low-cost hardware platforms.

Deep learning approaches, on the other hand, are associated with higher computational requirements during training, particularly when using deep architectures or large-scale datasets. However, recent studies emphasize that once trained, CNN-based models can achieve efficient inference, especially when lightweight architectures or optimized implementations are employed. Hybrid pipelines attempt to balance these considerations by combining deep feature extraction with simpler classifiers. Comparative analyses consistently highlight that scalability depends not only on model complexity but also on dataset size, class diversity, and deployment constraints.

6.3 Generalization Across Datasets and Contexts

Generalization remains a central challenge across all image-based waste segregation approaches. Classical machine learning models trained on handcrafted features often generalize poorly when applied to images captured under conditions different from those seen during training. Reviews attribute this limitation to feature sensitivity and dataset bias, particularly when training data lacks diversity.

Deep learning models are frequently reported to exhibit improved generalization, especially when transfer learning and data augmentation are employed. Surveys comparing models trained from scratch with transfer-learned CNNs consistently note that pre-trained representations enhance performance on new datasets with limited labelled data. Nevertheless, several studies caution that even deep models may struggle to generalize across geographic regions or waste compositions when training datasets are narrowly sourced. Hybrid and ensemble approaches are sometimes proposed to mitigate generalization issues, but their effectiveness is still constrained by the underlying data distribution.

6.4 Adaptability to Urban Waste Scenarios

Urban waste environments introduce additional complexity due to mixed waste streams, diverse packaging materials, and dynamic acquisition conditions. Classical machine learning-based approaches are often reported to perform adequately in simplified urban scenarios, such as single-object classification or binary categorization tasks. However, their reliance on engineered features limits adaptability when waste scenes become cluttered or when multiple objects are present.

Deep learning-based approaches demonstrate greater adaptability in urban contexts by learning more abstract and discriminative visual representations. Comparative studies suggest that CNN-based models are better suited to handle complex urban waste imagery, particularly when trained on datasets incorporating varied backgrounds and object configurations. Hybrid approaches, including lightweight CNNs and CNN–classifier combinations, are increasingly explored to address practical deployment constraints in urban settings. Overall, the literature indicates that while deep learning methods show greater adaptability, their success in urban environments is closely tied to dataset realism and annotation quality.

7. Challenges and Research Gaps

Despite significant progress in image-based waste segregation research, existing studies consistently report a range of unresolved challenges that limit practical deployment and generalization. These challenges span technical, dataset-related, and real-world dimensions. This section synthesizes the key issues highlighted across prior studies and identifies persistent research gaps without proposing specific solutions.

7.1 Technical Challenges

One of the most frequently reported technical challenges in image-based waste segregation is sensitivity to lighting variations. Waste images captured in real-world environments often exhibit uneven illumination, shadows, reflections, and varying light sources, which can substantially alter visual appearance. Classical feature-based methods are particularly affected by such variations, while deep learning models also demonstrate degraded performance when training data does not adequately capture lighting diversity.

Occlusion represents another major challenge, especially in urban waste scenarios where objects overlap or are partially visible. Many existing models are trained on images containing isolated waste items, limiting their ability to correctly classify partially occluded objects. Prior studies note that occlusion complicates feature extraction and increases ambiguity between visually similar waste categories.

The presence of mixed waste within a single scene further complicates classification. Real-world waste streams frequently contain multiple materials in close proximity, yet much of the existing literature focuses on single-object classification. This mismatch between training conditions and deployment environments remains a significant gap in current research.

Class imbalance is also widely acknowledged as a technical challenge. Waste datasets often contain uneven representation across categories, with certain materials appearing far more frequently than others. Such imbalance can bias learned models toward dominant classes, reducing reliability for underrepresented waste types. Survey papers consistently identify class imbalance as a persistent issue across both classical and deep learning-based approaches.

7.2 Dataset-Related Challenges

Dataset bias is a recurring concern in image-based waste classification research. Many publicly available datasets are collected under controlled conditions with uniform backgrounds, fixed viewpoints, and limited environmental variation. As a result, models trained on these datasets may implicitly learn background or acquisition-specific cues rather than material-specific characteristics. This bias reduces the validity of reported performance when models are applied beyond the original dataset context.

Small-scale datasets further constrain model development and evaluation. Numerous studies rely on datasets consisting of a few hundred images per class, which limits statistical robustness and increases the risk of overfitting. While data augmentation and transfer learning are commonly used to mitigate these issues, the underlying lack of large, diverse, and standardized datasets remains an open gap identified across multiple reviews.

In addition, inconsistencies in labeling schemes, category definitions, and annotation formats across datasets complicate cross-study comparison. The absence of standardized taxonomies and evaluation protocols makes it difficult to objectively assess progress across the literature and hampers reproducibility.

7.3 Generalization and Real-World Constraints

Generalization across domains and deployment contexts is one of the most critical challenges identified in the literature. Models trained on data from specific regions, waste compositions, or acquisition setups often struggle when applied to new environments. This domain shift arises from differences in packaging materials, waste handling practices, cultural consumption patterns, and environmental conditions.

Environmental variability further exacerbates generalization challenges. Urban waste images may be captured under diverse weather conditions, backgrounds, and camera perspectives, introducing noise and uncertainty not represented in training datasets. Review studies consistently note that performance reported under laboratory or controlled conditions does not reliably translate to real-world settings.

These factors collectively highlight a gap between academic evaluation and practical deployment. While existing research demonstrates the feasibility of image-based waste segregation, the literature acknowledges that current models often lack the robustness and adaptability required for consistent real-world performance. Addressing these gaps remains an open research challenge for the field.

8. Future Research Directions

Based on the synthesis of existing literature, several broad research directions emerge that may help advance image-based waste segregation research while addressing the limitations identified in prior studies. These directions are framed at a high level and focus on structural and methodological needs rather than specific technical solutions.

8.1 Need for Standardized Benchmarks and Evaluation Protocols

One recurring issue highlighted across survey and review studies is the lack of standardized benchmarks for image-based waste segregation. Existing research relies on heterogeneous datasets, varying class definitions, and inconsistent evaluation practices, which complicates objective comparison across studies. Future research would benefit from the development and wider adoption of standardized benchmark datasets and evaluation protocols that reflect realistic waste scenarios. Such standardization could improve reproducibility, facilitate fair comparison of methods, and provide clearer indicators of progress within the field.

8.2 Need for More Diverse and Representative Datasets

The literature consistently emphasizes the importance of dataset diversity for improving model robustness and generalization. Future efforts should prioritize datasets that capture greater variability in waste appearance, backgrounds, lighting conditions, and geographic contexts. Increased diversity in waste categories, including fine-grained material distinctions and mixed-waste scenes, is also identified as an important direction. Addressing these dataset limitations at a community level is widely regarded as a prerequisite for advancing real-world applicability of image-based waste classification research.

8.3 Importance of Interdisciplinary Collaboration

Image-based waste segregation lies at the intersection of computer vision, environmental science, urban planning, and waste management policy. Several reviews note that closer interdisciplinary collaboration could help align technical research with practical waste management needs. Engagement with domain experts may support more realistic problem formulations, dataset design, and evaluation criteria that reflect operational constraints. Such collaboration is viewed as an essential component for translating advances in image-based classification research into broader environmental and societal impact.

8.4 Bridging the Gap Between Academic Research and Practice

A broader direction identified in the literature is the need to better align academic research objectives with practical deployment considerations. While many studies demonstrate strong performance under controlled experimental conditions, fewer address the complexities of real-world waste management environments. Future research directions emphasize the importance of aligning datasets, evaluation settings, and problem scope with realistic operational contexts, while maintaining methodological rigor and transparency.

9. Conclusion

Image-based waste segregation has emerged as an important research direction within the broader domain of sustainable municipal solid waste management. This survey reviewed and organized existing literature on computer vision–based approaches for waste classification, tracing the evolution of methods from early feature-based techniques to classical machine learning models and, more recently, deep learning–based approaches. By synthesizing prior studies, this paper provides a consolidated overview of commonly adopted methodologies, datasets, and evaluation practices reported in the literature.

The review highlights that early research relied heavily on handcrafted visual features combined with traditional classifiers, achieving reasonable performance under controlled conditions. Subsequent studies increasingly adopted deep learning models, particularly convolutional neural networks and transfer learning, which reduced the need for manual feature engineering and improved robustness in many experimental settings. At the same time, the analysis shows that much of the existing work remains constrained by limited and simplified datasets, narrow evaluation scenarios, and variability in experimental design, making cross-study comparison and real-world generalization challenging.

Through comparative discussion, this survey emphasizes recurring trade-offs between classical and deep learning approaches in terms of robustness, scalability, generalization, and computational requirements. The analysis further identifies persistent technical and dataset-related challenges, including sensitivity to environmental variability, class imbalance, and domain shift, which continue to limit deployment in realistic waste management environments.

Overall, this paper contributes a structured synthesis of image-based waste segregation research, clarifying the current state of the field and the limitations that remain unresolved. This survey aims to provide a structured understanding of image-based waste segregation research and highlight open challenges for future investigation.

References

1. Thung G., Yang M., "Classification of Trash for Recyclability Status", Stanford University CS229 Project Report, 2016, 1–8.
2. Costa B.S., Pinto A.C., Sousa A.J., "Artificial intelligence in automated sorting in trash recycling", Anais do Encontro Nacional de Inteligência Artificial e Computacional, 2018, 1–6.
3. Flores M.G., Garcia N.M., Paredes H., "Literature review of automated waste segregation system using machine learning: A comprehensive analysis", International Journal of Simulation: Systems, Science and Technology, 2019, 20 (2), 1–10.
4. Sakr G.E., Mokbel M., Darwich A., "Comparing deep learning and classical machine learning approaches for waste classification", International Conference on Innovations in Computing and Technology, 2016, 1–6.
5. Gruber F., Woerl A., Schmidt V., "Classification of black plastics waste using fluorescence imaging and machine learning", Recycling, 2019, 4 (4), 1–16.
6. Adedeji O., Wang Z., "Intelligent waste classification system using deep learning convolutional neural network", Procedia Manufacturing, 2019, 35, 607–612.
7. Ozkan A., Ozkaya U., Seyfi L., "Fine-tuning of pre-trained convolutional neural networks for waste classification", International Journal of Engineering Research and Technology, 2019, 12 (8), 1211–1216.
8. Zhang Q., Yang J., Li X., "Waste image classification based on transfer learning and convolutional neural network", Waste Management, 2021, 120, 1–10.
9. Mao W.L., Wang J., Chen Y., "Recycling waste classification using optimized convolutional neural network", Resources, Conservation and Recycling, 2021, 164, 105–119.
10. Shi C., Zhang Y., Wang J., "A multilayer hybrid convolutional neural network for waste classification", Applied Sciences, 2021, 11 (5), 1–15.
11. Huang G., Liu Z., Van Der Maaten L., "Densely connected convolutional networks", IEEE Conference on Computer Vision and Pattern Recognition, 2017, 4700–4708.
12. He K., Zhang X., Ren S., Sun J., "Deep residual learning for image recognition", IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770–778.
13. Sandler M., Howard A., Zhu M., "MobileNetV2: Inverted residuals and linear bottlenecks", IEEE Conference on Computer Vision and Pattern Recognition, 2018, 4510–4520.
14. Tan M., Le Q., "EfficientNet: Rethinking model scaling for convolutional neural networks", International Conference on Machine Learning, 2019, 6105–6114.
15. Poudel S., Poudyal P., "Classification of waste materials using CNN based on transfer learning", Proceedings of the Annual Meeting of the FTRA, 2022, 1–6.
16. Li J., Zhang H., Liu J., "Automatic detection and classification system of domestic waste via multimodel cascaded convolutional neural network", IEEE Transactions on Industrial Informatics, 2021, 17 (6), 4164–4173.
17. Nahiduzzaman M., Islam M., Hasan M., "Hybrid deep learning framework for multi-class waste image classification", Expert Systems with Applications, 2025, 241, 122–135.
18. Ramos A., Santos R., Silva F., "Machine learning and deep learning approaches for municipal solid waste classification: A systematic review", Journal of Cleaner Production, 2024, 384, 135–150.
19. Islam M., Rahman M., Hasan M., "A systematic review of machine learning approaches for plastic waste classification", Waste Management, 2023, 158, 80–95.

20. Bianco S., Cadene R., Celona L., "Benchmark analysis of representative deep neural network architectures", IEEE Access, 2018, 6, 64270–64277.
21. Everingham M., Van Gool L., Williams C.K.I., "The Pascal Visual Object Classes challenge: A retrospective", International Journal of Computer Vision, 2015, 111 (1), 98–136.
22. Cordts M., Omran M., Ramos S., "The Cityscapes dataset for semantic urban scene understanding", IEEE Conference on Computer Vision and Pattern Recognition, 2016, 3213–3223.
23. Lin T.Y., Maire M., Belongie S., "Microsoft COCO: Common objects in context", European Conference on Computer Vision, 2014, 740–755.
24. World Bank, "What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050", World Bank Publications, 2018. https://openknowledge.worldbank.org/handle/10986/30317
25. Mohammadi M., Al-Fuqaha A., Sorour S., "Deep learning for IoT big data and streaming analytics: A survey", IEEE Communications Magazine, 2018, 56 (4), 26–33.