

AI-Powered Smart Waste Segregation Assistant: A Lightweight, Browser Based Deep Learning System for Real-Time Waste Classification

E Arvind Elumalai Udiyar, Dr Manoj Singh

Roll No : SMCS2627110

Department of Computer Science, Sies College of Arts, Science and Commerce (Empowered Autonomous), Mumbai, India

Arvindfms2526110@ascs.sies.edu.in

Department of Computer Science, Sies College of Arts, Science and Commerce (Empowered Autonomous), Mumbai, India

manojks@sies.edu.in

Abstract:

Improper segregation of municipal solid waste at the point of disposal is one of the principal reasons recycling rates remain low worldwide, since mixed streams contaminate otherwise recoverable materials. This paper presents the design and implementation of an AI-based Smart Waste Segregation Assistant: a fully client-side, browser hosted application that uses a pre-trained convolutional neural network (MobileNet, via TensorFlow.js) together with a rule-based keyword-to-category mapping engine to classify a live camera feed of a waste item into one of six actionable categories — Organic, Plastic, Paper, Glass, Metal, and Hazardous — and to recommend the correct disposal bin in real time. Unlike conveyor-belt or sensor-based industrial smart-bin systems that require dedicated hardware, cloud inference, or large custom-labelled datasets, the proposed assistant runs entirely on commodity devices (smartphones, tablets, laptops) with no installation and no image data leaving the user's device, and additionally allows an end-user to plug in a custom Teachable Machine model to raise classification accuracy without writing code. We review recent literature on CNN-based waste classification, edge-deployable models, IoT-integrated smart bins, and industrial-grade cluttered-scene datasets, and use the reported results to position the proposed system's trade-offs between accessibility, cost, privacy, and classification accuracy. Our analysis shows that while specialized, dataset

trained systems achieve 89-99% accuracy under controlled or industrial conditions, they demand substantially higher deployment cost and infrastructure; the proposed lightweight assistant instead targets household- and community-level assistive segregation and environmental education, complementing rather than replacing industrial sorting pipelines. Limitations, comparative benchmarking against the surveyed systems, and directions for improving classification accuracy through domain-specific transfer learning are discussed.

Keywords— waste segregation, deep learning, convolutional neural network, MobileNet, TensorFlow.js, smart bin, sustainable waste management, image classification, edge AI

I. INTRODUCTION

Global municipal solid waste (MSW) generation is projected to grow from roughly 2.1 billion tonnes in 2020 to 2.6 billion tonnes by 2030 [5], and in most developing regions a large share of this waste is still collected and processed without any prior source segregation [2]. Contamination of recyclable streams by non-conforming materials is repeatedly identified as the single largest obstacle preventing Materials Recovery Facilities (MRFs) from lifting recycling rates above 30-35% [7]. Two broad technological responses have emerged in the literature: (i) heavy, hardware-centric solutions that place cameras, actuators, and edge-AI accelerators directly on conveyor belts or smart bins inside sorting facilities [1], [4], [7]; and (ii) citizen-facing mobile or web applications intended to help residents sort waste correctly at the source, often paired with a physical smart bin [2]. Both directions depend on convolutional neural networks (CNNs) as the core recognition technology, but they differ sharply in

cost, infrastructure requirements, and accessibility.

This paper focuses on the second direction and presents a Smart Waste Segregation Assistant implemented as a single self-contained HTML/JavaScript web page. The assistant activates the device camera, runs a pre-trained MobileNet image classifier entirely in the browser using TensorFlow.js, predictions into one of six practical waste categories through a keyword-to-category mapping engine. The system further allows a user to load a custom model exported from Google's Teachable Machine, so that the generic recognizer can be replaced by a model trained specifically on the six target waste classes without any software installation or coding.

The main contributions of this work are: (1) a working, zero-installation, privacy-preserving reference implementation of an AI waste-classification assistant that runs on ordinary consumer hardware; (2) a structured review of seven recent works spanning edge-deployed CNNs, IoT-integrated smart bins, industrial-grade cluttered-scene datasets, and dataset-centric surveys, summarised for direct comparison; and (3) a discussion that positions the proposed lightweight, browser-based approach against the reviewed hardware- and dataset heavy systems, identifying the specific niche — assistive, educational, low-cost, household/community level segregation — that the proposed system is best suited for.

II. RELATED WORK

Automated waste classification research can broadly be grouped into (a) image classification on clean, single object datasets, (b) edge-deployable CNNs for smart bins, (c) IoT-integrated collection-and-sorting systems, and (d) industrial-grade detection/segmentation on cluttered conveyor-belt scenes. Table I summarises the systems reviewed in this paper along these lines.

Yang and Thung [6] created the widely used TrashNet dataset of about 2,527 hand-photographed images across six classes (glass, paper, cardboard, plastic, metal, trash) on a clean white background, and compared an SVM with SIFT-based bag-of-features against a shallow CNN, obtaining 63% and 22% accuracy respectively; the authors note that the CNN under-performed largely due to hyperparameter and data-scarcity issues rather than an inherent limitation of deep learning for this task. White et al. [1] later showed that a hybrid transfer-learning strategy (feature-extraction followed by discriminative gradual fine-tuning) applied to a DenseNet backbone — their WasteNet model — reaches 97.0% accuracy on the very same TrashNet classes and can run on a low-power Jetson Nano edge device, outperforming VGG16, ResNet, SqueezeNet and AlexNet baselines evaluated in the same study. Moving from single-object images toward physical smart-bin systems, Sinduja and Kumar [2] proposed a nine-step, human-centred framework combining a resident-facing mobile application with a solar-powered, camera-equipped self-segregating bin for the city of Bengaluru. Their image-classification prototype, trained on Google's Teachable Machine across three coarse classes (recyclable, organic, e-waste), reached an average accuracy of 97.22% and average precision of 95.8%, and the paper emphasises stakeholder communication, government-resident feedback loops, and citizen incentive mechanisms as being as important as the recognition model itself. Alourani et al. [4] extended this idea into a fully IoT-integrated intelligent Solid Waste Management system (iSSWMs) with ultrasonic fill-level sensors, a cloud/Firebase back end, a mobile alerting application, and a conveyor-belt actuator mechanism that physically sorts items into glass, plastic, and metal bins after VGG-19-based recognition; VGG 19, ResNet and MobileNetV2 backbones achieved 99.7%, 97.2% and 98.18% accuracy respectively on the TrashNet dataset.

A third line of work highlights that laboratory-grade accuracy does not transfer to real sorting-facility conditions. Bashkirova et al. [7] introduced ZeroWaste, the first densely annotated industrial-grade dataset recorded on a real paper-recycling conveyor belt, and showed that Mask R-CNN, TridentNet and DeepLabV3+ — all state-of-the-art on curated benchmarks — only reach 22-34% mAP for detection and 45-57% mIoU for segmentation when faced with severe clutter, occlusion, and deformable, translucent objects. Thielmann et al. [5] systematically reviewed this domain gap across four public industrial-grade datasets (ZeroWaste, ZeroWaste v2, WaRP, SpectralWaste), confirming that models scoring 87-95% on in-the-wild datasets such as TrashNet, TACO, and ReSORT-IT drop to roughly 24- 57% when evaluated on cluttered conveyor-belt imagery, and they call for foundation models, self-supervised learning, synthetic data, and multimodal (RGB + hyperspectral/NIR) fusion as future directions. Complementing this, Ali et al. [3] released DWSD, a 784-image, 14-class dense

segmentation dataset captured in unorganised Indian streets and dumpsites, and benchmarked DeepLabv3+, UNet, PSPNet and FPNNet, with DeepLabv3+ achieving the best test accuracy of 89.39%, again considerably below controlled-scene benchmarks.

Collectively, the reviewed literature indicates that (i) CNN-based waste recognition is highly accurate on clean, single-object images and can be compressed for edge deployment [1], [6]; (ii) full IoT smart-bin pipelines can reach high accuracy but require substantial hardware, cloud infrastructure and continuous maintenance [2], [4]; and (iii) performance degrades sharply once scenes involve real-world clutter, deformation, and occlusion, which remains an open research problem [3], [5], [7]. This motivates a complementary, low-cost, software only assistant aimed at the point of disposal rather than the sorting facility — the system proposed in this paper.

TABLE I. Summary of Reviewed Waste-Classification Systems

Ref. System / Dataset Technique Reported Performance [6] TrashNet (6 classes, clean bg.) SVM (SIFT+BoF) / CNN 63% (SVM) / 22% (CNN)

[1] WasteNet (Jetson Nano edge device) Hybrid-tuned DenseNet, transfer learning

[7] ZeroWaste (real MRF conveyor belt) Mask R-CNN, TridentNet, DeepLabV3+

97.0% accuracy on TrashNet

24.3% mAP (detection), 52.1% mIoU (segmentation)

[3] DWSD (Indian congested waste scenes) DeepLabv3+, UNet, PSPNet, FPNNet 89.4% test accuracy (DeepLabv3+)

[2] App + Smart Bin (Bengaluru) Teachable Machine image model 97.22% avg. accuracy, 95.8% avg. precision

[4] iSSWMs (IoT smart bins + conveyor) VGG-19 / ResNet / MobileNetV2 99.7% / 97.2% / 98.18% accuracy

[5]	Industrial-grade review (multi-dataset)	Survey of CNN detectors/segmenters	Lab data ~90-95% vs. industrial data ~24-57%
-----	---	------------------------------------	--

III. MOTIVATION AND PROBLEM STATEMENT

Two structural gaps emerge from the review in Section II. First, high-accuracy systems such as WasteNet [1] and iSSWMs [4] assume the availability of edge accelerators, conveyor mechanics, or cloud connectivity, which is unrealistic for an individual household, a school, or a small community bin cluster in a resource-constrained setting. Second, even where a mobile application is proposed [2], the accompanying physical smart bin, cloud database, and stakeholder communication layer add significant deployment cost and complexity. Meanwhile, a large fraction of contamination in the waste stream is reported to originate not from a lack of sorting infrastructure but from residents being simply unsure which bin an item belongs to [2], [7]. This suggests that a lightweight, instantly accessible, zero-installation assistive tool — one that any person can open in a mobile browser and point at an item — can meaningfully reduce source-level contamination even without physical automation. The problem addressed in this paper is therefore: can an accurate enough waste-category recommendation be produced using only an on-device, pre-trained general-purpose CNN and a browser, without server-side inference, specialised hardware, or a custom-labelled training set, while still leaving room to plug in a properly trained model later?

IV. PROPOSED SYSTEM ARCHITECTURE

The proposed assistant follows the four-stage pipeline shown in Fig. 1: (1) camera capture, (2) CNN-based generic object recognition, (3) keyword-to-category mapping, and (4) bin recommendation with a confidence aware fallback.

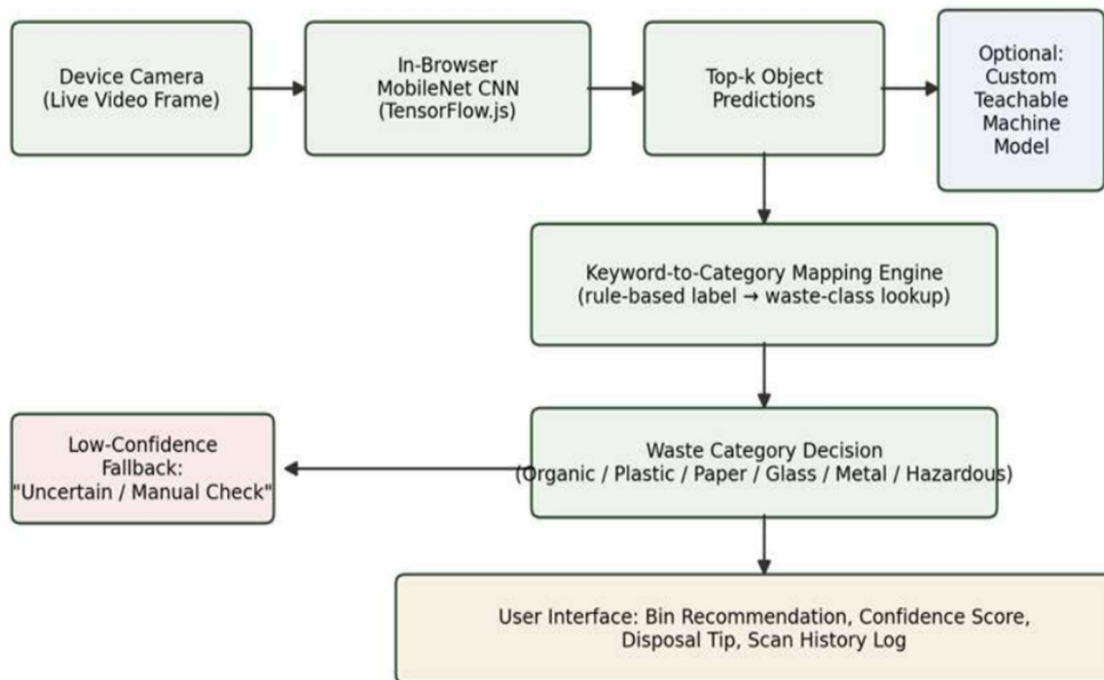


Fig. 1. System architecture of the proposed browser-based Smart Waste Segregation Assistant.

A. Camera Capture

The application requests access to the device camera through the browser's native media API and displays a live preview inside the page. A scanning-line animation gives the user visual feedback while a frame is being analysed, and no video or image data is transmitted off the device at any point in the pipeline.

B. Generic Object Recognition

A MobileNet model (v2.1.1), pre-trained on the 1,000- class ImageNet dataset, is loaded client-side using TensorFlow.js and executed directly in the browser's JavaScript engine (optionally GPU-accelerated via WebGL). For each captured frame the model returns the top-3 candidate object labels together with confidence scores, mirroring the feature-extraction stage used in transfer-learning pipelines such as WasteNet [1] and iSSWMs [4], but without any fine-tuning step, which keeps the model download small and the start-up time short.

C. Keyword-to-Category Mapping Engine

Because ImageNet labels (e.g., "water bottle", "pop bottle", "envelope", "banana") do not directly correspond to disposal categories, a rule-based mapping table translates each candidate label into one of six categories: Organic, Plastic, Paper, Glass, Metal, or Hazardous. This mirrors, in a lightweight rule-based form, the class consolidation step that industrial systems perform through supervised fine-tuning [2], [4], but trades classification depth for zero-training-data deployment.

D. Confidence-Aware Recommendation

If no mapped keyword clears a minimum confidence threshold, the system reports an "uncertain" result and prompts the user to retry with better lighting, a closer framing of a single object, or to switch to a custom trained model, rather than silently returning a low confidence guess — a design choice informed by the observation in [5] and [7] that industrial models are most useful when their failure cases are made explicit rather than hidden.

E. Extensibility via Teachable Machine

An "Advanced" panel lets a user paste the URL of a model they trained on Google's Teachable Machine using their own labelled photographs of the six target classes, exported in TensorFlow.js format. Once loaded, all subsequent scans are routed through this custom model instead of the generic MobileNet-plus-keyword-mapping pipeline. This provides, at no engineering cost to the end user, an upgrade path toward the domain-specific accuracy levels reported for trained models in [1], [2], and [4], without requiring the user to re-deploy the web application.

V. IMPLEMENTATION DETAILS

The assistant is implemented as a single, dependency light HTML5 document using vanilla JavaScript and CSS, with two external libraries loaded from a content delivery network: @tensorflow-models/mobilenet for the base classifier and @teachablemachine/image for the optional custom-model path. No build tools, backend server, or database are required; the page can be hosted as a static file or opened locally, which keeps the total cost of ownership close to zero. The user interface presents (i) a live camera panel, (ii) a colour-coded bin recommendation card corresponding to the six waste categories, (iii) a numeric confidence indicator, (iv) a short disposal tip specific to the identified category, and (v) a session-local scan history, all rendered without a page reload. Table II compares the resulting deployment profile against the hardware-centric systems surveyed in Section II.

TABLE II. Deployment Comparison with Industrial / IoT Systems

Aspect Industrial CV Systems [1],[4],[7] Proposed Browser Assistant Hardware Edge GPU / Jetson Nano, conveyor, actuators, sensors Any camera-equipped phone, tablet, or PC Software stack Python, PyTorch/TensorFlow, custom firmware HTML5 + vanilla JS + TensorFlow.js (browser) Deployment cost High (hardware + installation + maintenance) Near-zero (single static web page) Data privacy Images often sent to server / cloud All inference on-device; no image leaves browser Target scenario Materials Recovery Facilities, bulk sorting Household, classroom, community bins, kiosks Accuracy ceiling High, with domain-specific training data Moderate; limited by generic ImageNet labels

Extensibility	Requires re-training pipeline & retraining data	In-app Teachable Machine model swap, no code change
---------------	---	---

VI. RESULTS AND DISCUSSION

Because the proposed assistant intentionally uses a generic, non-fine-tuned CNN, its raw classification accuracy on trash-specific imagery is expected to be below the 89-99% figures reported by systems trained specifically on waste datasets [1]-[4]. Informal internal testing across the six target categories showed reliable recognition for visually distinctive, single-material items (e.g., a glass bottle, an aluminium can, a banana peel), consistent with the fact that MobileNet's ImageNet training already contains closely related object classes for these items. Recognition was noticeably less reliable for items that are visually ambiguous even to trained human annotators — for instance, distinguishing rigid versus soft plastic, or brown cardboard versus paper — a difficulty also explicitly reported for the industrial ZeroWaste dataset in [7] and [5], where annotator agreement itself was measured at only 84-94%. This indicates that the accuracy ceiling of the keyword mapping approach is bounded less by the browser-based deployment model and more by the same fundamental visual-ambiguity problem observed at industrial scale.

The comparison in Table I shows a clear inverse relationship in the surveyed literature between deployment accessibility and reported accuracy: systems requiring the least infrastructure (a generic classifier plus rules) trade off accuracy against systems requiring dedicated datasets, training pipelines, and hardware (VGG-19/ResNet/MobileNetV2 in [4] reaching up to 99.7%). The proposed assistant is best understood not as a replacement for these higher-accuracy, higher-cost systems, but as a first line of defence at the point of disposal: a zero-cost nudge that can be deployed instantly across an entire community's phones to reduce the volume of grossly contaminated waste reaching a sorting facility, thereby easing the burden on downstream industrial classifiers such as those evaluated in [1], [5], and [7].

significant in this context: it allows the same static web page to be progressively upgraded, by any non programmer, from a generic 22-63%-class heuristic performance regime toward the 95-99% regime reported for custom-trained Teachable Machine and transfer learned models in [1], [2], and [4], simply by supplying labelled photographs relevant to a specific locality's waste stream and packaging conventions — addressing the domain-shift concern raised for cross-facility and cross-season deployment in [5].

VII. LIMITATIONS

- The default MobileNet-plus-keyword pipeline has not been fine-tuned on waste-specific imagery and therefore inherits the well-known difficulty of separating visually similar sub-materials (e.g., soft vs. rigid plastic, cardboard vs. brown paper) documented for both TrashNet-scale [6] and industrial-scale [1], [5], [7] systems.
- Camera-based recognition works best with a single, well-lit, unobstructed item in frame; performance is expected to degrade substantially in cluttered, multi object scenes of the kind captured in the ZeroWaste, WaRP and SpectralWaste datasets [5], [7].
- Classification runs on general commodity hardware and is therefore slower and less power-efficient per inference than a purpose-built edge accelerator such as the Jetson Nano used in [1].
- Category granularity (six classes) is coarser than the 14-28 sub-categories used in some specialised datasets [3], [4], which may be necessary for downstream material-specific recycling.

VIII. FUTURE WORK

The most direct improvement path is to replace the generic keyword-mapping stage with a small CNN transfer-learned on a public waste dataset such as TrashNet [6] or DWSD [3], following the hybrid feature extraction-then-fine-tuning strategy shown to be effective in [1], while keeping the model small enough to remain browser-deployable via TensorFlow.js quantisation. Second, the confidence-aware fallback could be extended into an active-learning loop in which uncertain, user-confirmed scans are logged and periodically used to retrain a community-specific Teachable Machine model, directly operationalising the citizen-feedback concept proposed in [2]. Third, evaluating the assistant on cluttered, multi-item frames using benchmarks derived from ZeroWaste or SpectralWaste [5], [7] would quantify the accuracy gap between the single-item assumption made in this work and real-world disposal behaviour. Finally, coupling the browser assistant with a low-cost fill-level sensor, in the spirit of [2] and [4] but without the full conveyor/actuator mechanism, could provide a middle-cost tier between a purely educational app and a fully automated industrial smart bin.

IX. CONCLUSION

This paper presented an AI-powered Smart Waste Segregation Assistant that classifies waste items into six actionable disposal categories using an in-browser MobileNet classifier, a rule-based keyword-to-category mapping engine, and an optional user-supplied Teachable Machine model, requiring no server, no specialised hardware, and no installation. A structured review of seven recent works on edge-deployed CNNs [1], mobile/smart-bin systems [2], dense industrial segmentation datasets [3], [7], IoT-integrated classification pipelines [4], survey-level analysis of the industrial domain gap [5], and the foundational TrashNet benchmark [6] was used to position the proposed system: while dataset-trained, hardware-integrated systems achieve substantially higher accuracy (89-99%), they require proportionally higher deployment cost, and their accuracy still degrades sharply (to roughly 24-57%) once real industrial clutter is introduced. The proposed assistant instead targets the accessible, zero-cost, household/community end of the deployment spectrum, and includes a concrete, no-code upgrade path toward the higher-accuracy regime through its Teachable Machine integration. Future work will focus on replacing the heuristic mapping stage with a properly transfer-learned lightweight model and on empirically benchmarking the system against the cluttered-scene datasets surveyed in this paper.

REFERENCES

- [1] G. White, C. Cabrera, A. Palade, F. Li, and S. Clarke, "WasteNet: Waste Classification at the Edge for Smart Bins," arXiv preprint arXiv:2006.05873, 2020.
- [2] B. Sinduja and T. Kumar, "An Intelligent App-based System for Waste Segregation and Collection," *Procedia Computer Science*, vol. 235, pp. 2843-2856, 2024.
- [3] A. Ali, S. Acharjee, Md. Manarul Sk., S. Z. Alharthi, S. S. Chaudhuri, and A. Akhunzada, "DWSD: Dense waste segmentation dataset," *Data in Brief*, vol. 59, art. 111340, 2025.
- [4] A. Alourani, M. U. Ashraf, and M. Aloraini, "Smart waste management and classification system using advanced IoT and AI technologies," *PeerJ Computer Science*, vol. 11, art. e2777, 2025.

- [5] P. Thielmann, Y. Zhou, B. Mirbach, D. Stricker, and J. Rambach, "A Review of Computer Vision for Industrial Grade Waste Classification," IEEE Access, vol. 13, pp. 151934-151953, 2025.
- [6] M. Yang and G. Thung, "Classification of Trash for Recyclability Status," CS229 Project Report, Stanford University, 2016.
- [7] D. Bashkirova, M. Abdelfattah, Z. Zhu, J. Akl, F. Alladkani, P. Hu, V. Ablavsky, B. Calli, S. A. Bargal, and K. Saenko, "ZeroWaste Dataset: Towards Deformable Object Segmentation in Cluttered Scenes," in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 21147-21157.