sustainability

MDPI

Article

A PlastiVision-Edge Enabled Plastic Waste Bank Framework for Community Income Generation and Circular Economy Transition in Rural and Urban India

Madhini Balasundaram and Supraja Perumal *

Department of Networking and Communications, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai 603203, Tamil Nadu, India; mb7992@srmist.edu.in

* Correspondence: suprajap@srmist.edu.in

Abstract

India generates a large and growing volume of plastic waste each year, but community-level collection stays under-resourced: Self-Help Groups (SHGs) lack an affordable way to identify polymer type at the point of deposit. We present PlastiVision-Edge, a zero-shot, open-vocabulary smartphone detector that identifies plastic waste by material type without labelled examples of every new packaging design. Its core enabler is the Plastic Waste Vocabulary Ontology (PWVO), a 47-category, material-aware prompt set built on Basel Convention waste codes. PWVO prompts a standalone YOLOE-26 model that receives no additional fine-tuning, so recognition comes from the prompt design itself, not from further training. This ontology underpins a proposed Plastic Waste Bank framework involving seven stakeholders. The framework pays citizens per kilogram by polymer type, not a flat mixed-waste rate, and records each transaction as traceability evidence for Extended Producer Responsibility (EPR). On TACO (Trash Annotations in Context, a crowd-annotated litter-photo dataset; 459 images), PWVO-guided YOLOE-26 reaches 8.4% box and 8.3% mask average precision on estimable categories. For 13 packaging types with no labelled examples in any benchmark dataset, it succeeded in 80.0% of trials (95% confidence interval: 65.2–89.5%), tested on multiple author-captured photographs per type. On TrashNet (a studio-photo recyclables dataset), it flags plastic content in 25.7% of images. The model is light enough for field use, running at 41.5 FPS on a GPU and 440 ms per image on Android in a 10.7 MB offline-capable footprint, and its pixel-level output estimates scene-level plastic coverage, unlike box-only detectors. A 436-household urban field pilot shows a 67–108% per-kilogram value uplift over baseline, worth an estimated INR 46–75 per household per month in added income.

Keywords: plastic waste; circular economy; zero-shot detection; waste bank; extended producer responsibility; instance segmentation

Check for updates

Academic Editor: Abdelhadi Makan

Received: 27 August 2026
Revised: 15 October 2026
Accepted: 17 October 2026
Published: 23 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Picture a household in a Chennai suburb finishing a bottle of water, a packet of milk, and a bag of chips in the same week. All three are plastic, all three have some resale value once separated by polymer type, and all three, in most of India today, end up in the same bin. A *kabadiwala* (an informal scrap dealer who buys recyclable waste by weight) or a municipal collection vehicle picks up the mixed bag at a single flat rate, when it is paid for at all; PET (Polyethylene Terephthalate), HDPE (High-Density Polyethylene), PP

(Polypropylene), and multilayer sachets, whose per-kilogram values can differ by 5–10×, are treated as one undifferentiated commodity because no one at the point of collection has a fast, cheap way to tell them apart. This everyday gap in identification is one reason a large share of India's recyclable plastic ends up mismanaged instead of recycled, and it is also, we argue in this paper, a solvable one: not by asking households to change what they buy, but by giving the point of deposit a way to see what it is holding.

Scaled up, that household's bottle and sachet become a national problem. India releases an estimated 9.3 million tonnes of macroplastic pollution a year, close to one-fifth of the global total [1], while 30–60% of its plastic waste reaches formal or informal recycling [2]; the rest ends up in dumpsites, water bodies, or mixed Municipal Solid Waste (MSW). Reported generation figures, drawn from Central Pollution Control Board (CPCB) data, are lower, at about 4.1 million tonnes a year [3], though open burning and under-reported informal and rural collection make this an almost certain undercount [1]. None of this waste is worthless once it is identified correctly: local scrap dealers (kabadiwalas) currently quote about INR 10–18/kg for PET, INR 8–14/kg for woven PP sacks, and INR 1–3/kg even for multilayer sachets (indicative local rates averaged over quotes from six scrap dealers in and around Potheri), and India's blended recycling rate, estimated at 30–60% with the informal sector included, compares favourably with global figures, although the underlying data are unreliable [2,4]. What is missing is not the value of the material or the willingness to sort it, but a transaction mechanism at the point of deposit: something that rewards segregation, prices each polymer correctly, and leaves a verifiable record for Extended Producer Responsibility (EPR) compliance.

We see three separate reasons why closing this gap matters now, and not only as a long-term goal. Environmentally, plastic that is not recycled does not simply disappear: it can persist in dumpsites and water bodies for decades, and since India's true generation figure appears to be undercounted, the scale of the problem may well be larger than official statistics suggest. Economically, the market already places a real value on sorted plastic, as the per-kilogram rates above show; what seems to be missing is not demand for recycled polymer, but an affordable way to sort and verify it at the point of collection. Regulatorily, India's EPR framework is increasingly asking brand owners to demonstrate verified, traceable collection instead of simply contributing to a fund, which is gradually turning what was once a voluntary sustainability effort into a compliance requirement with real consequences for falling short. Taken together, these three pressures suggest that a circular economy transition, in this particular sense of moving plastic from single-use, unverified disposal into a priced and traceable material stream, is perhaps less an ideological choice than a natural next step, given where environmental need, market value, and regulation already seem to point. What appears to have been missing, in our view, is a simple technical layer that could make this transition possible at the point where it has to begin: the moment someone hands over a bag of mixed plastic.

1.1. The Missing Link: From Identification to Income

In India's waste-recycling landscape, the plastic value chain breaks at the earliest stage, where kabadiwala networks buy mixed plastics at a single flat rate regardless of polymer type. Manual sorting remains the usual method at municipal Dry Waste Collection Centres (DWCCs), a slow and unsafe method that persists despite SBM-U 2.0's (Swachh Bharat Mission Urban 2.0) mandatory source-segregation rule. Household compliance with that rule stays low, since there is neither a way to verify it nor a financial incentive to encourage it. The Plastic Waste Bank.

(PWB) model addresses this gap: citizens deposit plastic that has already been identified and are credited by polymer weight and type [5,6]. Indonesia has pioneered this

approach, operating more than 12,000 Bank Sampah units nationally and serving roughly 3.5 million households, with average earnings around USD 0.1/kg, or about USD 1.4 for a 15–20 kg drop-off [7,8]. The same model can be incorporated into India's EPR-governed economy, where the identification step would pay a citizen INR 12/kg while the brand owner separately claims a INR 2–5 EPR credit for the same transaction.

1.2. Our Approach: PlastiVision-Edge at the Point of Deposit

To our knowledge, no existing tool fills the gap of being fast, affordable, and general enough for an Indian waste bank to run on. Handheld Near-Infrared (NIR) spectrometers cost INR 2–8 lakhs; conveyor-grade industrial NIR systems cost INR 25–80 lakhs and need civil infrastructure neither an SHG nor a gram panchayat has; and closed-vocabulary classifiers trained on datasets such as TrashNet and TACO [9,10] often struggle to recognise Indian sachets and other regional packaging outside their training categories [11,12]. Recent zero-shot and few-shot vision-language approaches to plastic waste classification [13] report similar generalisation without task-specific retraining, which supports the case for a zero-shot approach here. None of them, however, address on-device deployment, real-time performance on a phone, or an India-specific packaging vocabulary—the three gaps this paper closes. PlastiVision-Edge is our answer: a zero-shot, open-vocabulary detection system that runs on an ordinary Android phone. It pairs YOLOE-26 [14] with the Plastic Waste Vocabulary Ontology (PWVO), a 47-category polymer-aware ontology, so that any packaging type can be identified from a natural-language prompt at the moment of scanning, with no retraining step in between. In practice, this hands an SHG member or waste bank operator a tool that can verify, classify, and price deposited plastic on the spot.

1.3. Plastic Waste Banks and Community Income

Turning waste disposal into a cash transaction is the core idea behind a waste bank. Indonesia's Bank Sampah model is the strongest evidence that this approach outperforms penalty-based enforcement at driving source segregation [5]; it has since become a recognised national strategic mechanism [7], and one community-facility study measured a roughly 5% rise in resident incomes as a result [15]. Digitisation pushed this further still: platforms like Rapel turned it into a mobile-app-mediated marketplace [8]. India, in principle, has an advantage Indonesia's system never had: an EPR framework that creates real demand for verified collection. In practice, however, existing Indian systems fall short of that model's effectiveness. The SWaCH model in Pune rests on household user fees, with members' monthly earnings in 2011–12 from under INR 500 to over INR 3500 [16] and, to our knowledge, no automatic polymer identification at intake. Recykal runs a digital waste marketplace and traceability platform [3] but, as far as we can determine, has no citizen-incentive mechanism at the point of initial collection. Proposed models such as buy-back and SHG-based schemes similarly lack a proper waste-identification mechanism [17]. This has led researchers to conclude that technology-mediated income, paired with supporting regulation, is what formalises the workforce [18], cutting across deeper structural barriers such as mistrust, municipal inaction, and the precarity of India's 1.5 million informal waste pickers.

Indonesia is not the only relevant precedent, and treating it as the sole external comparison understates how widely this problem, and variants of this solution, recur. A recent study of Ghana's informal settlements documents a structurally similar move: converting plastic collection into a tradeable-credit transaction (plastic credits sold against verified collection volumes) to fund waste-picker income in slum communities that have no formal municipal collection at all [19]. The mechanism differs in financing instrument (credits sold to buyers versus India's EPR-linked per-kilogram pricing), but the underlying

logic is the same one this paper relies on: informal collectors are already doing the physical work, and what is missing is a verified, priced transaction at the point of collection, not motivation or manpower. That Indonesia and Ghana arrive at broadly convergent designs independently, across very different regulatory and economic contexts, is itself evidence that the point-of-deposit verification problem, and not any single country's institutions, is the actual bottleneck this class of system is built to remove.

A companion qualitative study comparing plastic-focused and mixed-waste informal enterprises in Accra, Ghana, and Kisumu, Kenya, adds a caution that this paper's own design choices should be read against, and not as a further success story. Plastic-specialised collectors in both cities faced price volatility tied to seasonal demand and oil-linked resale prices, theft of segregated material once it visibly carried value, and storage insecurity; the collector-recycler relationships the study found to work best were built on a fixed, year-round flat-fee purchase price, not a spot-priced one [20]. This is directly relevant to why the framework proposed here is structured the way it is: sorted plastic is aggregated and scanned at a waste-bank-operator level (Section 5.1) instead of being left in an individual collector's possession, and the per-kilogram credit rate is a fixed, operator-set band (Section 5.6) instead of a spot-market price, closer to the flat-fee arrangement the Accra/Kisumu study found to work than to the volatile, individually-held arrangement it found to be problematic. Field evidence on whether this design choice prevents these documented problems in the Indian context is still to be collected; it is presented here as the reasoning behind the design, not as a validated outcome.

More broadly, waste banks and plastic-credit schemes are one design point within a larger circular-economy debate about how developing economies should handle plastic at all: whether to prioritise diversion and material recovery (the approach this paper and its Indonesian/Ghanaian precedents take), reduction and substitution away from plastic packaging entirely, or some blend of both. A recent systematic review of circular-economy approaches to plastic waste across developing economies finds that community-engagement and technology-enabled collection mechanisms (the category this paper's framework belongs to) are consistently identified as necessary but not sufficient on their own, and need to be paired with upstream policy (extended producer responsibility, packaging-design mandates) and downstream recycling-market capacity to close the loop, not merely formalise collection [21]. We position this paper accordingly: PlastiVision-Edge and the Plastic Waste Bank framework address the collection-and-identification bottleneck specifically, and are not a substitute for upstream reduction policy or downstream recycling infrastructure investment, both of which remain necessary regardless of how well the collection layer works.

1.4. Plastic Waste Policy: EPR, PWM Rules, and Circular Economy

India's Plastic Waste Management Rules 2022 [22] mandate EPR collection/recycling minimums of 30–50% by 2024–25, rising to 80% for rigid packaging by 2028–29. This is a distinct obligation from the recycled-content-by-weight mandate discussed in Section 5, which separately requires 30%→60% recycled content in new rigid packaging by the same year, under the 2026 Amendment Rules [23]. The Rules also require PIBOs (producers, importers, and brand owners) to register and report on India's Common EPR Portal (which succeeded the earlier standalone Plastic Waste EPR Portal on 28 June 2026, with prior registrant data migrated across), with 2025 amendments adding QR/barcode traceability [24]. Tradeable EPR certificates create a market price for verified collection that can flow to community waste banks [25], within a INR 22.17 billion (2023) Indian waste management market projected to reach INR 54.20 billion by 2030 [3]. On the recycling side, India's existing recycler and aggregator network still leans heavily on informal-broker arbitrage

for sorted-material pricing, and digital collection platforms with SHG-based incentive structures have been proposed specifically to reduce that sorting burden and give recyclers a more transparent, provenance-backed supply [2,26]; this is the same gap the per-deposit scan record in Section 5 is designed to close.

1.5. Related Work: AI-Enabled Waste Sorting

On controlled benchmarks, closed-vocabulary detectors already perform well: CFD-YOLO reports 93.3% mAP (mean Average Precision, the standard detection-accuracy metric used throughout this paper) [27], YOLO-11m 98.03% on WaDaBa [28], and a hybrid CNN-transfer model 97.12% on TrashNet [29]. The catch is that these models can only recognise packaging types they were trained on, which is a limitation in a market where packaging changes constantly. Open-vocabulary detection sidesteps this by not fixing the label set in the first place. CLIP [30], OWL-ViT [31], Grounding DINO [32], and YOLO-World [33] each pushed this idea forward, and YOLOE-26 [14] takes it further with NMS-free, DFL-free end-to-end segmentation across text-, visual-, and prompt-free modes. We had already shown, in earlier work, that prompt design matters a great deal for Indian packaging [11], and separately looked at how these models behave under edge quantisation [34], but neither piece of work put a structured ontology, instance segmentation, and an actual Android deployment together in one system, which is what this paper does.

A related finding worth flagging: fine-tuning the same underlying detector raises accuracy on the training classes substantially, but wipes out cross-domain generalisation unless adapter-based techniques are used to protect it [35]. PlastiVision-Edge takes the opposite bet. Given how fast South Asian packaging turns over, we favour keeping the vocabulary open instead of chasing peak accuracy on a fixed set, and we handle the resulting lower-confidence cases through operator-level manual review (Section 5) instead of through fine-tuning.

1.6. Gap Addressed by This Work

The review above reveals a clear and unaddressed intersection gap. Table 1 maps existing work across five axes that are simultaneously required for a viable community plastic waste bank, and suggests that, to our knowledge, no prior system addresses all of them together.

(1) Packaging diversity. The published systems we reviewed use fixed vocabularies (TrashNet has 6 categories, TACO has 60, WaDaBa has 5), and these tend to miss Indian sachets and regional packaging without warning [11,17], so they need retraining, which is hard to carry out at community level. Open-vocabulary detection avoids this: an operator can add a new category via a text prompt in seconds.
(2) The AI transaction primitive. Indian waste banks still rely on manual identification, which creates valuation uncertainty, rules out polymer-specific pricing, and leaves no EPR-compliant record [16,18,26]. This paper resolves that with a mobile AI running at 440 ms/frame on a commodity Android smartphone, with no NIR or conveyor infrastructure needed.
(3) Offline rural deployment. Existing digital platforms need continuous connectivity, which is not well suited to intermittent rural 4G [26]. The 10.7 MB INT8 YOLOE-26S model used here, paired with deferred ledger sync (Section 3), is offline-capable by design, not as an afterthought.

To the best of our knowledge, no prior work addresses all three of these together, alongside the citizen-incentive and edge-deployment columns in Table 1. This is the gap this paper is built to fill.

Table 1. Gap matrix: existing work vs. required capabilities for a community plastic waste bank. ✓ = addressed; ∼ = partially; × = not addressed.

Work	Waste Bank Model	AI Sorting	Open Vocab	Citizen Incentive	Edge/Offline
Bank Sampah/Rapel [7,8]	✓	×	×	✓	∼
SWaCH Pune [16]	✓	×	×	✓	×
Recykal/digital platforms [3]	∼	×	×	×	×
CFD-YOLO/YOLO-11m [27,28]	×	✓	×	×	×
YOLO-World/YOLOE [14,33]	×	✓	✓	×	×
Our own prior work (ZS-OVD for plastic waste) [11,34]	×	✓	∼	×	✓
This paper (PlastiVision-Edge)	✓	✓	✓	✓	✓

1.7. Contributions and Paper Structure

This paper makes four contributions: (1) a seven-stakeholder circular economy framework (Section 5) mapping physical, financial, and policy/data flows across consumers, waste bank operators, brand owners, recyclers, manufacturers, retailers, and government; (2) a community income model, grounded in this pilot's own measured collection and payout data, showing a 67–108% per-kilogram value uplift and INR 46–75/month additional urban household income (with a separate, unvalidated CPCB-proxy estimate for rural households); (3) a proposed EPR traceability linkage generating machine-verifiable, Common-EPR-Portal-aligned provenance records per scan; and (4) technical validation of PlastiVision-Edge: 7.7–8.4% Box mAP@0.5 on TACO under the full competitive vocabulary, macro-averaged over the 13 annotated benchmark categories (exceeding a same-subset YOLO-World-S/L evaluation under the identical protocol, 4.9–7.3%; see Section 4), with pixel-level masks, which the box-only YOLO-World does not provide; 80.0% pooled replicate-based zero-shot success (95% CI [65.2%, 89.5%]) on benchmark-uncovered South Asian packaging; 440 ms Android inference; 10.7 MB INT8 footprint.

Two parts of the system described here are already built and working: the AI detection model, and an Android application with live-camera and gallery-based detection (Section 5.2). The technical validation reported in Section 4 is based on these two components. The remaining platform components: Unified Payments Interface (UPI) linked wallet and payment integration, brand-partner premium pricing, Common EPR Portal connectivity, and the Potheri pilot itself, are presented as a proposed system design and a phased pilot plan.

The rest of the paper is organised as follows: Sections 1.3–1.6 review related work on plastic waste banks, policy, and AI-enabled sorting, and close with the specific gap this work addresses. Data and Study Site (Section 2) describes the Potheri/apartment-complex site and the TACO/TrashNet/WaDaBa benchmark datasets; Method (Section 3) presents the PWVO ontology, the YOLOE-26 detection backbone, and the evaluation/deployment pipeline; Results (Section 4) reports experimental findings; Framework (Section 5) presents the seven-stakeholder circular economy model, platform architecture, incentive/EPR mechanism, and behavioural design; Discussion and Conclusion follow in Sections 6 and 7.

2. Data and Study Site

This section describes the two data sources underpinning this paper: the Potheri/apartment-complex study site, comprising a proposed village-scale pilot and a completed preliminary apartment-complex study, and the TACO/TrashNet/WaDaBa benchmark datasets used to evaluate the detection system.

2.1. Study Site: Potheri Village, Chengalpattu District (Proposed Pilot Design/Plan)

The pilot is planned for Potheri village in Chengalpattu District, Tamil Nadu, a mixed rural/peri-urban geography with three distinct zones. The Potheri Gram Panchayat covers around 2800 lower-middle-income households, whose waste stream is dominated by multilayer milk sachets, multilayer snack pouches, and woven PP sacks. Active kabadiwalas are scarce, and there is no formal dry-waste collection. Four adjacent apartment complexes add roughly 1200 flats, where e-commerce packaging and PET bottles dominate the waste. Food stalls and canteens around SRM University generate EPS (Expanded Polystyrene) and multilayer beverage-pouch waste. Multiple provision stores and small supermarkets operating in Potheri stock packaged branded goods, providing existing retail infrastructure that could serve as a distribution channel for recycled-content products to reach local consumers without requiring new logistics.

The pilot plan is projected to run for 18 months across three stages. Phase 1 (Months 1–3) covers initial setup: registering 300 households, calibrating base point values against the existing kabadiwala tariff rate card, and fine-tuning the PWVO prompts for Potheri's local waste types. Phase 2 (Months 4–12) is the active collection trial: around 150 households join bi-weekly drives, and we track household income, diverted plastic volumes, uptake of premium subscription categories, and EPR record generation for at least one FMCG (Fast-Moving Consumer Goods) brand. Phase 3 (Months 13–18) assesses scaling and sensitisation, using pre- and post-intervention behavioural surveys to evaluate shifts in segregation knowledge and brand-specific willingness to return plastic. On the basis of the pilot's economic model (Section 5.6) and comparable Southeast Asian waste bank implementations [8,15], the pilot study is expected to show PlastiVision-Edge-enabled household monthly income of roughly INR 7 to INR 12 in the village (a CPCB-proxy estimate, pending field validation) and INR 115 to INR 144 in apartments (grounded in this paper's own observed apartment-complex data), a 60–108% premium over the respective baseline bulk rates. This can be enhanced by activating at least three premium categories by Month 4, which is expected to produce measurable gains as segregation knowledge becomes more consistent under digitally-mediated intervention, as reported for Indonesian waste-bank platforms [8].

2.2. Preliminary Field Data: A Nine-Week Apartment-Complex Collection Sample

Ahead of the Potheri pilot described above, we ran a small, non-randomised collection exercise at a six-block apartment complex (blocks A–E and G; block-level unit counts vary with floor count, 11 or 12 floors depending on the block, giving 436 households total) with an on-site supermarket and canteen, over a nine-week window (1 July–31 August 2026). This is explicitly preliminary field data, not the validated pilot itself: it covers a single site, not Potheri's mixed rural–peri-urban geography, was not randomised or powered for statistical inference, and its household-level participation and income outcomes have not yet been analysed. We report it here as an early empirical anchor for the framework's generation-rate and category-composition assumptions, distinct from the fuller 18-month Potheri validation in Section 2.1 above. Table 2 summarises the sample; the paragraphs below give the detail and caveats.

Daily housekeeping-log weighing across the six residential blocks totalled 12,795.9 kg of dry plastic waste over 61 collection days in the nine-week window (≈209.8 kg/day, or ≈0.48 kg/household/day at 436 households). Including the on-site supermarket and canteen, combined collection reached 14,269.4 kg. Across the eight complete weeks (plus a final six-day partial week), weekly block-level totals ranged from 917.5 kg to 1949.3 kg. A nine-week sample cannot yet separate this variability from genuine seasonal or collection-schedule effects.

Table 2. Nine-week apartment-complex collection sample. Preliminary field data, not the validated Potheri pilot (Section 2.2).

Metric	Value
Collection window	1 July–31 August 2026 (62 days; 61 collection days; 1 day with no log entry)
Households covered	436 (6 blocks, A–E and G; 11–12 floors each)
Residential dry plastic collected	12,795.9 kg
Total incl. on-site supermarket and canteen	14,269.4 kg
Daily rate (residential)	≈209.8 kg/day (≈0.48 kg/household/day)
Weekly range (8 complete weeks)	917.5–1949.3 kg
Composition, 9-sample hand-sort (2775.3 kg sorted)	HDPE 46.9%; multilayer 34.8%; PET/Polyester 14.7%; PP 3.6%
Composition, initial 5-sample reading	HDPE 47.5%; multilayer 36.3%; PET 13.0%; PP 3.2%
EPS/thermocol (tracked separately)	$\lesssim$2 kg/week ($\lesssim$0.7% of sorted total)

In parallel, a non-exhaustive weekly sampling exercise manually sorted 2775.3 kg of collected material across nine weekly samples (1 July–26 August) into four PWVO superclasses: HDPE containers (46.9%), multilayer pouches (34.8%), PET/Polyester (14.7%), and PP bags/sacks (3.6%). These shares stayed broadly stable as the sample grew from an initial five-sample reading (HDPE 47.5%, multilayer 36.3%, PET 13.0%, PP 3.2%), a useful if informal stability signal. The sort covered a subset of the collected weight, not every bag, so the percentages are indicative composition shares, not a verified material balance against the 12,795.9 kg total.

EPS/thermocol was tracked separately from this four-class sort. Site observation across the full collection window found it a consistently minor stream, at most ≈2 kg/week ($\lesssim$0.7% of total sorted material, had it been added to the sorted total above). This matches its typical profile at household level as a bulky but light, low-frequency item, not a high-mass stream.

All entries are primary daily records: the collected material was weighed each day and the weight written into the site's collection notebook, and the entries were then transcribed into the pilot spreadsheet. The digitised daily records and the aggregated tables behind every figure in this section are provided in the project repository (Data Availability Statement), so that each derived value can be re-checked against them. Only one day in the window (9 July) has no collection-log entry.

This preliminary sample shows that real-world generation rates at an urban apartment complex are broadly consistent with the order of magnitude assumed in this manuscript (Sections 2.1 and 5.6), and that HDPE and multilayer categories dominate collected volume at this site, which is directly relevant to prompt-vocabulary and Vocabulary Manager prioritisation. It does not yet demonstrate household-level income uplift, behavioural change, EPR-record generation, or app-mediated collection, since this exercise used the site's existing manual weighing process, not PlastiVision-Edge itself; that instrumentation (household enrolment, per-transaction app scans, and a Detection_Performance log comparing app output against manual sorting) is the next stage of data collection, needed before the framework's broader socio-economic claims can be considered field-validated.

2.3. Benchmark Datasets and Evaluation Protocol: TACO, TrashNet, and WaDaBa

We filtered the TACO annotation JSON (http://tacodataset.org, accessed on 15 October 2026) to a candidate list of 913 images with at least one annotation category mappable to PWVO (Section 4). TACO images are Flickr-hosted, not locally archived, so a fraction of candidate URLs resolve on any given attempt; the evaluation reported in Section 4 uses the 459 images successfully retrieved on one specific date (see Section 6.5 for the reproducibility caveat this implies). We converted the COCO polygon annotations to YOLO segmentation format and dropped instances under 300 px area. Two protocols do different

jobs here and are best read separately: quantitative mAP runs the full 47-class vocabulary at once, with categories competing through softmax normalisation and Pascal VOC 11-point interpolated AP, which is what tells you how the system performs under realistic inter-category competition; qualitative zero-shot evaluates each image against only its target category's prompt, isolating pure CLIP text-image alignment with no competition in the mix; this is the protocol underlying the replicate-based benchmark-uncovered-category success rate (80.0%, Wilson 95% CI [65.2%, 89.5%]) reported in Section 4.4.

TrashNet Dataset Preparation

TrashNet [9] is a 2527-image, single-object, studio-photographed dataset (a Stanford CS229 project) spanning six waste classes photographed against a plain white background: glass, paper, cardboard, plastic, metal, and general trash. Because it provides a single whole-image class label per photo instead of bounding-box or polygon annotations, we cannot compute mAP against it; we instead evaluate PlastiVision-Edge as a whole-image plastic-vs-non-plastic detection task on its 482 plastic-labelled images, reporting the fraction on which PlastiVision-Edge fires at least one plastic-category detection above the confidence threshold (Section 4). TrashNet's controlled, single-object, plain-background images are an easier and qualitatively different setting than TACO's cluttered, multi-object, real-world litter photographs; we report both to show that whole-image detection remains reliable in the controlled-photography case even where TACO's harder, cluttered scenes suppress mAP, and we do not treat the two datasets' numbers as comparable to one another.

3. Method: PlastiVision-Edge Detection Pipeline

Figure 1 gives a visual overview of the detection pipeline described in this section, from a single deposit-time photo to a recorded, priced transaction; each stage is specified in full in the subsection named on the diagram. Unless a caption states otherwise, every photograph used for illustration in this paper is author-captured, and is neither a TACO/WaDaBa/TrashNet dataset image nor from the Potheri pilot (Section 2.1), which has not yet been conducted.

3.1. Plastic Waste Vocabulary Ontology (PWVO): Full Specification

The PWVO is a structured 47-category, material-aware prompt ontology, built from three sources. The first is the Basel Convention Annex II/VIII plastic waste codes [36], used for polymer super-class assignment. We chose Basel over item-type litter taxonomies such as OSPAR's beach-litter categories [37] for two reasons. First, Basel classifies waste by polymer type and regulatory handling category, the same axis our EPR-traceability and payout logic uses (the "EPR category" column of Table 3 maps directly onto Basel Annex II/VIII listings). OSPAR classifies by visible item type (bottle caps, cotton buds, fishing line) for coastal litter surveys and has no polymer-code or EPR grouping. Second, Basel is the near-universal, India-ratified instrument already used in our regulatory discussion (Section 1.4), so the detection ontology and the regulatory narrative stay consistent. OSPAR has no formal standing in India. The second source is a cross-reference against the TACO [10], WaDaBa [28], and TrashNet [9] category lists, used to determine the benchmark-coverage split (see below). The third is South Asian regional packaging knowledge, covering emerging-market categories absent from TACO, WaDaBa, and TrashNet.

The 47 categories span six polymer super-classes: HDPE/LDPE (High- and Low-Density Polyethylene) film (11), Polypropylene (9), PET/Polyester (8), Polystyrene/EPS (7), Other (6), and Multilayer/composite (6). We use a 34/13 (72/28%) benchmark-coverage split, chosen to test generalisation while retaining enough benchmark-covered-class coverage for meaningful mAP estimation. A category is labelled benchmark-covered (BC) if

it appears in at least one annotated TACO/WaDaBa/TrashNet sample, and benchmark-uncovered (BU) otherwise. This label reflects only presence or absence in these three reference datasets. It says nothing about whether YOLOE-26 or MobileCLIP saw similar objects during their own general-purpose pre-training, which is not documented at the class level (Section 3.2).

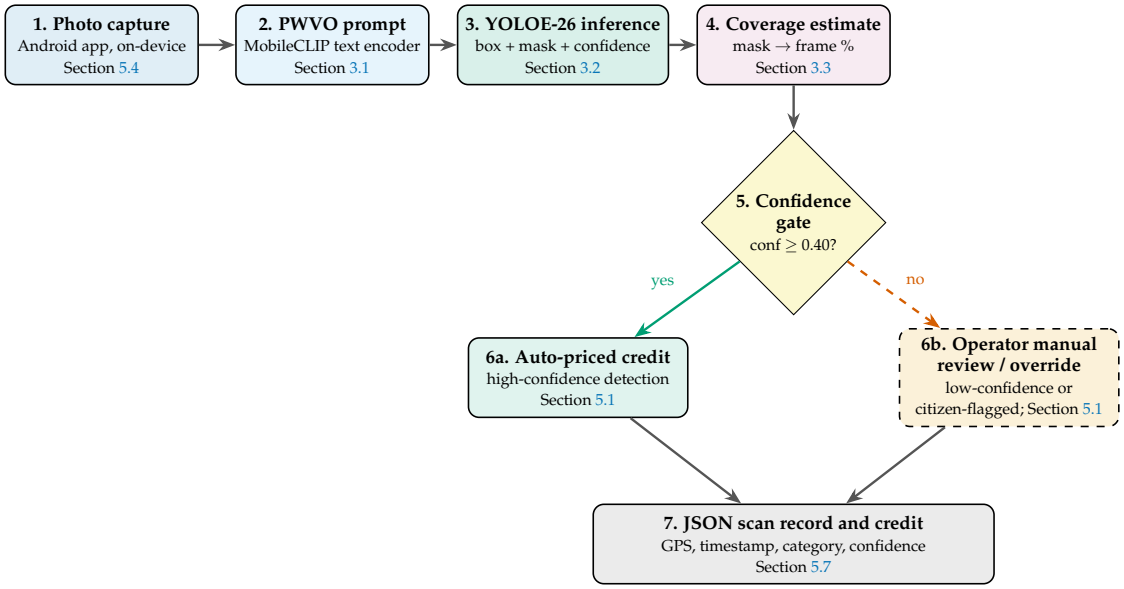


Figure 1. PlastiVision-Edge detection pipeline: from a single deposit-time photo to a recorded, priced transaction. Steps 1–4 (solid grey arrows) run identically for every scan; step 5 is a confidence threshold, not a separate model; step 6a/6b is the one point in the pipeline where a detection either credits automatically or is routed to a human operator before payment, per the correction mechanism described in Section 5.1. Only step 1 (citizen app) has been built and tested end-to-end on-device; steps 2–4 are implemented and evaluated offline/on-device as described in the cited subsections; the operator dashboard referenced at step 6b and the EPR Portal push referenced at step 7 are part of the proposed platform (Section 5.2), not yet deployed.

Each prompt names a material and a form and, where needed, adds visual-appearance and scene-context descriptors, following a material-form-context (MFC) template: `[polymer material] [physical form] [visual appearance descriptor] [scene context]`. For example, the `pet_bottle` prompt reads: "a plastic drinking bottle with cylindrical shape and screw cap, transparent or semi-transparent body…". This places each category in a distinct region of CLIP's embedding space, separate from generic object labels, without any domain-specific fine-tuning. Table 4 gives the full structure; the distribution is visualised, along with a waste-bank-relevance view (pricing, EPR category), in Figure 2 and Table 3 below.

Table 3. PWVO super-class structure with benchmark-coverage distribution and waste bank relevance. "Covered"/"Not covered" indicate presence or absence in the TACO/WaDaBa/TrashNet reference datasets. Rates are indicative local kabadiwala prices (average of six dealers in and around Potheri).

Super-Class	Total	Covered	Not Covered	Indicative Rate (INR/kg)	EPR Category
PET/Polyester	8	7	1	INR 10–18	Rigid packaging
HDPE/LDPE Film	11	8	3	INR 6–12	Flexible packaging
Polypropylene (PP)	9	8	1	INR 8–14	Rigid packaging
Polystyrene/EPS	7	6	1	INR 2–5	Rigid packaging
Multilayer/Composite	6	3	3	INR 1–3	Multilayer (Cat. 4)
Other/Unclassified	6	2	4	INR 3–8	Flexible/other
Total	47	34	13	—	—

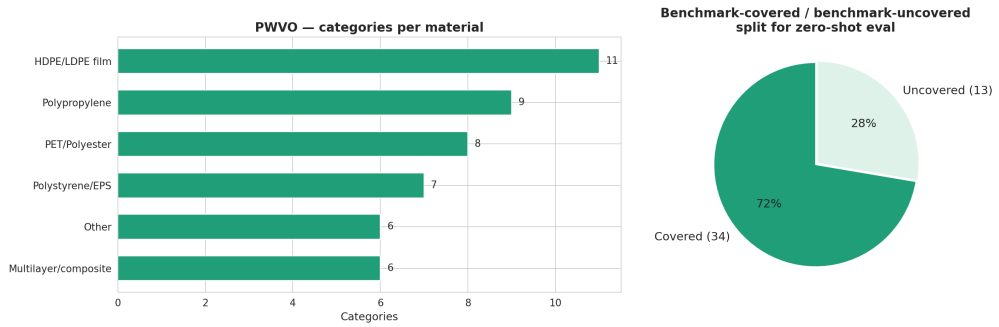


Figure 2. PWVO super-category distribution and benchmark-coverage split. **Left**: number of categories per polymer super-class (HDPE/LDPE film dominates with 11 entries). **Right**: the 34/13 benchmark-covered/benchmark-uncovered split (72%/28%) used for zero-shot generalisation evaluation. Benchmark-uncovered categories have zero annotations in TACO, WaDaBa, or TrashNet and serve as zero-shot transfer targets.

Table 4. PWVO super-category structure with benchmark-coverage distribution. BC = benchmark-covered, BU = benchmark-uncovered, as defined in Table 3. Colour codes are used in all detection visualisations.

Super-Class	Total	BC	BU	Example Prompt
PET/Polyester	8	7	1	hinged transparent plastic clamshell container for fruit
HDPE/LDPE Film	11	8	3	clear plastic bag used for bread packaging tied at top
Polypropylene (PP)	9	8	1	woven white polypropylene rice or grain sack
Polystyrene/EPS	7	6	1	white polystyrene foam supermarket meat or vegetable tray
Multilayer/Composite	6	3	3	shiny metallic-laminate chip or crisp snack packet discarded
Other	6	2	4	small black or white plastic zip tie cable fastener
Total	47	34	13	

3.2. YOLOE-26 Detection Backbone: Model and Inference Pipeline

PlastiVision-Edge uses YOLOE-26S (15.3 M parameters) and YOLOE-26L (39.3 M parameters) pre-trained on Objects365v1 [38], GQA, and Flickr30k (Ultralytics $\geq$ v8.4.0), in `-seg` format returning both boxes and binary segmentation masks per call. "Zero-shot" means no plastic-waste training or fine-tuning. Class-level exposure in Objects365v1, GQA, or Flickr30k is not documented and cannot be verified from the released model alone. YOLOE-26 supports three modes: (a) Text-prompted: PWVO prompts encoded via MobileCLIP's [39] text encoder into an embedding matrix passed to `model.set_classes()`; (b) Visual-prompted (SAVPE): one-shot detection from a reference crop; and (c) Prompt-free (LRPC): salient-object detection with no vocabulary. Vocabulary swap costs under 50 ms with no model reload. A full capability comparison against YOLO-World and YOLO11 baselines is given in Table 5 below.

Table 5. Model capability comparison. Open-vocab = supports text prompt at inference; Prompt-free = detects without any text input; Masks = returns pixel-level segmentation masks. ✓ = supported; × = not supported.

Model	Backbone	Masks	Open-Vocab	Prompt-Free	Params (M)
PlastiVision-Edge-S + PWVO	YOLO26	✓	✓	×	15.3
PlastiVision-Edge-L + PWVO	YOLO26	✓	✓	×	39.3
PlastiVision-Edge-S-PF	YOLO26	✓	✓	✓	16.2
YOLO-World-L (ref)	YOLOv8	×	✓	×	49.0
YOLO11-L (closed)	YOLO11	×	×	×	25.3

3.3. Plastic Coverage Area Estimation

Per-class plastic coverage is defined as

$$\text{coverage}_n = \frac{\sum_i \text{mask}_{i,n} \text{ pixels}}{\text{image total pixels}} \times 100\% \quad (1)$$

where n indexes the class, $\text{mask}_{i,n}$ is the i-th predicted mask of class n, and coverage_n is the resulting coverage of class n (in %). It is computed as a zero-cost post-processing step on existing masks, enabling per-class segmented frame-coverage estimation (a visual area proxy for relative pollution load, defined precisely below) unavailable to bounding-box-only detectors (single-frame illustration: Figure 3, 11.0% combined coverage).

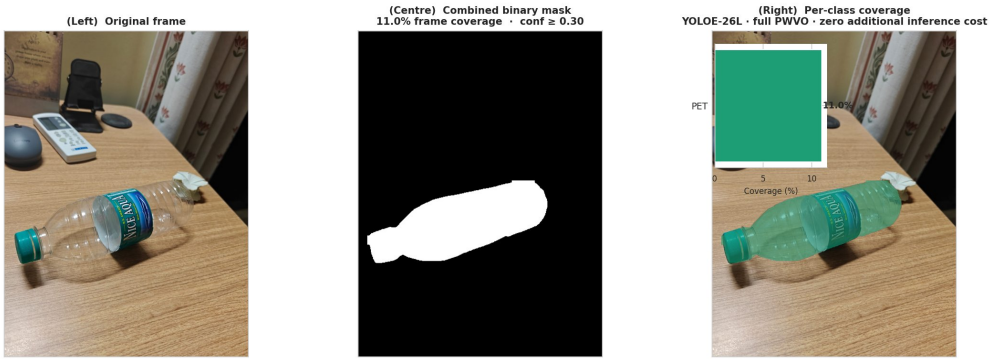


Figure 3. Plastic coverage area estimation. **Left**: original frame, an author-captured photograph (PET water bottle on indoor surface) used for illustration, not a TACO/WaDaBa/TrashNet dataset image. **Centre**: total binary segmentation mask (11.0% frame coverage, conf ≥ 0.30). **Right**: per-class breakdown by polymer super-class: PET/Polyester accounts for 11.0% of the frame (`pet_bottle` detection). This single frame is illustrative; the full 687-image dataset-scale evaluation is reported in the text above.

We evaluated coverage on the full 687-image TACO-PWVO subset. An initial pass showed a few near-full-frame, low-confidence (0.35–0.39) detections that sometimes dominated the per-image union mask, while genuine detections in the same images scored 0.66–0.82. We therefore added a sanity filter that excludes detections covering more than 80% of the frame unless their confidence is at least 0.60. The filter fired on 191 detections across 185 of the 687 images (26.9%). This is a dataset-scale, quantified instance of the EPS-dominance embedding attractor discussed in Section 4.4. After filtering, mean coverage is $19.3\% \pm 30.4\%$ (all images) and $25.5\% \pm 32.7\%$ across the 519 images (75.5%) with nonzero coverage. Per super-class (nonzero images), coverage is: Polystyrene/EPS 44.0%, HDPE/LDPE film 6.1%, Multilayer/composite 6.6%, PET/Polyester 5.8%, Polypropylene 2.6%, Other 2.2%. EPS remains disproportionately high even after filtering. This is a residual limitation of the confidence-based filter, so EPS-class coverage should be treated as an upper bound pending further mitigation.

We use the term frame coverage, not "pollution severity", throughout. The metric quantifies the 2D image-plane area occupied by detected plastic in a single photograph. It does not measure the underlying environmental severity of the pollution, which would also depend on factors not estimated here, such as mass, volume, material persistence, and ecological hazard. The $>80\%$/conf ≥ 0.60 sanity-filter threshold was chosen heuristically from the observed confidence gap between the near-full-frame artefacts (0.35–0.39) and genuine same-image detections (0.66–0.82) in this dataset. It was not tuned against a labelled coverage ground truth, so it is a disclosed methodological choice, not a validated threshold, and its sensitivity is listed as a limitation (Section 6.5).

3.4. On-Device Latency and Edge Packaging

GPU latency used a 10-frame warmup plus 100 timed inferences on a fixed 640 × 640 tensor (PyTorch 2.11.0, CUDA FP32); end-to-end latency includes pre-processing, inference, and mask decoding. This excludes camera capture and UI rendering, which are application-dependent. Android latency was measured on a flagship physical device as well as a budget model, averaged over 20 gallery-mode sessions and 30 live-mode sessions. Table 6 reports the profiled latency of PlastiVision-Edge and YOLO-World variants on the same T4 GPU (multi-trial means; see footnote for the one-time CUDA autotuning artifact affecting PlastiVision-Edge-L's raw trial spread).

Table 6. End-to-end latency and throughput (measured), GPU results on NVIDIA T4 (Colab, batch 1, FP32). PlastiVision-Edge-S/L and YOLO-World-S/L values are the mean of 5 independent timed trials (8-iteration warmup + 40 timed runs per trial) to control for run-to-run GPU timing variance. ‡ PlastiVision-Edge-L's first two of five trials (87.8, 92.8 ms) reflect a one-time CUDA kernel autotuning cost not fully absorbed by the per-trial warmup; the reported value is the mean of the remaining three trials (73.5, 57.7, 54.4 ms), consistent with independent single-run measurements taken in other sessions (44.8–54.9 ms). † Published TensorRT FP16 number on different hardware (V100); shown for reference only, not a same-hardware comparison.

Model	Device	Backend	ms	FPS
GPU benchmarks (same NVIDIA T4, measured this work)				
PlastiVision-Edge-N + PWVO	NVIDIA T4	CUDA FP32	15.7	63.6
PlastiVision-Edge-S + PWVO	NVIDIA T4	CUDA FP32	24.1	41.5
PlastiVision-Edge-L + PWVO	NVIDIA T4	CUDA FP32	61.9 ‡	16.2
YOLO-World-S	NVIDIA T4	PyTorch FP32	17.3	58.0
YOLO-World-L	NVIDIA T4	PyTorch FP32	55.5	18.0
Published reference (different hardware, context only)				
YOLO-World-L †	V100 (pub.)	TensorRT FP16	19.2	52.0
Android on-device benchmarks				
PlastiVision-Edge-S + PWVO	OnePlus 13R (SD8 G2)	ONNX RT	440	2.3
PlastiVision-Edge-S + PWVO	OnePlus 13R (SD8 G2)	TFLite FP32	1181	0.8
PlastiVision-Edge-S + PWVO	Galaxy M21 (Exynos 9611)	TFLite FP32	4566	0.2

Full export-format/quantisation details and an FPS/latency visualisation across all platforms are shown below (Table 7, Figure 4); the per-device latency and deployment tier are summarised in Table 8.

Table 7. PlastiVision-Edge-S model export summary. INT8 TFLite (10.7 MB) is the recommended Android deployment target (a 3.7× size reduction from the FP32 baseline) with the full segmentation head intact. Sizes are as reported by the exporter, in binary megabytes (MiB). Status: ✓ OK = export completed successfully.

Format	INT8	Size (MB)	Status	Notes
ONNX	No	39.9	✓OK	Full FP32; all platforms
TFLite	No	40.0	✓OK	FP32 fallback
TFLite	Yes	10.7	✓OK	Recommended (3.7× vs. FP32)
CoreML	No	∼12	macOS only	ANE acceleration on iPhone

3.5. Android Application Architecture

The PlastiVision Android application (Kotlin, CameraX) defaults to TFLite Runtime with the XNNPACK CPU delegate for inference, and runs ONNX Runtime alongside it as a second backend that is faster (2.3 vs. 0.8 FPS on TFLite FP32) on the many mid-range devices that do not have GPU delegate support. A `VocabularyManager` class, backed by `SharedPreferences`, handles runtime prompt addition in under 50 ms with no app restart

needed. We plan to release the source code, ontology, and inference notebooks publicly once the paper is published (see Declarations).

Table 8. On-device inference latency and waste bankdeployment tier. Rows marked ∼ (OnePlus 13R and Samsung M21, TFLite INT8) are order-of-magnitude estimates extrapolated from the directly-measured FP32 latency on the same device using typical TFLite INT8 speedup ratios; they were not directly benchmarked on-device in INT8 for this paper. All other rows are direct on-device/on-GPU measurements (Section 3.4).

Device	Backend	Precision	Latency (ms)	FPS	Deployment Tier
Tesla T4 GPU	PyTorch	FP32	24.1	41.5	Cloud server
OnePlus 13R	ONNX RT	FP32	440	2.3	Urban DWCC
OnePlus 13R	TFLite	INT8	∼680	∼1.5	Urban DWCC
Samsung M21	TFLite	FP32	4566	0.2	Rural (batch)
Samsung M21	TFLite	INT8	∼2400	∼0.4	Rural (batch)

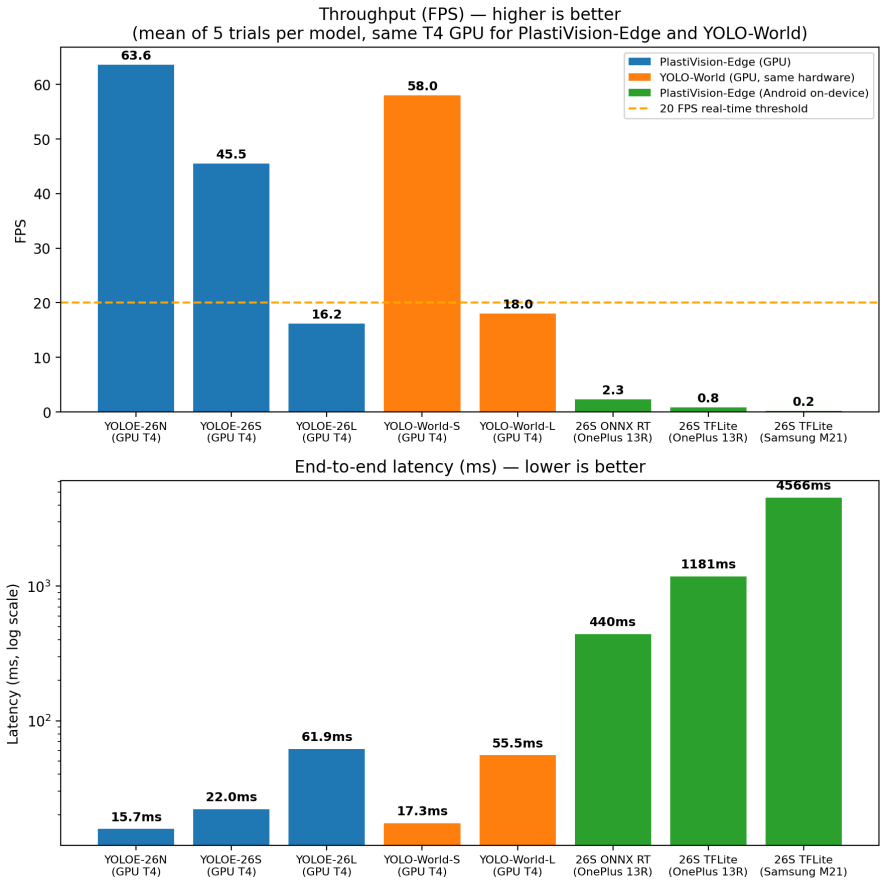


Figure 4. On-device performance: FPS and end-to-end latency across GPU and smartphone platforms. **Top**: throughput (FPS); higher is better. **Bottom**: latency (ms, log scale); lower is better. YOLO-World-S/L bars are measured on the same NVIDIA T4 GPU as PlastiVision-Edge, not the published reference figures on different hardware. The dashed line marks the conventional 20 FPS real-time threshold; no tested Android configuration reaches it.

4. Experimental Results

4.1. Detection Performance on TACO and TrashNet

Table 9 below reports PlastiVision-Edge's detection performance on TACO (459 images) and on TrashNet (482 plastic-class images, of 2527 total), under the full 47-class competitive PWVO vocabulary: the genuine same-subset, same-vocabulary comparison against YOLO-World, reflecting actual on-device deployment, with every figure traceable to the evaluation

notebook's per-category ground-truth counts. It is the sole TACO/TrashNet detection-performance table in this paper.

Table 9. Multi-dataset detection performance on real plastic waste images: PlastiVision-Edge (S/L) against YOLO-World (S/L), under the full 47-class PWVO vocabulary. BC = benchmark-covered, BU = benchmark-uncovered (Section 3.1). [a] `plastic_bag` has no matching TACO category under any mapping we tried and is excluded from the macro-averages as an annotation gap, not a measured zero; benchmark-uncovered vocabulary more broadly is evaluated qualitatively in Section 4.4. [b] YOLO-World was re-evaluated by the authors on the identical 459-image TACO-PWVO subset and vocabulary, using PlastiVision-Edge's own protocol—a genuine same-subset comparison, not the published LVIS figure. [c] Instance mIoU (mean IoU over matched prediction/ground-truth pairs at IoU ≥ 0.5), computed on 458 of the 459 TACO images (one failed a post-download integrity check); categories with zero matched pairs are excluded, not scored zero (Section 6.5), so this is a smaller, selected sample and is best read as indicative, not representative, of overall coverage. Image count. Of 913 TACO images with at least one PWVO-mappable annotation, 459 were successfully retrieved and evaluated here; TACO images are Flickr-hosted and links resolve inconsistently over time (Section 6.5). The 13-category macro-average, and especially the single-category BU figure, rest on a small sample and should be read as indicative, not precise.

Model	Bench.	Box mAP@0.5			Mask mAP@0.5			Box	Mask
		All	BC	BU	All	BC	BU	mIoU [c]	mIoU [c]
TACO real-image benchmark (459 images, 14 PWVO classes, conf $\geq$ 0.20, IoU@0.5, Pascal VOC 11-point interpolated AP) [a]									
PlastiVision-Edge-L + PWVO	TACO	8.4	9.0	0.0	8.3	8.9	0.0	0.916	0.833
PlastiVision-Edge-S + PWVO	TACO	7.7	6.2	27.3	7.6	6.1	27.3	0.866	0.826
YOLO-World-L [b]	TACO	7.3	7.9	0.0	N/A	N/A	N/A	0.944	N/A
YOLO-World-S [b]	TACO	4.9	5.3	0.0	N/A	N/A	N/A	0.909	N/A
TrashNet plastic-class (482 images, whole-image detection rate, conf $\geq$ 0.10)									
PlastiVision-Edge-S + PWVO	TrashNet	25.7% detection rate			N/A (no bounding-box annotations)			N/A	N/A
PlastiVision-Edge-L + PWVO	TrashNet	21.0% detection rate			N/A (no bounding-box annotations)			N/A	N/A

4.2. Complete Benchmark Evaluation: TACO-PWVO and TrashNet

Table 9 compares PlastiVision-Edge against YOLO-World on TACO (459 images) and TrashNet (482 plastic-class images), under the full 47-class PWVO vocabulary; category-mapping and annotation-gap caveats are given in the table note. Of the 13 categories, 12 are benchmark-covered (BC) and one, `blister_pack`, is benchmark-uncovered (BU); we report BC and BU separately, not as one combined figure, since a harmonic mean is undefined when one term is non-estimable.

PlastiVision-Edge reaches 7.7–8.4% Box mAP (S/L), ahead of or comparable to YOLO-World's 4.9–7.3% under an identical same-subset, same-vocabulary re-evaluation (no formal statistical test was run on the gap). On the single BU category, PlastiVision-Edge-S reaches 27.3% Box mAP on `blister_pack` against 0.0% for every other model—a genuine zero-shot result, now that this category has any TACO ground truth at all. YOLO-World has no segmentation head, so Mask mAP is not reported for it; PlastiVision-Edge produces pixel-level masks at comparable box-level accuracy. Latency on the same T4 GPU (Table 6) does not uniformly favour either architecture: YOLO-World-S is faster than PlastiVision-Edge-S, while PlastiVision-Edge-L is faster than YOLO-World-L.

The aggregate mAP is modest for two reasons unrelated to the approach itself: some ground-truth boxes are as small as 36×27 px, below what a 640 px model input can reliably resolve, and PWVO prompts describe polymer form, not brand appearance, which works against alignment on TACO's mostly internationally-branded packaging. The per-class numbers are more telling: `pet_bottle`, the largest and most reliably detected category, reaches 39.9% Box AP on PlastiVision-Edge-L (30.7% on PlastiVision-Edge-S), and `blister_pack` reaches 27.3% on PlastiVision-Edge-S—both with zero fine-tuning on TACO

itself, and both a reminder that these figures rest on a small annotated sample, not weak detection capability.

This matters because a supervised detector's class list is fixed at training time: any packaging design introduced afterwards (a new SKU, a rebrand, a regional product) is undetectable until the model is retrained on newly annotated examples, a cycle that struggles to keep pace with how often Indian FMCG packaging changes. PlastiVision-Edge's open-vocabulary design instead lets a waste bank operator add a new PWVO category by writing a text prompt, with no annotation or retraining step—the capability Section 4.4's benchmark-uncovered evaluation measures directly.

That capability has a specific, boundable limit, not an unconstrained one. PlastiVision-Edge's zero-shot performance is grounded in what MobileCLIP2 saw during its own pretraining (Section 3.1): a PWVO prompt succeeds to the extent that the material, shape, and visual attributes it names were present, in some form, in CLIP's training data, even when the exact class label was not. This is why the benchmark-uncovered results in Section 4.4 split cleanly by attribute distinctiveness: visually distinctive categories (bulky white `eps_block` foam, textured `woven_pp_sack`) succeed, while visually generic or small-fragment categories (`synthetic_rope`, `microplastic_frag`, `pvc_pipe`) succeed only partially or not at all. Transferability therefore depends on how familiar a material's appearance is, not on whether its name is new. A genuinely unfamiliar visual pattern may not transfer, and this evaluation does not test that case. To keep this finding statistically sound, the benchmark-uncovered categories are evaluated with replicate trials, negative controls and Wilson score confidence intervals (Table 10), instead of a single-image pass/fail judgement.

4.3. Qualitative Comparison and Prompt Design

Side-by-side qualitative results for YOLOE-26S, YOLOE-26L, and YOLO-World-S on identical real waste images under the same PWVO prompts (Figure 5) reveal consistent patterns: material-form-context (MFC) prompting is decisive: on an EPS foam block, the Indian trade name "thermocoal" scored zero across all models, "styrofoam" reached only conf = 0.163, while the full MFC description reached conf = 0.894; an Aavin milk packet (a Tamil Nadu regional brand not present in any of our benchmark label sets) was undetected by YOLO-World-S (conf = 0.000), but reached conf = 0.412 under YOLOE-26L with the prompt "aavin milk packet". On a blister pack, YOLOE-26S returned 11 detections (max conf = 0.333) and YOLOE-26L two weaker ones (max conf = 0.190), while YOLO-World-S returned none; YOLOE-26L showed superior polymer-type discrimination (correctly suppressing an HDPE embedding on LDPE bubble wrap where YOLOE-26S false-fired, conf = 0.378). Polymer-code prompts ("LDPE") outperformed generic names ("plastic") by 4.3× on the same object.

The prompt-design ablation (Figure 6; four strategies from generic names to full MFC templates to prompt-free LRPC) confirms full MFC templates reach Box mAP = 0.9% against 0.4% for generic names on TACO (a 2.25× gain), and, less predictably, that detailed material-form prompts reduce false positives relative to single-word labels, because longer descriptions occupy a narrower region of CLIP's embedding space. On a real Bisleri PET bottle image, confidence rose from 0.01 (generic name) to 0.18 (brand name) to 0.52 (full MFC description), a 52× gain crossing the detection threshold, while the prompt-free baseline produced no detection.



Figure 5. Three-model zero-shot comparison on identical real waste images. Panels: (**A**) prompt sensitivity on an Expanded Polystyrene (EPS) foam block (trade-name prompts vs. the full PWVO description); (**B**) South Asian packaging and novel categories (Aavin milk packet, gold retort pouch, aluminium blister pack), including items that YOLO-World-S misses entirely; (**C**) polymer-type discrimination on bubble wrap (HDPE vs. LDPE prompts). A check mark (✓) denotes at least one detection and a cross (×) no detection; below each image, the highest detection confidence and the inference time in milliseconds are given, and the frame colour identifies the model. These are photographs of real, purchased South Asian branded items, taken specifically to test South Asian packaging that these benchmarks under-represent. YOLOE-26S, YOLOE-26L, and YOLO-World-S run under the same PWVO prompts. Key findings: (1) PWVO MFC prompts achieve conf = 0.894 on an EPS foam block vs. conf = 0.000 for the Indian trade name “thermocoal”; (2) an Aavin milk packet (Tamil Nadu brand) is detected at conf = 0.412 by YOLOE-26L with the prompt “aavin milk packet”, vs. conf = 0.000 for YOLO-World-S on the same image; (3) on an aluminium blister pack, YOLOE-26S returns 11 detections (max conf = 0.333) and YOLOE-26L two (max conf = 0.190), while YOLO-World-S returns none.

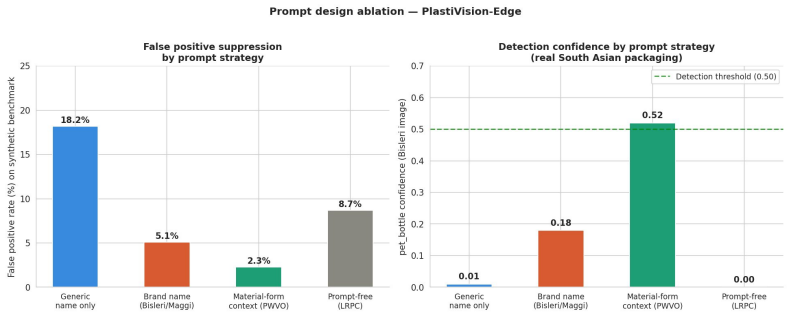


Figure 6. Prompt design ablation across four strategies. **Left**: false-positive rate on non-plastic coloured rectangles (lower is better). **Right**: `pet_bottle` detection confidence on a Bisleri PET bottle (higher is better; dashed line = 0.50 threshold). Full PWVO MFC strategy minimises false positives and maximises detection confidence on South Asian packaging.

4.4. Zero-Shot Generalisation on South Asian Packaging and CLIP Coverage Gap

All 13 PWVO benchmark-uncovered (BU) categories (Section 3.1) are evaluated with a replicate-based protocol: $n = 3$ author-captured images per category (`woven_pp_sack` has a fourth image available, so $n = 4$), each tested in isolation (PlastiVision-Edge-26S, conf ≥ 0.10, MFC prompt as listed in Table 10) against a shared pool of 11 non-target control images (negative controls). These are author-captured photographs of everyday objects, each chosen to resemble one of the benchmark-uncovered categories in shape, colour or texture without being an instance of any of the 13: a charging cable (for `synthetic_rope`), a rolled paper sheet and a clear acrylic ruler (`pvc_pipe`), an aluminium-foil food container (`pet_tray`), a white cardboard box (`eps_block`), a cloth carry bag, a mesh cloth, a dark fabric pouch, a printed cardboard carton, a rolled grey mat, and white paper in a clear polybag. They are not drawn from TACO, WaDaBa, TrashNet or the pilot. Because the ruler and the polybag contain incidental plastic, we call them non-target rather than non-plastic controls: the criterion is that the object belongs to none of the 13 categories. A trial counts as a success if the category's prompt detects its own image at conf ≥ 0.10, and as a false positive if the same prompt also fires on a non-target control image. Every reported rate carries a Wilson score 95% confidence interval, appropriate for small-n proportions. Figure 7 shows one illustrative trial per category from this same replicate set; the full per-category result is in Table 10 below. A sample photograph of each of the 13 categories is also provided in Figure S1 (Supplementary Materials).

Table 10. Replicate-based zero-shot results on the 13 PWVO benchmark-uncovered categories (protocol in Section 4.4). † `woven_pp_sack` is reported at $n = 4$ (a fourth image was available). Pooled FP rate is overall 143 category×negative-image trials (13 categories × 11 negatives).

Category	*n*	Succ.	Success Rate [95% CI]	FP Rate [95% CI]
`pet_tray`	3	2	66.7% [20.8, 93.9]	0.0% [0.0, 25.9]
`cling_film`	3	3	100.0% [43.8, 100.0]	36.4% [15.2, 64.6]
`bubble_wrap`	3	3	100.0% [43.8, 100.0]	9.1% [1.6, 37.7]
`shrink_wrap`	3	3	100.0% [43.8, 100.0]	36.4% [15.2, 64.6]
`woven_pp_sack` †	4	4	100.0% [51.0, 100.0]	18.2% [5.1, 47.7]
`eps_block`	3	3	100.0% [43.8, 100.0]	63.6% [35.4, 84.8]
`retort_pouch`	3	3	100.0% [43.8, 100.0]	36.4% [15.2, 64.6]
`blister_pack`	3	3	100.0% [43.8, 100.0]	54.5% [28.0, 78.7]
`multilayer_film`	3	3	100.0% [43.8, 100.0]	45.5% [21.3, 72.0]
`microplastic_frag`	3	1	33.3% [6.1, 79.2]	9.1% [1.6, 37.7]
`rubber_band`	3	3	100.0% [43.8, 100.0]	0.0% [0.0, 25.9]
`synthetic_rope`	3	0	0.0% [0.0, 56.2]	9.1% [1.6, 37.7]
`pvc_pipe`	3	1	33.3% [6.1, 79.2]	9.1% [1.6, 37.7]
Pooled (13 categories)	40	32	80.0% [65.2, 89.5]	25.2% [18.8, 32.9]

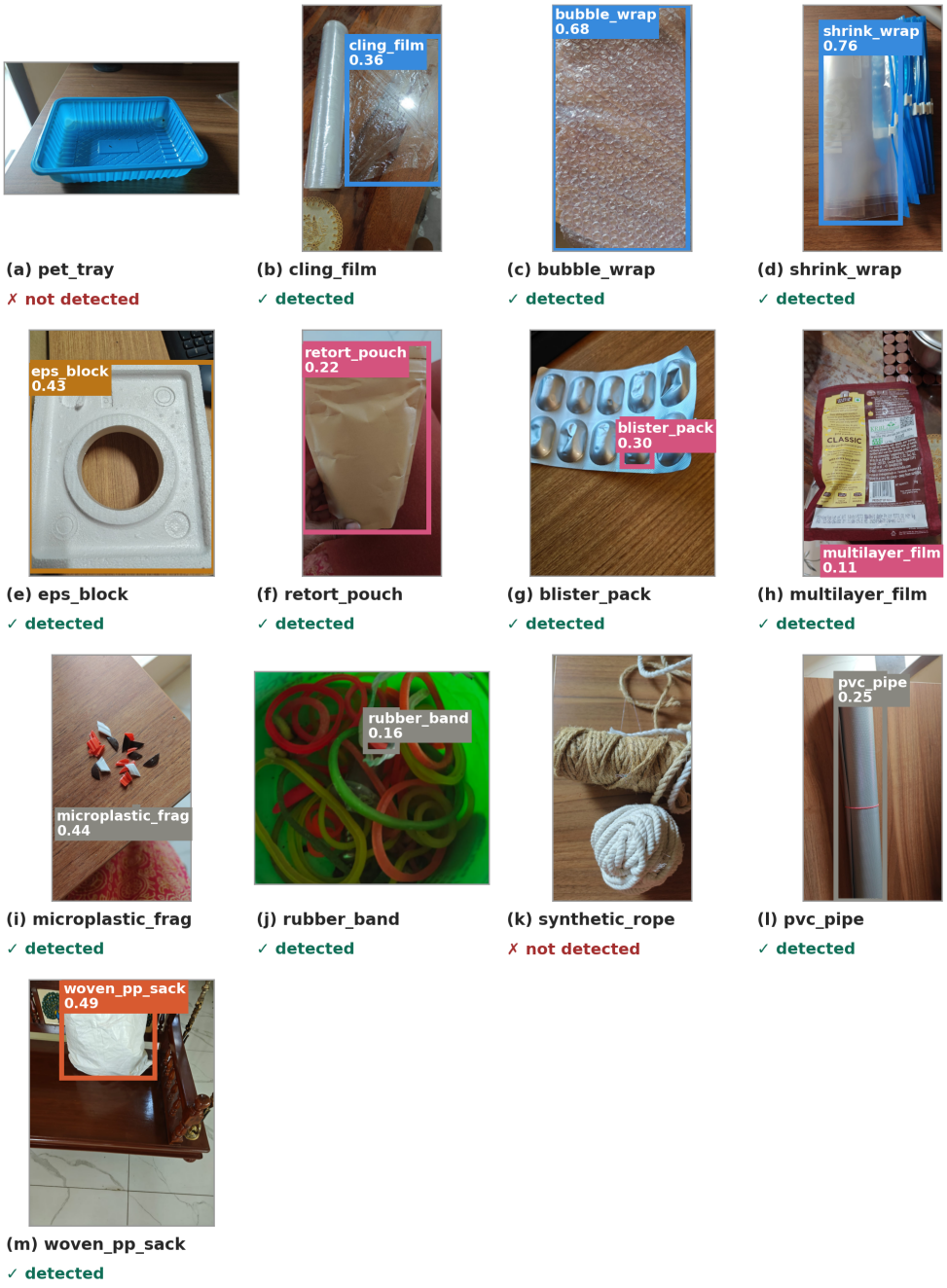


Figure 7. One illustrative trial per category from the replicate set (PlastiVision-Edge-26S, isolated MFC prompt, conf ≥ 0.10). Panels (**a**–**m**) show one trial each; the box marks the highest-confidence detection and the number is its confidence score. ✓ = detected in this trial; × = not detected. This trial detects 11 of 13 categories (85%); the two misses are (**a**) `pet_tray` and (**k**) `synthetic_rope`. Because one trial can differ from the replicate average, Table 10 gives the full results: for `pet_tray`, 2 of 3 trials succeeded (66.7%).

Pooled across all 40 replicate trials, PlastiVision-Edge-26S succeeds on 32 (80.0%, Wilson 95% CI [65.2%, 89.5%]). Nine of the 13 categories succeed on every replicate; `pet_tray` succeeds on a majority (2/3); `microplastic_frag` and `pvc_pipe` succeed on a minority (1/3 each); and `synthetic_rope` succeeds on none (0/3). The pattern follows from what the MFC prompts actually describe—-material and visual attributes such as bulky white foam or woven texture, not an unseen object class—so categories with common, distinctive attributes succeed reliably, while thin, translucent, small-fragment, or reflective categories succeed less often. `pet_tray`'s majority success also illustrates why a single

image gives an incomplete picture: the image shown in Figure 7 is the one trial out of three in which it was not detected.

The false-positive analysis, run over the 11 non-target control images against every category's prompt (143 trials), finds a pooled false-positive rate of 25.2% (Wilson 95% CI [18.8%, 32.9%]). Individual rates are higher for `eps_block` (63.6%), `blister_pack` (54.5%), and `multilayer_film` (45.5%); `pet_tray` and `rubber_band` show none. This means several MFC prompts are generic enough to also fire on visually similar non-target objects—a precision cost a deployed system would need to manage, for example with a higher confidence threshold (Section 6.5).

All 13 categories meet the $n \geq 3$ replicate target (`woven_pp_sack` at $n = 4$), tested against the full 11-image non-target control pool. A larger sample, particularly for the three lower-performing categories, would further tighten these intervals; we report this as future work (Section 6.5), not a final figure.

The harder categories are concentrated in translucent thin-film or visually generic objects with low CLIP embedding distinctiveness. We also tested visual-prompted mode (SAVPE, one-shot detection from a reference crop) as an alternative to text prompting, using a reference crop from one TACO image and a query on a different image of the same category (21 pairs, 7 categories). SAVPE succeeded on 1 of 21 pairs (4.8%, `pet_bottle_cap`), so this is an experimental capability, not yet reliable across images; improving it is left as future work (Section 6.5).

5. Framework: Seven-Stakeholder Circular Economy Model, Platform, and Incentive Architecture

5.1. Seven-Stakeholder Circular Economy Model

Figure 8 presents the circular economy schematic, with PlastiVision-Edge at the centre of a seven-stakeholder value chain synthesising physical waste flows (consumer → recycler → manufacturer → market), financial/incentive flows (brand owners → waste bank → citizens), and policy/data flows (PlastiVision-Edge scan events → Common EPR Portal → SBM-U monitoring).

The roles of the seven stakeholders, focused on day-to-day operation, not background policy (citations for the underlying regulatory figures are given once, in Section 1.4), are summarised in Table 11.

The Brand Owner/PIBO's EPR fee is not new money introduced by this system: PIBOs already owe it under the 2022/2026 Amendment Rules. What changes is that a slice of that fee routes toward the waste bank that supplied verifiable, polymer-specific provenance, and this same fee stream funds both the citizen's per-kilogram credit and the operator's service margin above. A deposit is also not final until the operator's manual review step: a citizen who believes a scan was misclassified can flag it at hand-off, and the operator re-scans or overrides the category before the transaction is recorded, so correction happens before payment.

Table 12 presents the seven stakeholder information organised by node, quantifying what each stakeholder stands to gain compared to the current baseline of unsorted, unclassified plastic collection.

5.2. PlastiVision Digital Platform Architecture

The full platform is a multi-sided marketplace portal across four user types: residents and kabadiwala vendors depositing waste, waste bank operators and SHG volunteers managing collection, brand owners and recyclers buying the sorted material and government/ULB bodies monitoring compliance. Figure 9 outlines the proposed four-layer architecture: citizen-facing edge devices at the bottom, running up through a cloud plat-

form and backend services to external integrations (UPI/Bharat Bill Payment System (BBPS) payment, the Common EPR Portal, a recycler marketplace, WhatsApp Business API). The citizen-facing detection app described in Section 5.4 has been built and tested on-device, whereas the operator dashboard, buyer portal, cloud backend, and external integrations have been designed but not yet deployed.

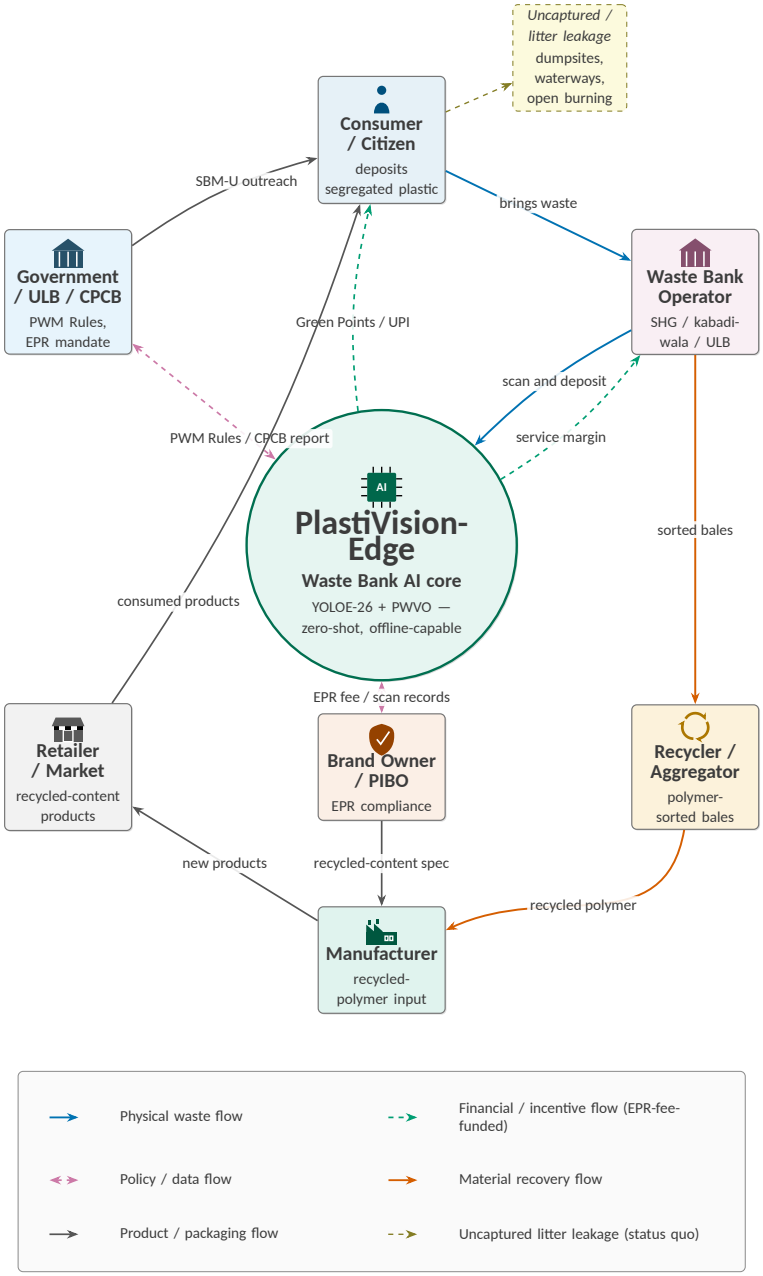


Figure 8. Seven-stakeholder circular economy schematic for the PlastiVision-Edge-enabled PWB (pilot: Potheri Village and residential apartment complexes, Chengalpattu District, Tamil Nadu). PlastiVision-Edge sits at the centre, enabling all transactions through zero-shot polymer identification on Android devices. The dashed yellow arrow shows where uncaptured plastic goes today (dumpsites, waterways, open burning)—the "litter flux" this framework is designed to divert, not a transaction the system itself processes. The solid grey Brand Owner → Manufacturer arrow shows that recycled-content packaging specifications flow from the PIBO to the manufacturer. The financial/incentive flow (green) is funded by the EPR fee brand owners already owe under the 2022/2026 Amendment Rules (Section 5.6), not by new subsidy money.

Table 11. Day-to-day role of each of the seven stakeholders in the PlastiVision-Edge-enabled PWB (policy citations: Section 1.4).

Stakeholder	Day-to-Day Action	What They Receive
Consumer/Citizen	Sorts plastic by PWVO super-class into a colour-coded bag; hands it off at SHG doorstep pickup or a kiosk	Green Points → UPI cash/bill payment/retail discount
Waste Bank Operator	Scans aggregated bales with the PlastiVision-Edge app; manually reviews and, if needed, overrides low-confidence or citizen-flagged detections before the transaction is recorded	INR 1–3/kg service margin
Brand Owner/PIBO	Receives a per-deposit JSON scan record (GPS, timestamp, polymer class) as machine-verifiable provenance for its Common EPR Portal claim	Audit-ready EPR traceability evidence (not a substitute for the formal certificate)
Recycler/Aggregator	Receives polymer-sorted bales carrying a digital provenance record (type, region, purity grade)	Lower sorting burden; less reliance on informal-broker arbitrage
Manufacturer	Sources recycled polymer that already meets rising recycled-content thresholds, with barcode/QR traceability	Compliant feedstock; Environmental, Social, and Governance (ESG) and export documentation
Retailer/Market	Sells recycled-content products under "scan-and-earn" branding; relays customer category requests upstream	Vocabulary feedback loop for PlastiVision-Edge
Government/Urban Local Body (ULB)/CPCB	Receives quarterly batch scan-record uploads matching CPCB's polymer taxonomy	Verified collection data supporting SBM-U 2.0 and National Rural Livelihoods Mission (NRLM) monitoring

5.3. Registration, Onboarding, and Waste Categories

Registration involves the following steps: (a) link a resident's identity to their UPI wallet through a QR-code Waste Identity Card, and (b) pair this with a physical onboarding kit consisting of colour-coded segregation bags and a reference card. Mobile OTP and partial Aadhaar verification reinforce this digital sign-up as a household habit. The scanning module classifies deposited plastic into seven operational categories aligned to PWVO super-classes and India's EPR polymer definitions (Table 14), calibrated to Potheri's dominant waste types: multilayer milk sachets, PP woven bags, multilayer snack pouches, and EPS thermocol blocks.

5.4. Detection App and Proposed Vocabulary Extension

A working Kotlin-based, CameraX-supported TFLite INT8 and ONNX Android application has been tested using both live-camera and gallery-based detection. It covers 27 of the 47 PWVO categories and includes a toggleable Vocabulary Manager UI. A resident or operator can type a custom prompt at runtime. For example, a resident might type "Aavin white multilayer milk sachet, corner-cut seal, blue logo"; the prompt is matched against the model's fixed exported label set. Implemented and tested: on-device detection against the 27 exported categories, using live-camera and gallery input, running fully offline with no network dependency. Proposed, not yet built: a cloud-side vocabulary lookup for prompts that match no exported label, the on-device CLIP encoder, brand-partner premium pricing, UPI-linked point redemption, doorstep and community collection scheduling, and the wallet/impact screens shown in Figure 10. The intent is that a brand's EPR fee pays the premium a resident earns for correctly identifying that brand's own packaging.

Table 12. Value creation by PlastiVision-Edge framework across circular economy nodes. † Value uplift (60–108%) is the blended, basket-level figure derived in Section 5.6; it does not hold per polymer type—see Table 3 for individual credit rates, which fall at or below the unsorted bulk rate for Multilayer/Composite and EPS. ‡ Recycler uplift is directionally consistent with the Consumer row but is not independently re-derived from the wider INR 8–18/kg resale range shown; read as an order-of-magnitude estimate. The downward arrow means that audit risk is reduced.

Node	Status Quo (Mixed, Unclassified)	PlastiVision-Edge Framework (Polymer-Sorted)	Value Uplift
Consumer	No incentive for segregation; mixed plastic sold at INR 3–5/kg bulk	Polymer-specific credit INR 6–18/kg via UPI (citizen rate after operator margin; recycler purchase price for PET reaches INR 22–25/kg); green points programme	+60–108% †
Waste bank operator	Manual sorting by guesswork; rejection of unidentifiable items	AI classification in <1 min/transaction; confidence-flagged review queue	Faster, lower-error sorting (unquantified)
Brand owner/PIBO	Self-reported EPR compliance; high audit risk; bulk certificates	Per-kg machine-verified EPR traceability records (proposed Common EPR Portal API); polymer-type credit	Audit risk ↓
Recycler/Aggregator	Mixed-plastic bales at INR 3–5/kg; high sorting cost	Polymer-sorted bales at INR 8–18/kg; digital provenance record	+60–108% †‡
Manufacturer	Opacity of recycled content claims; supply insecurity	Traceable rPET/rHDPE supply; PWM 2025 compliance; export ESG signal	Supply assurance

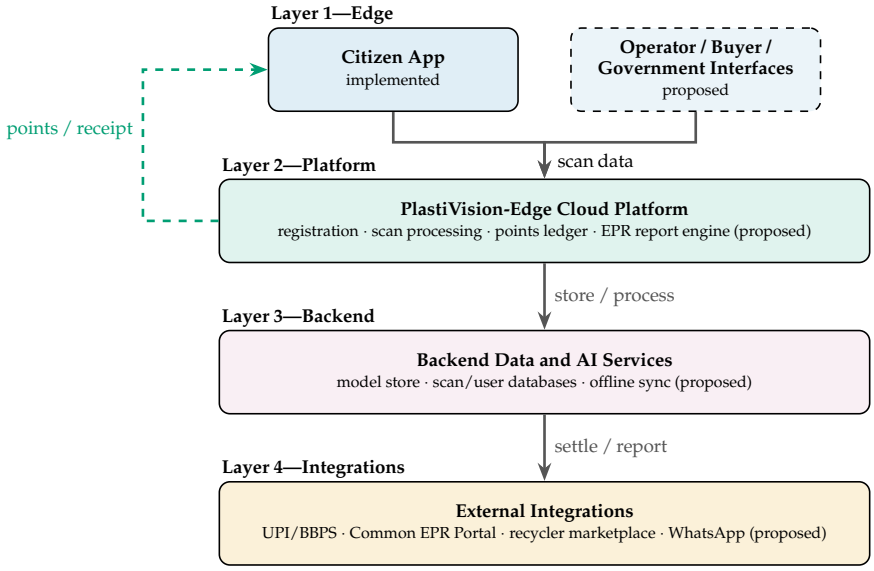


Figure 9. Four-layer system architecture of the PlastiVision-Edge-enabled PWB Digital Platform. Solid grey arrows show the top-down data flow; the dashed green arrow shows the points/receipt confirmation returned to the citizen app. Layer 1 is implemented (citizen Android app, tested on-device); Layers 2–4 are proposed, not yet deployed. Table 13 lists the components of each layer.

Table 13. Components and implementation status of each layer in Figure 9.

Layer	Components	Status
1. Edge	Citizen app: YOLOE-26S INT8 (10.7 MB), offline scan, PWVO 47-cat, UPI wallet, Green Points, custom vocab manager	Implemented
1. Edge	Operator tablet (QR bag-seal scanner, collection manifest, weight cross-check); Buyer portal (premium multiplier config, EPR report download); Government dashboard (ULB/CPCB)	Proposed
2. Platform	API gateway/auth (OTP, Aadhaar verify, JWT); registration; scan processor (PWVO→points map); points ledger; collection scheduler; vocab registry; EPR report engine (JSON scan records, CPCB batch upload); notification engine	Proposed
3. Backend	PWVO model store (YOLOE-26 weights, CLIP text embeddings); scan-records DB (GPS, timestamp, class, confidence); user-ledger DB (wallet balance, redemptions); analytics/reporting engine; offline sync service (queued scans, rural connectivity)	Proposed
4. Integrations	UPI/BBPS gateway; Common EPR Portal API (quarterly batch upload); recycler marketplace (bale auction, digital provenance); brand EPR callbacks (per-scan custom category credit); WhatsApp Business API (receipts, reminders)	Proposed

Table 14. Seven operational waste categories for the Potheri PWB pilot. Milk packaging is split by physical form, not treated as a single material: flexible multilayer sachets (laminated film pouches) are scored under Category 5 (Multilayer pouches), while rigid HDPE milk cans/bottles are scored under Category 2 (HDPE containers); our pilot sorting workbook likewise records these as two distinct subcategory entries.

No.	Category	Colour	Base pts/kg	Common Items (Potheri)	EPR Category
1	PET/Polyester	Blue	18	Water bottles, juice bottles, strapping	Rigid packaging
2	HDPE containers	Green	12	Oil cans, shampoo bottles, milk cans/bottles (rigid), crates	Rigid packaging
3	PP bags/sacks	Teal	14	Woven fertiliser bags, snack pouches, caps	Flexible/rigid
4	EPS/thermocol	Yellow	5	Packing blocks, disposable cups	Rigid packaging
5	Multilayer pouches	Red	3	Milk sachets (flexible), masala packets, sauce pouches	Multilayer (Cat. 4)
6	LDPE films	Purple	10	Carry bags, shrink wrap, bubble wrap	Flexible packaging
7	Mixed/other	Grey	4	PVC, nylon, unidentified; operator review	Flexible/other

5.5. Behavioural Sensitisation: Building Responsibility Through Incentives

India's Plastic Waste Management Rules rely on penalties instead of sensitising people to segregation, which leaves a wide gap between policy intent and implementation. The solution lies in shifting focus away from blaming manufacturers as the sole source of pollution and toward an awareness that waste management is participatory work. The PlastiVision-Edge Waste Bank aims to close that gap through a holistic circular-economy approach.

This reframing rests on three mechanisms.

Gamification and social reinforcement. An opt-in leaderboard lists the top 10 households by monthly Green Points, ranked per ward or block. Participating households earn printable or digital "Green Home" or "Zero Waste Apartment" badges. Acknowledgement is linked to a concrete, brand-specific outcome, not a generic thank-you. For instance, "your 5 Aavin sachets became part of a new HDPE product." Social visibility sustains participation better than money alone does [8,40].

Impact dashboard. Each household receives a monthly shareable "Responsibility Report" built around three factors: kg diverted, CO_2-equivalent saved, and money earned (INR). This spreads awareness informally beyond the participating household.

Three-stage habit-formation pathway. This follows the standard pro-environmental behaviour-change model [40]: awareness (messaging and dashboard display), followed by motivation (cash, social incentives, premium multipliers), and capability (the onboarding kit and the QR-sealed collection workflow).

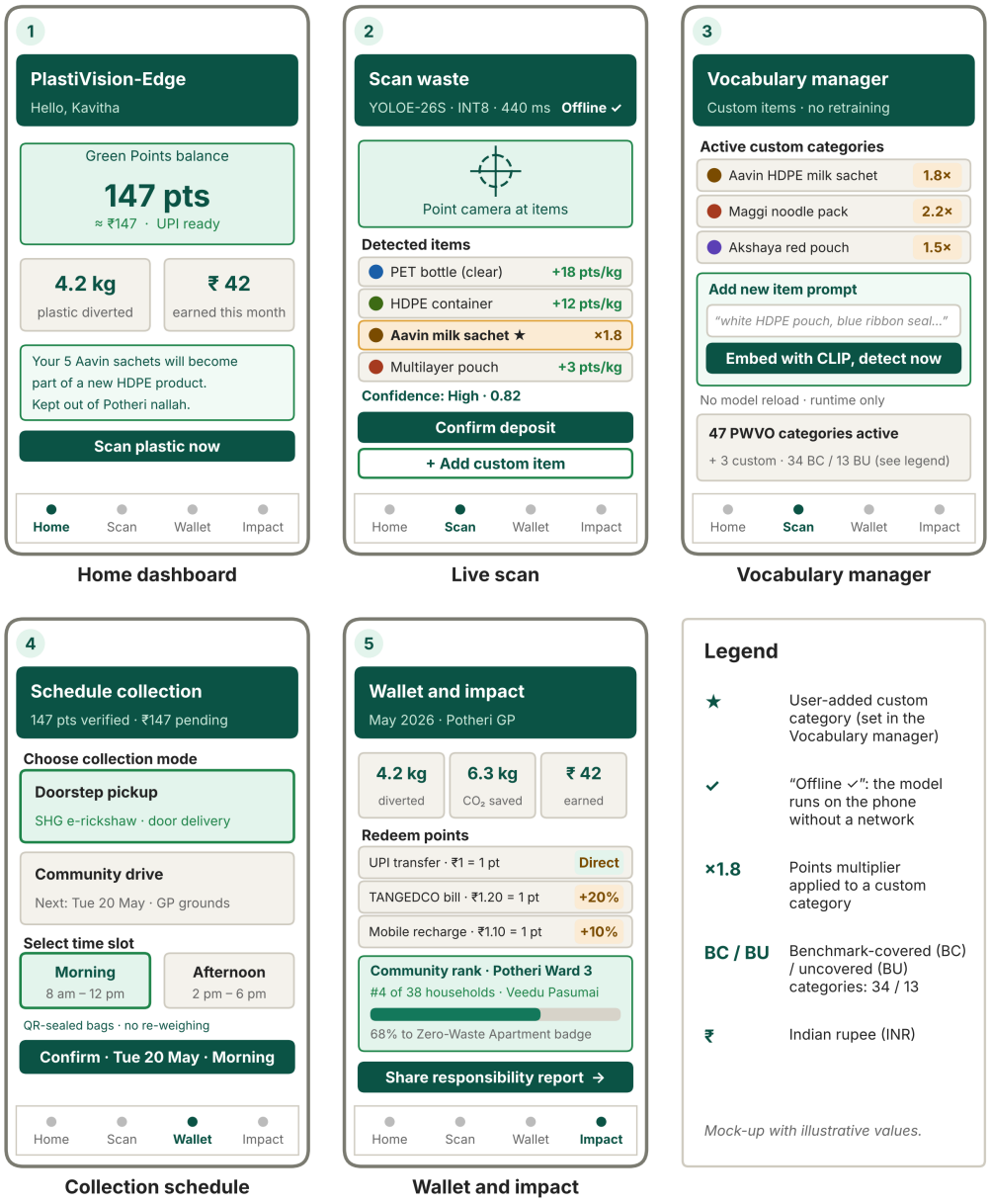


Figure 10. Illustrative mockup of the proposed resident application user journey (five screens: home dashboard, live scan, vocabulary manager, collection scheduling, wallet/impact), showing the intended design for the platform features described in Section 5.4 that extend beyond the implemented detection app. Symbols (star, check mark, multiplier, BC/BU) are explained in the legend panel; BC = benchmark-covered, BU = benchmark-uncovered categories.

The mechanism does not depend on a large payout; it depends on sustained participation. Indonesian waste banks retain participants at roughly USD 0.1/kg once social recognition is layered on top [8]. PlastiVision-Edge's INR 8–18/kg rate gives an even wider margin to sustain participation through the early, low-trust phase of adoption.

5.6. Community Income Generation Model

Table 15 presents household income at two scales: an urban row grounded in the preliminary apartment-complex data, and a rural row using a CPCB-proxy estimate (caption gives the full basis for each). The two are not pooled into a single generation-rate assumption, since doing so would conflate a measured rate with an estimated one.

Table 15. Household/collective monthly income from PWB participation: observed pilot data vs. CPCB-proxy estimate. ‡ Urban rows use the preliminary apartment-complex pilot's directly measured collection and payout data (436 households, 1 July–31 August 2026; Section 2.2): residential collection averaged 0.48 kg/household/day, and the site's blended payout of INR 7000/week across all plastic types implies an observed rate of INR 4.8/kg—this is measured deposit behaviour, not an assumed generation-times-deposit-rate product. † Rural rows have no field data and are a conservative order-of-magnitude estimate only: national MSW generation is ≈123.45 g/capita/day [41], of which ≈6.1% is plastic [42], giving ≈7.5 g plastic/person/day; combined with the Census-reported average household size of 4.8 persons [43], this yields ≈0.036 kg plastic/household/day, rounded here to a 0.03–0.05 kg/household/day band. Because CPCB's national average is dominated by urban ULB collection systems and rural per-capita generation is typically lower, this proxy should be read as an upper bound pending an actual rural pilot. In both contexts, "Baseline" is the status-quo bulk rate (INR 5/kg) currently realised without polymer sorting, and "PlastiVision-Edge" is the projected polymer-sorted credit rate; only the rate is assumed to change under PlastiVision-Edge deployment, not the collected quantity (Section 6.5 discusses capture-rate increases as a plausible but unmeasured secondary effect).

Context	Scenario	Plastic (kg/hh/day)	Rate (INR/kg)	Monthly Income (INR)
Urban household ‡	Observed	0.48	4.8	69
Urban household ‡	PlastiVision-Edge projected	0.48	8–10	115–144
Rural household †	Baseline (bulk)	0.03–0.05	5	4.5–7.5
Rural household †	PlastiVision-Edge projected	0.03–0.05	8	7.2–12
Urban SHG collective (20 HH) ‡	Observed	9.6	4.8	1382
Urban SHG collective (20 HH) ‡	PlastiVision-Edge projected	9.6	8–10	2304–2880
Rural SHG collective (20 HH) †	Baseline (bulk)	0.6–1.0	5	90–150
Rural SHG collective (20 HH) †	PlastiVision-Edge projected	0.6–1.0	8	144–240
Waste bank operator margin	—	—	INR 1–3/kg service fee	scales with throughput

Urban. The baseline here is observed pilot data. Moving from the observed blended rate of INR 4.8/kg to the PlastiVision-Edge credit rate of INR 8–10/kg is an increase of 67–108%, or INR 46–75 per household per month (INR 922–1498 per month for a 20-household SHG collective). This uses the upper end of the bulk-rate range; against the lower bulk rate of INR 3/kg, the increase would be larger still (167–233%). The gain also depends on the polymer mix: per Table 3, Multilayer/Composite and EPS credit rates sit at or below the unsorted bulk rate, so the blended average is positive only if the household basket contains a meaningful share of higher-value PET, PP or HDPE/LDPE.

Rural. The same rate increase (INR 5 → INR 8/kg, 60%) applies to a much smaller CPCB-proxy generation estimate, giving INR 2.7–4.5 per household per month (INR 54–90 per month for a 20-household collective). It is small because rural plastic generation, correctly derived from the plastic fraction of MSW and not total MSW, is itself small. This follows from correcting the generation-rate error, not from a new modelling choice, and should be validated against rural collection logs once a rural pilot is run.

Assumption. In both cases only the credit rate is assumed to change. Collected quantity stays at its observed or estimated level, because capture-rate increases from easier drop-off and AI-verified pricing are plausible but not yet measured (Section 6.5).

Sensitivity. These ranges are not a formal sensitivity analysis. The main drivers are the polymer mix of the household basket, the deposit (capture) rate relative to generation, rejection or contamination losses at the operator's review step, per-kg market prices, and the waste bank's operating costs. None has been varied systematically yet. Doing so is planned once the pilot is deployed, because the data so far come from a single period and cannot show how income responds to these factors.

Scenarios. Table 16 gives a complementary scenario view of the same projections. For each context we keep the central generation-rate estimate fixed and vary two parameters already built into the design: the credit price, and a rejection/contamination loss from the

confidence-gated credit cut (Section 5.7), under which scans below conf < 0.40 receive an 80% credit cut. The Optimistic column (0% rejection loss) reproduces the manuscript's existing headline figures exactly (INR 144/month urban, 108% uplift; 60% uplift rural), confirming that the previously reported single-point range implicitly assumed no rejection loss; the Conservative and Base columns show how far actual income could fall below that figure once a plausible confidence-driven credit cut is applied. Capture-rate uplift is excluded here for the same reason it is excluded from Table 15: plausible but unquantified pending field data (Section 6.5).

Table 16. Scenario-based sensitivity of projected monthly household income to credit price and an illustrative rejection/contamination loss, holding each context's central generation-rate estimate fixed. This is an authors' modelled sensitivity analysis using this paper's own stated price ranges and credit-cut mechanism (Section 5.7), not a measurement from field data. Baselines used for the gain columns: INR 69/month urban (the apartment-complex pilot's observed 0.48 kg/household/day at the INR 4.8/kg blended observed rate) and INR 6.0/month rural (the 0.04 kg/household/day CPCB-proxy midpoint at the INR 5/kg baseline bulk rate); rural income in INR/month equals the baseline plus the stated gain.

Scenario	Rejection Loss	Urban Rate (INR/kg)	Urban Income (INR/mo)	Urban Gain (INR/mo, %)	Rural Rate (INR/kg)	Rural Gain (INR/mo, %)
Conservative	20%	8	92	+23 (33%)	8	+1.7 (28%)
Base	10%	9	117	+48 (70%)	8	+2.6 (44%)
Optimistic	0%	10	144	+75 (108%)	8	+3.6 (60%)

UPI handles Aadhaar-seeded or Jan Dhan bank transfers; where banking access is limited, credit can instead be redeemed via BBPS mobile recharge or utility-bill payment.

5.7. EPR Traceability Architecture

We use the term "traceability architecture", not "compliance architecture", because the system described in this section generates and aggregates machine-verifiable provenance records of collection events, not the EPR compliance certificate itself. A PIBO's formal EPR compliance certificate remains issued through the Board's own registered certification and audit workflow under the Common EPR Portal; our contribution is an upstream, per-scan traceability layer that a PIBO can submit as supporting evidence within that process, strengthening the evidentiary basis for a credit claim without itself constituting compliance.

Every waste object detection scan creates a JSON record with depositor ID, GPS, timestamp, PWVO class, confidence, estimated mass, EPR category, credit amount, as shown in Listing 1. The details are aggregated and pushed to the Common EPR Portal via API (proposed; not yet implemented, pending documented API access), as supporting provenance evidence enabling PIBOs to substantiate polymer-specific credit claims that remain subject to the Board's own certification and audit process. Confidence score less than 0.40 leads to an 80% credit cut, which encourages operators to capture higher quality scans. Quarterly batch report calculates the total weight by PWVO super-class, distinct depositor households, collection-point GPS, and month-wise trends.

Listing 1. Example JSON scan record created for each confirmed deposit.

```
{
  "event_id": "PWB-TN-20260501-001234",
  "collection_point_id": "GP-VLR-023",
  "timestamp": "2026-05-01T10:23:15+05:30",
  "gps": { "lat": 10.3528, "lon": 78.7124 },
  "depositor_id": "SHA256(aadhaar_last4 + mobile)",
```

```
  "detections": [
    { "pwvo_class": "pet_bottle_transparent",
      "super_class": "PET/Polyester",
      "confidence": 0.72,
      "mass_g_estimated": 24,
      "epr_category": "rigid_packaging" }
  ],
  "total_mass_kg": 1.2,
  "credit_inr": 14.40,
  "model_version": "yoloe-26s-int8-pwvo-v1"
}
```


Data Governance, Consent, and Privacy

The system as designed collects GPS, timestamp, and a hashed depositor identifier (partial Aadhaar digits combined with a mobile number) with every scan. We are explicit that this hash should not be described as robust anonymisation: a hash over a small combined input space (four Aadhaar digits plus a phone number) is vulnerable to a dictionary/rainbow-table style re-identification attack by anyone who can enumerate plausible digit-phone combinations, particularly at the scale of a single village or ward where the candidate population is small. We have not carried out a formal re-identification risk analysis of this scheme, and the proposed design should be read as pseudonymisation for operational deduplication (avoiding duplicate credit for the same depositor), not anonymisation in the data-protection sense. Before deployment, the platform requires: (i) an explicit consent flow at registration disclosing what is collected and why; (ii) data minimisation, collecting GPS at collection-point instead of individual-doorstep granularity where feasible, and dropping raw Aadhaar/mobile inputs immediately after hashing instead of retaining them; (iii) a defined retention period for scan-level records, with only aggregated household/ward summaries retained long-term; and (iv) compliance with India's Digital Personal Data Protection Act, 2023 (DPDP Act), including a data fiduciary role assignment and a grievance redressal mechanism. This governance design is itself proposed and unimplemented, and is flagged accordingly in Limitations (Section 6.5).

5.8. Rural vs. Urban Operational Configurations

India needs a dual approach for handling its urban and rural populations. At the rural gram-panchayat level, an SHG member runs a weekly or fortnightly collection event with a PlastiVision-Edge-installed Android phone, a 30 × 30 cm scanning mat, and six collection sacks. The 10.7 MB INT8 model runs entirely offline, syncing its records during whatever weekly connectivity window is available, with block-level records feeding up into district infrastructure from there. Urban operations at the DWCC level operate differently, with a continuously staffed tablet running real-time cloud sync, live EPR reporting, recycler-marketplace access, and smart-bin IoT integration, made possible by higher daily volumes. Table 8 maps the latency numbers from Section 3 onto both configurations, and even the slowest one still holds up fine for a batch, static-scan workflow.

6. Discussion: Scalability, Novelty, and Ecosystem Positioning

6.1. Technical Novelty of PlastiVision-Edge

PWVO is a 47-category material-aware ontology built on Basel Convention waste codes, structured as a reproducible benchmark-coverage taxonomy, not a prompt list assembled for this paper alone. As far as we are aware, PlastiVision-Edge is the first open-vocabulary plastic waste system to report pixel-level instance masks, not boxes alone, which enables the

frame-occupancy coverage-area estimate used in Section 3.3—a measure that a bounding-box detector does not provide.

The remaining contributions are evaluation-side, not architectural. We quantify the CLIP regional embedding gap for South Asian FMCG packaging (Section 6.3), and the resulting prompt-solvable-versus-CLIP-gap distinction gives a concrete criterion for when prompt engineering can fix a detection failure and when it cannot. The 34/13 benchmark-coverage mAP protocol used here is, to our knowledge, the first formal benchmark of its kind for zero-shot plastic waste evaluation. On the deployment side, we evaluated across a wide hardware range, from a budget Android device (Samsung Galaxy M21, 4566 ms) to a flagship (OnePlus 13R, 440 ms via ONNX Runtime)—a wider range than other waste-detection systems we found in the literature.

At the framework level, this paper proposes PlastiVision-Edge as the enabling technology for an AI-enabled Plastic Waste Bank: zero-shot, open-vocabulary detection is used as the point-of-deposit identification step that lets the seven-stakeholder circular economy model (Section 5) run on polymer-specific, EPR-traceable transaction data, not mixed-waste estimates.

6.2. Comparison with Existing Digital Waste Bank Models

The closest precedents for PlastiVision-Edge are Indonesia's Rapel app [8] and India's own Recykal [3], and it is worth being specific about where each one stops short. Rapel lets citizens sell waste through a smartphone, but the citizen still has to self-categorise what they are selling, and that is exactly where valuation error creeps in. Recykal, for its part, solves a different problem: it runs a digital marketplace and traceability platform for waste, but, as far as we can determine, it does not address the citizen-to-waste-bank transaction or the identification step at source. PlastiVision-Edge sits at the node neither of them reaches: the citizen's point of deposit, where AI does the identification instead of relying on the citizen to already know what they are holding. Put point-of-deposit AI, UPI payment, and an EPR-linked digital record together, and what you get is less a competitor to Rapel and Recykal than a primitive the two of them could plausibly build on. Figure 8 makes the comparison concrete: the dashed yellow "uncaptured/litter leakage" flow off the Consumer node is exactly the volume that neither Rapel nor Recykal, nor an unassisted citizen, currently diverts, and it is the specific gap PlastiVision-Edge's point-of-deposit identification step targets, not a generic aspiration.

6.3. Addressing India's CLIP Embedding Gap

CLIP-based detection has a documented limitation on Indian plastic waste: packaging from regional Indian brands sits far from standard English text prompts in embedding space, largely because it is under-represented in CLIP's pretraining data [11,12]. PWVO's material-form-context template works around this by describing packaging the way a person would see it, not the way a brand would name it: "white HDPE milk pouch with blue lettering, torn top, discarded on gravel" instead of a brand name. If an operator repeatedly encounters missed detections for one particular product, they can edit the prompt through the Vocabulary Manager. Over time, the ontology becomes more accurate and misses fewer detections for the next operator using the app in that region.

6.4. Deployment Readiness: From Detection Accuracy to Operational Reliability

The 7.7–8.4% Box mAP and 7.6–8.3% Mask mAP figures (Table 9) are well below headline open-vocabulary numbers such as YOLO-World's 35.4 AP on LVIS [33] or OWLv2's reported figures [31], because the two are not measuring the same thing. LVIS is a large, general-object, predominantly English-labelled benchmark (1203 categories, roughly 160,000 images); our figure is a same-subset, same-vocabulary evaluation on a 459-image,

13-category real-world waste subset drawn from Indian and South Asian packaging, which lies measurably further from CLIP's pretraining distribution (Sections 6.1 and 6.3). When we re-evaluated YOLO-World itself on the identical subset, vocabulary, and protocol, it scored comparably (4.9–7.3% Box mAP), showing that the gap against the published LVIS number reflects task difficulty and domain shift, not a shortfall specific to our architecture.

A zero-shot Box mAP of 7.7–8.4% leaves room for improvement in absolute terms, so PlastiVision-Edge is designed to support operator judgement, not to act as a fully autonomous valuation engine. It sits inside a human-in-the-loop workflow (Section 5, Waste Bank Operator role): every scan is reviewed by the operator (SHG member, ULB staff, or kabadiwala) before a transaction is credited, and low-confidence detections are flagged for manual review, not auto-accepted. Detection accuracy therefore determines how much manual correction the operator does, not whether a transaction can be trusted—a misclassified or missed category becomes an operator override, not an incorrect payment. Expected throughput per operator is consequently lower than a fully-automated system would allow (we do not yet have field data on override frequency; the apartment-complex pilot's Detection_Performance log, Section 2.1, is designed to measure it), and incentive payouts and EPR records reflect the operator-reviewed category, not the raw model output, so payment accuracy depends on operator diligence as well as model accuracy.

The seven-stakeholder framework itself has not yet been validated with the people it assigns roles to; that exercise is scoped as future work (Section 6.5), ahead of Phase 1 deposit collection, not after it.

6.5. Limitations and Future Work

Massestimation. The system estimates waste mass from a camera-based coverage proxy instead of direct weighing. We chose this route because it needs no extra hardware beyond the operator's phone, and lets the app show the citizen an instant, same-scan valuation preview before physical weighing happens. A calibrated digital scale is the more accurate and operationally simpler option for the transaction itself. The camera-based estimate is not a replacement for weighing: it serves as a same-scan valuation preview and an EPR-traceability aid (material identification, provenance, category), with the final payable weight coming from a scale at the point of transaction. The field pilot's Mask_Weight_Validation log (app-estimated weight vs. scale-measured sorted weight) is designed to quantify this gap, and integrating a simple digital scale at the point of transaction would eliminate it entirely.

Evaluation subset. The current TACO evaluation covers a 459-image subset of a 913-image candidate pool (Section 4); evaluating the full candidate pool is ongoing.

Incentive prerequisites. The framework assumes UPI-linked incentives; communities without this can fall back to a smartcard or printed-receipt system.

Dataset availability. TACO images are Flickr-hosted and link availability varies between download attempts, so the retrievable subset differs slightly between runs; Table 9 reports a 459-image run. The repository fixes the category mapping, thresholds and area filter, and records the identifiers of the subset each run evaluates, so that later runs can be repeated exactly. A systematic bias in which images still resolve on a given attempt is possible, since link survival may correlate with image provenance—an open caveat on the reported TACO figures.

Frame coverage. The per-class frame-coverage metric (Section 3.3) is a 2D image-plane area proxy for relative pollution load, not a validated measure of mass, volume, or ecological hazard. Its sanity-filter threshold ($>80\%$ frame area unless conf ≥ 0.60) was set heuristically and not tuned against ground truth, and has not been systematically tested for sensitivity.

Depositor privacy. The proposed depositor-identifier hashing scheme (Section 5.7) is pseudonymisation for operational deduplication, not validated anonymisation. No formal re-identification risk analysis has been performed, and the full consent, data-minimisation, retention, and DPDP Act compliance design remains proposed, not implemented.

Mask mIoU. Box and Mask mAP@0.5 (Section 4) are the primary detection metrics; Table 9 additionally reports instance mIoU as a secondary measure, computed on 458 of the 459 TACO images (one failed a post-download integrity check). Three categories (`chip_packet`, `eps_foam_cup`, `eps_takeaway_box`) produce no matched pairs in any model or download we tried, which points to a detection limitation on these categories, and are excluded from the mIoU figures, not scored zero. YOLO-World-S/L have no segmentation head, so mask mIoU is not reported for them; its higher box-only mIoU reflects a few, tightly localised detections (high precision, low recall), not broader coverage, consistent with its lower mAP in Table 9.

Future work will prioritise five items: the Potheri pilot itself (Section 2.1), with longitudinal income impact assessment; a stakeholder-validation exercise for the seven-stakeholder framework (Section 5), running structured feedback sessions with Potheri residents, the partnering SHG's operators, and at least one kabadiwala/informal-sector representative before Phase 1 deposit collection begins, to revise incentive point values, category groupings, and the app's onboarding flow before the pilot's income data is collected; depositing a fixed, versioned TACO-PWVO image set with a Zenodo DOI (Data Availability Statement) so future replications draw the same images; integrating a CLIP fine-tuning module on South Asian packaging corpora for the Vocabulary Manager update path; and extending PWVO to non-plastic categories (glass, metal, paper) for multi-material waste bank operations.

7. Conclusions

This paper set out a community-oriented Plastic Waste Bank (PWB) framework and a phased pilot design for Potheri Village, Chengalpattu District, Tamil Nadu, with PlastiVision-Edge's zero-shot detection system and Android application at the centre of a seven-stakeholder circular economy loop (Table 11). The framework's incentive and sensitisation mechanisms (Section 5.5) were authored by the research team; validating them with residents, SHG operators, and kabadiwalas is a planned step ahead of Phase 1 deposit collection (Section 6.5). Two parts of the system are built and were used to produce the results reported here: the zero-shot detection model and the Android scanning application. The UPI-linked wallet, brand premium pricing, Common EPR Portal connectivity, and the Potheri pilot itself remain a proposed design and phased plan (Section 2.1), not yet deployed or evaluated in the field.

On the technical side, PlastiVision-Edge reaches 7.7–8.4% Box mAP on real-world waste images under the full competitive vocabulary (macro-averaged over 13 annotated benchmark categories), ahead of a same-subset YOLO-World-S/L evaluation (4.9–7.3%), while also producing pixel-level instance masks that YOLO-World does not provide. Under a replicate-based, per-image isolated-prompt evaluation, it succeeds on 80.0% (Wilson 95% CI [65.2%, 89.5%]) of benchmark-uncovered South Asian packaging trials. It runs on Android in 440 ms with a 10.7 MB INT8 footprint, fully offline. On the economic side, a nine-week collection exercise at a 436-household apartment complex (Section 2.2) shows, using measured data, an incremental gain of INR 46–75/month per household (a 67–108% premium over the observed baseline rate); the parallel rural estimate is CPCB-proxy-based and not yet field-validated.

The remaining work is scoped as future work: validation of the seven-stakeholder framework with residents, SHG operators, and kabadiwalas; the Potheri pilot study itself,

with longitudinal income tracking; an mIoU evaluation pass alongside the existing mAP figures; and CLIP fine-tuning on South Asian packaging corpora—all detailed in Section 6.5.

PlastiVision-Edge-enabled Plastic Waste Banks are designed as reusable infrastructure: a zero-shot system that can absorb new packaging as it appears, funded by EPR obligations that already exist independently of the waste bank, and able to run on the hardware roughly 700 million Indian smartphone users already carry.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su182110768/s1, Figure S1: Sample images of the 13 Plastic Waste Vocabulary Ontology (PWVO) benchmark-uncovered categories, one author-captured photograph per category, corresponding to the categories evaluated in Table 10.

Author Contributions: Conceptualisation, M.B. and S.P.; methodology, M.B.; software, M.B.; validation, M.B. and S.P.; formal analysis, M.B.; investigation, M.B.; resources, S.P.; data curation, M.B.; writing—original draft preparation, M.B.; writing—review and editing, S.P.; visualisation, M.B.; supervision, S.P.; project administration, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The TACO and TrashNet datasets used in this study are publicly available at http://tacodataset.org (accessed on 15 October 2026) and https://github.com/garythung/trashnet (accessed on 15 October 2026), respectively. TACO images are Flickr-hosted and links resolve inconsistently over time (Section 6.5). The code (the notebook that regenerates the reported results, with the TACO-PWVO category mapping and evaluation parameters), the PWVO ontology, the author-captured photographs (zero-shot replicates and control images) and the digitised urban pilot records are available at https://github.com/Madhini81/PlastiVision-Edge (accessed on 15 October 2026); code is released under the MIT licence, and photographs and pilot data under CC BY 4.0. Once the proposed Potheri pilot (Section 2.1) and any resulting adaptation of the ontology are complete, the validated PWVO mapping will be archived permanently on Zenodo with a minted DOI.

Acknowledgments: The authors thank the Ultralytics, Roboflow, and SAM open-source communities and the TrashNet, TACO, and Drinking Waste Classification dataset contributors. During the preparation of this manuscript, the authors used Claude Sonnet 5 (Anthropic) for the purposes of LaTeX editing assistance and consistency checking. The authors have reviewed and edited all output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cottom, J.W.; Cook, E.; Velis, C.A. A local-to-global emissions inventory of macroplastic pollution. *Nature* **2024**, *633*, 101–108. [CrossRef] [PubMed]
2. Council on Energy, Environment and Water (CEEW). *How Can India Scale Plastic Waste Recycling Process & Management?* Technical Report; CEEW: New Delhi, India, 2025. Available online: https://www.ceew.in/plastic-waste-recycling (accessed on 14 October 2026).
3. Aspen Network of Development Entrepreneurs (ANDE). Investing in the Waste and Circularity Sector in India: An Introductory Guide. Technical Report, ANDE, 2024. Available online: https://andeglobal.org/publication/investing-in-the-waste-and-circularity-sector-in-india-an-introductory-guide/ (accessed on 14 October 2026).
4. The Energy and Resources Institute (TERI). *Towards a Circular Plastics Economy: India's Actions to #BeatPlasticPollution*; Technical Report; TERI: New Delhi, India, 2023. Available online: https://www.teriin.org/article/towards-circular-plastics-economy-indias-actions-beatplasticpollution (accessed on 14 October 2026).
5. Budiyarto, A.; Clarke, B.; Ross, K. Overview of waste bank application in Indonesian regencies. *Waste Manag. Res.* **2025**, *43*, 306–321. [CrossRef] [PubMed]

6. Kaza, S.; Yao, L.; Bhada-Tata, P.; Van Woerden, F. *What a Waste 2.0*; World Bank: Washington, DC, USA, 2018. [CrossRef]
7. Ratnawati, R.V.; Tahar, N.; Sidik, U.S.; Onogawa, K.; Gamaralalage, P.J.D.; Tsukiji, M. *National Plastic Waste Reduction Strategic Actions for Indonesia*; Technical Report; Ministry of Environment and Forestry: Jakarta, Indonesia, 2020. Available online: https://www.iges.or.jp/en/pub/national-plastic-waste-reduction-strategic-actions-indonesia/en (accessed on 14 October 2026).
8. Kurniawan, T.A.; Othman, M.H.D.; Hwang, G.H.; Gikas, P. Unlocking digital technologies for waste recycling in Industry 4.0 era: A transformation towards a digitalization-based circular economy in Indonesia. *J. Clean. Prod.* **2022**, *357*, 131911. [CrossRef]
9. Yang, M.; Thung, G. Classification of Trash for Recyclability Status. In *Proceedings of the CS229 Project Report*; Stanford University: Stanford, CA, USA, 2016. Available online: http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf (accessed on 14 October 2026).
10. Proença, P.F.; Simões, P. TACO: Trash Annotations in Context for Litter Detection. *arXiv* **2020**, arXiv:2003.06975.
11. Madhini, B.; Supraja, P. Analysis of Using Zero-Shot Open Vocabulary Detection Methods for Plastic Waste Classification. In *Proceedings of the 5th International Conference on Communication, Computing & Industry 6.0 (C2I6)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–9. [CrossRef]
12. Mewada, D.; Grua, E.M.; Eising, C.; Denny, P.; Van de Ven, P.; Scanlan, A. Zero-Shot Learning for Sustainable Municipal Waste Classification. *Recycling* **2025**, *10*, 144. [CrossRef]
13. Ranjbar, I.; Ventikos, Y.; Arashpour, M. Zero-shot and few-shot multimodal plastic waste classification with vision-language models. *Waste Manag.* **2025**, *202*, 114815. [CrossRef] [PubMed]
14. Wang, A.; Liu, L.; Chen, H.; Lin, Z.; Han, J.; Ding, G. YOLOE: Real-Time Seeing Anything. *arXiv* **2025**, arXiv:2503.07465. [CrossRef]
15. Budihardjo, M.A.; Ardiansyah, S.Y.; Ramadan, B.S. Community-driven material recovery facility (CdMRF) for sustainable economic incentives of waste management: Evidence from Semarang City, Indonesia. *Habitat Int.* **2022**, *119*, 102488. [CrossRef]
16. Chikarmane, P. Integrating Waste Pickers into Municipal Solid Waste Management in Pune, India. Technical Report 8, WIEGO, 2012. Available online: https://www.wiego.org/research-library-publications/integrating-waste-pickers-municipal-solid-waste-management-pune-india/ (accessed on 14 October 2026).
17. Hossain, R.; Islam, M.T.; Shanker, R.; Khan, D.; Locock, K.E.S.; Ghose, A.; Schandl, H.; Dhodapkar, R.; Sahajwalla, V. Plastic Waste Management in India: Challenges, Opportunities, and Roadmap for Circular Economy. *Sustainability* **2022**, *14*, 4425. [CrossRef]
18. Juárez Pastor, L.; Subramanian, V.; Cucurachi, S.; Ghorbani, A. Caste, mistrust and municipal inaction: The interwoven barriers for the integration of waste pickers in India. *J. Environ. Manag.* **2024**, *356*, 120513. [CrossRef] [PubMed]
19. Agyabeng, A.N.; Chimhowu, A.; Desouza, M.N.M.; Swanzy, E.K.; Angenu, B.B.; Yamoah, M.; Mensah, S. Plastic credits for profit and planet: Unlocking a circular economy in the Ghanaian slums. *J. Mater. Cycles Waste Manag.* **2026**, *Epub ahead of print*. [CrossRef]
20. Hornuvo, R.; Boafor, E.E.; Okotto, L.G.; Okotto-Okotto, J.; Dzodzomenyo, M.; Asamoah, M.; Wright, J. Challenges and opportunities for plastic versus mixed waste enterprises in Greater Accra and Kisumu: A qualitative study. *PLoS ONE* **2026**, *21*, e0350670. [CrossRef] [PubMed]
21. Ramli, A.H.M.; Manaf, L.A.; Zulkeflee, Z.; Andriyono, S. Advancing circular economy approaches in plastic waste management: A systematic literature review in developing economies. *Sustain. Prod. Consum.* **2024**, *51*, 420–431. [CrossRef]
22. Ministry of Environment, Forest and Climate Change. *Plastic Waste Management (Amendment) Rules 2022*; Technical Report; Gazette of India: New Delhi, India, 2022.
23. Ministry of Environment, Forest and Climate Change (MoEFCC). Plastic Waste Management (Amendment) Rules, 2026. Technical Report G.S.R. 237(E), Government of India, Gazette of India. Available online: https://www.livelaw.in/articles/plastic-waste-management-amendment-rules2026-india-plastic-pollution-530015 (accessed on 14 October 2026).
24. Ministry of Environment, Forest and Climate Change. *Plastic Waste Management (Amendment) Rules 2025*; Gazette of India: New Delhi, India, 2025.
25. Invest India. What Are Plastic Waste Rules and Its Amendment? Invest India, 2022. Available online: https://www.investindia.gov.in/blogs/circular-economy-potential-and-opportunities-under-plastic-waste-management-rules (accessed on 14 October 2026).
26. Shanker, R.; Khan, D.; Hossain, R.; Islam, M.T.; Locock, K.; Ghose, A.; Sahajwalla, V.; Schandl, H.; Dhodapkar, R. Plastic waste recycling: Existing Indian scenario and future opportunities. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 5895–5912. [CrossRef] [PubMed]
27. Ma, W.; Chen, H.; Long, R. An integrated detection-semantic fusion and near-infrared system for food-delivery packaging waste. *Waste Manag.* **2025**, *207*, 115125. [CrossRef] [PubMed]
28. Kunwar, S.; Owabumoye, B.R.; Alade, A.S. Plastic Waste Detection Using Deep Learning: Insights from the WaDaBa Dataset. *Ind. Domest. Waste Manag.* **2025**, *5*, 1–11. [CrossRef]
29. Lilhore, U.K.; Simaiya, S.; Dalal, S.; Radulescu, M.; Balsalobre-Lorente, D. Intelligent waste sorting for sustainable environment: A hybrid deep learning and transfer learning model. *Gondwana Res.* **2025**, *146*, 252–266. [CrossRef]

30. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning Transferable Visual Models from Natural Language Supervision. In Proceedings of the ICML 2021, Virtual, 18–24 July 2021; pp. 8748–8763.
31. Minderer, M.; Gritsenko, A.; Stone, A.; Neumann, M.; Weissenborn, D.; Dosovitskiy, A.; Mahendran, A.; Arnab, A.; Dehghani, M.; Shen, Z.; et al. Simple Open-Vocabulary Object Detection. In Proceedings of the ECCV 2022, Tel Aviv, Israel, 23–27 October 2022. [CrossRef]
32. Liu, S.; Zeng, Z.; Ren, T.; Li, F.; Zhang, H.; Yang, J.; Jiang, Q.; Li, C.; Yang, J.; Su, H.; et al. Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection. *arXiv* **2023**, arXiv:2303.05499. [CrossRef]
33. Cheng, T.; Song, L.; Ge, Y.; Liu, W.; Wang, X.; Shan, Y. YOLO-World: Real-Time Open-Vocabulary Object Detection. In Proceedings of the CVPR 2024, Seattle, WA, USA, 16–22 June 2024; pp. 16901–16911. [CrossRef]
34. Madhini, B.; Supraja, P. Optimization Techniques for Edge Deployment of Zero Shot Open Vocabulary Detection Methods for Plastic Waste Classification. In *Proceedings of the 3rd International Conference on Climate Change and Ocean Renewable Energy (CCORE)*; Springer: Cham, Switzerland, 2025, pp. 148–159. [CrossRef]
35. Balasundaram, M.; Perumal, S. EcoSortBin: Accuracy–Generalisation Trade-Offs in Open-Vocabulary Campus Waste Detection on Raspberry Pi 4. *AI* **2026**, *7*, 305. [CrossRef]
36. Secretariat of the Basel Convention, United Nations Environment Programme. Basel Convention Plastic Waste Amendments: Listing of Plastic Wastes Under Annex II and Annex VIII. UNEP, 2019. Available online: https://www.basel.int/Implementation/Plasticwaste/PlasticWasteAmendments/Overview/tabid/8426/Default.aspx (accessed on 14 October 2026).
37. Wenneker, B.; Oosterbaan, L. *Guideline for Monitoring Marine Litter on the Beaches in the OSPAR Maritime Area*; Technical Report; OSPAR Commission: London, UK, 2010. [CrossRef]
38. Shao, S.; Li, Z.; Zhang, T.; Peng, C.; Yu, G.; Zhang, X.; Li, J.; Sun, J. Objects365: A Large-Scale, High-Quality Dataset for Object Detection. In Proceedings of the IEEE/CVF ICCV, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8430–8439. [CrossRef]
39. Vasu, P.K.A.; Pouransari, H.; Faghri, F.; Vemulapalli, R.; Tuzel, O. MobileCLIP: Fast Image-Text Models through Multi-Modal Reinforced Training. In Proceedings of the IEEE/CVF CVPR, Seattle, WA, USA, 16–22 June 2024; pp. 15963–15972. [CrossRef]
40. Zahrah, Y.; Yu, J.; Liu, X. How Indonesia's Cities Are Grappling with Plastic Waste: An Integrated Approach towards Sustainable Plastic Waste Management. *Sustainability* **2024**, *16*, 3921. [CrossRef]
41. Central Pollution Control Board. Annual Report on Solid Waste Management, 2021–2022. Technical Report, Ministry of Environment, Forest and Climate Change, Government of India, 2022. Available online: https://cpcb.nic.in/uploads/MSW/MSW_AnnualReport_2021-22.pdf (accessed on 14 October 2026).
42. Krishnan, A. [Explainer] The Cost of Plastic Waste. Mongabay-India, 2021. Available online: https://india.mongabay.com/2021/12/explainer-the-cost-of-plastic-waste/ (accessed on 14 October 2026).
43. Office of the Registrar General & Census Commissioner, India. Houselisting and Housing Census Data Highlights—2011. Technical Report, Ministry of Home Affairs, Government of India, 2011. Available online: https://censusindia.gov.in/ (accessed on 14 October 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.