International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Real-Time Plastic Waste Classification and Segregation Using ESP32-CAM

Kumaravel V[1], Kavinraj R[1], Kavinbalakumar T M[1], Dinesh Kumar S[1], Mani P[2]

[1]student, [2]Assistant Professor
[1,2]Department of Electronics and Communication Engineering
[1,2]K.Ramakrishnan College of Technology, Samayapuram, Trichy.

Abstract - Improper segregation of plastic waste reduces recycling efficiency, and manual sorting is slow, labour-intensive and error-prone. This paper presents a low-cost, vision-based system that classifies plastic waste items in real time into three categories: Polyvinyl Chloride (PVC), Polyethylene Terephthalate (PET) and Other Plastics. An ESP32-CAM module captures images of waste items and streams the frames over Wi-Fi to a laptop running a FastAPI and OpenCV backend, which extracts, crops, resizes (224 × 224 pixels) and normalizes each frame. A Convolutional Neural Network (CNN) based on a pretrained MobileNetV2 backbone, trained by transfer learning on images captured with the ESP32-CAM itself, predicts the category. Predictions whose confidence is below 0.6 are reported as UNKNOWN to avoid unreliable decisions. A web dashboard displays the detected category and confidence score and maintains separate counters for each class, while an optional servo-driven bin mechanism can physically separate the item. The model is evaluated using validation accuracy and a confusion matrix. The proposed design offers an affordable and scalable approach for automated plastic waste segregation.

Key Words: plastic waste classification, ESP32-CAM, MobileNetV2, convolutional neural network, transfer learning, FastAPI.

1. INTRODUCTION

Plastics are widely used because they are lightweight, durable and inexpensive, but the same properties make plastic waste persistent in the environment. Global production and disposal of plastics have grown rapidly, and only a small fraction of the plastic ever produced has been recycled [1]. Effective recycling depends on separating waste into material types, since mixed plastics reduce the quality and value of recycled material.

In most facilities, plastic waste is still segregated by hand. Manual sorting is slow, tiring and inconsistent, and it exposes workers to hygiene risks. Computer vision and deep learning offer an alternative: a Convolutional Neural Network (CNN) can learn visual cues such as shape, colour, transparency and surface texture directly from images [2]. Public datasets such as TrashNet have demonstrated that image-based waste classification is feasible [3], and later studies have applied deep learning specifically to plastic waste [4].

This work focuses on three practical categories: Polyvinyl Chloride (PVC), Polyethylene Terephthalate (PET) and Other Plastics. PVC is common in pipes, fittings and sheets; PET is common in water and soft-drink bottles and food containers; the remaining plastics (such as HDPE, LDPE, PP and PS) are grouped into a single class. The system is built from inexpensive hardware, an ESP32-CAM camera module and a laptop, so that it can be reproduced in laboratories, schools and small recycling units.

The main contributions of this paper are: (i) an end-to-end pipeline from ESP32-CAM image capture to a live classification dashboard; (ii) a MobileNetV2 transfer-learning classifier trained on a self-collected dataset for PVC, PET and Other Plastics; (iii) a confidence-threshold rule that returns UNKNOWN for uncertain predictions; and (iv) an optional actuation interface for automatic separation.

2. RELATED WORK

Image-based waste classification has been studied extensively with CNNs. Early datasets such as TrashNet provided labelled images of common waste categories and encouraged the use of deep models for recyclability classification [3]. Transfer learning, in which a network pretrained on a large dataset is adapted to a new task, is widely used when the target dataset is small, because the early layers learn general visual features that remain useful across domains [5].

Lightweight architectures are attractive for low-cost deployment. MobileNetV2 uses inverted residual blocks with linear bottlenecks to obtain good accuracy with a small number of parameters and low computation [6]. Data augmentation is a standard technique to improve generalization when training images are limited [7]. These ideas are combined in the present work with a low-cost camera front end and a simple web-based backend.

3. SYSTEM ARCHITECTURE AND WORKFLOW

The proposed system consists of an ESP32-CAM image-acquisition unit, a laptop backend, a web dashboard and an optional separation mechanism. The overall workflow is shown in Fig. 1 and is described below.

3.1 Image Capture

The ESP32-CAM captures images of plastic waste items and streams the frames to the laptop through Wi-Fi.

3.2 Frame Reception

The laptop backend, developed using FastAPI and OpenCV, receives the video stream and extracts individual frames for processing.

3.3 Preprocessing

Each captured frame is cropped to isolate the plastic waste item, resized to 224 × 224 pixels, and normalized. Resizing to 224 × 224 matches the input size expected by MobileNetV2.

3.4 Classification and Rejection of Uncertain Predictions

A trained CNN model classifies the item as PVC, PET or Other Plastics. The final layer produces a probability for each class. If the highest probability is below the selected threshold of 0.6, the system returns UNKNOWN instead of forcing a decision. The decision rule is: label = argmax(p) if max(p) ≥ 0.6, otherwise UNKNOWN. This reduces the risk of misrouting items that are blurred, partly hidden or unlike the training data.

3.5 Serving and Dashboard

The backend sends the predicted category and confidence score to the dashboard. The dashboard displays the detected category and keeps separate counters for PVC, PET and Other Plastics.

3.6 Optional Automatic Separation

Based on the predicted category, the laptop can trigger a servo motor or an automatic bin mechanism through serial communication or Wi-Fi to separate the item into the corresponding bin.

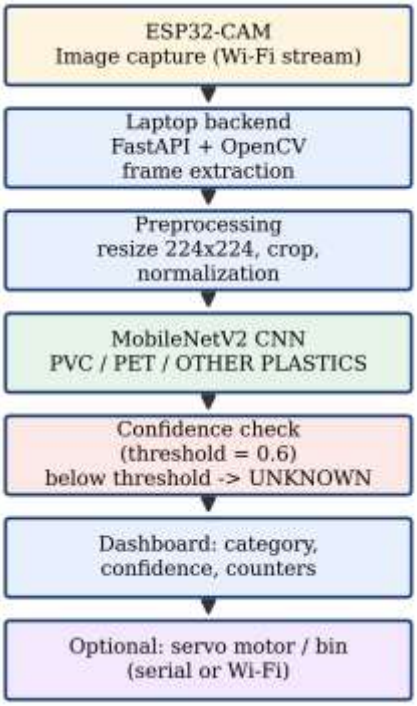


Fig -1: System architecture and workflow of the proposed plastic waste classifier

4. CNN MODEL DEVELOPMENT

The training procedure is summarized in Fig. 2 and described step by step in this section.

4.1 Dataset Collection

Real-world images of plastic waste are captured with the ESP32-CAM so that the training data match the image quality of the deployed camera. The three classes and their typical items are listed in Table -1. Approximately 300–500 images are collected for each category, using different viewing angles, distances, backgrounds and lighting conditions. This gives roughly 900–1,500 images in total.

Table -1: Classes and representative items

Class	Representative items
PVC	PVC pipes, PVC fittings, PVC sheets and other PVC plastic products
PET	PET water bottles, soft-drink bottles and PET food containers
OTHER PLASTICS	HDPE containers, LDPE plastic bags, PP containers, PS products and other plastic materials

4.2 Dataset Organization

Images are stored in one folder per class (dataset/PVC, dataset/PET and dataset/OTHER_PLASTICS), so that the folder name serves as the class label when the data are loaded.

4.3 Data Augmentation

Random rotation, horizontal flipping, brightness adjustment and zooming are applied to the training images. These transformations simulate changes in orientation, lighting and distance, and improve the ability of the model to recognize plastic materials under different conditions.

4.4 Dataset Splitting

The collected images are divided into 80% for training and 20% for validation. For the expected dataset size this corresponds to about 720–1,200 training images and 180–300 validation images.

4.5 Model Architecture

MobileNetV2 pretrained on a large image dataset is used as the feature extractor [6]. Its base layers are initially frozen so that the pretrained features are preserved, and a small classification layer with three output classes (PVC, PET and Other Plastics) is added on top. Only this new layer is trained in the first stage.

4.6 Training

The model is trained using the Adam optimizer [8] and the categorical cross-entropy loss for approximately 10–20 epochs. If required, the final layers of the base network are then unfrozen and fine-tuned using a lower learning rate. The training settings are summarized in Table -2.

Table -2: Training configuration

Parameter	Value
Backbone	MobileNetV2 (pretrained)
Input size	224 × 224 × 3
Output classes	3 (PVC, PET, OTHER PLASTICS)
Images per class	≈ 300–500
Train / validation split	80% / 20%
Optimizer	Adam
Loss function	Categorical cross-entropy
Epochs	≈ 10–20
Fine-tuning	Final layers, lower learning rate
Confidence threshold	0.6 (below: UNKNOWN)

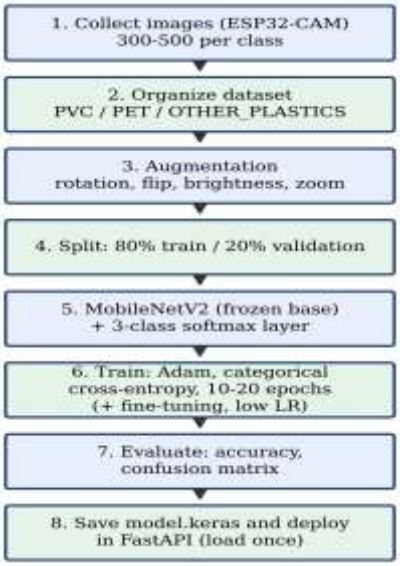


Fig -2: CNN training and deployment pipeline

4.7 Evaluation

The model is evaluated using validation accuracy and a confusion matrix. The confusion matrix shows which classes are mistaken for one another, and is particularly useful for identifying errors between PVC, PET and other plastic materials. Precision, recall and F1-score can be derived from it for each class.

4.8 Saving and Deployment

The trained model is saved as model.keras. The FastAPI backend loads the model once at startup, which avoids repeated loading delays, and then predicts the category of each incoming image. If the highest prediction probability is below 0.6, the dashboard displays UNKNOWN.

5. RESULTS AND DISCUSSION

[Authors to complete: insert the measured results from the trained model in Tables -3 and -4 and update this section.]

Performance is reported on the 20% validation set using overall validation accuracy, per-class precision, recall and F1-score (Table -3), and a confusion matrix (Table -4). Training and validation accuracy curves over the epochs should also be reported to show whether the model is over-fitting.

Table -3: Per-class performance on the validation set

Class	Precision	Recall	F1-score
PVC	[ ]	[ ]	[ ]
PET	[ ]	[ ]	[ ]
OTHER PLASTICS	[ ]	[ ]	[ ]
Overall accuracy	[ ]		

Table -4: Confusion matrix on the validation set (counts)

Actual \ Predicted	PVC	PET	OTHER
PVC	[ ]	[ ]	[ ]
PET	[ ]	[ ]	[ ]
OTHER PLASTICS	[ ]	[ ]	[ ]

Expected sources of error. Several factors are likely to affect accuracy. Items made of PVC and PET can look similar in colour and shape, and transparent PET bottles can change appearance with the background and lighting. The Other Plastics class mixes several polymers (HDPE, LDPE, PP, PS and others), so it is visually diverse. A camera also sees only the surface of an item and cannot measure the polymer directly. The confusion matrix will show which of these effects matter most in practice.

Effect of the confidence threshold. The 0.6 threshold trades coverage for reliability: a higher value rejects more uncertain items as UNKNOWN and lowers the number of wrong sorting actions, but increases the number of items that need manual handling. The threshold can be tuned on the validation set.

6. LIMITATIONS AND FUTURE WORK

The system classifies a single item at a time, and its accuracy depends on how well the training images represent the real conditions of use. Visual features alone cannot reliably identify every polymer type, so the three-class scheme is a practical simplification. Future work includes enlarging the dataset, handling several items in one frame with object detection, running the model on an edge device to remove the laptop, adding more plastic categories, and combining the camera with additional sensing to improve material identification.

7. CONCLUSIONS

This paper presented a low-cost system for real-time classification of plastic waste into PVC, PET and Other Plastics using an ESP32-CAM, a FastAPI and OpenCV backend and a MobileNetV2 transfer-learning model. Images are captured over Wi-Fi, preprocessed to 224 × 224 pixels, classified, and displayed on a dashboard with per-class counters, with an UNKNOWN output for predictions below a 0.6 confidence threshold. An optional servo or bin mechanism allows automatic separation. The approach uses inexpensive hardware and a lightweight network, which makes it suitable for small-scale and educational recycling set-ups.

ACKNOWLEDGEMENT

[Authors to add acknowledgement, e.g. guide, institution or funding agency.]

REFERENCES

[1] Geyer, R., Jambeck, J.R., Law, K.L.: Production, use, and fate of all plastics ever made. Science Advances 3(7), e1700782 (2017)
[2] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521, 436–444 (2015)
[3] Yang, M., Thung, G.: Classification of trash for recyclability status. Stanford University CS229 Project Report (2016)
[4] Bobulski, J., Kubanek, M.: Deep learning for plastic waste classification system. Applied Computational Intelligence and Soft Computing 2021, 6626948 (2021)
[5] Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? In: Advances in Neural Information Processing Systems 27 (2014)
[6] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C.: MobileNetV2: Inverted residuals and linear bottlenecks. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)
[7] Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. Journal of Big Data 6, 60 (2019)
[8] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: Proc. International Conference on Learning Representations (2015)