

Received 5 November 2022, accepted 25 November 2022, date of publication 5 December 2022, date of current version 12 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3226682

The logo consists of a series of vertical bars of varying heights on the left, followed by the word "SURVEY" in a blue, sans-serif font inside a rounded rectangular border.

A Survey on Waste Detection and Classification Using Deep Learning

HARUNA ABDU AND MOHD HALIM MOHD NOOR 

School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia

Corresponding author: Mohd Halim Mohd Noor (halimnoor@usm.my)

This work was supported in part by the Ministry of Higher Education Malaysia, and in part by the Universiti Sains Malaysia (USM) for Long-Term Research Grant Scheme (LRGS) Malaysian Research University Network (MRUN) under Project LRGS MRUN/F2/01/2019.


ABSTRACT Waste or trash management is receiving increased attention for intelligent and sustainable development, particularly in developed and developing countries. The waste or trash management system comprises several related processes that carry out various complex functions. Recently, interest in deep learning (DL) has increased in providing alternative computational techniques for determining the solution to various waste or trash management problems. Researchers have concentrated on this domain, and as a result, significant research has been published, particularly in recent years. According to the literature, a few comprehensive surveys have been done on waste detection and classification. However, no study has investigated the application of DL to solve waste or trash management problems in various domains and highlight the available datasets for waste detection and classification in different domains. To this end, this survey contributes by reviewing various image classification and object detection models, and their applications in waste detection and classification problems, providing an analysis of waste detection and classification techniques with precise and organized representation and compiling over twenty benchmarked trash datasets. Also, we backed up the study with the challenges of existing methods and the future potential in this field. This will give researchers in this area a solid background and knowledge of the state-of-the-art deep learning models and insight into the research areas that can still be explored.

INDEX TERMS Deep learning survey, trash datasets, waste detection, waste classification.

I. INTRODUCTION

Waste generation has risen dramatically in recent years. According to World Bank data, the global solid waste generation in 2016 was approximately 2.01 billion tonnes per year. By 2030 and 2050, the world is expected to produce 2.01 and 3.40 billion tonnes, respectively [1], [2]. Trash management failure can have disastrous consequences for almost every environment. As a result of a large amount of waste, waste detection and sorting should be done early in the waste management process to maximize the number of recyclable items and reduce the possibility of environmental contamination by other items.

The daily increase in solid waste in all environments endangers both human and animal health and life. Poorly managed and openly deposited trash harms the environment,

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo .

endangers local residents' health, causes water and air pollution, land contamination/degradation, and has numerous other consequences [3]. In areas that are not technically designated as toxic waste dump sites, such as cultivable land, highways, buildings, and construction sites, as well as occasionally inside homes or nearby, illegal trash burying happens.

Due to the challenges posed by improper garbage/trash depositions in undesignated locations [4], many have been using various techniques to detect and classify trash [5], [6]. Some research such as in [7] focuses on the direct detection of waste through its spectral signature using satellite imagery and remote sensing methods. But the satellite images varied in characteristics, and they will have different resolutions at different distances, and the objects are taken at different angles. Despite the fact that variations in light absorption allow satellites to locate objects in space, acquisition will take place in inaccessible regions with limited transportation

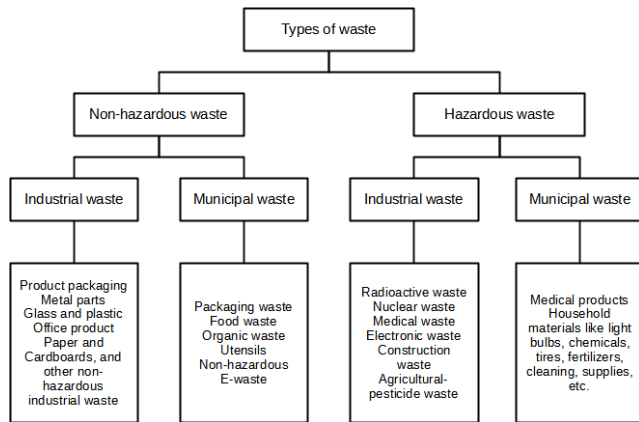


FIGURE 1. Classification of waste. [9].

options [8]. In contrast to features seen in terrestrial litters, those found in marine litters will be observed from the most advantageous vantage position. However, most of the methods used to classify trash depend on human expertise, which is sometimes very challenging and tedious to classify the waste accurately and the satellite imaging methods are computationally cost, with inability to perfectly separate trash objects, especially in case of occlusion and variation or light. Different kinds of waste are shown in Figure 1. The waste can be hazardous or non-hazardous with further subdivision in a different environment. The physical state, technical elements, reusable potentials, biodegradable potential, manufacturing source, and the degree of environmental effects are some of the specific features considered in the classification of garbage [9]. After considering these characteristics, waste, by material nature, can be commonly divided into three primary types: liquid, solid and gaseous waste [10], [11]. Domestic waste is also called municipal solid waste, although some of its content may be associated with commercial and industrial waste.

Computer vision is a field of study that enables computers to analyze and derive information from visual data. Object detection and image classification are the two most common applications of computer vision. Identifying objects in digital images is referred to as object detection whereby it locates the object of interest in an image and creates a bounding box around it. Predicting an object's class is referred to as image classification. Face detection, and pedestrian detection, are common examples of object detection.

Machine learning and deep learning techniques are subsets of artificial intelligence that automatically without being explicitly programmed or learning without any direct human intervention learned from any input data. Machine learning and deep learning technique have been widely used in object detection and image classification. The same is applied to waste detection and classification.

Several survey papers have been written by different researchers in relation to waste management. There are some surveys focusing on object detection in general such as [12],

[13], [14], [15], [16], and [17] all focusing on general object detection only. Illegal waste disposal surveys were conducted by many researchers such as [18], the authors survey gaseous and water waste in borehole drilling. The paper [19] discusses the conventional techniques used to dispose of waste. It also discusses the shortcomings of the current systems and how to fix them. A bibliometric-based review was conducted by [20] on the classification of domestic waste covering the years between 2000 and 2019. The author(s) reported that European countries are in the lead in doing research in this field, and plastics and metal wastes were the existing focus of automating trash detection and classification. However, this survey/review is limited to only Engineering, Environmental Science Economics, and Chemistry. Another review by [21] focuses on reviewing images underwater for better detection and classification. The authors introduce existing research on underwater target image recognition and primarily present deep learning-based underwater image recognition technology. The current problems of underwater image recognition are then summarized in this review. A systematic literature review was conducted by [22]. The authors examine disaster waste management research systematically from nine perspectives: planning, waste, waste treatment options, environment, economics, social considerations, organizational aspects, legal frameworks, and funding. Table 1 shows the existing survey papers, their contributions, and topics not well discussed.

All of the reviewed surveys focus on object detection and a few on waste detection and classification. However, none of them comprehensively surveyed the available benchmarked dataset and the deep learning models for single and multi-object detection on the waste detection and classification.

This survey paper aims to contribute by reviewing existing deep learning models for detecting and classifying waste. This work offers an organized and thorough review of several existing waste detection, classification, and DL approaches. Furthermore, it readily explained existing waste detection and classification datasets in various environments. The benefits and drawbacks of current approaches and datasets and the potential future research are highlighted to support the study.

In order to give a comprehensive technical study of the trash object identification and classification techniques as part of this survey review, many articles from several digital libraries, including IEEE Xplore, Science Direct, Scopus, ACM Digital Library, and many more, were retrieved. The most appropriate papers for the review have been identified and organized logically in this portion of the paper after careful consideration of the paper title, abstract, introduction, experiment, and future scope. Automatic and manual search strategies were conducted. The automatic search was carried out by inputting keywords on online scientific database. The search keywords are ("Waste Detection" AND "Waste Classification", "Garbage Detection", "Garbage Classification"). The manual search was carried out by scanning the primary study references of the automatic search. This

TABLE 1. Existing Survey papers summary.

Survey Papers	Year	Survey focusses and topics not well discussed
Ferrara L., et, al., [18]	2013	Survey on illegal geochemical disposable sites, and not considering technological solutions.
Prasanna M. A., et, al., [19]	2018	Survey on the drawbacks of the existing waste classification methods, but only conventional machine learning methods were reviewed leaving out deep learning methods.
Yang, T., et, al., [20]	2021	This paper focuses on applied bibliometrics to assess publications related to waste classification technology. However, recent papers are not captured, and the papers in different environments like water were ignored.
Wang, Y., et, al., [21]	2022	Focuses on underwater trash image detection, excluding other environments and most deep learning models, were not discussed.
Zhang F., et, al. [22]	2019	Surveyed on the disaster management methods between 2011 to 2019. However, recent technologies such as ML and DL were not discussed.

study collected primary studies from journals and conference proceedings written in English, that are related to DL-based waste detection and classification. Studies that are not related to the domain and not written in English are excluded.

The remaining of this paper is organized into sections as follows: Section 2 is the related literature review of the image classification and object detection models. In Section 3, the deep learning methods for waste detection and classification are discussed. While in section 4, the benchmark dataset for waste detection and classification is explained vividly. In section 5, the challenges of waste detection and classification are summarized. The conclusion is provided at section 6.

II. IMAGE CLASSIFICATION AND OBJECT DETECTION MODELS

Nearly all object detection and classification were done using traditional machine learning methods before deep learning gained popularity in the previous decade. Typical ones were the histogram of directed gradients [23], scale-invariant feature transforms [24], and the viola-jones object detection algorithm [25]. These would identify a number of recurring features in the image and categorize their clusters using random forests, logistic regression, or color histograms. Image identification or classification through the use of machine learning makes use of the potential of algorithms to learn previously unknown information from a dataset containing both structured and disorganized samples (Supervised Learning). Deep learning, a kind of machine learning that employs

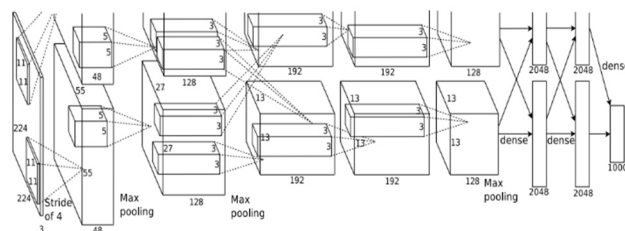


FIGURE 2. Alexnet architecture [26].

numerous concealed layers within a model, is now the most widely utilized approach to the field. Deep learning, when used in conjunction with powerful AI technology and graphics processing units (GPUs), has made it possible to obtain exceptional results in image classification tasks. As a result, deep learning has been the driving force behind several recent breakthroughs in the fields of image identification, facial recognition, and the development of image classification algorithms that attain above-human-level performance and real-time image/object detection. Deep learning-based methods used now do far better than these. This section reviews the widely used or adapted deep learning models for waste detection and classification.

The process of image classification occurs when a computer examines an image in order to determine the “class” that the image belongs to. (Or a likelihood that the image belongs to a certain class.) A class is essentially a label, such as vehicle,” “animal,” “house,” and so on, among many more examples. Some of the deep learning classification models includes, AlexNet, VGG16, followed by ResNet, MobileNet, Inception-ResNet and DenseNet.

A. AlexNet

Five Convolution layers and three Fully Connected layers are included in AlexNet, along with ReLU nonlinearity and Local Response Normalization (LRN). Alexnet will accept a maximum input image dimension of 227 by 227 pixels. Alexnet was trained with ImageNet, which consists of images of 1000 classes. Figure 2 depicts the architecture of the alexnet network.

B. VGG16

The results of the ImageNet competition in 2014 served as the basis for the development of this model. During the training phase, the ConvNets receive RGB images of a fixed size of 224 by 224 as their input. The only thing that is done in the way of pre-processing is taking each pixel and removing the mean RGB value that was derived using the training set. The model was trained using a total of a thousand different classes. The VGG16 architecture can be shown in Figure 3.

C. Inception-ResNet,(InceptionResNetV2)

The Inception-v4 model is an improved version of the Inception-v3 model, and both are a hybrid of inception modules and residual connections. Residual connections improve the efficiency of deeper and wider inception

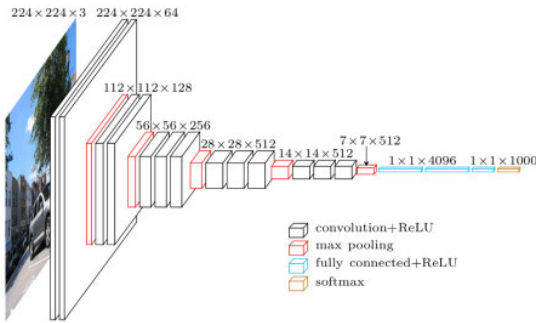


FIGURE 3. VGG16 architecture [27].

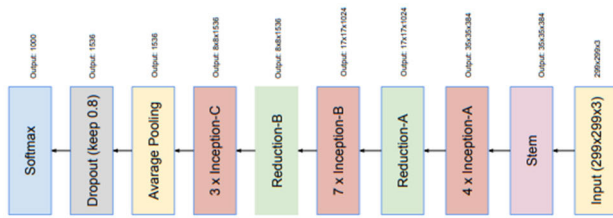


FIGURE 4. Overall schema of InceptionNetv4 [28].

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5 × Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

FIGURE 5. MobileNet model body architecture [25].

networks with fewer hyperparameters. Previously, module implementation required more training resources and time; however, with residual connection improvement, requirements were reduced, and the model proved to be more efficient to train. In contrast to previous Inception ResNet works, batch normalization in inception-v4 occurs on top of traditional convolutional layers rather than before residual blocks; as a result, the inception block size is increased. Aside from inception-v3, inception-v4 includes residual scaling before layer accumulation, which results in more stable training and higher accuracy, as well as the ability to build a larger model size [28]. The input size for Inception-v4 is 299 by 299 pixels.

D. MobileNet

MobileNets is a model proposed by Google’s research team for efficient mobile device usage. MobileNets perform depth-wise separable convolution after full convolution, allowing

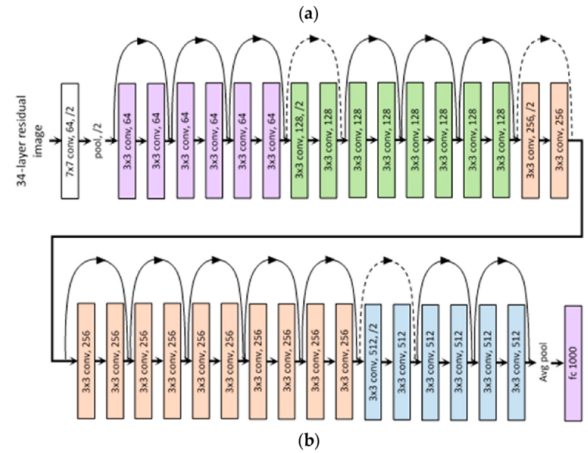


FIGURE 6. Deep residual networks [31].

for higher accuracy with a small number of hyper-parameters. MobileNets, in addition to depth-wise convolution, are thinner with fewer parameters because they use reduced representations of the input and rely on the model shrinkage parameter to keep the model from producing additional hyper-parameters. Because of the smaller size of MobileNets, they can train faster with fewer resources, which is one of their most useful properties for versatility [29]. The input size for MobileNets is 224 by 224.

E. DEEP RESIDUAL NETWORKS

Deep residual networks, or ResNet, won the 2015 ILSRVC (ImageNet Large Scale Visual Recognition Competition) on the ImageNet dataset, and ResNet50 is one of the most commonly used residual network architectures. Deep residual networks scientifically solved inaccuracy problems while increasing training and test error as the layers stacked for deeper networks. The use of residual blocks enabled the construction of extremely deep convolutional neural networks. ResNet proposes an elegant formation of residual blocks with three convolutional layers, batch normalization, and a rectified linear unit (ReLU) activation function, in addition to simply stacking convolutional layers. The recurrent use of these residual blocks constructs ResNET as deeply as possible with fewer hyperparameters. Because of this deep structure and batch normalization between residual blocks, ResNET achieves better feature extraction than previous versions [30]. For our classification task in this study, we used the ResNet50 variation. ResNet50 input dimensions are 224 by 224.

F. DENSELY CONNECTED CONVOLUTIONAL NETWORKS

Densely connected Convolutional Networks, or DenseNets for short, are among the most efficient deep convolutional neural network structures because they have shorter connections between layers near the input and output. DenseNets outperform previous research by reducing parameters, strengthening feature propagation, and solving the

vanishing gradient problem as the network grows. DenseNet’s main layer architecture is a dense connectivity layer, and its feature map concatenates multiple inputs into a single tensor. DenseNet works more efficiently with fewer parameters than ResNet variants. Identity mapping, deep supervision, and depth diversification are all natural components of DenseNets. It is worth noting that DenseNet claimed to work better without data augmentation due to large margin properties [32]. Despite the limited number of data points, our findings appear to support this claim. DenseNet input dimensions are 224 by 224.

Deep learning-based methods identify the labels of the objects based on their features using neural network architectures such as RetinaNet, YOLO (You Only Look Once), CenterNet, SSD (Single Shot Multibox detector), and Region proposals (R-CNN, Fast-RCNN, Faster RCNN, Cascade R-CNN).

Modern object detection architectures often have two stages (single-stage and two-stage object detectors), and many of them have already been pre-trained using the COCO dataset [33]. The COCO picture dataset includes 90 distinct item classifications (e.g., person, airplane, car, bicycle, handbag, tv, door, etc.). Many image classification models have been in existence, leveraging its potentials from neural networks architecture. Some of the existing deep learning models are discussed below:

G. R-CNN, OR REGIONS WITH CNN FEATURES

is an object detection model that employs top-down region proposals and high-capacity CNNs to localize and segment objects. [34]. Multiple bounding-box object region candidates are found using selective search (“regions of interest”), and each region’s features are extracted separately and used for classification. Depending on the version, multiple input shape sizes are accepted by R-CNN. Some of the permitted sizes include 640 by 640, 1024 by 1024, and 800 by 1333.

H. YOLO

You Only Look Once, more often known as YOLO, is a well-known real-time object identification technique. YOLO takes what was traditionally a multi-step process and integrates it into a single operation by employing a single neural network to conduct both classification and prediction of bounding boxes for objects that have been spotted. The image that is being input is scaled down to 448 square pixels bytes. A PASCAL VOC consisting of 20 labeled classes was utilized for the purpose of analyzing YOLO. Figure 7 shows the YOLO model architecture. The YOLOR [35] technique, which was launched in 2021, achieves inference times of 12 milliseconds on the same MS COCO dataset, thus surpassing the popularly deep learning algorithms YOLOv3 [36] and YOLOv4 [37]. There are some other versions of YOLO like YOLOv5 and YOLOv6, that are considered unofficial and still performing wonderfully well. However, YOLOv5 and YOLOv6 are not better than YOLOv4 in terms of performance [38]. Meanwhile, the latest version of YOLO that was recently released

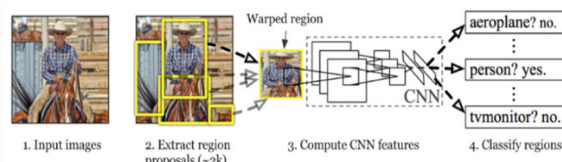


FIGURE 7. RCNN architecture.

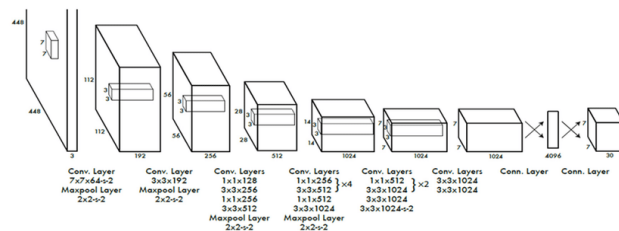


FIGURE 8. YOLO model architecture [40].

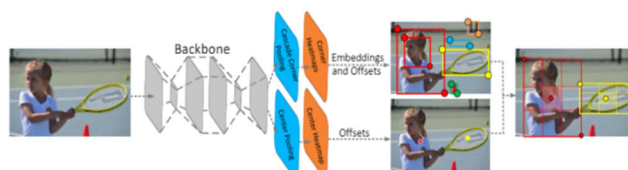


FIGURE 9. CenterNet architecture [42].

is YOLORv7 [39] which addressed some of the issues that the previous versions are having.

I. CenterNet

[41] is a single-stage object detector that recognises each object as a triplet of keypoints rather than a pair. In order to enrich the information gathered by the top-left and bottom-right corners and to provide more recognisable information at the central regions, it makes use of two customized modules called cascade corner pooling and centre pooling. In CenterNet, the likelihood that the centre keypoint in its central area will be predicted as belonging to the same class is high if the predicted bounding box and the ground-truth box have high IoUs. CenterNet accept the input size of 512 by 512.

J. EfficientDet

EfficientDet was developed after careful study of the neural network architecture design choices for better object detection and proposals of several key optimizations to improve efficiency. [43] Weighted bi-directional features pyramid network (BiFPN) which allows easy and fast multi-scaling feature fusion, and a compound scaling method that uniformly scales the resolution, depth, and width for all backbone, feature network, and the box was proposed and formed the efficientDet network architecture, and it accept the input image size of 512 by 512.

K. ExtremeNet

is a bottom-up object detection framework that recognizes an object’s four extreme points (top, left, bottom, and right). In order to identify extreme points, it predicts four multi-peak heatmaps for each item category using a key point estimation technique. Additionally, it averages two bounding box edges

in the x and y dimensions to forecast the object center using one heatmap for each category. A strictly geometric method is used to classify extreme points into objects. The four extreme points were combined, one from each map, if and only if their geometric center was predicted in the center heatmap with a score greater than a predetermined threshold. The predictions for the extreme points were listed, and the legitimate ones were chosen [44].

L. Mask-RCNN

Over the last few years, a lot of giant breakthroughs was achieved in image classification and detection, with an inference time of 330 milliseconds per frame in 2017, the Mask R-CNN [45] technique was the real-time object detector that proved to be the most successful in the MS COCO benchmark.

III. WASTE DETECTION AND CLASSIFICATION

A. WASTE DETECTION

Some researchers concentrate on detecting and reporting the presence of abandoned waste through real-time video stream analysis. The research by [46] uses an improved YOLOv3 network model to perform waste detection and recognition. The network was fine-tuned using the dataset gathered for this purpose. The findings indicate that the proposed approach could make a significant contribution to more efficient waste management in smart cities.

[47] examines a variety of deep-learning algorithms for visually detecting trash in realistic underwater environments, with the goal of exploring, mapping, and extracting such debris using AUVs. A large, publicly available dataset of actual debris in open-water locations is annotated and used to train a variety of convolutional neural network architectures for object detection. The four selected algorithms tested are YOLOv2, Tiny-YOLO, Faster RCNN with Inception v2, SSD with MobileNet v2. With corresponding results of YOLOv2 - mAP=47.9, Tiny-YOLO - mAP=31.6, Faster RCNN with Inception v2 - mAP=81, SSD with MobileNet v2 - mAP=67.4.

The author [48] proposed an automatic trash detection system based on deep learning and the narrowband Internet of things. The system detects and identifies decoration trash directly in the front-end embedded monitoring module and manages thousands of monitoring front ends via the narrowband Internet of Things and background server. an improved YOLOV2 was used for the experimentation.

[49] use a deep learning strategy to detect trash automatically. FastRCNN was the model trained, and a data fusion and augmentation strategy is proposed to improve the method's accuracy. As a result of the experiments, the method has a good generalization ability and a high-precision detection function.

Three different pieces of waste classes were experimented and reported by [50] using Fast RCNN. On the overall

classification of the trash images, the authors achieved a mean Average Precision (mAP) of 0.683.

A smartphone app, called SpotGarbage was proposed and developed by [51] which detects and coarsely segments garbage regions in a geo-tagged image clicked by the user for detecting garbage in images, the app employs the proposed deep architecture of fully convolutional networks. The model was trained on a Garbage In Images (GINI) dataset, with a mean accuracy of 87.69%.

Aquatic animals do also experience serious health issues that can lead to death straight or environmental contamination by floating trash that can easily lead to their death. Research by [52] proposes a method for detecting visible trash floating on the water surface of urban canals. The authors also provide a large dataset of trash in water channels, the first of its kind, with object-level annotations. A novel attention layer that improves the detection of smaller objects is proposed. Another research in the same environment by [53] Aqua-Vision, a cutting-edge deep learning-based object detection model, was proposed over the AquaTrash dataset. With a mean Average Precision (mAP) of 0.8148, the proposed model detects and classifies various pollutants and hazardous waste items floating in the oceans and on the seashores. The proposed method localizes waste objects, which aids in the cleaning of water bodies and contributes to the environment by preserving the aquatic ecosystem.

Research by [54] has proposed a garbage detection algorithm for underwater environments that is based on an enhanced version of the YOLOv5s algorithm. The feature extraction module of the YOLOv5s network is replaced by the MobileNetv3 network, which is a lightweight network, thanks to the algorithm. While this is going on, the enhanced network is pruned in order to cut down on the number of redundant parameters and further compress the model. The findings of the experiments indicate that the approach's detection accuracy can reach 97.5% based on one-ninth of the parameters of YOLOv5s, and that the real-time detection speed on the CPU is 2.5 times that of YOLOv5s. [55] has developed and implemented an image-based detection system that can distinguish between various garbage cans for the sake of categorization..

Research conducted by [56] presented a strategy with the goal of reducing the costs associated with monitoring urban waste and better coordinating the data acquired with the essential information requirements of cities. The authors used cameras mounted on vehicles and a deep convolutional neural network model to quantify the amount of urban waste that accumulated along roadsides. The model was used to identify trash in the images that were captured. Using data collected along 84 road segments in two California cities, they compared the performance of three different models for trash detection, with the highest performing model (Mask R-CNN) obtaining 91% recall, 83% precision, and 77% accuracy.

Research by [48] proposes a deep learning-based automatic garbage detection system and a narrowband Internet of things. The system detects and identifies decoration

garbage directly in the front-end embedded monitoring module, and it manages thousands of monitoring front ends via the narrow-band Internet of Things and the background server. The improved YOLOv2 network model is used in the system's front-end embedded module for garbage detection and recognition. As a result of low image resolution, a research by [57] a new and innovative feature fusion module that is lightweight was proposed as part of an algorithm that is an improved single-shot multibox detector (SSD). In the course of this study, the backbone network of VGG16 was upgraded to ResNet-101 in order to accomplish more precise detection.

A Semi Smart Trash Separator to detect and classify garbage and trash was proposed by [58], precycling techniques was used by assigning a barcode or QR code to each material, which will enable the separation process as per assigned code; Magnetic separator helps in collecting conductive metal, then the non-conductive materials are classified according to their hardness. The material recognition accuracy rate from the obtained results on AlexNet and GoogLeNet are 75% and 83% respectively. The lightweight detection network GhostNet is used in [59], which detects trash in real-time outside using robots, as the backbone of the detection network. The network was trained using a dataset that was created by researchers, and it contains four different categories of items. The findings of the experiments reveal that the upgraded version of the YOLOv4 method that was proposed has better detection performance compared to the YOLOV4 algorithm that was used initially, and that it has created adequate generalization performance in a variety of various sorts of trash, similar waste detection research for robotics applications was conducted by [60].

YOLO-Green is a waste identification model that was proposed by [61]. The model was trained on a dataset that was acquired from real-world trash and was then categorized into seven of the most prevalent forms of solid waste. YOLO-Green has an amazing mAP of 78.04% after only undergoing training for a total of 100 epochs. A fresh and lightweight waste identification system was suggested in the research carried out by [62]. The system makes use of a modified version of the yolov5 algorithm. In addition to this, the researchers came up with two approaches that they named tracking object transmission and video backtracking. These methods, together with a tracking algorithm that was based on a kernelized correlation filter, were proposed by the researchers. Table 2 summarizes some of the existing waste detection methods/models.

B. WASTE CLASSIFICATION

Many scholars have begun research in this field in the context of promoting waste sorting and recycling and its effect [63]. The research by [64] uses Trashnet dataset that consists of 6 classes of trash objects for the trash image classification. Support vector machines (SVM) with scale-invariant feature transform (SIFT) features and a convolutional neural network (CNN) were used as models. In their experiment, the SVM outperformed the CNN; however, the CNN was not

TABLE 2. Summary of the existing waste detection methods.

Paper	Models	Accuracy	Comment
2022 [37]	YOLOV4	67%	The dataset used was plastic and paper in the river. Dataset Used: (TACO)
2022 [54]	YOLOv5+Mobile Netv3	97.7%	Underwater images was used and the device is immersed in to water. Data set Used: (Under water trash)
2022 [56]	Mask RCNN	83%	Other models were tested but Mask RCNN is the best. Dataset Used: (Roadside Trash)
2021 [61]	YOLO-Green	78.04%	The model is very large to work on the low resource devices such as smart phones,it also needs to be train on a more diverse dataset. Dataset Used: (self-collected)
2020 [52]	SSD, YOLO-v3-Tiny, RetinaNet, Peleenet (+attention)	48.1%	Focus on detecting waste floating in the river using self-collected data. Dataset Used: (Self-collected)
2020 [53]	ResNet50	81.48%	The dataset focuses on four trash objects only. Dataset Used: (Aqua trash)
2020 [57]	L-SSD	83.48%	The model is tested with trashnet dataset, there is need for more powerful model to test. Dataset Used: (TrashNet)
2019 [47]	SSD MobileNetV2, Tiny-YOLO, Faster RCNN	Highest mAP is 67.4%	Marine debris dataset was used, that is images in water. Dataset Used: (J-EDI)
2018 [48]	YOLOV2	89.71	The model doesn't detect individual object in the targeted trash. Dataset Used: (Self Collected)
2016 [50]	FastRCNN	68%	The images are in clear background. Dataset Used: (Trashnet)

trained to its full potential due to difficulties in determining optimal hyperparameters. The SVM outperformed the Neural Network in terms of performance. It achieved a 63 % using a 70/30 train/test data split. A neural network with a 70/30 train/test split achieved a 27 % testing accuracy.

The model RecycleNet developed by [32] is a carefully optimized deep convolutional neural network architecture for the classification of selected recyclable object classes. Trashnet dataset was also used, and many deep learning models were tested to classify waste with both saved model weight and training from scratch.

Deep Learning models can be hybridized to improve the accuracy of object classification models. In a study by [65] uses 5000 images with a resolution of 640 by 480 pixels and a plain grey background are used. When the investigated items have strong image features, both the Multilayer Hybrid System and CNN perform well. CNN, on the other hand, performs poorly when waste items lack distinguishing image features, particularly “other” waste. Under two different testing scenarios, MHS achieves significantly higher classification performance: the overall performance accuracies are 98.2 percent and 91.6 percent, respectively (the accuracy of the reference model is 87.7 percent and 80.0 percent). The item is positioned in both fixed and random orientations.

As trash can belong to different environments, [66] propose a deep learning approach for medical waste identification and classification. The authors propose ResNeXt, a deep learning-based classification method that was applied to 3480 images and successfully identified 8 types of medical waste with an accuracy of 97.2 percent; the average F1-score of five-fold cross-validation was 97.2 percent.

DSCR-Net model was proposed by [67] The study creates an open-source dataset with a large sample size. The dataset's classification is based on the Shanghai Municipal Household Waste Management Regulations, and it is the first open-source dataset to use this classification method. To facilitate migration, the study's new algorithm borrows from Inception-V4 and the ResNet network, and some layers of the model have been adjusted. With a 94.38 % accuracy rate, the new algorithm was optimized and tested on the dataset.

Most of the trash classification models focus on a single object in an image. A paper by [68] attempts to identify a single trash object in an image and classify it into one of the recycling categories. Support vector machines (SVM) with HOG features, simple convolutional neural networks (CNN), and CNN with residual blocks are among the models used. According to the results of the evaluation, they conclude that simple CNN networks with or without residual blocks perform well. Besides a single object detection, a single trash class was investigated by [69].

Different types of waste necessitate different management techniques; thus, proper waste segregation according to type is essential to facilitate proper recycling. The current method of segregation still relies on manual hand-picking. In the paper [70], a method for classifying wastes using images into six different waste types (glass, metal, paper, plastic, cardboard, and others) based on deep learning and computer vision concepts is proposed. For waste classification, a multiple-layered Convolutional Neural Network (CNN) model, specifically the well-known Inception-v3 model, was

used, with a trained dataset obtained from online sources. The proposed method achieves a high classification accuracy of 92.5%. A model that can realize intelligent decision-making for garbage categorization for big data in the scene in a complicated system is proposed in [71]. This model also includes certain conditions for promotion and landing. The findings of the tests model indicate that the suggested model has a greater level of accuracy in terms of both detection and classification in comparison to the original YOLOv5 model, and that it is also capable of meeting the actual application requirements in terms of its real-time performance. In a study by [72] makes a suggestion for an algorithm that is based on InceptionV3 networks and tests the model on a garbage classification dataset that is quite huge in scale. Transfer learning was used in the dataset, which was then segmented into training sets consisting of 80 %, validation sets consisting of 10 %, and test sets consisting of 10 %. The accuracy of the model was determined to be 93.125 %.

In paper [73], the authors present a novel garbage image recognition model called Garbage Classification Net (GCNet), which is based on transfer learning and model fusion. Following the extraction of trash image features, the neural network models EfficientNetv2, Vision Transformer, and DenseNet are successively integrated to create the GCNet model of the garbage classification neural network. The dataset is expanded with the help of data augmentation, and the expanded dataset contains 41,650 images that are considered to be trash. The authors of [74] work on constructing a deep CNN that is tailored specifically for garbage image classification. The authors came up with the idea for the attention module known as DSCAM, which offers an original method to build attention weights. a large number of other classification models, including VGG16, Xception, MobileNet-V3, and GNet, among others, were evaluated, and the proposed DSCAM models were found to have the highest accuracy of 98.9%.

A deep neural network model for garbage classification was developed by [75] and given the name DNN-TC. This model is an improvement on the ResNext model and was developed to increase the predicted performance. After the global average pooling layer, the authors of this study changed the original ResNext-101 model by adding two fully connected layers with outputs of 1024 and N class dimensions respectively. This was done to reduce the amount of redundancy in the model. In order to evaluate the model, both the VN-trash dataset and the Trashnet dataset were utilized. [76] proposes a potential solution by creating AlphaTrash, a machine that can be fitted to conventional curb-side trashcan and used to sort out deposit trash automatically. The researchers use a pre-trained convolutional neural network (Inception- v1), the machine can classify trash with an accuracy of 94%, while using 4.2 seconds per classification.

Due to the scarcity of trash data, for the purpose of data regeneration, [77] make use of both the two-stage variational autoencoder (VAE) and the binary classifier (augmentation). An evaluation of the effect that the augmentation procedure

has is carried out with the use of a multi-class classifier. This is done by determining how well an object detector was educated using a mixture of actual and simulated trash image. [78] focuses on the classification of garbage using meta-data and evaluated the strategy using multiple deep learning algorithms such as VGG16, ResNet50, and DenseNet169 to compare it with the recently developed model ThanosNet, which achieved an accuracy of 94%. A lot of more research focuses on trash image classification from different devices such as in [79] for robotics, and those purely works with CNN with low accuracy such as in [80] and [81] using different benchmark datasets. The summary of the trash classification-based research is in Table 3.

IV. BENCHMARK TRASH DATASET

Trash-based object detection is one of the trending topics in the field of object detection of computer vision. In a polluted, controlled environment, these systems enhance living things’ quality of life and health. However, it is vital to assess the efficacy of the automated waste detector or classifier in detecting and classifying trash using dataset benchmarks in experimental scenarios, especially in the wild, before making it available to end users. As a result, the scientific community has created and made available a number of databases for the detection and classification of refuse. Therefore, selecting the ones to employ in the evaluation process and the strategies that are most suited for a specific environment or waste class is a difficult challenge and essential to moving this field of study forward.

The detection, classification, and segmentation of waste using deep learning have been the subject of numerous recent attempts. A few available datasets have been used to try and classify litter/trash according to typical waste types using images. Table 4 summarizes the available datasets. Figure 10 and Figure 11 show the image samples from the trash net dataset and the collected images in the wild, respectively. From Figure 11, some challenges of the presence of other objects and an unstable background can be observed that will leads to the inability of the deep learning models to generalize in the trash detection and classification tasks. While in figure 10 all images have a clear background that also will not make the trained model with that kind of image perform well in real life.

A. TRASHNET DATASET [64]

The dataset includes six categories: waste, glass, paper, cardboard, plastic, and metal. The dataset comprises 2,527 photos labeled with a category (501 glass, 594 paper, 403 cardboard, 482 plastic, 410 metal, and 137 rubbish/trash). The dataset consists of images of trash taken on a white backdrop using various exposure and lighting settings (mainly one object per photo). The authors investigate the SVM and CNN algorithms to sort waste into six recycling categories. They employed an architecture resembling AlexNet but with fewer and smaller filters. Results from the SVM were superior to those from the Neural Network. Using a 70/30 train/test data split, it attained

TABLE 3. Summary of the existing waste classification methods.

Year	Models	Accuracy	Comment
2022 [73]	GCNet	97.54%	Combine dataset collected by hand and online with clear background was used. Dataset Used: (self-collected)
2022 [74]	DSCAM	98.9%	Three different kind of trash dataset was merged and used in testing the model. Dataset Used: (Huawei-40 and Baidu-214)
2022 [83]	ResNet + Raspberry pi	96%	The dataset used is very limited and depend on the IoT device. Dataset Used: (UAVWaste)
2022 [82]	Mask-RCNN, YOLO, Custom VGG16	90%	The dataset used are complete trash and couldn’t classify all type of waste. Dataset Used: (TACO)
2021 [84]	EfficientDet	92.87%	Multiple datasets are combined to form 8135 records. Dataset Used: (WaDaBa)
2021 [71]	YOLOv5-Attention-KG	73.2%	It is computationally expensive as it requires very large hardware to run the model. Dataset Used: (Self-collected)
2021 [72]	InceptionV3	93.13%	Household dataset was used in the training and testing the model. Dataset Used: (Kaggle)
2020 [77]	VAE-Model	96% accuracy	Under water images are used with limited number of classes. Dataset Used: (Marine Litter)
2020 [79]	MarineTrash	90.60 accuracy	A lot of misclassifications happened as a result of minimal data preprocessing. Dataset Used: (Marine Trash)
2020 [78]	ThanosNet	94.7%	Minimal number of images and classes are considered. Dataset Used: (ISBNet)
2019 [75]	DNN-TC	98%	Trashnet dataset only was used, and it hasn’t been tested on images in the wild. Dataset Used: (Trashnet)
2019 [76]	AlphaTrash	94%	Uses RFID technology which is costly and not perfect to use in real-time. Dataset Used: (AlphaTrash)
2018 [37]	ResidualNet (RecycleNet)	DenseNet121 gave the best result with 95% test accuracy.	The network is slower to train. Dataset Used: (TrashNet)
2018 [60]	CNN and AlexNet	92.2%	Images are in a clear background and slower to train. Dataset Used: (TrashNet)
2016 [59]	SVM and Convolution Neural Network (CNN)	SVM archived 63% and CNN archived 27%	The dataset used is not evenly distributed and is limited to a single object in an image. CNN is expected to outperform SVM. Dataset Used: (Self-collected)

an accuracy of 63 percent. Testing accuracy for a neural network with a 70/30 train/test split was 27%.

Garbage in Image (GINI) Dataset [51]: The Garbage in Photos (GINI) collection consists of 2,561 images of an

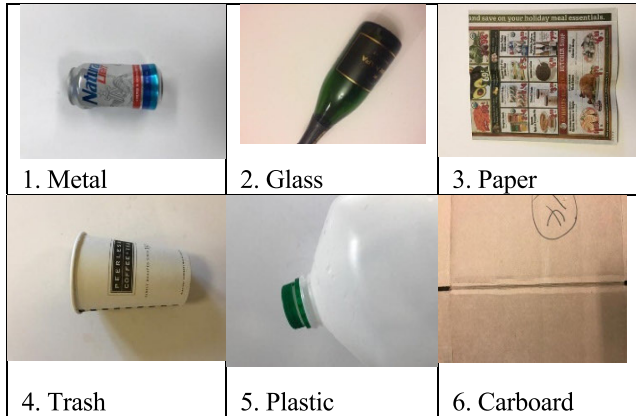


FIGURE 10. Trashnet image samples.

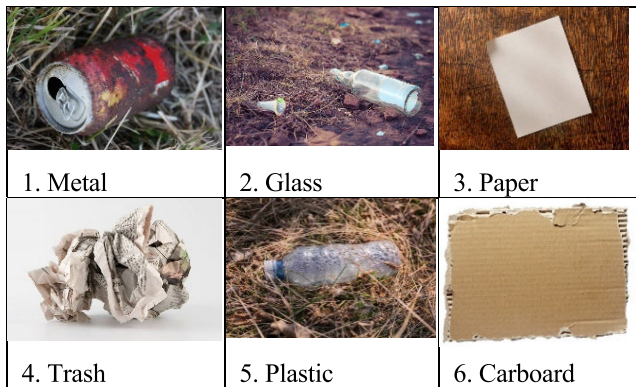


FIGURE 11. Images collected from the wild samples.

unknown resolution, 1,496 of which have bounding boxes annotated (one class - trash). They created their dataset using the Bing Image Search API. The authors use a pre-trained AlexNet, and their method focuses on segmenting a pile of trash in an image without giving any information regarding the kinds of waste that are contained in that segment. Their approach, which is based on extracting image patches and aggregating predictions, is unable to capture the more minute characteristics of object boundaries. GarbNet achieved an accuracy of 87.69% for the task of garbage detection; however, it made incorrect predictions when it found waste-like things in an image or when they were far away.

B. TrashICRA19 Dataset [47]

This data set originated from the marine debris J-EDI dataset. The quality, depth, scene items, and cameras used in the videos that make up that dataset vary widely. They provide a variety of things in various levels of decay, occlusion, and overgrowth. They include photographs of many distinct forms of marine trash, taken from actual locations and containing a wide range of objects. Additionally, the water’s transparency and the quality of the light differ greatly from one video to the next. 5,700 photos total, all annotated with bounding boxes on instances of trash, biological items including plants and

TABLE 4. Summary of the benchmark trash dataset.

Ref.	Dataset Name	No. of Images	Comment
[64]	TrashNet	2,527	Clear background images
[51]	Garbage in Image (GINI)	1,496	In the wild
[46]	TrashICRA19 Dataset	5,700	Underwater Images
[82]	TACO	1,500	In the wild images
[83]	UAVVaste	772	Drone Dataset
[84]	Wadaba Dataset	4,000	Plastic dataset, clear background
[85]	GLASSENSE-VISION dataset	2,000	Home supplies, clear background
[86]	MJU-Waste	2,475	Plain background, indoor RGB images
[89]	Waste Classification data	22,500	Scrapped from google search
[90]	Waste Images from Sushi Restaurant	500	Clear Background
[91]	Open litter map	Over 1000	Waste in the wild
[92]	Litter	Almost 14,000	Waste in the wild
[93]	Drinking Waste dataset	9,640	Cans and Bottles with clear background
[94]	Spot Garbage	Almost 24,000	Scrapped from web
[95]	Deep Sea Waste	3,055	Under water images
[96]	Domestic Trash Dataset	9,000	Waste in the wild
[97]	Cigarette butt dataset	2,200	Waste in the wild, synthetic images
[98]	Trash box dataset	1,7785	Scrapped from web
[88]	Waste Pictures dataset	Almost 24,000	Scrapped from web
[99]	Wade.AI	2,200	Scrapped from web
[100]	NWNU-TRASH	18,911	Scrapped from web
[101]	Classify-Waste	21,000	Other dataset combination
[97]	Cigarette Butt Dataset	2,200	Images in the wild

animals, and ROVs, were extracted from these films through processing to make up this dataset. The ultimate objective is to create effective and precise waste-detecting techniques suited for onboard robot deployment.

C. TACO [82]

The Trash Annotations in Context (TACO) is an open image dataset for litter detection and segmentation, which is growing through crowdsourcing. TACO contains high-resolution images, taken mostly by mobile phones. 1500 annotated images with approximately 5000 objects made up the TACO dataset. All trash has been categorized into one of 60 classifications, which are subdivided into 28 super (top) categories, including Unlabeled litter for objects that are difficult to identify or that are extensively concealed. On the

instance segmentation level, the annotations are provided in the well-known COCO format [33] with an additional background description for Trash, Vegetation, Sand, Water, Indoor, and Pavement. Additionally, TACO provides about 3,000 unannotated images, of which more than 3,000 were annotated on the detection level, resulting in a total of over 14,000 instances. The fact that TACO is distinguished by a wide range of litter types and a sizable diversity of backgrounds, from tropical beaches to London streets, is a huge benefit. Although labels may contain some user-induced bias and inaccuracies due to the dataset's crowdsourcing nature, not all objects in TACO may be categorically classified strictly as litter as their category is frequently reliant on context.

D. UAVWaste [83]

The public UAVWaste dataset currently has 772 photos and 3718 annotations and is expected to be updated. The primary motivation for developing this dataset was a lack of domain-specific data. As a result, this image set is recommended not only for benchmarking object detection evaluations, but also for building solutions connected to UAVs, remote sensing, and even environmental cleaning.

E. WADABA [84]

The WADABA dataset contains images of plastic trash collected from households. A minimum of 100 objects were planned for capture, with each object receiving forty photographs under various situations. Two types of lighting were used: fluorescent lamps and LEDs. The image settings are as follows: 1920×1277 pixels, 300 dpi resolution, RGB 24 bit colour palette, and JPG file format.

F. GLASSENSE-VISION [85]

The Glassense-Vision dataset is a collection of image data that has been collected from different objects. The collection contains 505 photos representing several object types (banknotes, cereals, medications, cans, tomato sauce, water bottles, and deodorant sticks). All the images in the collection have been manually annotated. The many use cases (object categories) can be classified into three geometrical types: flat items, boxes, and cylindrical things. All photos were saved with a resolution of 665×1182 pixels.

G. MJU-Waste [86]

This dataset was developed by collecting waste items from a university campus, transporting them to a lab, and photographing people carrying waste objects. The images in the collection were all taken by the author. with a Microsoft Kinect RGBD camera. This dataset's current version, MJU-Waste V1, contains 2475 co-registered RGB and depth picture pairs. The dataset was specifically divided into a training set, a validation set, and a test set of 1485, 248, and 742 photos, respectively. The depth frames contain missing data at reflecting surfaces, occlusion boundaries, and remote locations due to sensor limitations. In order to obtain

high-quality depth photographs, the median filter was employed to fill in the missing values. MJU-Trash annotates each image with a pixel-wise mask of waste elements.

H. OPEN LITTER MAP [87]

Over 100k images from phone cameras make up the free, public, and crowd-sourced dataset known as Open Litter Map. Each image includes details such as the kind of litter it was taken from, the coordinates, the timestamp, or the phone model. Images came from all across the world and were captured by various individuals. Consequently, they are very different from one another.

I. WASTE PICTURES [88]

Nearly 24000 trashes images from Google searches are collected in Waste Pictures, which is split into 34 classes. Even x-rays and drawing of trash are included in the wide variety of images. The image sizes vary greatly as well. However, the majority of the images are smaller than 2000×2000 pixels. Images should be carefully examined before being used in a categorization task due to their provenance.

J. Wade-AI [99]

Images of trash in a natural setting are available in the Wade-AI dataset thanks to Google Street View. It has 2200 manually labelled instance mask annotations on around 1400 photos in COCO format, all of which belong to the same class, garbage. The source of the photographs affects the environment and size of the images. The majority of photos are less than 1000×1000 .

K. NWNUTRASH [100]

The web crawler technology, Python code, and manual photography were used to create a recyclable waste image dataset named NWNUTRASH, which includes waste glass (3845), waste fabric (3862), wastepaper (3766), waste plastic (3865), and waste metal (3573), with a total of 18911 images. Different backgrounds are chosen for the images, and the number of different types of waste images is balanced, as is the data diversity, which is more in line with the needs of the real background.

L. CLASSIFY-WASTE [101]

Over 21000 waste instances from Extended TACO, drinking-waste, waste pictures, Google search, TrashNet, and Places are included in the classify-waste dataset. The majority of trash is made up of metal and plastic, or an unknown category that is closely related to the distribution of waste types produced by humans. Nonetheless, it contains a diverse set of trash that ensures the generalizability of a model trained on this dataset. The waste classification dataset contains eight labels the categories include: Fruit, vegetables, herbs, used paper towels, and tissues are examples of biowaste. Glass objects include glass bottles, jars, and cosmetic packaging. Scrap metal and nonferrous metal, beverage cans, plastic beverage bottles, plastic shards, plastic food packaging, or plastic

straws are examples of metals and plastics. Non-recyclable items include disposable diapers, string, polystyrene packaging, polystyrene elements, blankets, clothing, and used paper cups. Other types of waste include construction and demolition waste, large-sized waste (such as tyres), used electronics and household appliances, batteries, paint, and varnish cans, and expired medicines. Paper, cardboard packaging, receipts, newspapers, catalogs, and books are all examples of paper. Unknown waste (highly decomposed and difficult-to-identify litter), and extra class background label (no litter): a sidewalk, a forest trail, and a lawn.

M. CIGARETTE BUTT DATASET [97]

This dataset contains 2200 images of cigarettes on the ground that were created synthetically. It is intended for CNN training (convolutional neural networks). The images were generated automatically using custom code that used the Python Imaging Library to apply random scale, rotation, brightness, and other parameters to the foreground cutouts. iPhone 8 is used in taking pictures, and the original pixel resolution was 3024 by 4032.

There are currently few open waste datasets, with the TrashNet dataset being the most widely used. It is a small collection of recyclable waste images, including glass, paper, cardboard, plastic, and metal, with 2,527 photos in total. Currently, the majority of waste classification research based on image recognition is based on the TrashNet dataset, which has a high classification accuracy rate. This dataset, however, has some flaws: 1. The amount of sample data is insufficient; 2. The number of different types of waste is unevenly distributed; 3. The background of the image is single or clear, which does not meet the needs of real scenes and is detrimental to the training model's generalization ability; and 4. The number of items is insufficient to represent the majority of objects in a community or domain.

V. CHALLENGES OF WASTE DETECTION AND CLASSIFICATION

Even though deep learning-based models have emerged as an extremely powerful framework for dealing with various types of vision problems, such as image classification [26], object detection [26], [102], and more relevantly single object tracking. Despite the contributions of DL models, there are still challenges that remain in trash or garbage detection and classification.

A. LIGHTNING CONDITION

Because of illumination changes, problems become more complex, as different lighting conditions change the visibility of an object or should alter its appearance which leads to serious difficulty.

B. INSUFFICIENT DATA

lack of available trash data is a major obstacle affecting the implementation of AI systems. AI models are primarily driven by large data sets for training and calibration. Current

research is frequently hampered by a lack or inadequacy of waste data. This is partly because waste and trash management industries are mostly out-of-date, with few reliable records and scarce sensory data, particularly in developing countries.

C. OBJECT SIZE AND LOCATION

Objects in low visibility should not be visible enough to be recognized. The system may fail if the object is too small or the distance from the system is too great. Various lighting conditions and shadows should also make it difficult to identify the images.

D. OBJECT LOCALIZATION OR IN THE WILD

Image classification to determine the class of the images is a major problem in object detection and identification. The system was unable to predict the location of the object in the images. As a result, image classification is a major issue.

E. OCCLUSION OR TRASH IN THE WILD

Some objects are blocked or hidden by the image of another object's presence, which leads to the blockage of most or some parts of the targeted object, which will cause serious low recognition accuracy.

VI. CONCLUSION

In conclusion, this paper discusses a vast number of research papers, to be exact on the subject of deep learning in trash detection and classification, as well as object recognition, with a primary emphasis on the most recently published articles in the field. In the references, you will find a list of the papers that were utilized for the purpose of this survey. The papers that were collected were from reputable and reliable publishers such as IEEE, Scopus, Google Scholar, and Springer, amongst others. The purpose of this survey study is to investigate the many different uses of deep learning for recognizing and classifying waste. This study provides an orderly and comprehensive assessment of numerous available methods for the detection and classification of garbage using machine learning and deep learning. In addition to this, it easily explained benchmarked datasets on the detection and classification of trash in a variety of settings. In order to support the study, both the benefits and the drawbacks of the existing methods and datasets, as well as the possibilities for future research, are highlighted. In addition, we are considering performing a systematic literature review on this subject, and also experimenting with different machine learning and deep learning algorithms.

REFERENCES

- [1] K. N. Sami, Z. M. A. Amin, and R. Hassan, "Waste management using machine learning and deep learning algorithms," *Int. J. Perceptive Cogn. Comput.*, vol. 6, no. 2, pp. 97–106, Dec. 2020, doi: 10.31436/ijpcc.v6i2.165.
- [2] S. Shahab, M. Anjum, and M. S. Umar, "Deep learning applications in solid waste management: A deep literature review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 3, pp. 381–395, 2022, doi: 10.14569/IJACSA.2022.0130347.

- [3] M. Triassi, R. Alfano, M. Illario, A. Nardone, O. Caporale, and P. Montuori, "Environmental pollution from illegal waste disposal and health effects: A review on the 'triangle of death,'" *Int. J. Environ. Res. Public Health*, vol. 12, no. 2, pp. 1216–1236, Jan. 2015, doi: [10.3390/ijerph120201216](https://doi.org/10.3390/ijerph120201216).
- [4] A. A. Namen, F. da Costa Brasil, J. J. G. Abrunhosa, G. G. S. Abrunhosa, R. M. Tarré, and F. J. G. Marques, "RFID technology for hazardous waste management and tracking," *Waste Manage. Res., J. Sustain. Circular Economy*, vol. 32, no. 9, pp. 59–66, Sep. 2014, doi: [10.1177/0734242X14536463](https://doi.org/10.1177/0734242X14536463).
- [5] S. S. Chandra, M. Kulshreshtha, and P. Randhawa, "Garbage detection and path-planning in autonomous robots," in *Proc. 9th Int. Conf. Rel., Infocom Technol. Optim. Trends Future Directions (ICRITO)*, Sep. 2021, pp. 1–4, doi: [10.1109/ICRITO51393.2021.9596382](https://doi.org/10.1109/ICRITO51393.2021.9596382).
- [6] N. Sarker, S. Chaki, A. Das, and M. S. A. Forhad, "Illegal trash thrower detection based on HOG/SVM for a real-time monitoring system," in *Proc. 2nd Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST)*, Jan. 2021, pp. 483–487, doi: [10.1109/ICREST51555.2021.9331183](https://doi.org/10.1109/ICREST51555.2021.9331183).
- [7] M. Didelija, N. Kulo, A. Mulahusić, N. Tuno, and J. Topoljak, "Segmentation scale parameter influence on the accuracy of detecting illegal landfills on satellite imagery. A case study for Novo Sarajevo," *Ecol. Informat.*, vol. 70, Sep. 2022, Art. no. 101755, doi: [10.1016/j.ecoinf.2022.101755](https://doi.org/10.1016/j.ecoinf.2022.101755).
- [8] E. Morin, P.-A. Herrault, Y. Guinard, F. Grandjean, and N. Bech, "The promising combination of a remote sensing approach and landscape connectivity modelling at a fine scale in urban planning," *Ecological Indicators*, vol. 139, Jun. 2022, Art. no. 108930, doi: [10.1016/j.ecolind.2022.108930](https://doi.org/10.1016/j.ecolind.2022.108930).
- [9] S. Bhandari, "Automatic waste sorting in industrial environments via machine learning approaches," Tampere Univ., Tampere, Finland, Oct. 2020. [Online]. Available: <https://trepo.tuni.fi/handle/10024/123574>
- [10] P. White, M. Franke, and P. Hindle, "Lifecycle inventory: A part of lifecycle assessment," in *Integrated Solid Waste Management: A Lifecycle Inventory*, 1st ed. Boston, MA, USA: Springer, 1999, pp. 25–36, doi: [10.1007/978-1-4615-2369-7_3](https://doi.org/10.1007/978-1-4615-2369-7_3).
- [11] A. Demirbas, "Waste management, waste resource facilities and waste conversion processes," *Energy Convers. Manage.*, vol. 52, no. 2, pp. 1280–1287, Feb. 2011, doi: [10.1016/j.enconman.2010.09.025](https://doi.org/10.1016/j.enconman.2010.09.025).
- [12] J. Wang, T. Zhang, Y. Cheng, and N. Al-Nabhan, "Deep learning for object detection: A survey," *Comput. Syst. Sci. Eng.*, vol. 38, no. 2, pp. 165–182, 2021, doi: [10.32604/CSSE.2021.017016](https://doi.org/10.32604/CSSE.2021.017016).
- [13] M. S. Swetha, M. S. Veena, M. S. Muneshwara, and M. Thungamani, "Survey of object detection using deep neural networks," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 7, no. 11, pp. 19–24, Nov. 2018, doi: [10.17148/ijarce.2018.71104](https://doi.org/10.17148/ijarce.2018.71104).
- [14] M. Yewange, K. Gaikwad, R. Kamble, S. Maske, and R. Shahu. (2022). *Real-Time Object Detection by using Deep Learning: A Survey*. [Online]. Available: <http://www.ijisrt.com>
- [15] M. Ahmed, K. A. Hashmi, A. Pagani, M. Liwicki, D. Stricker, and M. Z. Afzal, "Survey and performance analysis of deep learning based object detection in challenging environments," *Sensors*, vol. 21, no. 15, p. 5116, Jul. 2021, doi: [10.3390/s21155116](https://doi.org/10.3390/s21155116).
- [16] B. S. Rekha, A. Mariam, G. N. Srinivasan, and S. A. Shetty, "Literature survey on object detection using YOLO," *Int. Res. J. Eng. Technol.*, vol. 7, no. 6, pp. 3082–3088, 2020. [Online]. Available: <https://www.irjet.net/archives/V7/i6/IRJET-V7I6576.pdf>
- [17] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling, and R. Yang, "Salient object detection in the deep learning era: An in-depth survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 6, pp. 3239–3259, Jun. 2022, doi: [10.1109/TPAMI.2021.3051099](https://doi.org/10.1109/TPAMI.2021.3051099).
- [18] L. Ferrara, M. Iannace, A. M. Patelli, and M. Arienzo, "Geochemical survey of an illegal waste disposal site under a waste emergency scenario (Northwest Naples, Italy)," *Environ. Monit. Assessment*, vol. 185, no. 3, pp. 2671–2682, Mar. 2013, doi: [10.1007/s10661-012-2738-2](https://doi.org/10.1007/s10661-012-2738-2).
- [19] A. Prasanna M, S. V. Kaushal, and P. Mahalakshmi, "Survey on identification and classification of waste for efficient disposal and recycling," *Int. J. Eng. Technol.*, vol. 7, no. 2.8, p. 520, Mar. 2018, doi: [10.14419/ijet.v7i2.8.10513](https://doi.org/10.14419/ijet.v7i2.8.10513).
- [20] T. Yang, J. Xu, Y. Zhao, T. Gong, R. Zhao, M. Sun, and B. Xi, "Classification technology of domestic waste from 2000 to 2019: A bibliometrics-based review," *Environ. Sci. Pollut. Res.*, vol. 28, no. 21, pp. 26313–26324, Jun. 2021, doi: [10.1007/s11356-021-12816-x](https://doi.org/10.1007/s11356-021-12816-x).
- [21] Y. Wang, Q. Wang, S. Jin, W. Long, and L. Hu, "A literature review of underwater image," in *Frontiers in Artificial Intelligence and Applications*, vol. 347. Amsterdam, The Netherlands: IOS Press, 2022, pp. 42–51, doi: [10.3233/FAIA220009](https://doi.org/10.3233/FAIA220009).
- [22] F. Zhang, C. Cao, C. Li, Y. Liu, and D. Huisingh, "A systematic review of recent developments in disaster waste management," *J. Cleaner Prod.*, vol. 235, pp. 822–840, Oct. 2019, doi: [10.1016/j.jclepro.2019.06.229](https://doi.org/10.1016/j.jclepro.2019.06.229).
- [23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, 2005, pp. 886–893, doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177).
- [24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004, doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).
- [25] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Feb. 2001, pp. 511–518.
- [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton. *ImageNet Classification With Deep Convolutional Neural Networks*. Accessed: Jul. 2, 2022. [Online]. Available: <http://code.google.com/p/cuda-convnet/>
- [27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Sep. 2014, *arXiv:1409.1556*.
- [28] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," Feb. 2016, *arXiv:1602.07261*.
- [29] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, N. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," Apr. 2017, *arXiv:1704.04861*.
- [30] K. He, X. Zhang, S. Ren, and J. Sun. *Deep Residual Learning for Image Recognition*. Accessed: Jul. 2, 2022. [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [32] C. Bircanoglu, M. Atay, F. Beser, O. Genc, and M. A. Kizrak, "RecycleNet: Intelligent waste sorting using deep neural networks," in *Proc. Innov. Intell. Syst. Appl. (INISTA)*, Jul. 2018, pp. 1–7, doi: [10.1109/INISTA.2018.8466276](https://doi.org/10.1109/INISTA.2018.8466276).
- [33] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, in Lecture Notes in Computer Science, vol. 8693, 2014, pp. 740–755.
- [34] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," Nov. 2013, *arXiv:1311.2524*.
- [35] C.-Y. Wang, I.-H. Yeh, and H.-Y. Mark Liao, "You only learn one representation: Unified network for multiple tasks," May 2021, *arXiv:2105.04206*.
- [36] J. Redmon and A. Farhadi. *YOLOv3: An Incremental Improvement*. Accessed: Jul. 2, 2022. [Online]. Available: <https://pjreddie.com/yolo/>
- [37] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," Apr. 2020, *arXiv:2004.10934*.
- [38] (2022). *YOLOv7: The Fastest Object Detection Algorithm*. Accessed: Sep. 22, 2022. [Online]. Available: <https://viso.ai/deep-learning/yolov7-guide/>
- [39] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," Jul. 2022, *arXiv:2207.02696*.
- [40] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," Jun. 2015, *arXiv:1506.02640*.
- [41] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "CenterNet: Keypoint triplets for object detection," Apr. 2019, *arXiv:1904.08189*.
- [42] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "CenterNet: Keypoint triplets for object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6568–6577, doi: [10.1109/ICCV.2019.00667](https://doi.org/10.1109/ICCV.2019.00667).
- [43] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," Nov. 2019, *arXiv:1911.09070*.
- [44] X. Zhou, J. Zhuo, and P. Krähenbühl, "Bottom-up object detection by grouping extreme and center points," Jan. 2019, *arXiv:1901.08043*.
- [45] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," Mar. 2017, *arXiv:1703.06870*.
- [46] B. D. Carolis, F. Ladogana, and N. Macchiarulo, "YOLO TrashNet: Garbage detection in video streams," in *Proc. IEEE Trans. Evolving Adapt. Intell. Syst. (EAIS)*, May 2020, pp. 1–7, doi: [10.1109/EAIS48028.2020.9122693](https://doi.org/10.1109/EAIS48028.2020.9122693).
- [47] M. Fulton, J. Hong, M. J. Islam, and J. Sattar, "Robotic detection of marine litter using deep visual detection models," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 5752–5758, doi: [10.1109/ICRA.2019.8793975](https://doi.org/10.1109/ICRA.2019.8793975).
- [48] Y. Liu, Z. Ge, G. Lv, and S. Wang, "Research on automatic garbage detection system based on deep learning and narrowband Internet of Things," *J. Phys., Conf. Ser.*, vol. 1069, Aug. 2018, Art. no. 012032, doi: [10.1088/1742-6596/1069/1/012032](https://doi.org/10.1088/1742-6596/1069/1/012032).

- [49] Y. Wang and X. Zhang, "Autonomous garbage detection for intelligent urban management," in *Proc. MATEC Web Conf.*, vol. 232, Nov. 2018, p. 01056, doi: [10.1051/mateconf/201823201056](https://doi.org/10.1051/mateconf/201823201056).
- [50] O. A. Mengistu, "Smart trash Net: Waste localization and classification," pp. 1–6, 2017. [Online]. Available: <https://www.semanticscholar.org/paper/Final-Report-%3A-Smart-Trash-Net-%3A-Waste-Localization-Awe-Mengistu/581fb0f0405c7f0e60610d88ceaceb9af44d8569>
- [51] G. Mittal, K. B. Yagnik, M. Garg, and N. C. Krishnan, "SpotGarbage: Smartphone app to detect garbage using deep learning," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2016, pp. 940–945, doi: [10.1145/2971648.2971731](https://doi.org/10.1145/2971648.2971731).
- [52] M. Tharani, A. W. Amin, M. Maaz, and M. Taj, "Attention neural network for trash detection on water channels," Jul. 2020, *arXiv:2007.04639*.
- [53] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, S. Sharma, and I. H. Sarker, "AquaVision: Automating the detection of waste in water bodies using deep transfer learning," *Case Stud. Chem. Environ. Eng.*, vol. 2, Sep. 2020, Art. no. 100026, doi: [10.1016/j.csee.2020.100026](https://doi.org/10.1016/j.csee.2020.100026).
- [54] C. Wu, Y. Sun, T. Wang, and Y. Liu, "Underwater trash detection algorithm based on improved YOLOv5s," *J. Real-Time Image Process.*, vol. 19, no. 5, pp. 911–920, Oct. 2022, doi: [10.1007/S11554-022-01232-0](https://doi.org/10.1007/S11554-022-01232-0).
- [55] Y. Yu, "A computer vision based detection system for trash bins identification during trash classification," *J. Phys., Conf. Ser.*, vol. 1617, no. 1, Aug. 2020, Art. no. 012015, doi: [10.1088/1742-6596/1617/1/012015](https://doi.org/10.1088/1742-6596/1617/1/012015).
- [56] G. Conley, S. C. Zinn, T. Hanson, K. McDonald, N. Beck, and H. Wen, "Using a deep learning model to quantify trash accumulation for cleaner urban stormwater," *Comput., Environ. Urban Syst.*, vol. 93, Apr. 2022, Art. no. 101752, doi: [10.1016/J.COMPENVURBSYS.2021.101752](https://doi.org/10.1016/J.COMPENVURBSYS.2021.101752).
- [57] W. Ma, X. Wang, and J. Yu, "A lightweight feature fusion single shot multibox detector for garbage detection," *IEEE Access*, vol. 8, pp. 188577–188586, 2020, doi: [10.1109/ACCESS.2020.3031990](https://doi.org/10.1109/ACCESS.2020.3031990).
- [58] F. Alzyoud, W. Maqableh, and F. Al Shrouf, "A semi smart adaptive approach for trash classification," *Int. J. Comput. Commun. CONTROL*, vol. 16, no. 4, pp. 1–13, Jul. 2021, doi: [10.15837/ijccc.2021.4.4172](https://doi.org/10.15837/ijccc.2021.4.4172).
- [59] P. Zhou, Z. Zhu, X. Xu, X. Liu, B. He, and J. Zhang, "Towards the urban future: A novel trash segregation algorithm based on improved YOLOV4," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2021, pp. 1526–1531, doi: [10.1109/ROBIO54168.2021.9739288](https://doi.org/10.1109/ROBIO54168.2021.9739288).
- [60] S. Hossain, B. Debnath, A. Anika, M. Junaed-Al-Hossain, S. Biswas, and C. Shahnaz, "Autonomous trash collector based on object detection using deep neural network," in *Proc. IEEE Region 10 Conf. (TENCON)*, Oct. 2019, pp. 1406–1410, doi: [10.1109/TENCON.2019.8929270](https://doi.org/10.1109/TENCON.2019.8929270).
- [61] W. Lin, "YOLO-green: A real-time classification and object detection model optimized for waste management," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2021, pp. 51–57, doi: [10.1109/BIG-DATA52589.2021.9671821](https://doi.org/10.1109/BIG-DATA52589.2021.9671821).
- [62] Z. Yu, J. Liu, and X. Li, "LTDTS: A lightweight trash detecting and tracking system," in *Proc. Int. Conf. Adapt. Intell. Syst.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 13338, 2022, pp. 240–250, doi: [10.1007/978-3-031-06794-5_20](https://doi.org/10.1007/978-3-031-06794-5_20).
- [63] C. Yu, J. Xu, A. Zhao, P. Xiao, J. Tai, Z. Bi, and G. Li, "The generation and effects for recyclable waste from households in a megapolis: A case study in Shanghai," *Sustainability*, vol. 14, no. 13, p. 7854, Jun. 2022, doi: [10.3390/su14137854](https://doi.org/10.3390/su14137854).
- [64] M. Yang and G. Thung, "Classification of trash for recyclability status," CS229 Project Rep. 3, 2016. [Online]. Available: <http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf>
- [65] Y. Chu, C. Huang, X. Xie, B. Tan, S. Kamal, and X. Xiong, "Multilayer hybrid deep-learning method for waste classification and recycling," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–9, Nov. 2018, doi: [10.1155/2018/5060857](https://doi.org/10.1155/2018/5060857).
- [66] H. Zhou, X. Yu, A. Alhaskawi, Y. Dong, Z. Wang, Q. Jin, X. Hu, Z. Liu, V. G. Kota, M. H. A. H. Abdulla, S. H. A. Ezzi, B. Qi, J. Li, B. Wang, J. Fang, and H. Lu, "A deep learning approach for medical waste classification," *Sci. Rep.*, vol. 12, no. 1, pp. 1–9, Feb. 2022, doi: [10.1038/s41598-022-06146-2](https://doi.org/10.1038/s41598-022-06146-2).
- [67] F. Song, Y. Zhang, and J. Zhang, "Optimization of CNN-based garbage classification model," in *Proc. 4th Int. Conf. Comput. Sci. Appl. Eng.*, Oct. 2020, pp. 1–5, doi: [10.1145/3424978.3425089](https://doi.org/10.1145/3424978.3425089).
- [68] S. Meng and W.-T. Chu, "A study of garbage classification with convolutional neural networks," in *Proc. Indo Taiwan 2nd Int. Conf. Comput., Anal. Netw. (Indo-Taiwan ICAN)*, Feb. 2020, pp. 152–157, doi: [10.1109/Indo-TaiwanICAN48429.2020.9181311](https://doi.org/10.1109/Indo-TaiwanICAN48429.2020.9181311).
- [69] J. Bobulski and M. Kubanek, "Deep learning for plastic waste classification system," *Appl. Comput. Intell. Soft Comput.*, vol. 2021, pp. 1–7, May 2021, doi: [10.1155/2021/6626948](https://doi.org/10.1155/2021/6626948).
- [70] F. A. Azis, H. Suhaimi, and E. Abas, "Waste classification using convolutional neural network," in *Proc. 2nd Int. Conf. Inf. Technol. Comput. Commun.*, Aug. 2020, pp. 9–13, doi: [10.1145/3417473.3417474](https://doi.org/10.1145/3417473.3417474).
- [71] Y. Wu, X. Shen, Q. Liu, F. Xiao, and C. Li, "A garbage detection and classification method based on visual scene understanding in the home environment," *Complexity*, vol. 2021, pp. 1–14, Nov. 2021, doi: [10.1155/2021/1055604](https://doi.org/10.1155/2021/1055604).
- [72] Y. Chen, W. Han, J. Jin, H. Wang, Q. Xing, and Y. Zhang, "Clean our city: An automatic urban garbage classification algorithm using computer vision and transfer learning technologies," *J. Phys., Conf. Ser.*, vol. 1994, no. 1, Aug. 2021, Art. no. 012022, doi: [10.1088/1742-6596/1994/1/012022](https://doi.org/10.1088/1742-6596/1994/1/012022).
- [73] W. Liu, H. Ouyang, Q. Liu, S. Cai, C. Wang, J. Xie, and W. Hu, "Image recognition for garbage classification based on transfer learning and model fusion," *Math. Problems Eng.*, vol. 2022, pp. 1–12, Aug. 2022, doi: [10.1155/2022/4793555](https://doi.org/10.1155/2022/4793555).
- [74] F. Liu, H. Xu, M. Qi, D. Liu, J. Wang, and J. Kong, "Depth-wise separable convolution attention module for garbage image classification," *Sustainability*, vol. 14, no. 5, p. 3099, Mar. 2022, doi: [10.3390/su14053099](https://doi.org/10.3390/su14053099).
- [75] A. H. Vo, L. H. Son, M. T. Vo, and T. Le, "A novel framework for trash classification using deep transfer learning," *IEEE Access*, vol. 7, pp. 178631–178639, 2019, doi: [10.1109/ACCESS.2019.2959033](https://doi.org/10.1109/ACCESS.2019.2959033).
- [76] P. Tiyajamorn, P. Lorprasertkul, R. Assabumrungrat, W. Poomarin, and R. Chanchaoren, "Automatic trash classification using convolutional neural network machine learning," in *Proc. IEEE Int. Conf. Cybern. Intell. Syst. (CIS) IEEE Conf. Robot., Autom. Mechatronics (RAM)*, Nov. 2019, pp. 71–76, doi: [10.1109/CIS-RAM47153.2019.9095775](https://doi.org/10.1109/CIS-RAM47153.2019.9095775).
- [77] J. Hong, M. Fulton, and J. Sattar, "A generative approach towards improved robotic detection of marine litter," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 10525–10531, doi: [10.1109/ICRA40945.2020.9197575](https://doi.org/10.1109/ICRA40945.2020.9197575).
- [78] A. Sun and H. Xiao, "ThanosNet: A novel trash classification method using metadata," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2020, pp. 1394–1401, doi: [10.1109/BigData50022.2020.9378287](https://doi.org/10.1109/BigData50022.2020.9378287).
- [79] X. Dong, "Research and design of marine trash classification robot based on color recognition," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 514, no. 3, Jul. 2020, Art. no. 032043, doi: [10.1088/1755-1315/514/3/032043](https://doi.org/10.1088/1755-1315/514/3/032043).
- [80] H. Liu, Z. Guo, J. Bao, and L. Xie, "Research on trash classification based on artificial intelligence and sensor," in *Proc. 2nd Int. Conf. Intell. Comput. Hum.-Comput. Interact. (ICHCI)*, Nov. 2021, pp. 274–279, doi: [10.1109/ICHCI54629.2021.00062](https://doi.org/10.1109/ICHCI54629.2021.00062).
- [81] A. Patil, A. Tatke, N. Vachhani, M. Patil, and P. Gulhane, "Garbage classifying application using deep learning techniques," in *Proc. Int. Conf. Recent Trends Electron., Inf., Commun. Technol. (RTEICT)*, Aug. 2021, pp. 122–130, doi: [10.1109/RTEICT52294.2021.9573599](https://doi.org/10.1109/RTEICT52294.2021.9573599).
- [82] P. F. Proença and P. Simões, "TACO: Trash annotations in context for litter detection," 2020, *arXiv:2003.06975*.
- [83] M. Kraft, M. Piechocki, B. Ptak, and K. Walas, "Autonomous, onboard vision-based trash and litter detection in low altitude aerial images collected by an unmanned aerial vehicle," *Remote Sens.*, vol. 13, no. 5, pp. 1–17, Mar. 2021, doi: [10.3390/rs13050965](https://doi.org/10.3390/rs13050965).
- [84] J. Bobulski and J. Piatkowski, "PET waste classification method and plastic waste database—WaDaBa," in *Advances in Intelligent Systems and Computing*, vol. 681. Springer Verlag, 2018, pp. 57–64, doi: [10.1007/978-3-319-68720-9_8](https://doi.org/10.1007/978-3-319-68720-9_8).
- [85] *Glassense-Vision Dataset*. Accessed: Aug. 19, 2022. [Online]. Available: http://slipguru.unige.it/Data/glassense_vision/
- [86] T. Wang, Y. Cai, L. Liang, and D. Ye, "A multi-level approach to waste object segmentation," *Sensors*, vol. 20, no. 14, pp. 1–22, Jul. 2020, doi: [10.3390/s20143816](https://doi.org/10.3390/s20143816).
- [87] S. Lynch, "OpenLitterMap.com—Open data on plastic pollution with blockchain rewards (littercoin)," *Open Geospatial Data, Softw. Standards*, vol. 3, no. 1, 2018, doi: [10.1186/s40965-018-0050-y](https://doi.org/10.1186/s40965-018-0050-y).
- [88] *Waste_Pictures* | Kaggle. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/wangzhang/waste-pictures>
- [89] *Waste Classification Data* | Kaggle. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/techsash/waste-classification-data>
- [90] *Waste Images From Sushi Restaurant* | Kaggle. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/arthurcen/waste-images-from-sushi-restaurant>

- [91] *OpenLitterMap*. Accessed: Aug. 21, 2022. [Online]. Available: <https://openlittermap.com/>
- [92] *Litter Dataset*. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.imageannotation.ai/litter-dataset>
- [93] *Drinking Waste Classification | Kaggle*. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/arkadiyhacks/drinking-waste-classification>
- [94] *Garbage | Kaggle*. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/apremeyan/garbage>
- [95] *DeepSeaWaste | Kaggle*. Accessed: Aug. 21, 2022. [Online]. Available: <https://www.kaggle.com/datasets/henryhaefliger/deepseawaste>
- [96] *Datacluster-Labs/Domestic-Trash-Dataset*. Accessed: Aug. 21, 2022. [Online]. Available: <https://github.com/datacluster-labs/Domestic-Trash-Dataset>
- [97] *Cigarette Butt Dataset—Immersive Limit*. Accessed: Aug. 23, 2022. [Online]. Available: <https://www.immersivelimit.com/datasets/cigarette-butts>
- [98] N. V. Kumsetty, A. B. Nekkare, S. S. K. Kamath, and M. A. Kumar, “TrashBox: Trash detection and classification using quantum transfer learning,” in *Proc. 31st Conf. Open Innov. Assoc. (FRUCT)*, Apr. 2022, pp. 125–130, doi: [10.23919/FRUCT54823.2022.9770922](https://doi.org/10.23919/FRUCT54823.2022.9770922).
- [99] *Letsdoitworld/Wade-AI: An AI Algorithm for Detecting Trash in Geolocated Images*. Accessed: Aug. 21, 2022. [Online]. Available: <https://github.com/ltsdoitworld/wade-ai>
- [100] Q. Zhang, Q. Yang, X. Zhang, Q. Bao, J. Su, and X. Liu, “Waste image classification based on transfer learning and convolutional neural network,” *Waste Manag.*, vol. 135, pp. 150–157, Nov. 2021, doi: [10.1016/j.wasman.2021.08.038](https://doi.org/10.1016/j.wasman.2021.08.038).
- [101] S. Majchrowska, A. Mikołajczyk, M. Ferlin, Z. Klawikowska, M. A. Plantygow, A. Kwasigroch, and K. Majek, “Deep learning-based waste detection in natural and urban environments,” *Waste Manag.*, vol. 138, pp. 274–284, Feb. 2022, doi: [10.1016/j.wasman.2021.12.001](https://doi.org/10.1016/j.wasman.2021.12.001).
- [102] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” Jun. 2015, *arXiv:1506.01497*.



HARUNA ABDU was born in Mani, Katsina, Nigeria. He received the B.Sc. and M.Sc. degrees in computers science from Umaru Musa Yar’adua University, Katsina, Katsina State, Nigeria, in 2011 and 2016, respectively. He is currently pursuing the Ph.D. degree with Universiti Sains Malaysia. He is also working as a Lecturer II at Federal University Lokoja, Kogi, Nigeria. His research interests include signal processing, applications of machine learning and deep learning in solving local community problems, and quantum computing. He served as coach in some international programming contests. He was a recipient of Vice Chancellor Award of Academic Excellence during his undergraduate studies and Award of Merit from Nigeria Mathematical Olympiad Pre-PAMO/IMO Competition.



MOHD HALIM MOHD NOOR is currently an academician with the School of Computer Sciences, Universiti Sains Malaysia (USM). Prior to this, he worked at Universiti Teknologi MARA Pulau Pinang as a Senior Lecturer in computer engineering. His research interests include the fields of machine learning and deep learning for computer vision and pervasive computing.

...