Computer Vision for Solid Waste Sorting: A Critical Review of Academic Research

Weisheng Lu [a] and Junjie Chen [a, *]

[a] Department of Real Estate and Construction, The University of Hong Kong, Pokfulam Road, Hong Kong, China

[*] Correspondence information: Junjie Chen, Department of Real Estate and Construction, The University of Hong Kong, Pokfulam Road, Hong Kong, China, chenjj10@hku.hk

This is the peer-reviewed post-print version of the paper:

Lu, W., & Chen, J. (2022). Computer vision for solid waste sorting: A critical review of academic research. *Waste Management*, 142: 29-43. Doi: 10.1016/j.wasman.2022.02.009.

The final version of this paper is available at: https://doi.org/10.1016/j.wasman.2022.02.009.

The use of this file must follow the Creative Commons Attribution Non-Commercial No Derivatives License, as required by Elsevier's policy.

Abstract

Waste sorting is highly recommended for municipal solid waste (MSW) management. Increasingly, computer vision (CV), robotics, and other smart technologies are used for MSW sorting. Particularly, the field of CV-enabled waste sorting is experiencing an unprecedented explosion of academic research. However, little attention has been paid to understanding its evolvement path, status quo, and prospects and challenges ahead. To address the knowledge gap, this paper provides a critical review of academic research that focuses on CV-enabled MSW sorting. Prevalent CV algorithms, in particular their technical rationales and prediction performance, are introduced and compared. The distribution of academic research outputs is also examined from the aspects of waste sources, task objectives, application domains, and dataset accessibility. The review discovers a trend of shifting from traditional machine learning to deep learning algorithms. The robustness of CV for waste sorting is increasingly enhanced owing to the improved computation powers and algorithms. Academic studies were unevenly distributed in different sectors such as household, commerce and institution, and construction. Too often, researchers reported some preliminary studies using simplified environments and artificially collected data. Future research efforts are encouraged to consider the complexities of real-world scenarios and implement CV in industrial waste sorting practice. This paper also calls for open sharing of waste image datasets for interested researchers to train and share their CV algorithms.

Keywords: Municipal solid waste; Waste sorting; Computer vision; Image recognition; Machine learning; Deep learning.

Word count: 11,685 words (excluding references)

List of abbreviations:

AI	Artificial intelligence	ML	Machine learning
ANN	Artificial neural network	MLP	Multilayer perceptron

C&D	Construction and demolition	MSW	Municipal solid waste
CCD	Charge-coupled device	NN	Nearest neighbor
CMOS	Complementary metal oxide semiconductor	PCA	Principal Component Analysis
CNN	Convolutional neural network	R-CNN	Region-based convolutional neural network
CV	Computer vision	RM	Residential and municipal services
DCNN	Deep convolutional neural network	RP	Region proposals
DL	Deep learning	SIFT	Scale-invariant feature transform
DT	Decision tree	SSD	Single shot multibox detector
ELVs	End-of-life vehicles	SVM	Support vector machine
GPC	Gaussian process classification	TACO	Trash annotation in context
HOG	Histogram of oriented gradient	TL	Transfer learning
HSI	Hyperspectral imaging	WM	Waste management
ICI	Industrial, commercial and institutional	WoS	Web of Science
LDA	Linear discriminant analysis	YOLO	You only look once

1. Introduction

Rapid industrialization has caused the skyrocketing municipal solid waste (MSW) in the past few decades. In 2016, for example, 2.01 billion tonnes of MSW were generated globally, which was projected to surge to 2.59 billion tonnes by 2030 (Kaza et al., 2018). The definition and composition of MSW vary from case to case. This paper accepts the definition used by the 2012 World Bank report (Hoornweg and Bhada-Tata, 2012), in which MSW encompasses a wide range of waste generation sources, e.g., residential and municipal services (RM), industrial, commercial and institutional sources (ICI), and construction and demolition (C&D) activities. A typical MSW management cycle is comprised of waste generation, collection, treatment, and disposal as appropriate (Hoornweg and Bhada-Tata, 2012).

Waste sorting is a practice highly recommended in MSW management (Wang et al., 2020a; Xia et al., 2021). The term "waste sorting" usually appears along with two other closely related terminologies: "waste segregation" and "waste separation". While there is no universal, explicit definition on waste sorting, it is widely accepted in the field of waste management that: (a) waste segregation refers primarily to the grouping of waste materials into different categories when they are generated at source (e.g., households, workplace, or construction sites) or at the point of collection/dumping (Christensen and Matsufuji, 2011); (b) waste sorting, using interchangeably with waste separation, can either occur manually at source (Wang et al., 2020a), or be implemented in a relatively central place (Gundupalli et al., 2017b). In this research, we accept this conventional definition, and use the term "waste sorting" to encompass the waste separation behaviors occurring both at source and in central treatment facilities.

When the waste is generated at source, waste generators are encouraged to separate their waste according to its type (Christensen and Matsufuji, 2011; Gundupalli et al., 2017b), e.g., "wet" food waste and "dry" recyclables for household residue in Shanghai (Zuo and Yan, 2019), inert and non-inert materials for construction waste in Hong Kong (Lu et al., 2015; Lu and Yuan, 2021), or separate collection bins used for different waste types in public commercial areas (Keramitsoglou and Tsagarakis, 2018). However, with the increasing complication of waste taxonomies, it becomes more and more difficult for both citizens and regulators to distinguish among different MSW materials (Chen et al., 2021). The application of computer vision (CV) can potentially assist the sorting of MSW at source or when it is collected. With sufficient data, it is viable to train a CV model to identify various MSW materials. Applied to smart devices such as mobile phones, the CV model can then help users determine types of their generated waste materials for proper classification. The CV model can even be deployed to robotic platforms such as unmanned ground vehicles (UGV), enabling automatic waste collection in indoor built environments (Paulraj et al., 2016) or open construction sites (Wang et al., 2020b; Wang et al., 2019b) .

Waste sorting can be conducted in a relatively central place as well. There, the collected solid waste is transported by a conveyor belt through a series of sorting machinery and robots (Faibish et al., 1997; Gundupalli et al., 2017b; Huang et al., 2010; Mattone et al., 2000). During this process, two lines of approaches, i.e., direct and indirect sorting, are adopted (Gundupalli et al., 2017b; Huang et al., 2010). Direct sorting separates waste materials directly by applying forces such as gravity, magnetic force, or manual picking. In contrast, indirect sorting first uses sensors (e.g., optical sensor, spectroscope, inductive sensors, and thermal camera) to detect specific types of waste materials, and then sorts the detected materials with machinery or robots (Gundupalli et al., 2017b; Huang et al., 2010). The potential of using CV for indirect waste sorting has long been acknowledged. For example, Faibish et al. (1997) presented a robotic system with stereo vision to detect and separate paper objects for recycling. Mattone et al. (2000) formalized the problem of sorting items on a moving conveyor belt, and provided a solution based on optical devices. Compared with other sensing techniques such as hyperspectral imaging (HSI) and X-Ray, visual sensors, e.g., CCD (charge-coupled device) cameras, are cost-effective, easy for maintenance, and versatile for a wide range of waste (Rahman et al., 2014; Zulkifley et al., 2014). While traditional approaches rely more on huge investments on expensive sensing hardware (e.g., HSI cameras), sorting based on CV only requires simple installation of RGB cameras and harnesses the power of algorithms for waste material detection. In such waste sorting systems, CV serves as the "eye" and "brain", which is used to detect and identify waste materials on the conveyor belts, enabling robots to execute sorting operations autonomously.

Despite the promising prospects, the role of CV in MSW sorting had been limited and remained relatively marginal for a long time. The sluggish development is attributable to the tremendous manual efforts required for feature handcrafting and the relatively low robustness in the early stage. The situation has been improved with the development of deep learning (DL), signaled by AlexNet's great success in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012 (Krizhevsky et al., 2012). The DL techniques can address the limitations of traditional CV algorithms via two aspects. Firstly, with DL's power in feature extraction, the tiresome feature handcrafting can be avoided since visual features displayed by different waste types can be automatically learned from big data. Secondly, the robustness of waste classification algorithms can be enhanced by feeding DL model with a massive amount of visual data captured in different environments. The potential benefits of DL drove the surge of relevant research efforts. There have been reviews on the applications of artificial intelligence in waste management (WM) (Abdallah et al., 2020) or on the general topic of automated waste sorting (Gundupalli et al., 2017b). However, to the best of the authors' knowledge, no previous study has focused specifically on the applications of CV in waste sorting.

This research aims to conduct a comprehensive literature review of academic studies concerning the evolvement path, status quo, and prospects and challenges of applying CV for waste sorting. The primary focus is the processing of RGB images captured by ordinary digital cameras, as it is the area where most of the recent progress has been achieved, and represents a major trend of future development. The review sets out to address the following questions: Firstly, what are the main CV algorithms for waste sorting? How are the DL algorithms different from the traditional ones, and what makes them surpass the others to drive the recent publication boom? Secondly, what is the current status of academic research? How is previous research output distributed over different application domains and waste generation sectors? How was the performance evaluated, and how has it evolved over the years? Thirdly, by a comprehensive review, what lessons can we draw from previous research efforts? What challenges and issues shall we foresee in the future?

The rest of this paper is organized as follows. Subsequent to this introductory section is Section 2 to delineate the material search, selection, and some preliminary analyses. Section 3 is a detailed analysis of CV algorithms for waste sorting, and Section 4 is an analysis of research outputs from four aspects: waste sources, task objectives, application domains, and dataset accessibility. Section 5 is to articulate the prospects and challenges of CV for waste sorting, and conclusions are drawn in Section 6.

2. Bibliographic material collection

The materials included for this review were confined to academic literature in English and published between 1997 and 2021. We used two keyword combinations: "waste sorting AND computer vision" and "waste sorting AND image recognition", to search for relevant publications on the Web of Science (WoS) platform, as they directly reflect the two main themes of the review topic, i.e., “waste sorting” and “computer vision”. The searching results were manually screened to exclude irrelevant papers. After the suitability check, the remaining records were merged, which resulted in a total of 17 papers. Many other studies might have been left out as the initial search was only confined to the WoS database with two keyword combinations. To find the missing literature, references cited by the 17 papers were successively checked out to find other articles that have addressed the review topic. This “snowball” technique drastically expanded the literature collection to 86 publications, which consist of 51 journal articles, 32 conference papers, and 3 reports/preprints. The expanded literature collection includes not only publications from other notable databases such as Scopus, but also many conference papers that are highly regarded, particularly in computer science.

Based on the collected literature, the evolvement of research productivity over the past two decades was analyzed. As shown by Fig. 1 (a), the academic research outputs have experienced an explosion since 2016, when the historic triumph of AlphaGo against Lee Sedol further promoted the concept of DL to the general public. It is observed that the number of publications related to DL presents an upward trend while the numbers of those using other CV algorithms gradually shrink. Note that the number in 2021 only takes into account literature published before July of the year, when this review paper is drafted.

Fig. 1 (b) and (c) shows composition of the literature collection by publication names/types and countries. As shown in Fig. 1 (b), the most productive journals during the past 20 years are *Waste Management* and *Resources, Conservation and Recycling*, which have published 14 papers in total. As shown in Fig. 1 (c), the three leading countries in the number of relevant publications are, respectively, *Malaysia*, *China*, and *India*, which, interestingly, are all developing countries. This might be related to the rising concerns on the increasing waste generation brought by the rapid economic development in these countries.

3. Computer vision algorithms for waste sorting

Many CV algorithms have been used for waste sorting, which involves image preprocessing (e.g., denoising, thresholding, and segmentation), feature extraction such as scale-invariant feature transform (SIFT), histogram of oriented gradient (HOG) and principal component analysis (PCA), and machine learning (ML) algorithms for classification. Among all, ML algorithms are a critical component, for it directly affects the accuracy and efficiency of waste classification. This section summarizes the basic rationales and application scenarios of prevalent ML algorithms, and compares their performance in differentiating different waste materials. Note that although computation time is an important part for performance evaluation, it has not been reported by most previous studies; thus, accuracy is used as the primary evaluation metric here.

3.1. Traditional ML algorithms

Before the emergence of DL, research efforts sought to use traditional ML models to classify different waste materials from visual data. These models usually have simple structures, and lack the ability to automatically extract high-level representations from raw images; thus, they need to be fed with hand-engineered features, e.g., HOG and intensity histogram. Table 1 listed prevalent traditional ML algorithms used by previous studies, which include linear discriminant analysis (LDA), nearest neighbor (NN), decision tree (DT), Bayesian network, artificial neural network (ANN), support vector machine (SVM), and rule-based classifier.

3.1.1. Linear discriminant analysis

LDA is a frequently used linear classification algorithm. Given a bunch of data samples, the algorithm seeks to optimize a linear mapping from the original high dimensional feature space to a lower-dimensional subspace, based on which a classifier can then be designated to separate the samples into different classes (Leitner et al., 2003). Originated from Fisher's research on dichotomous discriminant analysis in 1936, LDA has been developed into many variants and extensions, e.g., Multiple Discriminant Analysis and Quadratic Discriminant Analysis. As indicated by Table 1, six articles have used LDA as classifiers in their studies for the recognition/detection of RM and ICI wastes. The prediction performance of the algorithm varies with the specific types of waste, task objectives, and used features, ranging from as low as 53% to over 98%.

3.1.2. Nearest neighbor

NN is a non-parametric ML algorithm that does not make strong assumptions about the distribution of the mapping between the input variables and the output class labels (Brownlee, 2016). The rationale of NN algorithm is straightforward: given a set $\boldsymbol{S}$ of sample points from different classes in the feature space and a query point $\boldsymbol{q}$ with unknown class, the algorithm iteratively calculates the distance (Euclidean or other distance metrics) between $\boldsymbol{q}$ and all the points in $\boldsymbol{S}$, and designates the class label of the closest point in $\boldsymbol{S}$ as the class of $\boldsymbol{q}$. A variant of NN algorithm is k-NN, which sets out to find the k closest points, and then decide the class label by majority vote. The number of studies based on NN surpasses all the other traditional ML algorithms. Despite the simplicity of the algorithm, it performed quite well in existing studies, most of which have attained an accuracy of over 85%.

3.1.3. Decision tree

DT algorithm transforms given data samples (i.e., the training set) into a representation of a tree structure consisting of nodes, branches, and leaves. A node represents a feature, and a leaf represents a class label; the branches stretching out from the nodes represent the specific values of corresponding features. DT is easy and straightforward to interpret and explain by human beings (Tachwali et al., 2007). Random forest is a variant algorithm of DT, which intends to address the preference of DT for overfitting to the training set. Although the number of studies that have applied DT is relatively small, its performance is significant, with an accuracy of over 90% in all the studies.

3.1.4. Bayesian network

A Bayesian network refers to a probabilistic graphical model that exploits the causal-effect relationship between variables. In the specific case of waste classification, a Bayesian network gives the conditional probability of a sample belonging to a type of waste with its given visual features, e.g., color (Gokyuu et al., 2011; Liu et al., 2010; Zulkifley et al., 2014), shape such as area ratio and aspect ratio (Gokyuu et al., 2011), and texture such as gray level co-occurrence matrix (GLCM) (Xiao et al., 2020) and Jet (Liu et al., 2010). Naïve Bayes classifier is a subset of Bayesian networks, which presumes the strong independence between features. The model has primarily been used for C&D and RM waste sorting, and the resulting accuracy is less than 80%.

3.1.5. Artificial neural network

An ANN learns the input-output mapping via a structure that mimics the biological neural networks (Abdallah et al., 2020; Guo et al., 2021). One of the earliest ANN models is a single-layer perceptron, which is only capable of learning linear separable patterns. A more competent and widely-used ANN structure is multilayer perceptron (MLP), which consists of an input layer, one or multiple hidden layers and an output layer. With such a multilayer structure and the use of non-linear activation functions (e.g., sigmoid and tanh), MLP can distinguish data that is not linearly separable. ANNs have been used for the classification of various waste materials, ranging from plastics to metal waste from the automobile industry. Many studies based on ANN achieved a classification accuracy of more than 95%.

3.1.6. Support vector machine

SVM can produce significant classification accuracy with a small amount of training data. The objective of SVM training is to find the optimal hyperplane that can best separate the data samples of different classes. The optimal hyperplane is the one that has the maximum distance from the nearest points of all classes (support vectors) (Rogers and Girolami, 2016). Despite its original focus on linear binary classification problems, SVM has been widely extended to address multi-classification problems with non-linear hyperplanes by the strategy of one-against-all and kernel operations. SVM is a prevalent technique in CV-based waste classification. The attained performance is remarkable: all but one of the studies have achieved an accuracy of over 90%.

3.1.7. Rule-based classifier

Rule-based classifiers are a type of classification models that decide the class of a given example by following a set of "if … then …" rules. While in some studies (Pothula et al., 2015; Rahman et al., 2009a; Zhu et al., 2018), the rules were specified as the relationship of a numerical feature regarding a given threshold (e.g., "*if* the aspect ratio of the regions of interest is over a threshold, *then* it is likely to be a slim bottle"), others (Mattone et al., 2000; 1998) set rules based on empirical observations by employing Fuzzy techniques to translate mathematical numbers to natural languages.

3.2. Deep learning algorithms

DL harnesses the power of big data via its deep structure to enable the so-called "end-to-end" training. In a typical DL architecture, original raw images, instead of the handcrafted features, are directly fed to the network consisting of multiple convolutional, pooling and fully-connected layers, through which hidden features from the images can be automatically learnt and extracted, and finally be used to predict the class label. The end-to-end learning mechanism avoids the tiresome process of feature handcrafting and thus greatly expands the applicability of CV-enabled waste sorting. In addition, with a dataset encompassing a wide range of waste samples, the resulting DL models tend to be more robust than those trained with traditional ML algorithms. A convolutional neural network (CNN) is a state-of-the-art DL algorithm that has been successfully applied in a broad spectrum of CV-related tasks. Fig. S1 in the Supplementary Material shows the structure of VGG-16, a well-known CNN architecture with 16 hidden layers proposed by the Visual Geometry Group (VGG). Table 2 summarizes previous researches that have applied DL in CV-enabled waste sorting according to their used CNN architectures.

3.2.1. Prevalent backbone networks

In deep learning, a backbone network refers to a CNN structure used to extract features from the input images. On top of the backbone networks, the extracted features can be leveraged by subsequent structures to accomplish various tasks. One of such tasks is waste recognition, which only aims to tell if given images belong to one of the predetermined categories. In such cases, the backbone networks are usually followed by a fully connected layer and a Softmax activation layer to output a vector indicating the probabilities of the input being certain waste materials, e.g., Mao et al. (2021), Yang and Thung (2016), and Meng and Chu (2020). By integrating patch-based classification or sliding window, the extracted features can be directly used to locate wastes on images as well (Anjum and Umar, 2018; Mittal et al., 2016).

Rows 1-X to 7-X in Table 2 list backbone networks that are frequently used in waste classification. The number of researches based on AlexNet and ResNet exceeds all the other backbone architectures. For AlexNet, the obtained accuracy varies from case to case, ranging from as low as 22% for the classification of six common MSWs (Yang and Thung, 2016) to the significant 96.41% for the binary classification of plastics/paper (Bobulski and Kubanek, 2019). As for ResNet, all but one studies achieve classification accuracy higher than 85%. Direct comparison among the reported accuracies is not recommended, as their performances might have been evaluated on different datasets and some datasets might be more challenging than the others. However, by comparing studies based on the same datasets and manually checking out the level of difficulties of relevant datasets, a general pattern can be identified: ResNet, Inception, DenseNet, and VGG tend to yield higher performance. In a general sense, MobileNet performed poorer than the others; however, the model requires less computation power and hence might be a preferable choice in actual industrial deployment.

The training of DL models relies on a massive amount of data, which is usually difficult to collect or have access to in the field of WM. A common practice to address the problem is the use of transfer learning (TL), a technique that exploits the model structure and parameters learned from the source domain, and adapts them to a new domain where only a few data are needed for parameter fine-tuning (Goodfellow et al., 2016). The technique has been used by most of previous studies.

3.2.2. The R-CNN series and other object detection networks

Based on features extracted by the backbones, more advanced tasks can be accomplished. Object detection is one such task, aiming not only to determine if given images contain objects of interest, but also locate the objects on images with bounding boxes (Zou et al., 2019). Some CV tasks such as semantic/instance segmentation go even further to extract pixel areas corresponding to the objects (Garcia-Garcia et al., 2017). In industrial practice, the waste streams are usually in a highly cluttered state where different materials are overlapped with each other. Therefore, compared with waste classification, waste object detection represents a more promising research direction by pinpointing the specific locations of wastes on images. Row 8-X to 12-X in Table 2 show 5 prevalent object detection networks in the field of waste sorting.

R-CNN (region-based convolutional neural network) (Girshick et al., 2014) uses an algorithm called "selective search" to find candidate region proposals (RP) from the original image. As the feature extraction is repeated many times for all the RP on the image, R-CNN does not perform well regarding efficiency (Ku et al., 2020). A series of algorithms have been developed based on R-CNN. Fast R-CNN improves the original R-CNN in a sense that, instead of extracting features separately for each individual RP, it produces a unified feature map from the input image and detect candidate regions from the feature map (Girshick, 2015). The improvement significantly reduces the required time to process each image. In Chen et al. (2017), Fast R-CNN was employed to detect and locate waste objects on conveyor belts, which demonstrated a false negative rate (FNR) of 3%, a false positive rate (FPR) of 9%, and an computation efficiency of 0.22 s/image. Ren et al. (2015) proposed Faster R-CNN, for which the computation time is reduced to sub-second level. Awe et al. (2017), Wang et al (2019b), and Nowakowski and Pamuła (2020) applied Faster R-CNN for the detection of RM waste, C&D waste, and electronic waste, respectively, which achieved desired performance. Another variant of R-CNN is Mask R-CNN, which, as compared to Faster R-CNN, has an additional branch to extract pixels corresponding to each individual instance of the detected objects (Proença and Simões, 2020; Wang et al., 2020b).

There is another stream of networks called single-stage detectors that treat objection detection as a single regression problem, e.g., YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and RetinaNet. Among them, RetinaNet surpasses the accuracy of many two-stage detectors such as the R-CNN series while still maintains its advantages on efficiency (Lin et al., 2017). Panwar et al. (2020) trained their waste detection model based on RetinaNet to identify contaminants in water body, which reached a mean average precision (mAP) of 0.814. Defined as the mean of average precision for each class, mAP is a popular metric for object detection evaluation. There are also studies that devised new CNN frameworks according to the domain-specific problem of waste sorting. Liang and Gu (2021) developed a CNN-based multi-task learning architecture in order to simultaneously classify and locate wastes in contexts. The architecture integrated components such as the attention mechanism, multi-level feature pyramids, and joint learning sub-networks, and reached a mAP of 0.815.

3.2.3. Integration with traditional ML algorithms

In the above studies, CNN models acted as both feature extractors and classifiers to recognize or detect wastes from images. However, there are studies where CNNs were only used for feature extraction, and classification was performed by traditional ML models such as SVM, MLP, and *k*-NN. Table 3 provides a summary of these studies. The integration of CNN with tradition ML models has the following benefits:

Take advantages of different machine learning models. CNNs are good at learning features via its deep networks while SVM, a traditional ML algorithm, is a powerful tool for classification. By combining the two, presumably higher performance can be obtained. Adedeji and Wang (2019) employed such a technique, which extracted features learnt by the ResNet-50 and performed waste classification with SVM. With the continuous improvement of CNN, the advantages of such "CNN+SVM" is diminishing. Thus, this is no longer the main consideration to decide whether the integration with tradition ML should be adopted.

Flexibility to fuse with different features. The combined use of CNN and traditional ML models allows easy integration of deep features extracted by different networks (Toğaçar et al., 2020) or the fusion with sensing data from other modalities (e.g., weight and magnetism) (Chu et al., 2018), leading to accuracy improvement. In order the automate the gauging of C&D waste composition, Chen et al. (2021) developed a hybrid model that integrated visual features extracted by a DenseNet-169 network and physical features such as weight and depth collected by other sensors. The hybrid features were input to a SVM for waste composition classification, which resulted in an accuracy improvement of over 20%.

Scalability to accommodate new input categories. In industrial practice, it is not uncommon that new waste categories required to be classified will dynamically increase. In such cases, the original DL model will have to be re-trained, which is a prolonged process. To address the problem, Yang et al. (2021) adopted a "ResNeXt+*k*-NN" structure. When new waste categories are added, retraining of the CNN is not required. Instead, *k*-NN can classify samples of the new categories according to the similarity of ResNeXt features.

4. Analysis of academic research outputs

The academic research outputs during 1997 and 2021 are analyzed from four aspects, i.e., waste sources, task objectives, application domains, and dataset accessibility. The pie charts in Fig. 2 show the distribution of publication numbers over these four aspects.

4.1. Waste sources

MSW can be generated from three sources, i.e., RM, ICI, and C&D. Fig. 2 (a) shows the distribution of academic output over the three sectors, among which the RM sector emerges as the most productive one with 61 papers (or reports) published, far outnumbering the ICI sector's 9 papers and the C&D sector' 11 papers.

4.1.1. Residential and municipal waste

The compelling imbalance of the number of publications implies the major focus on enabling the RM waste sorting with CV. To segregate inorganic wastes, Salmador et al. (2008) developed an intelligent garbage classifier system that comprises webcam vision, robotic arms, user interface, and conveyor belt. The CV module of the system first employed thresholding and watershed segmentation to extract waste items, and then performed classification with handcrafted features such as Fourier descriptors and moments. Rahman et al. (2011; 2010; 2009a, 2009b) conducted a line of work for the sorting of waste paper, where co-occurrence features (Rahman et al., 2009b) and window features (Rahman et al., 2010) were integrated with *k*-NN, template matching, and rule-based classifiers. As an important component in waste streams, facilitating the separation of plastic with CV has attracted the attention of many researchers (Rahman et al., 2009b; Ramli et al., 2008; Tachwali et al., 2007; Wang et al., 2019c; Zulkifley et al., 2014). Özkan et al. (2015) investigated five different feature extraction algorithms, and fed the extracted features to SVM for plastic classification. The study attained an average accuracy of 88% based on majority voting. Some researchers focused on using CV to distinguish between common RM waste, e.g., glass, paper, metal, plastic, and cardboard; examples of such studies include Bobulski and Kubanek (2019), Liu et al. (2019), and Toğaçar et al. (2020).

4.1.2. Industrial, commercial and institutional waste

Research attempts to apply CV for ICI waste sorting scatters among the different sectors of the national economy, e.g., the nuclear industry (Shaukat et al., 2016; Sun et al., 2019), the automobile industry (Koyanaka and Kobayashi, 2010, 2011; Wang et al., 2019a), and agriculture and food production (Guttormsen et al., 2016; Pothula et al., 2015; Verheyen et al., 2016; Zhu et al., 2018). Due to the hazardous nature of nuclear radiation, automated robots are in high demand to separate the nuclear waste materials after the decommissioning of nuclear facilities. Vision-based autonomous recognition is deemed as an indispensable module of such robotic systems. Shaukat et al. (2016) stressed the superiority of passive cameras with CMOS (complementary metal oxide semiconductor) sensors over other sensors (e.g., X-Ray, laser scanner, and HSI) regarding the cost-effectiveness and the reliability to be exposed to the extreme nuclear environment. Sun et al. (2019) integrated Gaussian process classification (GPC) and deep convolutional neural network (DCNN) to train a nuclear waste detection model with a limited amount of labeled RGB-D data.

In the automobile industry, the recycling of secondary raw materials from end-of-life vehicles (ELVs) is economically beneficial (Wang et al., 2019a). Koyanaka and Kobayashi (2010, 2011) used a 3D imaging camera system to measure shape parameters of metal pieces of ELVs, which were then fed to a discriminant analyzer and a neural network for the sorting of cast aluminum, wrought aluminum, and magnesium. Wang et al. (2019a) developed a computer vision-based system for the separation of non-ferrous metals from ELVs, where the performance of multiple ML algorithms, e.g., SVM, *k*-NN, and DT, were compared. As for the agriculture and food production industry, several works have focused on enabling effective CV-based sorting of the different components of crops and biomass for better usability (Pothula et al., 2015; Verheyen et al., 2016; Zhu et al., 2018).

4.1.3. Construction and demolition waste

C&D waste is an important component in MSW that accounts for up to 40% of the total waste stream in some cities (Hoornweg and Bhada-Tata, 2012). Traditional sorting methods based on wind selection, screening, and manual sorting are inefficient and tend to be inaccurate (Xiao et al., 2020); thus, researchers seek to incorporate CV into the sorting lines for better efficiency and accuracy. Lukka et al. (2014) and Kujala et al. (2015) presented a robotic system called ZenRobotics Recycler for the C&D waste sorting, of which CV is an important module to tackle the issue of material classification and object grasping. Xiao et al. (2020) developed an approach to classifying five typical categories of C&D waste, i.e., wood, brick, rubber, rock, and concrete. Chen et al. (2022) proposed a monocular vision approach to estimating composition of C&D wastes loaded by haul trucks. The classification of particles (Brisola et al., 2010) and aggregates also forms a challenging task in the recycling of C&D waste. Lau Hiu Hoong et al. (2020) employed CNN to develop a (near) real-time CV-based method to determine the composition of recycled aggregates. In Wang et al. (2020b; 2019b), a Faster R-CNN and a Mask R-CNN were respectively trained on a collection of images showing construction wastes on the ground. The trained models achieved high performance in detecting C&D wastes (e.g., nails, screws, and residual pipes and cables) scattered around the construction site, which were then used to enable a robot prototype to perform waste collection.

Despite the research attempts, the exploitation of CV to empower C&D waste sorting is still limited—only 11 publications over the span of more than two decades. The situation seems even worse and puzzling when considering the huge proportion (up to 40%) of C&D waste in the total waste streams and the high demand for new technologies for better recycling practices.

4.2. Task objectives

According to the tasks they aimed to accomplish, existing research can be divided into two categories, namely waste recognition and waste detection (Fig. S2 in the Supplementary Material). The former aims to classify images of waste materials into one of the predetermined categories, while the latter not only recognizes the waste categories but also locates them on the images with bounding boxes or pixelwise labels.

4.2.1. Waste recognition

As shown in Fig. 2 (b), much of existing research attention has been paid to the task of waste recognition. Before the prevalence of DL, early studies for waste recognition either simplified the problem to allow initial extraction of waste areas on images (Nawrocky et al., 2010; Ramli et al., 2008; Tachwali et al., 2007), or assumed waste materials on the conveyer belt can only appear in the camera field of view one by one (Sreelakshmi et al., 2019). Nawrocky et al. (2010) assumed images with bounding boxes around waste items were readily available, and applied SVM for waste classification. Based on the assumption that the waste items are not overlaid with each other, Tachwali et al. (2007) extracted bottles from conveyor belt images by using background subtraction techniques.

Recent development of DL makes it possible to directly process images without prior background subtraction. A major line of such works (Aral et al., 2018; Bircanoğlu et al., 2018; Rabano et al., 2018) originate from the TrashNet project by Yang and Thung (2016). In the project, the authors simplified waste sorting as a problem to classify a given single-object image into a waste type. The significant amount of research efforts has resulted in the high accuracy performance on public datasets such as TrashNet (Yang and Thung, 2016). Yang and Li (2020) developed a lightweight neural network for garbage classification called WasNet. The WasNet incorporates attention mechanism to force the network to pay more attention to sensitive area related with waste recognition. With data augmentation applied, WasNet realized a 96.10% classification accuracy on TrashNet. Zhang et al. (2021) proposed a residual network with a self-monitoring module for recyclable waste classification on the same dataset. By the application of a genetic algorithm for hyperparameter optimization, Mao et al. (2021) improved the classification accuracy on TrashNet to 99.60%.

Despite the remarkable progress, the nature of waste recognition has determined that it will have very limited application scenarios in the WM industry. Firstly, since waste recognition can only classify given images into one of the predefined categories, it is not suitable for automated waste sorting with robotics, which requires not only waste category information, but also position and geometry of the waste materials to guide the robot operations. Secondly, waste recognition tends to require individual waste items appearing against a relatively simple background, which is not the case in most real-life scenarios where the waste materials usually scatter or even overlap with each other in varying contexts. Based on the above analysis, such waste recognition techniques should be primarily considered for source separation at the waste collection stage, where smart phones can be used to assist citizens to distinguish different household wastes (Srinilta and Kanharattanachai, 2019; Yang et al., 2021) or be used by the authority to solicit waste collection information from the general public (Singh et al., 2017; Yang and Li, 2020).

4.2.2. Waste detection

Compared with its counterpart, the proportion of research focusing on waste detection merely exceeds a quarter. In practical engineering applications, it is quite common to have multiple waste items appear on the same image. Hence, sorting operations rely on not only the identified categories but also the exact location and geometry of the wastes (Awe et al., 2017; Ku et al., 2020; Lau Hiu Hoong et al., 2020). Waste detection not only recognizes the different waste types, but also identifies their positions and geometric boundaries on the images, providing critical information to infer their actual 3D positions for subsequent grasping or sorting with robot arms (Shaukat et al., 2016). Waste detection presents a promising research direction to address the need of waste sorting industry.

In the waste sorting process, the bulk of waste materials are randomly thrown onto the conveyor belt, inevitably leading to some waste items cluttered with each other. Overcoming the issue is a challenging prerequisite for the effective detection of individual waste items. Wang et al. (2019c) touched on the problem from the specific case of plastic bottle classification. The study integrated morphological operations, convex hull analysis, and concave points calculation to identify “adjacent” or “overlapping” waste items. The development of DL techniques makes it possible to directly train end-to-end models to identify wastes from complex clustered environments (Awe et al., 2017; Ku et al., 2020; Mittal et al., 2016; Nowakowski and Pamuła, 2020; Rad et al., 2017; Sun et al., 2019). Awe et al. (2017) pointed out that previous works based on TrashNet (Yang and Thung, 2016) can only recognize a single object image, and stressed the importance of waste detection from clusters of trash. With a fine-tuned Fast R-CNN, they achieved a mAP of 0.683 for the detection of typical MSW such as paper.

In recent years, more and more researchers have realized the limitations of waste recognition, and turned to the problem of waste detection (Anjum and Umar, 2018; Liang and Gu, 2021; Panwar et al., 2020; Proença and Simões, 2020; Wang et al., 2020b; Wang et al., 2019b). Proença and Simões (2020) realized the importance of image datasets with wastes in context. Therefore, they created the TACO (Trash Annotation in Context) dataset, and implemented Mask R-CNN on the dataset as a benchmark for waste detection. Similarly, Mask R-CNN was employed by Koskinopoulou et al. (2021) for waste detection in industrial scenarios. Panwar et al. (2020) applied RetinaNet to detect wastes in water body. Liang and Gu (2021) released a new dataset with bounding box annotations and multiple labels in each image, and developed a multi-task learning framework that can detect organic, recyclable, hazardous, and other wastes with high performance.

4.3. Application scenarios

Waste collection at source and waste sorting at disposal facility are two application scenarios where previous research attention has been paid to. Fig. 2 (c) shows academic publication distribution over the two scenarios.

4.3.1. Waste collection at source

Applications of CV in waste collection stage set out to recognize/detect waste objects from the source, which provides information for WM departments to make collection plans (Mittal et al., 2016; Nowakowski and Pamuła, 2020; Yang and Li, 2020), enable initial source separation with robotic systems (Paulraj et al., 2016; Rad et al., 2017), or assists citizens to classify their generated wastes in the household (Srinilta and Kanharattanachai, 2019; Yang et al., 2021). Nowakowski and Pamuła (2020) proposed a CNN classifier to identify e-waste (e.g., refrigerator, washing machine, and TV set) with smartphones, which can facilitate information exchange between waste generators and collection companies. Mittal et al. (2016) developed an Android app, SpotGarbage, to automatically detect and localize garbage in unconstrained real-world images. Singh et al. (2017) developed a system to solicit information on uncollected roadside garbage from the general public, in which CV was used to determine if the uploaded information is valid.

Another noticeable line of research is around the development of "smart bins". One prominent feature of such smart bins is that they use CV to detect the level of trash bins (e.g., empty, occupied, or full), which can then inform the municipal departments for waste collection (Abdallah et al., 2020; Aziz et al., 2018; Islam et al., 2014; Sarc et al., 2019). Hannan et al. (2012; 2016) and Aziz et al. (2015; 2018) are among the most productive researchers in the field, and have developed a series of ML waste level detection methods based on features such as gray level aura matrix, GLCM and Hough line detection. Some "smart bins" research goes further, and intends to automatically segregate trash in the bins by the integration of CV and robotics (Jacobsen et al., 2020). Most of existing researches on smart bins are based on traditional ML models; research on the applications of DL and its performance comparison with previous methods is not well established.

In recent years, increasing research attentions are paid to marine debris, a type of human-created solid waste that is discarded at sea or reach the sea through waterways or domestic and industrial outfalls (Ribic et al., 1992). The collection of marine debris requires information on its category and position, which can be obtained with CV technologies. Fulton et al. (2019) compared the performance of various DL algorithms in detecting trash in underwater environments, paving the way for automated waste collection with autonomous underwater vehicles. Hong et al. (2020) proposed a generative approach to augmenting underwater images for visual detection of marine debris. Aside from underwater trash, a stream of research aims to identify floating or near surface marine debris from remote sensing images, such as Mace (2012), Taddia et al. (2021), and Hu (2021).

4.3.2. Waste sorting at disposal facility

As shown by Fig. 2 (c), research output in automated waste sorting with robotics (Chu et al., 2018; Gundupalli et al., 2017a, 2018) is overwhelmingly higher than that in waste collection. In waste sorting facilities, a CV-enabled waste classification system consists of both hardware and software: the former is usually a low-cost camera, which acts as “eyes” of the system; the latter is essentially a bunch of computer algorithms that serve as the system’s “brain” to enable the identification of waste objects. In a review of automated sorting techniques, Gundupalli et al. (2017b) compared the performance of various sensing techniques, and concluded that optical-based sensors are applicable to a diverse range of materials and can attain acceptable accuracy with very little time consumption.

As early as the end of 20th century, researchers have already conceptualized such systems that use machine vision to guide robots for waste sorting (Faibish et al., 1997; Mattone et al., 2000; Mattone et al., 1998). The early-year CV algorithms usually require deliberate feature handcrafting for specific waste materials, which confined their applicability to a very limited range of wastes with simple features. In addition, the hand-engineered features are usually not sufficiently robust to adapt to the complexity of practical sorting tasks. The development of AI, especially DL techniques, significantly improves the robustness of CV algorithms, and expands its applicability to a wide range of waste materials including RM, ICI and C&D wastes. However, as mentioned in previous section, most existing studies focus on the problem of waste recognition, which deviates from the ultimate goal of automated waste sorting with robotics. The practical deployment of CV for automatic waste sorting requires significantly more research efforts to tackle the waste detection problems.

4.4. Dataset accessibility

According to their accessibility, existing waste datasets are either "public" ones that can be freely accessed by researchers or "private" ones that are only used by their owners. Figs. 2 (d) and 3 (a) show the distribution of previous publications by the accessibility of their datasets.

4.4.1. Private datasets

Most of previous studies evaluated their models on private datasets. Table 4 lists details of two notable private datasets. In the table, the “Task” column indicates whether the corresponding dataset is used for “waste recognition (R)” or “waste detection (D)”. The “Background” column reflects the complexity of background in the images: “Simple” means the background is simple and plain while “In context” means the wastes in the images were captured in real-life contexts that are random and complex. Fig. 3 (b) demonstrates the evolvement of waste classification accuracy on private datasets over the past two decades. The graph presents a fluctuation pattern, which contradicts a common perception that a line of research in the same area shall display a pattern of progressive improvement over time. A fundamental problem behind this is the use of private datasets makes it difficult to provide a unified standard and benchmark for meaningful performance comparison. For example, a collection of waste images acquired from an outdoor and cluttered environment is more challenging than one collected in a well-controlled lab environment; thus, directly comparing models trained from the two datasets is meaningless as performance resulting from the latter is presumably higher. The problem calls for the creation and sharing of public waste datasets.

4.4.2. Public datasets

Several waste datasets have been collected and publicized by researchers. On Kaggle, a well-known online platform for data science and machine learning, data scientists, ML engineers and industrial practitioners have also published some useful datasets. Table 5 lists information of representative public datasets.

The most widely used public dataset is TrashNet, collected and publicized by Yang and Thung (2016). The dataset includes 2,527 single-object photos of six waste types (i.e., paper, glass, plastic, metal, cardboard, and trash) with a white poster board as background. Fig. 3 (c) depicts the accuracy evolvement on the TrashNet. A clear trend of growth can be observed. This demonstrates that, with a unified publicly available dataset, research efforts can be concentrated to enable continuous performance improvement. However, despite the significant performance (99.60% accuracy by 2021), the TrashNet dataset seems rather too idealistic to allow for practical applications. As emphasized by (Meng and Chu, 2020), "it is unrealistic to get a picture of an object with a clean background" in reality. Similar concerns have been expressed by (Liang and Gu, 2021) and (Proença and Simões, 2020) on the oversimplicity of previous waste images. Therefore, datasets with waste images collected in real-life contexts are required. TACO is a notable initiative to move towards that direction. It includes 1,500 images of RM wastes in context, encompassing a wide range of waste items (28 categories and 60 subclasses) such as plastic bags, cigarette and bottles (Proença and Simões, 2020). The dataset has high-quality bounding box annotations and pixel-level waste labels, which can be used for object detection or even instance segmentation. A similar research thrust is WasteRL (Liang and Gu, 2021), which includes 57,000 images of four RM waste categories. The dataset is available by reasonable request to the corresponding author, and thus is not considered a "public" one as listed in Table 4.

A noticeable observation from Table 5 is that all but one public datasets are oriented to RM wastes. This is not surprising as a predominant portion of research efforts have been focused on RM waste sorting (Fig. 2 (a)). However, as the other two wastes types, especially C&D wastes, are also important sources of MSW, more attention should be paid to the development of public ICI or C&D waste datasets to facilitate relevant research.

5. Challenges and prospects

5.1. Challenges ahead

Driven by prevailing DL, the application of CV for waste sorting is gaining increasing attention. However, our review identifies multiple lessons from existing research, which might pose challenges to transferring current academic efforts into practical applications.

First, current studies lack publicly available real-world datasets that are oriented to the industrial practices of waste sorting. This has led to the difficulties in rigorous comparison among different studies. Most existing datasets were privately owned by respective research teams. There are open-access datasets, but most of them tend to oversimplify the industrial need as a waste recognition problem, with waste objects appearing on plain and well-control background (Mohamed, 2021; Sekar, 2019; x670783915, 2019; Yang and Thung, 2016). Some studies have embarked on collecting waste images in contexts. However, they are either limited in the number (Panwar et al., 2020; Proença and Simões, 2020) or lack high-quality annotations (DataCluster Labs, 2021). In addition, a predominant proportion of public datasets are for the sorting of RM wastes. The absence of high-quality public datasets oriented to the diverse MSW types makes it difficult to allow meaningful performance evaluation based on a unified standard. This would distract research efforts from the core line of the continuous improvement of waste sorting.

Second, existing studies simplified, more or less, the objectives or working conditions of their proposed algorithms, which might pose a challenge to their future deployment to industrial applications. Before the popularization of DL, it was a common practice to simplify the problem of waste recognition with certain boundary conditions, e.g., assuming waste materials can only be thrown onto the conveyor belt one by one (Sreelakshmi et al., 2019), or cannot overlap with others (Tachwali et al., 2007). Most recent DL-related studies confine their scope in recognizing single objects from images (Aral et al., 2018; Yang and Thung, 2016). Such simplification with tight constraints can make the developed approaches incompatible with the unstructured environments in actual sorting facilities, where the waste is usually randomly distributed in a cluttered background (Chen et al., 2021; Lu et al., 2022).

Third, while C&D waste accounts for a large proportion in MSW steams, only a limited amount of research has been dedicated to applying CV for its sorting. C&D waste can represent as much as 40% of the total waste streams generated by some cities (Hoornweg and Bhada-Tata, 2012); In Hong Kong, C&D waste takes up at least one quarter of the materials that end up being landfilled (HKEPD, 2020). The disproportionately limited number of studies, as compared to the large amount of C&D waste, poses a great challenge ahead, which signifies an urgent need of more research attentions. A failure in addressing the need will not only continue to allow the recyclable construction materials not being properly reused, but also cause a series of environmental problems for producing new materials that could have been recycled and taking up valuable land resources for landfills.

Fourth, despite the versatility of CV in distinguishing a wide range of materials, vision-based methods are inherently incapable of characterizing physiochemical properties of waste objects. Materials with different physiochemical properties can present similar visual features. For example, glass might look similar to a piece of transparent plastic sheet, but clearly, they are two different types of materials. In such cases, it would be challenging to rely on CV to implement effective sorting without other useful information provided.

5.2. Prospects and future directions

In view of the above challenges, several directions are suggested as follows for the reference of future CV-based waste sorting research.

Industry-oriented public datasets for various waste sources. Publicly available datasets with high-quality annotations are required to provide benchmarks for rigorous performance evaluation. The datasets should, on the one hand, orient to the practical needs of industry by providing real-life photos in context, and on the other hand, ensure free and open access to the research community. Some latest studies have realized the limitations of previous oversimplified datasets, and strived to create datasets that captures the variations of reality, e.g., TACO (Proença and Simões, 2020) and WasteRL (Liang and Gu, 2021). However, as these studies were only published recently, their influences remain relatively low. More research efforts should be made to create and share datasets for ICI and C&D wastes. The access to such waste photos is usually monopolized by stakeholders of the respective sectors, and thus the engagement of industry practitioners is essential. While an ideal and more sustainable way is to construct a centralized database where images of various waste types and their annotations can be solicited by crowdsourcing (PEER, 2018), it is not a realistic goal in the short term due to a series of legal and managerial issues such as copy right and confidentiality. In the foreseeable future, waste datasets will still be collected, managed and publicized by separate parties, i.e., via a decentralized manner. Thus, it is important to formulate a common protocol to guide how the datasets should be constructed. For examples, since different countries/regions and sectors generate different types of waste, it is suggested to include the geolocation (e.g., countries or geographical coordinates) and sources (e.g., RM, ICI, or C&D) as an annotation data field or metadata when constructing relevant datasets.

Address the need of engineering practices. Future researches are expected to focus more on the practical problems encountered by the WM industry. For example, sorting waste from conveyor belts requires more than only recognizing the waste type from a single-object image. Hence, waste detection, which aims to detect and classify multiple waste items, should attract more research attention in the future. The efficiency of CV algorithms is also an indispensable factor to consider in practical applications. While efficiency is comparatively less important in certain scenarios such as smartphone-assisted household waste classification, demanding time performance is required in other scenarios like automated waste sorting with robotics. This is especially true when considering that a waste recovery facility can receive thousands of tonnes of MSW per day, and the sorting speed directly affects the overall throughput. However, the great majority of existing studies failed to report the time performance of their algorithms. Future researches are suggested to share the inference time as well as accuracy of their algorithms. Note that hardware with different computing power can result in different time performance of the same algorithm. Hence, the specific hardware configuration with which the respective algorithm is implemented should also be mentioned for readers' reference.

Computer vision for C&D waste sorting. As a major component of MSW, efficiency improvement of C&D waste sorting is beneficial from both economic and environmental perspectives. More research efforts are needed to investigate how CV and its relevant algorithms can be integrated into the sorting process given the bulky and heterogeneous characteristics of C&D wastes. In fact, stimulated by the economic benefits and recent technological development, notable change is happening. By 2018, only five out of the reviewed studies are relevant to C&D wastes, but since 2019, up to six studies have been published on the topic in less than three years. The level of increase reached 120%. The publication boom reflects a trend that more and more research is directed to the use of CV in C&D waste sorting. In the near future, the trend is expected to continue.

Multi-modal sensing data fusion. Although it is expected the performance of CV in waste sorting will continue to improve in the near future, visual sensors have inherent weakness in sensing materials' physiochemical properties. Fusion of sensing data from different modalities (Chu et al., 2018; Kuritcyn et al., 2015; Tachwali et al., 2007) can take advantages of both the cost-effectiveness of visual sensors and the capability of other sensors, e.g., weight meters, near-infrared spectroscopy (NIR) and inductive sensor, to detect the physiochemical properties of waste items. It can increase the classification accuracy and improve the system robustness. However, the emerging DL poses new challenges, as the features extracted by CNN can be hundreds of thousands compared with the sing-digit number of physical features. Hence, future research is suggested to further investigate how visual features extracted by CV can be effectively fused with other features from different domains.

6. Conclusions

Waste sorting is a critical step towards efficient MSW management. The application of CV for waste sorting has been conceptualized and under investigation for over two decades. Driven by the emerging DL techniques, the field is currently experiencing an unprecedented development. Against this background, this paper provides a critical review of academic research to understand the past, present, and future of the field of CV-based waste sorting. Prevalent CV algorithms for waste sorting can be categorized into two types: traditional ML and DL algorithms. Traditional ML algorithms require handcrafted visual features as input, while DL algorithms can automatically extract hidden features from raw images. For its advantages on robustness and automated end-to-end training, DL has become the predominant CV algorithms to enable solid waste sorting. It is found that the academic studies distributed disproportionately among the RM, ICI, and C&D sectors. While existing studies primarily focused on employing CV for waste sorting at central disposal facilities, there were also studies aiming to facilitate waste collection by machine vision. Previous studies tend to confine their research scope to the task of waste recognition with simplified working conditions (e.g., using a small set of artificially collected images in lab environments). Only a few studies used publicly available, secondary datasets for model training and evaluation.

The critical review identified several challenges confronting the further promotion of CV for waste sorting. These include (a) the lack of comprehensive, sharable datasets that can be used by interested researchers to train their models, (b) oversimplified working conditions leading to detached links between research and real-life practice, (c) less attention to C&D waste sorting, and (d) limitation of CV in distinguishing materials with similar appearance. Future research efforts should strive to create and share industry-level datasets oriented to the diversity of waste sources. Researchers are also suggested to accommodate the actual industrial needs by focusing on the waste detection problems while attaching equivalent importance to efficiency of their algorithms. C&D waste sorting and multimodal feature fusion are two other promising directions that should gain more attention.

Acknowledgment

This research is jointly supported by the Strategic Public Policy Research (SPPR) Funding Scheme (Project No.: S2018.A8.010.18S) and the Environment and Conservation Fund (ECF) (Project No.: ECF 2019-111) of the Government of the Hong Kong Special Administrative Region.

References

Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., Mahfood, B., 2020. Artificial intelligence applications in solid waste management: A systematic research review. Waste Management 109, 231-246.

Adedeji, O., Wang, Z., 2019. Intelligent waste classification system using deep learning convolutional neural network, 2nd International Conference on Sustainable Materials Processing and Manufacturing, pp. 607-612.

Anjum, M., Umar, M.S., 2018. Garbage localization based on weakly supervised learning in deep convolutional neural network, Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018, pp. 1108-1113.

Aral, R.A., Keskin, Ş.R., Kaya, M., Hacıömeroğlu, M., 2018. Classification of TrashNet dataset based on deep learning models, 2018 IEEE International Conference on Big Data (Big Data), pp. 2058-2062.

Arebey, M., Hannan, M.A., Begum, R.A., Basri, H., 2012. Solid waste bin level detection using gray level co-occurrence matrix feature extraction approach. Journal of Environmental Management 104, 9-18.

Awe, O., Mengistu, R., Sreedhar, V., 2017. Smart trash net: Waste localization and classification, arXiv preprint.

Aziz, F., Arof, H., Mokhtar, N., Mubin, M., Abu Talip, M.S., 2015. Rotation invariant bin detection and solid waste level classification. Measurement: Journal of the International Measurement Confederation 65, 19-28.

Aziz, F., Arof, H., Mokhtar, N., Shah, N.M., Khairuddin, A.S.M., Hanafi, E., Talip, M.S.A., 2018. Waste level detection and HMM based collection scheduling of multiple bins. PLoS ONE 13.

Bircanoğlu, C., Atay, M., Beşer, F., Genç, Ö., Kızrak, M.A., 2018. RecycleNet: Intelligent waste sorting using deep neural networks, 2018 Innovations in Intelligent Systems and Applications (INISTA), pp. 1-7.

Bobulski, J., Kubanek, M., 2019. Waste classification system using image processing and convolutional neural networks, International Work-Conference on Artificial Neural Networks. Springer, pp. 350-361.

Brisola, D.F., Cunha, B.M., Gomes, O., Lima, P., Paciornik, S., 2010. Automatic classification of particles from construction and demolition waste through digital image analysis, 65th ABM International Congress, 18th IFHTSE Congress and 1st TMS/ABM International Materials Congress 2010, pp. 3046-3052.

Brownlee, J., 2016, Parametric and nonparametric machine learning algorithms. https://machinelearningmastery.com/parametric-and-nonparametric-machine-learning-algorithms/ (Accessed Dec. 6 2021).

Chen, J., Lu, W., Xue, F., 2021. "Looking beneath the surface": A visual-physical feature hybrid approach for unattended gauging of construction waste composition. Journal of Environmental Management 286, 112233.

Chen, J., Lu, W., Yuan, L., Wu, Y., Xue, F., 2022. Estimating construction waste truck payload volume using monocular vision. Resources, Conservation and Recycling 177, 106013.

Chen, Z., Hebin, Z., Yanbo, W., Binyan, L., Yu, L., 2017. A vision-based robotic grasping system using deep learning for garbage sorting, 2017 36th Chinese Control Conference (CCC), pp. 11223-11226.

Christensen, T.H., Matsufuji, Y., 2011. Source segregation and collection of source-segregated waste, Solid waste technology and management. Wiley, pp. 296-310.

Chu, Y., Huang, C., Xie, X., Tan, B., Kamal, S., Xiong, X., 2018. Multilayer hybrid deep-learning method for waste classification and recycling. Computational Intelligence and Neuroscience 2018, 5060857.

DataCluster Labs, 2021, Domestic trash/garbage dataset. https://www.kaggle.com/dataclusterlabs/domestic-trash-garbage-dataset (Accessed Aug. 2 2021).

Faibish, S., Bacakoglu, H., Goldenberg, A.A., 1997. An eye-hand system for automated paper recycling, Proceedings of International Conference on Robotics and Automation, pp. 9-14 vol.11.

Fulton, M., Hong, J., Islam, M.J., Sattar, J., 2019. Robotic detection of marine litter using deep visual detection models, 2019 International Conference on Robotics and Automation (ICRA), pp. 5752-5758.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857.

Girshick, R., 2015. Fast R-CNN, Proceedings of the IEEE international conference on computer vision, pp. 1440-1448.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587.

Gokyuu, T., Nakamura, S., Ueno, T., Nakamura, M., Inoue, D., Yanagihara, Y., 2011. Sorting system for recycling of construction byproducts with Bayes' theorem-based robot vision, Journal of Robotics and Mechatronics, pp. 1066-1072.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep learning. MIT press Cambridge.

Gundupalli, S.P., Hait, S., Thakur, A., 2017a. Multi-material classification of dry recyclables from municipal solid waste based on thermal imaging, Waste Management, pp. 13-21.

Gundupalli, S.P., Hait, S., Thakur, A., 2017b. A review on automated sorting of source-separated municipal solid waste for recycling. Waste Management 60, 56-74.

Gundupalli, S.P., Hait, S., Thakur, A., 2018. Classification of metallic and non-metallic fractions of e-waste using thermal imaging-based technique. Process Safety and Environmental Protection 118, 32-39.

Guo, H., Wu, S., Tian, Y., Zhang, J., Liu, H., 2021. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review. Bioresource Technology 319, 124114.

Guttormsen, E., Toldnes, B., Bondø, M., Eilertsen, A., Gravdahl, J.T., Mathiassen, J.R., 2016. A machine vision system for robust sorting of herring fractions, Food and Bioprocess Technology, pp. 1893-1900.

Hannan, M.A., Arebey, M., Begum, R.A., Basri, H., 2012. An automated solid waste bin level detection system using a gray level AURA matrix. Waste Management 32, 2229-2238.

Hannan, M.A., Arebey, M., Begum, R.A., Basri, H., Al Mamun, M.A., 2016. Content-based image retrieval system for solid waste bin level detection and performance evaluation, Waste Management, pp. 10-19.

HKEPD, 2020, Hong Kong waste treatment and disposal statistics. https://www.epd.gov.hk/epd/english/environmentinhk/waste/data/stat_treat.html (Accessed April 7 2021).

Hong, J., Fulton, M., Sattar, J., 2020. A generative approach towards improved robotic detection of marine litter, 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 10525-10531.

Hoornweg, D., Bhada-Tata, P., 2012. What a waste: A global review of solid waste management. World Bank, Washington, DC.

Hu, C., 2021. Remote detection of marine debris using satellite observations in the visible and near infrared spectral range: Challenges and potentials. Remote Sensing of Environment 259, 112414.

Huang, G.-L., He, J., Xu, Z., Huang, G., 2020. A combination model based on transfer learning for waste classification. Concurrency and Computation: Practice and Experience 32, e5751.

Huang, J., Pretz, T., Bian, Z., 2010. Intelligent solid waste processing using optical sensor based sorting technology, 2010 3rd International Congress on Image and Signal Processing, pp. 1657-1661.

Islam, M.S., Hannan, M.A., Basri, H., Hussain, A., Arebey, M., 2014. Solid waste bin detection and classification using dynamic time warping and mlp classifier. Waste Management 34, 281-290.

Jacobsen, R.M., Johansen, P.S., Bysted, L.B.L., Skov, M.B., 2020. Waste wizard: Exploring waste sorting using ai in public spaces, ACM International Conference Proceeding Series.

Kaza, S., Yao, L., Bhada-Tata, P., Van Woerden, F., 2018. What a waste 2.0: A global snapshot of solid waste management to 2050. World Bank Publications.

Keramitsoglou, K.M., Tsagarakis, K.P., 2018. Public participation in designing the recycling bins to encourage recycling. Sustainability 10, 1240.

Kim, J., Nocentini, O., Scafuro, M., Limosani, R., Manzi, A., Dario, P., Cavallo, F., 2019. An innovative automated robotic system based on deep learning approach for recycling objects, ICINCO (2), pp. 613-622.

Koskinopoulou, M., Raptopoulos, F., Papadopoulos, G., Mavrakis, N., Maniadakis, M., 2021. Robotic waste sorting technology: Toward a vision-based categorization system for the industrial robotic separation of recyclable waste. IEEE Robotics & Automation Magazine 28, 50-60.

Koyanaka, S., Kobayashi, K., 2010. Automatic sorting of lightweight metal scrap by sensing apparent density and three-dimensional shape, Resources, Conservation and Recycling, pp. 571-578.

Koyanaka, S., Kobayashi, K., 2011. Incorporation of neural network analysis into a technique for automatically sorting lightweight metal scrap generated by elv shredder facilities. Resources, Conservation and Recycling 55, 515-523.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, pp. 1097-1105.

Ku, Y., Yang, J., Fang, H., Xiao, W., Zhuang, J., 2020. Deep learning of grasping detection for a robot used in sorting construction and demolition waste. Journal of Material Cycles and Waste Management.

Kujala, J.V., Lukka, T.J., Holopainen, H., 2015. Picking a conveyor clean by an autonomously learning robot, arXiv preprint arXiv:1511.07608.

Kuritcyn, P., Anding, K., Linß, E., Latyev, S.M., 2015. Increasing the safety in recycling of construction and demolition waste by using supervised machine learning. Journal of Physics: Conference Series 588, 012035.

Lai, K., Bo, L., Ren, X., Fox, D., 2011. A large-scale hierarchical multi-view RGB-D object dataset,

2011 IEEE International Conference on Robotics and Automation, pp. 1817-1824.
Lau Hiu Hoong, J.D., Lux, J., Mahieux, P.-Y., Turcry, P., Aït-Mokhtar, A., 2020. Determination of the composition of recycled aggregates using a deep learning-based image analysis. Automation in Construction 116, 103204.
Leitner, R., Mairer, H., Kercek, A., 2003. Real-time classification of polymers with nir spectral imaging and blob analysis, Real-Time Imaging, pp. 245-251.
Liang, S., Gu, Y., 2021. A deep convolutional neural network to simultaneously localize and recognize waste types in images. Waste Management 126, 247-257.
Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection, Proceedings of the IEEE international conference on computer vision, pp. 2980-2988.
Liu, C., Sharan, L., Adelson, E.H., Rosenholtz, R., 2010. Exploring features in a Bayesian framework for material recognition, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 239-246.
Liu, H., Owolabi, G.O., Kim, S.H., 2019. Automatic classifications and recognition for recycled garbage by utilizing deep learning technology, Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City, pp. 1-4.
Lu, W., Chen, J., Xue, F., 2022. Using computer vision to recognize composition of construction waste mixtures: A semantic segmentation approach. Resources, Conservation and Recycling 178, 106022.
Lu, W., Peng, Y., Webster, C., Zuo, J., 2015. Stakeholders' willingness to pay for enhanced construction waste management: A Hong Kong study, Renewable and Sustainable Energy Reviews, pp. 233-240.
Lu, W., Yuan, L., 2021. Investigating the bulk density of construction waste: A big data-driven approach. Resources, Conservation & Recycling 169, 105480.
Lukka, T.J., Tossavainen, T., Kujala, J.V., Raiko, T., 2014. Zenrobotics recycler–robotic sorting using machine learning, Proceedings of the International Conference on Sensor-Based Sorting (SBS), pp. 1-8.
Mace, T.H., 2012. At-sea detection of marine debris: Overview of technologies, processes, issues, and options. Marine Pollution Bulletin 65, 23-27.
Mao, W., Chen, W., Wang, C., Lin, Y., 2021. Recycling waste classification using optimized convolutional neural network. Resources, Conservation and Recycling 164, 105132.
Mattone, R., Campagiorni, G., Galati, F., 2000. Sorting of items on a moving conveyor belt. Part 1: A technique for detecting and classifying objects, Robotics and Computer-Integrated Manufacturing, pp. 73-80.
Mattone, R., Campagiorni, G., Wolf, A., 1998. Fuzzy-based processing of 3d information for items localization in the automated sorting of recyclable packaging, 1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World Congress on Computational

Intelligence (Cat. No.98CH36228), pp. 1613-1618 vol.1612.

Meng, S., Chu, W., 2020. A study of garbage classification with convolutional neural networks, 2020 Indo – Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), pp. 152-157.

Mittal, G., Yagnik, K.B., Garg, M., Krishnan, N.C., 2016. Spotgarbage: Smartphone app to detect garbage using deep learning, Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 940-945.

Mohamed, M., 2021, Garbage classification (12 classes). https://www.kaggle.com/mostafaabla/garbage-classification (Accessed Aug. 2 2021).

Nawrocky, M., Schuurman, D.C., Fortuna, J., 2010. Visual sorting of recyclable goods using a support vector machine, CCECE 2010, pp. 1-4.

Nowakowski, P., Pamuła, T., 2020. Application of deep learning object classifier to improve e-waste collection planning. Waste Management 109, 1-9.

Özkan, K., Ergin, S., Işık, Ş., Işıklı, İ., 2015. A new classification scheme of plastic wastes based upon recycling labels, Waste Management, pp. 29-35.

Panwar, H., Gupta, P.K., Siddiqui, M.K., Morales-Menendez, R., Bhardwaj, P., Sharma, S., Sarker, I.H., 2020. Aquavision: Automating the detection of waste in water bodies using deep transfer learning. Case Studies in Chemical and Environmental Engineering 2, 100026.

Paulraj, S.G., Hait, S., Thakur, A., 2016. Automated municipal solid waste sorting for recycling using a mobile manipulator, Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2016.

PEER, 2018, Call for uploading images for phi (peer hub imagenet) challenge. https://apps.peer.berkeley.edu/spo/ (Accessed Aug. 8 2021).

Pothula, A.K., Igathinathane, C., Kronberg, S., 2015. Profile based image analysis for identification of chopped biomass stem nodes and internodes, Industrial Crops and Products, pp. 374-382.

Proença, P.F., Simões, P., 2020. Taco: Trash annotations in context for litter detection. arXiv preprint arXiv:2003.06975.

Rabano, S.L., Cabatuan, M.K., Sybingco, E., Dadios, E.P., Calilung, E.J., 2018. Common garbage classification using MobileNet, 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, HNICEM 2018. Institute of Electrical and Electronics Engineers Inc.

Rad, M.S., von Kaenel, A., Droux, A., Tieche, F., Ouerhani, N., Ekenel, H.K., Thiran, J.P., 2017. A computer vision system to localize and classify wastes on the streets, International Conference on computer vision systems. Springer, pp. 195-204.

Rahman, M.O., Hussain, A., Basri, H., 2014. A critical review on waste paper sorting techniques, International Journal of Environmental Science and Technology, pp. 551-564.

Rahman, M.O., Hussain, A., Scavino, E., Basri, H., Hannan, M.A., 2011. Intelligent computer vision system for segregating recyclable waste papers, Expert Systems with Applications, pp. 10398-10407.

Rahman, M.O., Hussain, A., Scavino, E., Ezlin, N., Basri, H., Hannan, M.A., 2010. Waste paper grade identification system using window features, Journal of Computational Information Systems, pp. 2077-2091.

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2009a. Recyclable waste paper sorting using template matching, Recyclable Waste Paper Sorting Using Template Matching, pp. 467-478.

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2009b. Segregating recyclable waste papers using co-occurrence features, Proceedings of the 9th WSEAS International Conference on Applied Computer Science, ACS '09.

Ramli, S., Mustafa, M., Hussain, A., Abd Wahab, D., 2008. Histogram of intensity feature extraction for automatic plastic bottle recycling system using machine vision, American Journal of Environmental Sciences.

Ramli, S., Mustafa, M.M., Wahab, D.A., Hussain, A., 2010. Plastic bottle shape classification using partial erosion-based approach, 2010 6th International Colloquium on Signal Processing & its Applications, pp. 1-4.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in neural information processing systems 28, 91-99.

Ribic, C.A., Dixon, T.R., Vining, I., 1992. Marine debris survey manual.

Rogers, S., Girolami, M., 2016. A first course in machine learning. CRC Press.

Sakr, G.E., Mokbel, M., Darwich, A., Khneisser, M.N., Hadi, A., 2016. Comparing deep learning and support vector machines for autonomous waste sorting, 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET), pp. 207-212.

Salmador, A., Pérez Cid, J., Rodríguez Novelle, I., 2008. Intelligent garbage classifier, International Journal of Interactive Multimedia and Artificial Intelligence. International Journal of Interactive Multimedia and Artificial Intelligence ….

Sarc, R., Curtis, A., Kandlbauer, L., Khodier, K., Lorber, K.E., Pomberger, R., 2019. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management - a review. Waste Management 95, 476-492.

Scavino, E., Wahab, D.A., Hussain, A., Basri, H., Mustafa, M.M., 2009. Application of automated image analysis to the identification and extraction of recyclable plastic bottles, Journal of Zhejiang University-SCIENCE A, pp. 794-799.

Sekar, S., 2019, Waste classification data| kaggle. https://www.kaggle.com/techsash/waste-classification-data (Accessed Aug. 2 2021).

Sharan, L., Rosenholtz, R., Adelson, E., 2009. Material perception: What can you see in a brief glance? Journal of Vision 9, 784-784.

Shaukat, A., Gao, Y., Kuo, J.A., Bowen, B.A., Mort, P.E., 2016. Visual classification of waste material for nuclear decommissioning, Robotics and Autonomous Systems, pp. 365-378.

Singh, S., Mehta, K.S., Bhattacharya, N., Prasad, J., Lakshmi, S.K., Subramaniam, K.V., Sitaram, D., 2017. Identifying uncollected garbage in urban areas using crowdsourcing and machine learning, 2017 IEEE Region 10 Symposium (TENSYMP), pp. 1-5.

Sreelakshmi, K., Akarsh, S., Vinayakumar, R., Soman, K.P., Ieee, 2019. Capsule neural networks and visualization for segregation of plastic and non-plastic wastes, 2019 5th International Conference on Advanced Computing & Communication Systems, pp. 631-636.

Srinilta, C., Kanharattanachai, S., 2019. Municipal solid waste segregation with CNN, 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST), pp. 1-4.

Sudha, S., Vidhyalakshmi, M., Pavithra, K., Sangeetha, K., Swaathi, V., 2016. An automatic classification method for environment: Friendly waste segregation using deep learning, 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), pp. 65-70.

Sun, L., Zhao, C., Yan, Z., Liu, P., Duckett, T., Stolkin, R., 2019. A novel weakly-supervised approach for rgb-d-based nuclear waste object detection, Ieee Sensors Journal, pp. 3487-3500.

Tachwali, Y., Al-Assaf, Y., Al-Ali, A.R., 2007. Automatic multistage classification system for plastic bottles recycling, Resources, Conservation and Recycling, pp. 266-285.

Taddia, Y., Corbau, C., Buoninsegni, J., Simeoni, U., Pellegrinelli, A., 2021. UAV approach for detecting plastic marine debris on the beach: A case study in the po river delta (italy). Drones, 140.

Toğaçar, M., Ergen, B., Cömert, Z., 2020. Waste classification using autoencoder network with integrated feature selection method in convolutional neural network models, Measurement, p. 107459.

Verheyen, M., Beckers, W., Claesen, E., Moonen, G., Demeester, E., 2016. Vision-based sorting of medium density fibreboard and grade a wood waste, 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1-6.

Vo, A.H., Son, L.H., Vo, M.T., Le, T., 2019. A novel framework for trash classification using deep transfer learning. IEEE Access 7, 178631-178639.

Vrancken, C., Longhurst, P., Wagland, S., 2019. Deep learning in material recovery: Development of method to create training database, Expert Systems with Applications, pp. 268-280.

Wang, C., Hu, Z., Pang, Q., Hua, L., 2019a. Research on the classification algorithm and operation parameters optimization of the system for separating non-ferrous metals from end-of-life vehicles based on machine vision, Waste Management, pp. 10-17.

Wang, S., Wang, J., Yang, S., Li, J., Zhou, K., 2020a. From intention to behavior: Comprehending residents' waste sorting intention and behavior formation process. Waste Management 113, 41-50.

Wang, Z., Li, H., Yang, X., 2020b. Vision-based robotic system for on-site construction and demolition waste sorting and recycling. Journal of Building Engineering 32, 101769.

Wang, Z., Li, H., Zhang, X., 2019b. Construction waste recycling robot for nails and screws: Computer vision technology and neural network approach. Automation in Construction 97, 220-228.

Wang, Z., Peng, B., Huang, Y., Sun, G., 2019c. Classification for plastic bottles recycling based on image recognition, Waste Management, pp. 170-181.

x670783915, 2019, Huaweiyun garbage classify learning. https://github.com/x670783915/huaweiyun_garbage_classify__learning (Accessed Aug. 2 2021).

Xia, Z., Zhang, S., Tian, X., Liu, Y., 2021. Understanding waste sorting behavior and key influencing factors through internet of things: Evidence from college student community. Resources, Conservation and Recycling 174, 105775.

Xiao, W., Yang, J., Fang, H., Zhuang, J., Ku, Y., 2020. Classifying construction and demolition waste by combining spatial and spectral features, Proceedings of the Institution of Civil Engineers - Waste and Resource Management. ICE Publishing, pp. 79-90.

Yang, J., Zeng, Z., Wang, K., Zou, H., Xie, L., 2021. GarbageNet: A unified learning framework for robust garbage classification. IEEE Transactions on Artificial Intelligence, 1-1.

Yang, M., Thung, G., 2016. Classification of trash for recyclability status, CS229 Project Report.

Yang, Z., Li, D., 2020. WasNet: A neural network-based garbage collection management system. Ieee Access 8, 103984-103993.

Zhang, Q., Zhang, X., Mu, X., Wang, Z., Tian, R., Wang, X., Liu, X., 2021. Recyclable waste image recognition based on deep learning. Resources, Conservation and Recycling 171, 105636.

Zhu, W., Chen, L., Wang, B., Wang, Z., 2018. Online detection in the separation process of tobacco leaf stems as biomass byproducts based on low energy x-ray imaging, Waste and Biomass Valorization, pp. 1451-1458.

Zou, Z., Shi, Z., Guo, Y., Ye, J., 2019. Object detection in 20 years: A survey. arXiv preprint arXiv:1905.05055.

Zulkifley, M.A., Mustafa, M.M., Hussain, A., Mustapha, A., Ramli, S., 2014. Robust identification of polyethylene terephthalate (pet) plastics through bayesian decision, PLoS ONE.

Zuo, M., Yan, A., 2019, Shanghai begins new waste sorting era, as china eyes cleaner image. https://www.scmp.com/news/china/society/article/3016801/shanghai-begins-new-waste-sorting-era-china-eyes-cleaner-image (Accessed 2 November 2020).

Figures

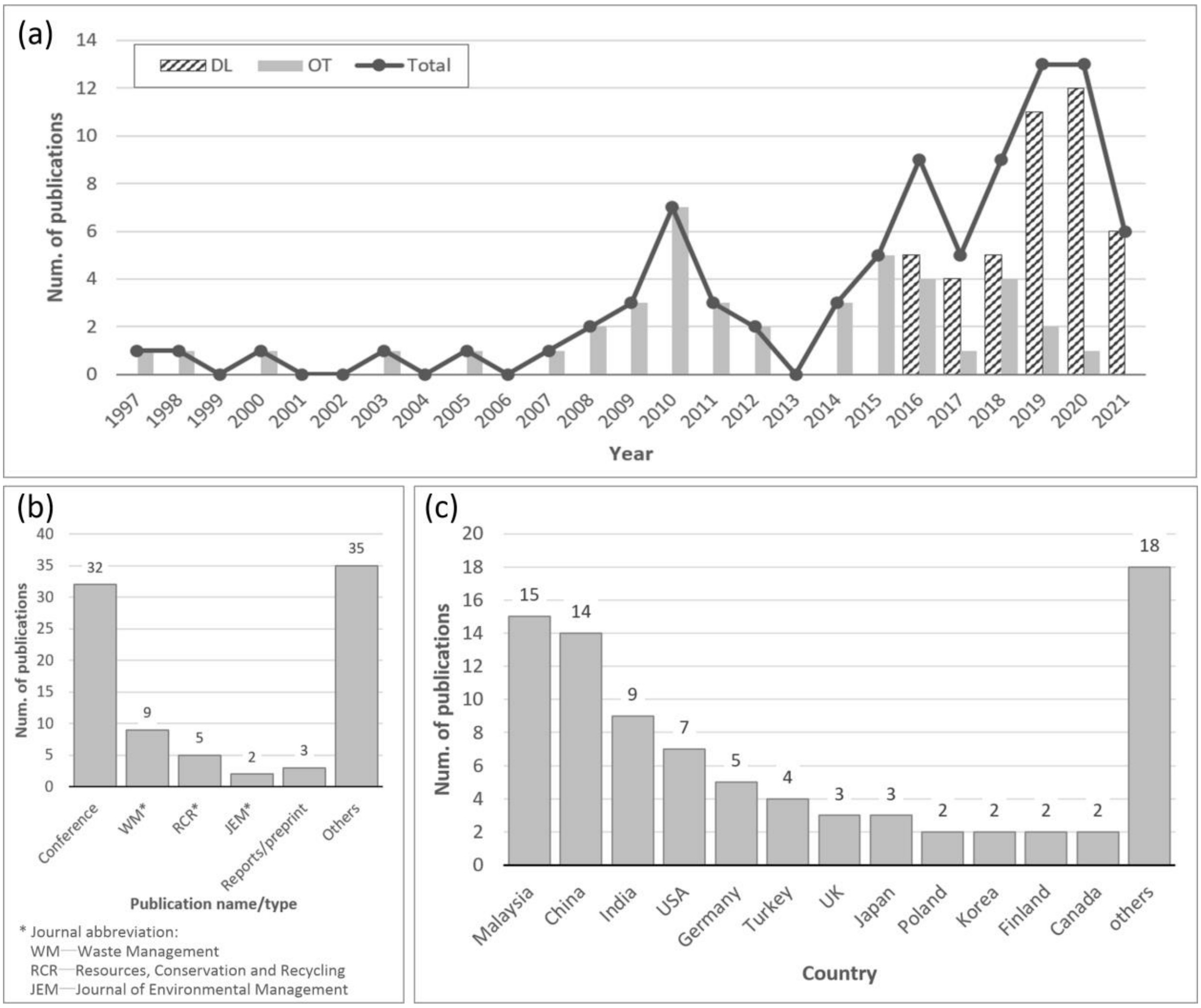

Fig. 1. (a) Yearly number of publications of CV in waste separation, where DL represents papers that have used or partially used deep learning, and OT represents papers that used other CV algorithms; Distribution of the literature collection of CV-based waste sorting (b) over different publication types/names; (c) over different countries.

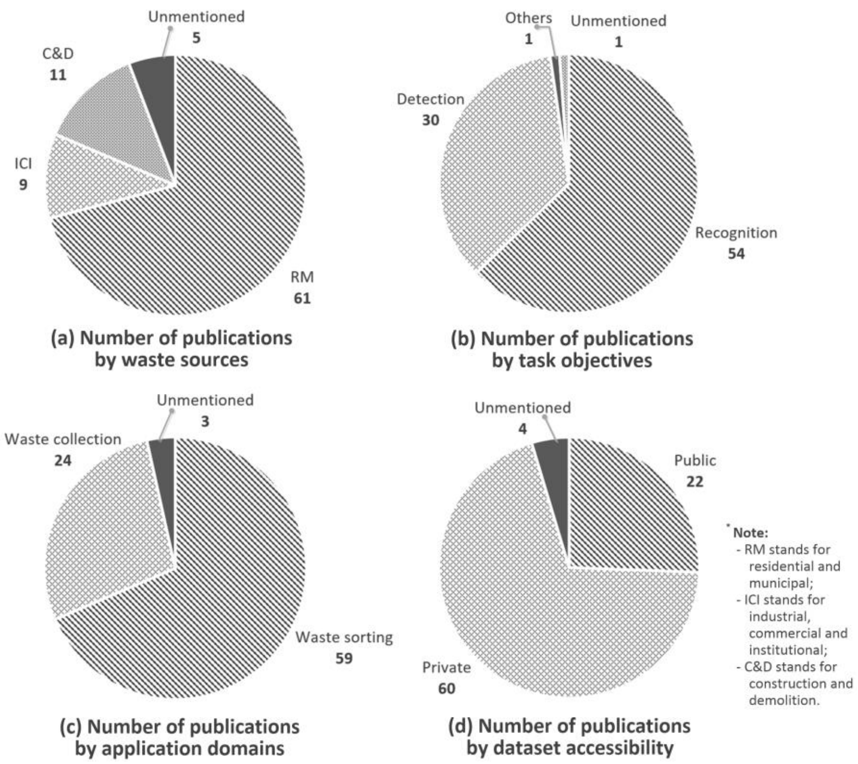

Fig. 2. Distribution of the number of publications over (a) waste sources, (b) task objectives, (c) application domains, and (d) dataset accessibility.

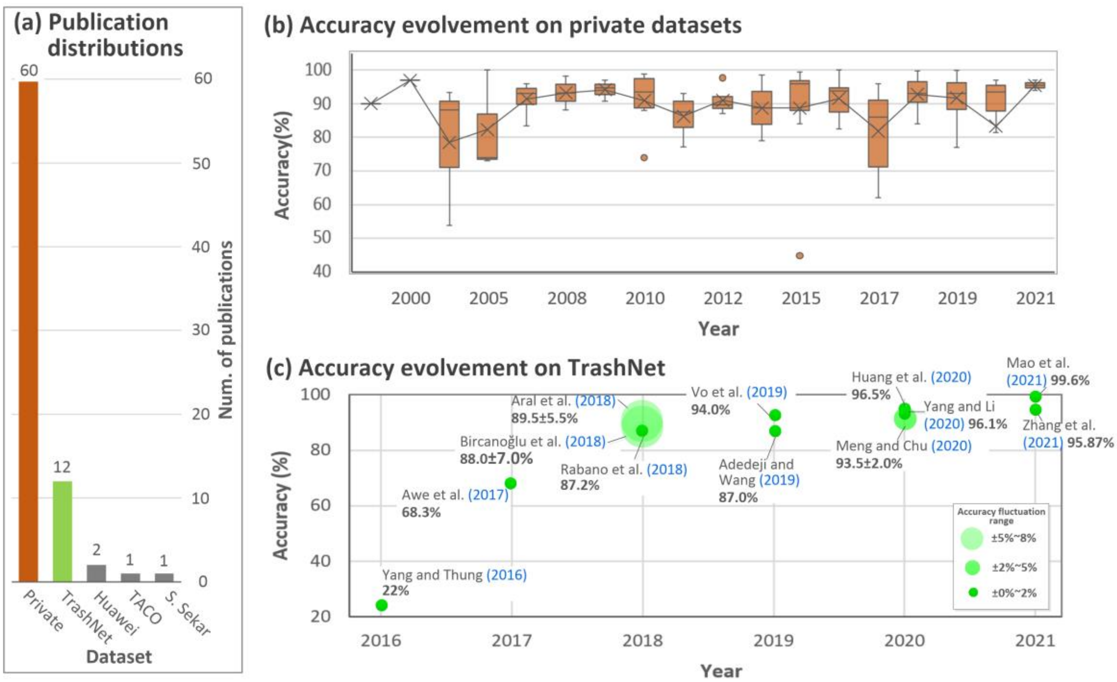

Fig. 3. (a) Numbers of publications based on private and representative public datasets; (b) accuracy evolvement on private dataset: each bin represents the range of accuracy achieved by studies in the respective year; (c) accuracy evolvement on TrashNet dataset: each bubble represents a study, and location and size of the bubbles indicate the range of accuracy.

Tables

Table 1. Summary of researches based on traditional ML techniques.

No.	Algorithms	Num. of studies	Study	Waste materials	Task* (R/D)	Accuracy
1-1	Linear discriminant analysis	6	Faibish et al. (1997)	RM	R	90%
1-2			Leitner et al. (2003)	RM	D	53.89%; 88.16%
1-3			Tachwali et al. (2007)	RM	R	94.14%
1-4			Ramli et al. (2008)	RM	R	88.15%; 98.3%
1-5			Koyanaka and Kobayashi (2010)	ICI	R	88%
1-6			Koyanaka and Kobayashi (2011)	ICI	R	85%
2-1	Nearest neighbor	11	Faibish et al. (1997)	RM	R	< 90%
2-2			Leitner et al. (2003)	RM	D	93.41%
2-3			Salmador et al. (2008)	RM	D	/
2-4			Scavino et al. (2009)	RM	R	90%
2-5			Rahman et al. (2011)	RM	R	90%; 93%
2-6			Arebey et al. (2012)	RM	D	97.67%
2-7			Hannan et al. (2012)	RM	D	89.14%
2-8			Kuritcyn et al. (2015)	C&D	R	84.8%
2-9			Gundupalli et al. (2017a)	RM	D	96%; 94%; 85%; 90%
2-10			Gundupalli et al. (2018)	RM	D	96%; 84%; 87%; 93%
2-11			Wang et al. (2019a)	ICI	R	96%
3-1	Decision tree	4	Tachwali et al. (2007)	RM	R	92%
3-2			Kuritcyn et al. (2015)	C&D	R	88%
3-3			Shaukat et al. (2016)	ICI	R	98.15%
3-4			Wang et al. (2019a)	ICI	R	94%
4-1	Bayesian network	5	Brisola et al. (2010)	C&D	R	73.96%
4-2			Liu et al. (2010)	RM	R	/
4-3			Gokyuu et al. (2011)	C&D	R	77.2%
4-4			Zulkifley et al. (2014)	RM	R	79%
4-5			Kuritcyn et al. (2015)	C&D	R	44.8%
5-1	Artificial neural network	8	Faibish et al. (1997)	RM	R	< 90%
5-2			Mattone et al. (2000)	RM	R	97%
5-3			Scavino et al. (2009)	RM	R	95%
5-4			Ramli et al. (2010)	RM	R	98.8%
5-5			Koyanaka and Kobayashi (2011)	ICI	R	85%
5-6			Arebey et al. (2012)	RM	D	87.03%
5-7			Hannan et al. (2012)	RM	D	90.19%
5-8			Islam et al. (2014)	RM	D	98.5%

6-1	Support vector machine	10	Nawrocky et al. (2010)	RM	R	96%
6-2			Aziz et al. (2015)	RM	D	99.4%
6-3			Özkan et al. (2015)	RM	R	96%
6-4			Kuritcyn et al. (2015)	C&D	R	96.5%
6-5			Guttormsen et al. (2016)	ICI	R	82.5%; 93.4%; 94.5%; 100%
6-6			Paulraj et al. (2016)	RM	R	94.3%
6-7			Singh et al. (2017)	RM	R	72%
6-8			Aziz et al. (2018)	RM	D	99.73%
6-9			Wang, C. et al. (2019a)	ICI	R	96.64%
6-10			Wang et al. (2019c)	RM	D	94.7%
7-1	Rule-based classifier	5	Mattone et al. (1998)	RM	/	/
7-2			Mattone et al. (2000)	RM	R	/
7-3			Rahman et al. (2009a)	RM	R	90.7%
7-4			Pothula et al. (2015)	ICI	D	97.76%
7-5			Zhu et al. (2018)	ICI	D	93%

[*] The column indicates the task objective a study intended to tackle, which includes two options:
- R (recognition): recognize category of a single waste in an image;
- D (detection): recognize and locate multiple waste items in an image.

Table 2. Summary of researchers based on end-to-end deep learning.

No.	Network	Num. of studies	Study	Waste materials	Task [a] (R/D)	TL[b]? (Y/N)	Accuracy
1-1	AlexNet	6	Mittal et al. [c] (2016)	RM	D	Y	87.69%
1-2			Sakr et al. (2016)	RM	R	N	83%
1-3			Yang and Thung (2016)	RM	R	N	22%
1-4			Singh et al. (2017)	RM	R	/	87%
1-5			Vrancken et al. (2019)	RM	R	Y	77%
1-6			Bobulski and Kubanek (2019)	RM	R	/	96.41%
2-1	MobileNet	4	Rabano et al. (2018)	RM	R	Y	87.2%
2-2			Bircanoğlu et al. (2018)	RM	R	Y	76%
2-3			Aral et al. (2018)	RM	R	Y	84%
2-4			Srinilta and Kanharattanachai (2019)	RM	R	Y	~88%
3-1	ResNet	6	Bircanoğlu et al. (2018)	RM	R	Y	75%
3-2			Srinilta and Kanharattanachai (2019)	RM	R	Y	91.3%
3-3			Lau Hiu Hoong et al. [d] (2020)	C&D	R	N	97%
3-4			Huang et al. (2020)	RM	R	Y	86.1%
3-5			Meng and Chu (2020)	RM	R	Y	91~95%

3-6			Zhang et al. (2021)	RM	R	Y	95.87%
4-1	Inception	3	Bircanoğlu et al. (2018)	RM	R	Y	90%
4-2			Aral et al. (2018)	RM	R	Y	94%
4-3			Huang et al. (2020)	RM	R	Y	80
5-1	DenseNet	5	Bircanoğlu et al. (2018)	RM	R	Y	95%
5-2			Aral et al. (2018)	RM	R	Y	95%
5-3			Srinilta and Kanharattanachai (2019)	RM	R	Y	~89.9%
5-4			Huang et al. (2020)	RM	R	Y	82.2%; 88.6%; 84.2%
5-5			Mao et al. (2021)	RM	R	Y	99.6%
6-1	Xception	3	Bircanoğlu et al. (2018)	RM	R	Y	85%
6-2			Aral et al. (2018)	RM	R	Y	80%
6-3			Huang et al. (2020)	RM	R	Y	84.7%
7-1	VGG	3	Sun et al. (2019)	ICI	D	Y	95~100%
7-2			Srinilta and Kanharattanachai (2019)	RM	R	Y	~88%
7-3			Huang et al. (2020)	RM	R	Y	89.7%
8-1	R-CNN	1	Ku et al. (2020)	C&D	D	N	90%
9-1	Fast R-CNN	1	Chen et al. (2017)	RM	D	/	FNR: 3%; FPR: 9%[e]
10-1	Faster R-CNN	3	Awe et al. (2017)	RM	D	Y	68.3%
10-2			Wang et al. (2019b)	C&D	D	Y	mAP[f]: 0.891
10-3			Nowakowski and Pamuła (2020)	RM	D	/	93.3%; 96.7%
11-1	Mask R-CNN	2	Proença and Simões (2020; 2020)	RM	D	/	mAP[f]: 0.194
11-2			Wang et al. (2020b)	C&D	D	Y	mAP[f]: 0.937
12-1	RetinaNet	1	Panwar et al. (2020)	RM	D	Y	mAP[f]: 0.814
13-1	Others [g]	10	Sudha et al. (2016)	/	R	N	65~70%
13-2			Rad et al. (2017)	RM	D	Y	62%; 69%
13-3			Anjum and Umar (2018)	RM	D	/	Score 4.1 out of 5.0
13-4			Sreelakshmi et al. (2019)	RM	R	N	95.7%; 96.3%
13-5			Liu et al. (2019)	RM	R	/	83.87%
13-6			Kim et al. (2019)	RM	R	/	96%
13-7			Vo et al. (2019)	RM	R	T	94%; 98%
13-8			Yang and Li (2020)	RM	R	Y	96.1%
13-9			Meng and Chu (2020)	RM	R	Y	93.75%
13-10			Liang and Gu (2021)	RM	D	Y	81.5%

[a] The column indicates the task objective a study intended to tackle, which includes two options:

- R (recognition): recognize the category of a single waste appeared on an image;

- D (detection): recognize and locate multiple waste items on an image.

[b] TL stands for transfer learning

[c] The study only aimed at detecting curb-side garbage in a general sense, instead of classifying it into specific waste types.

[d] The study aimed at determining the composition of recycled aggregates.

[e] FNR and FPR stand for false negative rate and false positive rate, respectively; the measurements in the study evaluate the accuracy of robotic sorting based on waste detection results.

[f] mAP stands for mean average precision, which is frequently used to evaluate object detection performance.

[g] Other network structures, such as Capsule-Net, OverFeat-GoogLeNet, and some self-designed CNNs, etc.

Table 3. Summary of researches integrating CNN with traditional ML models.

No.	**Study**	**Extractor**	**Classifier**	**Waste materials**	**Task[*] (R/D)**	**Accuracy**
1	Chu et al. (2018)	AlexNet	MLP	RM	R	91.6%; 98.2%
2	Adedeji and Wang (2019)	ResNet-50	SVM	RM	R	87%
3	Toğaçar et al. (2020)	AlexNet ResNet Inception	SVM	RM	R	99.95%
4	Xiao et al. (2020)	/	ELM	C&D	R	95%
5	Chen et al. (2021)	DenseNet169	SVM	C&D	R	94%
6	Yang et al. (2021)	ResNeXt-101	*k*-NN	RM	R	96.96%

[*] The column indicates the task objective a study intended to tackle, which includes two options:
- R (recognition): recognize category of a single waste item in an image;
- D (detection): recognize and locate multiple waste items in an image.

Table 4. Information of two notable private datasets of waste images.

Dataset	**Num. of img**	**Waste materials**	**Background[a]**	**Task[b] (R/D)**	**Waste types**
VN-trash (Vo et al., 2019)	5,904	RM	Simple	R	3 categories: medical, organic and inorganic waste
WasteRL[c] (Liang and Gu, 2021)	57,000	RM	In context	D	4 categories: organic waste, recyclables, hazardous waste, and other wastes

[a] Background complexity of wastes on images:
- "Simple" means the waste images have a simple and plain background
- "In context" means the wastes images were taken in complex real-life context

[b] CV tasks that the datasets are primarily used for, i.e., "R (recognition)" or "D (detection)"

[c] The dataset is available from the corresponding author by reasonable request

Table 5. Information of representative public datasets of waste images.

Dataset	**Num. of img**	**Waste materials**	**Background[a]**	**Task[b]**	**Waste types**
TrashNet (Yang and Thung, 2016)	2,527	RM	Simple	R	6 types: Glass, Paper, Cardboard, Plastic, Metal, Trash

Flickr Material Dataset [c] (Sharan et al., 2009)	1,000	/	/	R	10 types: Fabric, Foliage, Glass, Leather, Metal, Paper, Plastic, Stone, Water, Wood
Huawei Trash Classification Dataset (x670783915, 2019)	38,918	RM	Simple	R	4 types, and 43 classes: - Recyclable: Package, power bank, etc. - Kitchen: fruit peels, leftovers, etc. - Harmful: Dry battery, ointment, etc. - Others: toothpicks, fouled plastics, etc.
Washington [d] (Lai et al., 2011)	250,000	/	In context	D	51 categories of common objects, e.g., fruit and vegetable, and device and container
Birmingham [d] (Sun et al., 2019)	217	ICI	In Context	D	10 categories: Plastic bottles, cans, chains, cleaning cloths, gloves, metal objects, plastics pipes, pipe joints, sponges, wooden blocks
The Sekar's (Sekar, 2019)	25,077	RM	Simple	R	Two classes: Organic and recyclable wastes
TACO [e] (Proença and Simões, 2020)	1,500	RM	In context	D	28 categories and 60 subclasses, e.g., plastic bags, cigarette, bottle, and can
AquaTrash (Panwar et al., 2020)	369	RM	In context	D	4 categories: glass, metal, paper, plastic
ReSORT-IT [e] (Koskinopoulou et al., 2021)	21,600	RM	In context	D	4 categories: aluminum, paper and cardboard, bottles, nylon
Garbage Classification [f] (Mohamed, 2021)	15,150	RM	Simple	R	12 classes: Paper, cardboard, biological, metal, plastic, green-glass, brown-glass, white-glass, clothes, shoes, batteries, and trash
Domestic Garbage Dataset [f, g] (DataCluster Labs, 2021)	>9,000	RM	In Context	D	Paper/plastic cups, razor, plastic bags, etc.

[a] Background complexity of wastes on images:
- "Simple" means the waste images have a simple and plain background
- "In context" means the wastes images were taken in complex real-life context

[b] CV tasks that the datasets are primarily used for, i.e., "R (recognition)" or "D (detection)"

[c] Material dataset that focuses on relatively micro texture features, and thus the "Waste sources" and "Background" fields are not applicable

[d] Datasets of RGB-D images

[e] The dataset also includes pixel-level annotation that can be used for instance segmentation

[f] Datasets from Kaggle platform

[g] The dataset only includes waste images, but they have not been properly annotated with labels

1137