

Review

Analyzing mixed construction and demolition waste in material recovery facilities: Evolution, challenges, and applications of computer vision and deep learning



Adrian Langley ^{a,b}, Matthew Lonergan ^{b,c}, Tao Huang ^a, Mostafa Rahimi Azghadi ^{a,*}

^a College of Science of Engineering, James Cook University, Townsville, Queensland, Australia

^b Bingo Recycling, Sydney, New South Wales, Australia,

^c Machine Vision AI, New South Wales, Australia

ARTICLE INFO

Keywords:

Artificial intelligence
Circular economy
Computer vision
Construction and demolition waste
Deep learning

ABSTRACT

Improving the automatic and timely recognition of construction and demolition waste composition is crucial for enhancing business returns, economic outcomes and sustainability. While deep learning models show promise in recognizing and classifying homogenous materials, the current literature lacks research assessing their performance for mixed, contaminated material in commercial material recycling facility settings. Despite the increasing numbers of deep learning models and datasets generated in this area, the sub-domain of deep learning analysis of construction and demolition waste piles remains underexplored. To address this gap, recent deep learning algorithms and techniques were explored. This review examines the progression in datasets, sensors and the evolution from object detection towards real-time segmentation models. It also synthesizes research from the past five years on deep learning for construction and demolition waste management, highlighting recent advancements while acknowledging limitations that hinder widespread commercial adoption. The analysis underscores the critical requirement for diverse and high-fidelity datasets, advanced sensor technologies, and robust algorithmic frameworks to facilitate the effective integration of deep learning methodologies into construction and demolition waste management systems. This integration is envisioned to contribute significantly towards the advancement of a more sustainable and circular economic model.

1. Introduction

The amount of construction and demolition waste (CDW) sent to landfills or material recycling facilities (MRFs) is a dynamic issue influenced by a combination of economic, environmental and regulatory factors. The global trend of increasing CDW poses significant challenges for existing MRFs in maintaining efficient sorting, processing and recovery rates. In Australia, 29 million tons (38 % of total waste) of CDW were generated between 2020–21, representing a 24 % increase over the previous 4-year period (Pickin et al., 2022). The sheer volume and diversity of waste materials is a challenge for the scalability of recycling systems and the efficiency of processing materials. Significant investment in MRF infrastructure, equipment, advanced sorting technologies and or sensors or regulatory changes may be required to meet this demand (Ali and Courtenay, 2014).

Identifying the types of waste in the waste stream has been a

significant obstacle in improving the recycling rate (proportion of materials recycled or recovered). Current research in sorting and segregation largely focuses on advances in optical sensor technology (image, spectroscopic, spectral, etc.), artificial intelligence (AI) or a combination of both (Pučnik et al., 2024). Deep learning (DL) models for recognition and classification of CDW show promising results with high levels of accuracy; however, these are usually trained on datasets without complex backgrounds or heterogeneity of source material (Shahab et al., 2022). Challenges to successful implementation for commercial use include varying accuracy, lack of high-quality datasets for training and dynamic nature of construction sites (Majchrowska et al., 2022).

Intelligent detection and or segmentation using optical sensors or DL of waste in highly cluttered collections remains only a small fraction of the CDW research in the past decade; even fewer studies have examined these in actual commercial MRFs (Prasad and Arashpour, 2024). Adding to the complexity is the diversity of the waste streams, which may also

* Corresponding author at: College of Science and Engineering, James Cook University, Townsville, QLD, Australia.
E-mail address: Mostafa.rahimiazghadi@jcu.edu.au (M. Rahimi Azghadi).

include a wide range of contaminants, making generalization across different sorting scenarios or locations difficult. Real-world sorting or transfer centers can be harsh, with physical and environmental factors influencing accuracy. Sensor degradation can occur due to high dust, debris and moisture levels, lighting variability and interference from other systems and equipment, for example, vibration, affecting performance (Kroell et al., 2022).

In Australia, most companies use skip bins to collect CDW material. A top-down or surface assessment can be problematic (Driouache et al., 2024). Bins may have lighter material, such as pallets, stacked on top of heavier mixed waste (Davis et al., 2021). Relying on surface features alone may also introduce bias; contractors may load trucks in a particular way depending on accessibility to the waste or stage of demolition (Chen et al., 2021). Fig. 1(a) shows a skip bin being delivered to an MRF, with its contents ready to be tipped. The waste pile is then spread using an excavator (Fig. 1(b)) before being loaded onto a conveyor belt by a hydraulic grab. Fig. 2 demonstrates the subsequent processes of that waste in the MRF. The colored box is the focus of this review. This step often requires manual inspection to check for prohibited or bulky items that pose a safety risk to personnel or equipment downstream. Analysis of this pile not only informs operational decisions but also highlights the critical need for advanced technologies, such as deep learning, to enhance the accuracy and efficiency of material identification and processing. Given the importance of these advancements, a comprehensive review is necessary to understand the current state of the field and identify pathways for future progress.

This review aims to systematically examine and synthesize existing literature on DL applications for CDW management, with a particular

focus on object detection and segmentation for real-time implementation in MRFs. Current research mainly focuses on discrete, homogenous waste piles or mixed waste streams on conveyor belts. In contrast, this paper reviews advancements in DL technology and examines case studies of DL-assisted CDW recycling, highlighting successes and persistent challenges limiting commercial adoption. This review underscores the need for future research to harness DL methods, driving significant improvements in CDW recycling processes and supporting the industry's shift towards a circular economy.

2. Research methodology

This paper reviews DL technologies in CDW management to better understand the progress towards real-time composition analysis and object detection/segmentation. This literature review relied heavily on the quality of its initial search terms. To identify the most relevant keywords, a preliminary review was conducted on construction and demolition waste recycling, artificial intelligence, circular economy, computer vision, sensors, material recycling facilities, and deep learning. Fig. 3 illustrates the flowchart of the paper selection process.

Scopus was selected as the primary database due to its extensive coverage of peer-reviewed journals in relevant fields and its frequent use in construction engineering research (Pal and Hsieh, 2021). The search encompassed article titles, abstracts, and keywords within the specified timeframe (2004-2025). The same search criteria were also used for Web of Science (WoS) core collection. Google Scholar was also searched to identify reports, technical papers, and other materials that may not be indexed in traditional databases like Scopus or Web of Science. The bibliographies of these reports were examined to find relevant references. Although only peer-reviewed sources were analyzed, this was crucial for accessing potentially valuable research in this rapidly evolving area.

The search field in Scopus was set as "Article title, Abstract, Keywords". The first set of searches consisted of "construction AND demolition AND waste (management OR recycling)" and deep learning OR CNN OR transformer OR object detection OR segmentation OR conv* OR artificial intelligence. The second keyword set consisted of "construction waste AND learning" AND sensor OR dataset. The keywords within each set were combined with 'AND' or 'OR' operators with '*' symbol used for related word variants. The third key word search consisted of "material recovery facility" AND recycling AND deep learning OR sensors. The same searches were conducted in WoS using 'All Fields' and Google Scholar.

The initial search retrieved 692 papers from Scopus and 94 from WoS. The lists were integrated by removing duplicate entries, conditional formatting applied to the 'Article Title' and a manual review of each abstract for relevance performed. Document types such as articles, conference papers, conference reviews, reviews, and book chapters written in English were used for this review. The integration process retained 121 papers for analysis, excluding those on solid or municipal waste and deep learning studies unrelated to CDW recycling or management. These 121 papers served as a foundation for the review helping to understand the landscape of CDW research; however, given the focus on deep learning, only research from the past five years on segmentation and object detection was considered, resulting in 27 papers listed in Table 1.

Table 1 presents the information retrieved from the relevant papers organized for clarity and comparison. The fields include type of CDW materials, sensors used to collect the data, method or models investigated, research setting, image resolution (if reported), and dataset size. Each publication year has been noted to highlight the increasing volume of publications in recent years. Analysis of Table 1 highlights the diverse CDW material datasets, sensors, and deep learning models used for object detection and segmentation. The following sections examine each of these topics, exploring their opportunities and challenges.



Fig. 1. Real world example of CDW waste. (a) an example of a typical mixed CDW-containing skip bin. (b) Pile waste being spread and examined with an excavator.

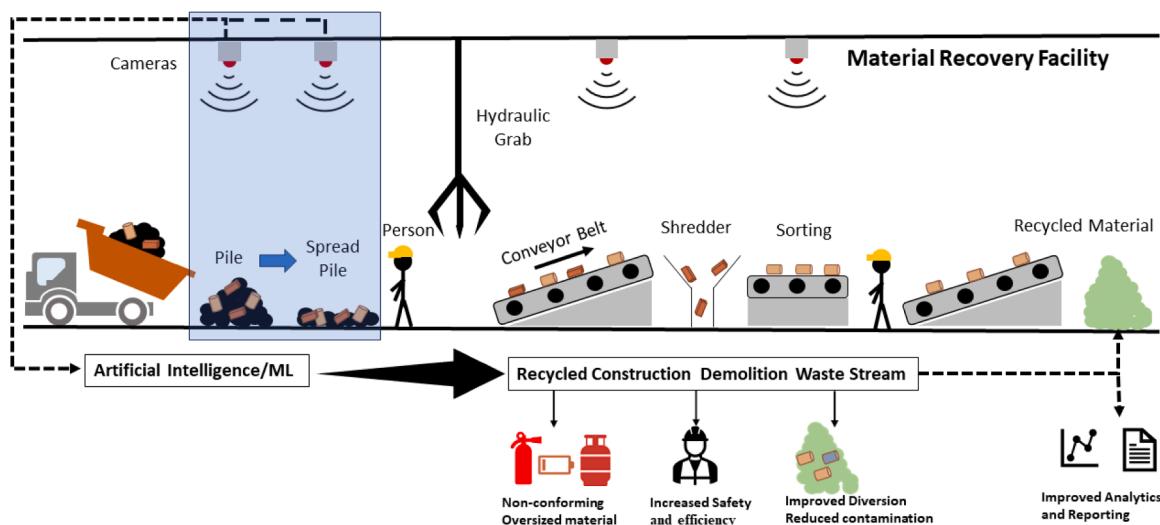


Fig. 2. Overview of the CDW recycling stream with various processing stages, including sorting and shredding. The colored box is the focus of this review.

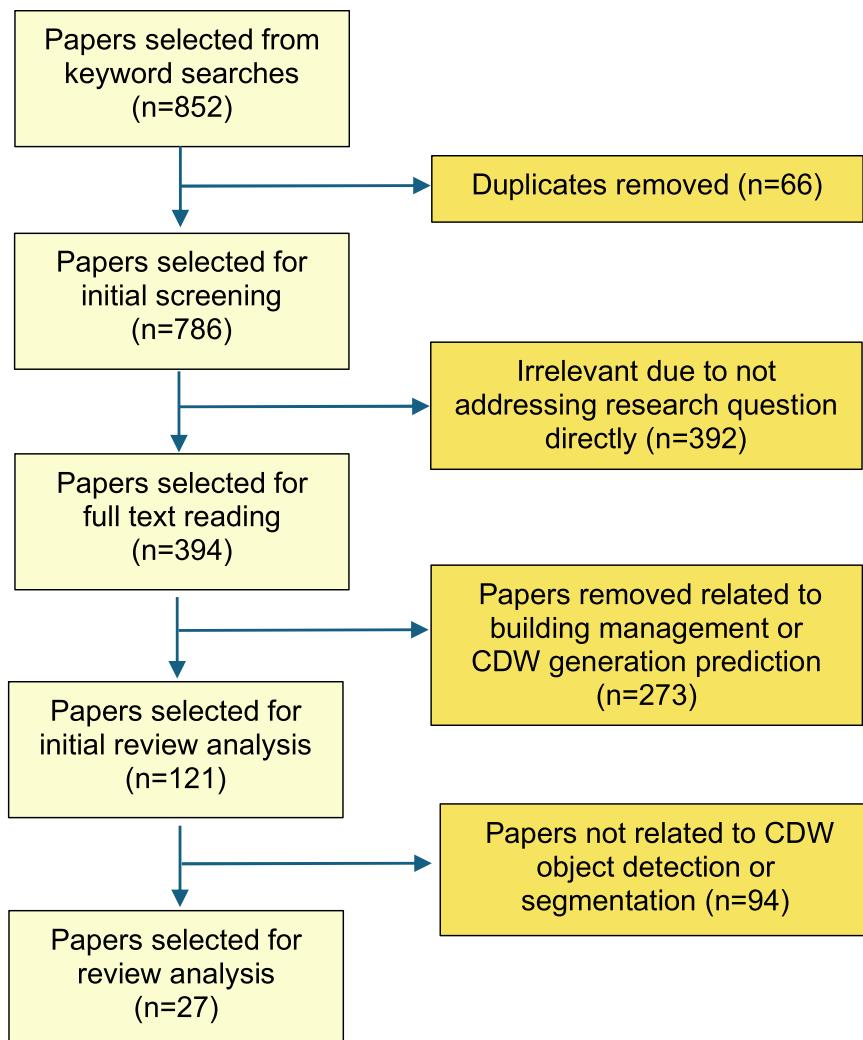


Fig. 3. Flowchart of paper selection process.

Table 1

Recent studies (since 2019) for C&DW object detection and segmentation assessment.

Reference	Year	CDW Materials	Camera/Sensor Specification	Method/Model(s)	Setting	Image Resolution	Dataset Size
(Prasad and Arashpour, 2024)	2024	concrete, rock, stone, bricks; deformed cardboard, high density polyethylene (HDPE) materials; metals (copper, aluminum, steel, iron); translucent polyethylene (LDPE) materials; waste timber and wood	RGB-D	Object segmentation/RGB-DL depth fusion strategy	Material recovery facility	1280 × 720	train: 2500, validation 713, test 355 (3568 images)
(Sirimewan et al., 2024)	2024	concrete and aggregates; wood and timber; hard plastic; soft plastic; steel; cardboard and paper; mixed waste	RGB	Object detection; segmentation/ DuoSeg++ DeepLab3+ Unet	Material recovery facility	Non-disclosed	train: 75 %, validation 15 %, test: 10 % (430 images)
(Demetriou et al., 2024)	2024	concrete; brick; tile; foam; general waste; plaster board; pipes; plastic; wood; stone	RGB	Object detection; segmentation/ YOLOv8	Conveyor belt	1920 × 1200	train: 1984, test: 573, validation: 570 (3127 images)
(Kronenwett et al., 2024)	2024	brick; sand-lime brick	Line-Scan Camera	Object detection/ SSD, YOLOv3 Faster R-CNN	Conveyor belt	300 × 300	train: 5000, test: 500, validation: 500
(Demetriou et al., 2023)	2023	concrete; brick; tile	RGB	Object detection/ SSD, YOLO Faster R-CNN	Conveyor belt	1920 × 1200	train: 4230, test 1727 (4230 images)
(Wang et al., 2023)	2023	rebar; bricks; PVC pipes, plastic wires; debris.	RGB	Object segmentation/ Swin Transformer, Twins Transformer, K-Net	Construction site	620 × 770 (average)	train: 1696, test: 200, validation 23. (1919 images - different data sources, synthetic images)
(Lux et al., 2023)	2023	concrete grains; natural stones; ceramics; bituminous grains; glass and others	RGB	Object classification; segmentation; mass regression/ RACNET	Conveyor belt	8192 × 4096 (database 2)	train: 8000, test: 800, validation: 800 (images – from 2 datasets)
K. (Lin, Kunsen et al., 2022)	2022	concrete; brick; stone; ceramic tile; glass; metal scrap; gypsum board; wood; plastic and paper	RGB	Object classification/ ResNet based	Construction site	Non-disclosed (scaled to 224 × 224)	train: 36,711, validation: 4080, test: 10,203 (2836 original images -web crawling/Image augmentation)
(Bashkirova et al., 2021)	2022	cardboard; soft plastic; rigid plastic; metal	RGB	Object detection; segmentation/ RetinaNet, Mask-RCNN, TridentNet	Conveyor belt	1920 × 1080	train: 3002, test: 929, validation 572, unlabeled: 6212 (10,715 images)
(Zhou et al., 2022)	2022	brick; wood; stone; plastic	RGB	Object detection/ Faster-RCNN, YOLO	Construction site	Non-disclosed	train: 80 %, test: 10 %, validation 10 % (3046 images)
(Na et al., 2022)	2022	concrete; brick; lumber; board; mixed waste	RGB	Object detection; segmentation/ YOLACT	Construction site	512 × 512	(500 images construction site, 288 web crawling)
(Sunwoo et al., 2022)	2022	concrete; brick; board; mixed waste; wood	RGB	Object classification; detection; segmentation/ YOLACT	Construction site	512 × 512	(599 images construction site, web crawling)
(Dong et al., 2022)	2022	rock; gravel; earth; packaging; wood; other non-inert, and mixed	RGB	Object segmentation/ Boundary aware Transformer	Weigh bridge top-down view	1980 × 1080	(5366 images)
(Li et al., 2022)	2022	concrete; machine made brick; fired brick; wood; plaster; plastic; ceramic; carton	RGB-Depth	Object classification; detection/ Mask R-CNN based fusion models	Conveyor belt	Non-disclosed (scaled to 640 × 640)	train: 70 %, test: 30 % (3367 images)
(Ko et al., 2022)	2022	rebar; bricks; PVC pipes; wires; cementitious debris	RGB	Object segmentation/ Detectron2, YOLACT and MMDetection	Construction site	Non-disclosed	(858 images)
(Lu et al., 2022)	2022	rock; gravel; earth; packaging; wood; other non-inert, and mixed	RGB	Object segmentation/ DeepLabv3+	Weigh bridge top-down view	1980 × 1080	train: 3515, validation: 754, test: 753 (5366 images)
(Chen et al., 2022)	2022	cotton gloves; wood blocks; small ferrous; plastic pipe, bamboo; corrugated paper; rebar	RGB-Depth; 3D LiDAR.	Object detection; segmentation/ Mask R-CNN	Construction site	640 × 480	Train: 454, validation: 151, test: 151 (756 images)

(continued on next page)

Table 1 (continued)

Reference	Year	CDW Materials	Camera/Sensor Specification	Method/Model(s)	Setting	Image Resolution	Dataset Size
(Song et al., 2022)	2022	brick; woods; plastics; concretes; foams	RGB	Object classification/ VGG-16, ResNet-50, and Transformer	Conveyor belt	Non-disclosed	125 pictures of each class. train: 100 samples, test: 25 samples of each type
(K. Lin, K. et al., 2022)	2022	concrete; brick; stone; ceramic tile; glass; metal scrap; gypsum board; wood; plastic and paper	RGB	Object classification/ VGG based models	Construction site	Non-disclosed (scaled to 224 × 224)	train: 36,711, validation: 4080, test: 10,203 (2836 original images- web crawling/Image augmentation)
(Davis et al., 2021)	2021	Second fix timbers; shuttering/ formwork timbers; shuttering/ formwork ply and particle boards; hard plastics; soft plastics; brick; concrete; cardboard; polystyrene	RGB	Object classification/ VGG based models	Skip bin top-down view	3000 × 2250	(525 images; 84 images empty skip bins)
(Chen et al., 2021)	2021	Inert (concrete and bricks) and non-inert (wood, plastic and bamboo)	RGB	Percentage of inert waste exceeds certain level (e.g. 50 %)/ DenseNet169	Weigh bridge top-down view	Non-disclosed	train: 70 %, validation 15 %, test: 15 % (1127 records: images, physical properties such as net weight, weight depth)
(Lau Hiu Hoong et al., 2020)	2020	concrete grains; white stones; grey stones; light colored stones; slate; clay bricks; ceramic tiles; bituminous grains; glass; wood; plastics; steel; paper and cardboard; other	RGB	Object classification/ ResNet	Laboratory setting	6000 × 4000 (scaled to 256 × 256)	train: 2000, test: 500 in each of 9 subclasses (36,000 labelled database images)
(Xiao et al., 2020)	2020	wood; rubber; brick; concrete	NIR hyperspectral	Object classification/ CNN (class not defined)	Conveyor belt	100 × 100/ wavelength 900–1700 nm.	(750 samples)
(Wang et al., 2019)	2020	Concrete; bricks; plastic bottles; rubber; wood	Hyperspectral camera, laser beam, 3D camera	Object detection/ RCNN	Conveyor belt	160 × 160	train: test 4:1 ratio, (2500 grasping rectangles)
(Ku et al., 2020)	2020	Bricks; concrete; plastic; metal; wood; rubber	NIR hyperspectral/ 3D camera	Autoencoder	Conveyor belt	640 × 640	train: 75 %, test 25 % (2500 samples)
(W. Xiao, Wen et al., 2019)	2019	foam; plastic; brick; concrete; wood	RGB/NIR hyperspectral	Object detection	Conveyor belt	640 × 480/ wavelength 900–1700 nm.	train: 250, test 150 (samples)
(W. Xiao, W. et al., 2019)	2019	Wood; plastic; bricks; concrete; rubber; black bricks	NIR hyperspectral	Object classification/ Single hidden-layer forward neural network	Conveyor belt	wavelength 900–1700 nm.	train: 150 samples each type, test: 166 pieces of woods, 130 pieces of plastics, 142 pieces of red bricks, 150 pieces of concretes, 198 pieces of rubbers, and 158 pieces of black bricks

3. Challenges and opportunities in developing CDW in-the-wild datasets

DL algorithms require large, varied datasets to extract high-level, complex abstractions; no single algorithm can guarantee the same results across all datasets. The growth of datasets designed for CDW recycling has evolved over the years since the early benchmark for waste analysis were proposed in 2016 (Yang and Thung, 2016). Early approaches dealt with single-stream datasets, which were too simple for commercial applications and did not reflect the variety of CDW materials. Table 1 highlights not only the evolution and complexity of these datasets but also the challenges of comparing research results. There exists a wide range of dataset diversity including vary image quality and resolution, location settings and types of material included (Lu and Chen, 2022).

MRF specific datasets remain a small part of the current research domain and have only started to appear in recent years. The Construction and Demolition Waste Object Detection (CODD) (Demetrou, 2022) and ZeroWaste Datasets (Bashkirova et al., 2021) attempt to capture the composition complexities of CDW through high levels of clutter and

variation. However, these may still fail to provide realistic performance of model assessment in an MRF due to lack of variation in recyclable classes. These datasets may have limited utility in assessing CDW in mixed piles (Prasad and Arashpour, 2024). The annotation and expertise-related expenses associated with developing fully supervised DL algorithms can be prohibitive (Munappy et al., 2022). To fill this gap, techniques such as data augmentation and synthetic data creation may be necessary to develop commercially viable algorithms.

3.1. Data augmentation

Data augmentation is a technique to artificially expand a training dataset by creating additional, slightly modified versions of existing data. The process involves applying a set of transformations or manipulations of the original data that preserves its label or class while creating new data points that are similar but not identical to the original. Image augmentation can include transformations such as rotation, flips, color jittering, geometric transformations, cropping and changes in brightness or contrast. The technique helps deep learning by increasing the data to train models effectively, reducing overfitting and improving

robustness through training a model on a more extensive and diverse dataset (Shorten and Khoshgoftaar, 2019). In one study, increasing the amount of CDW data through augmentation using transfer learning increased the mAP by 16 % (Na et al., 2022).

3.2. Synthetic CDW data

Synthetic data refers to artificially generated data that mimics the statistical properties of a real-world dataset. It is often used when real-world data is challenging to obtain, expensive, or private and sensitive (Nikolenko, 2021). It can also be used to augment existing datasets or to create simulations for testing and training purposes. Privacy is less of a consideration for CDW datasets than in other fields, such as medicine; however, these may be commercially sensitive (Rajotte et al., 2022).

Software, including Blender 3D and Unity, have been used to create 3D synthetic data for DL models for PET recycling and waste sorting (Fesčenko et al., 2023). A gaming engine (Unreal Engine 4) was used to generate synthetic data with similar distance, orientation, camera rotation, texture and light source to balance classes for image detection using CNN-based YOLOv5. The model utilized purely synthetic data to identify several categories of objects, including pallets and crates, on a portable Jetson Nano single-board computer with a RealSense Depth Camera D435i (Rasmussen et al., 2022). Although this work shows promising results using synthetic data in the CDW visual analysis, some limitations should be considered. Models trained purely on synthetic data may not generalize well if the data does not capture the complexity and variability of real-world conditions; however, fine-tuning models trained on large synthetic datasets on a few real images may increase real-world performance (Baaz et al., 2022).

3.3. Large language models (LLM) and segment anything (SAM)

Recently, the release of several computer vision foundation models, such as Segment Anything (SAM) (Kirillov et al., 2023), SAM-2 (Nikhila et al., 2024), DINOv2 (Oquab et al., 2023), and CLIP (Radford et al., 2021), have greatly stimulated research in the CV community and have tremendous implications for CDW modelling. Although some examples focus on text and language, these models can also assist segmentation through zero-shot recognition or generate images through prompts such as speech-to-text-to-image/video. Given their strong generalization capabilities, LLMs (Frei and Kramer, 2023) may offer significant advantages for developing or utilizing assisted segmentation techniques to efficiently create datasets for contaminated CDW.

Meta's SAM is a foundation CV model for segmentation and has been applied to numerous domains, from medical (Zhang et al., 2023) to the construction industry (Ahmadi et al., 2023). The potential to automate labelling CDW datasets makes model development easier and may improve the time and accuracy of resultant datasets. SAM's zero-shot abilities across various tasks are impressive as an out-of-the-box tool, but CDW introduces nuances not seen in everyday images. It is currently unknown if domain-specific fine-tuning, incorporating expert annotations, domain-adaptive techniques or modification to the algorithm structure, through methods such as patch inference, is needed to improve for specific circumstances or datasets (Xie et al., 2024).

To address this question, Panizza et al. studied five classes of material (bricks, concrete, PVC pipes, plastic wire and rebars) obtained from the Site Object Detection Dataset (SODA), Google and synthetic model generators (Shutterstock, 2024) and Sketchfab (Sketchfab Inc, 2024). The dataset consisted of 1276 images, divided into training (80 %) and testing (20 %) fractions. The images were labelled using LabelMe (Wada, 2018) and SAM. The authors concluded that SAM improved the ease of labelling but with a maximal IoU loss ranging from 6.6 % to 28.35 %, depending upon the material (Panizza et al., 2024). This is consistent with similar studies from other domains, for example, medicine (Ferreira and Arnaout, 2023).

LLMs came into prominence in 2018 with subsequent iterations,

including Open AI ChatGPT and Google's Gemini (Schalkwyk et al., 2023). These perform well on diverse tasks in domains including business and education; however, little work has been done in the field of CDW recycling (Saka et al., 2024). Multi-modal LLMs can also generate images; text prompting can allow precise control over image content, allowing modification or customization of backgrounds, styles and content. This may enable the rapid generation of larger CDW datasets whilst reducing costs and allowing the visualization of desired mixtures or components. While improving dataset creation times, these synthetic images might still have subtle artifacts or unrealistic elements that source content experts can identify (Cao et al., 2024). Fig. 4(a) illustrates an image generated with Google's Gemini. The text prompt was "show an image, using an overhead top-down view, of a waste tipped from a skip bin onto the floor, which has been spread out, with construction waste consisting of 50 % concrete and 50 % wood". Fig. 4(b) demonstrates the same text prompt using DeepSeek (Lu et al., 2024). Other settings included natural lighting, realistic art style and neutral mood.

4. Sensors for CDW analysis

The use of different sensors in DL systems for CDW recycling has progressed notably. Initially, the emphasis was on traditional imaging methods, where approaches centered around basic visual checks for recycling processes. However, as technology evolved, more advanced sensors have been utilized to boost the accuracy of material identification in MRFs (Chen, L. et al., 2024). As early as 2002, edge detector algorithms were used to detect pixels that belong to the target object (concrete, steel, timber, aluminum) by comparing the RGB value of the pixel with the RGB range of the material from which the target is made.



(a)



(b)

Fig. 4. Synthetic images generated by (a) Google's Gemini, and (b) DeepSeek, asked to create an image of CDW waste pile containing concrete and wood.

Once the edges have been detected, the pixels were grouped inside them to form an object (Abeid Neto et al., 2002).

In 2010, the detection of concrete in construction site images relied upon predefined color/texture value ranges for material recognition with varying thresholds. With manual color and texture features, automatic concrete detection correctly identified concrete using neural networks and SVMs with an average precision and recall of around 80 %. Missing classification of small regions (< 200 pixels) occurred as part of the pre-processing and image segmentation process (Zhu and Brilakis, 2010); however, small object identification remains a problem today with even more complex and sophisticated object detection algorithms (Chen, X. et al., 2024).

By 2014, accuracy of greater than 97.1 % were obtained using the SVM classifier on images from the Construction Materials Library (20 typical construction classes) for high-quality 200×200 -pixel color images. It averaged above 90 % for small 30×30 pixels and 92 % for highly compressed, low-quality images under real-world conditions, improving upon previous results of 70 % (Dimitrov and Golparvar-Fard, 2014). The relationship between sensor technologies and DL suggested the potential of integrating real-time data gathering with DL models to enhance sorting and recovery processes in CDW recycling. This trend reflects an increasing acknowledgment of the significance of advanced sensor technologies in boosting deep learning capabilities, ultimately leading to better material recovery solutions (Choi et al., 2023).

4.1. RGB-depth cameras

RGB cameras are widely used visual sensors in many DL applications; in many scenarios, providing enough information for a DL model to perform a recognition task properly (Qiao et al., 2024). In MRFs, dust may cover object surfaces, reducing camera recognition accuracy. Adding in depth information, which is not easily affected by dust, color, or lighting changes, may improve detection rates. However, in one model using laser line scanning for depth information, classification accuracy was only modestly increased by 1.92 ~ 2.27 % (Li et al., 2022).

A further problem for RGB cameras in CDW identification is heterogeneous sample composition within a single category. Wood, for example, can be derived from natural or engineered sources. Phenotypically similar materials such as aerated, lightweight and porous concrete may have different optical properties. Similar-looking materials, such as natural aggregates, have higher water adsorption with a lower grain strength than concrete (Anding et al., 2011). Therefore, these similarly looking materials cannot be easily identified by visible spectrum (i.e., RGB) necessitating other sensor types.

4.2. Near-infrared

Near-infrared (NIR) spectroscopy, encompassing wavelengths from 700 to 2500 nanometers, elucidates the chemical composition of a sample by analyzing the characteristic vibrational transitions induced by the absorption of infrared radiation (Emsley et al., 2022). NIR may improve plastic identification compared to standard image identification but may have some shortcomings, particularly in identifying black materials, as these are less distinguishable. Adding NIR spectral recognition increases the cost and complexity of systems. In the context of CDW, many inorganic substances in construction waste do not contain distinct functional groups absorption spectra, making NIR spectroscopy difficult without initial pre-processing of the spectral curve (W. Xiao, W. et al., 2019).

4.3. Hyperspectral imaging

Hyperspectral imaging (HSI) is an emerging rapid and non-destructive technology that may have promising applications for the identification of CDW. HSI offers a potential advantage over NIR by integrating spectroscopic and visible imaging capabilities within a single

system. This unique approach enables simultaneous acquisition of both spectral and spatial information (Tahmasbian et al., 2021). It has been used to separate different types of plastics in municipal solid wastes and to recover and recycle concrete, mortar aggregates, bricks, tiles and wood (Castro-Díaz et al., 2023). HIS images are costly due to the complicated apparatus required to acquire a wide continuous spectrum. Still, they are also more robust to solar reflections, a significant concern for RGB cameras. Reconstruction of HIS from single RGB images using convolutional neural networks (CNNs) offers a portable, low-cost alternative (Gao et al., 2021), which can be adopted in the CDW domain but requires significant research and development.

4.4. Multi-sensor fusion

The composition of CDW may be complex, resulting in equally complex spectral features obtained from NIR or HIS techniques. In such cases, multi-sensor fusion and the internet of things (IoT) may provide agile solutions for classification and real-time monitoring of municipal solid waste (MSW) (Mookkaiah et al., 2022). Such techniques may also be introduced into CDW recognition. Emerging 5G and 6G networks enable IoT infrastructure that will allow higher resolution images and integration of wireless sensor networks to track the CDW at the source and direct it to the appropriate MRF (Jagan and Jayarin, 2022).

4.5. Other sensors

Other sensors that can be fused for CDW recognition include infrared, ultrasonic, line and laser scanning, weight and chemicals. Lidar sensor fusion leverages detailed 3D information, whereas ultrasonic can provide high precision distance measurement capability to improve waste sorting and efficiency (Aliew, 2022). Combining multiple sensors can provide a more comprehensive understanding of waste streams but at increased expense to integrate into existing infrastructure and may be subject to calibration, data synchronization and complexity or processing issues. Algorithms using sensor fusion can require significant technical expertise. Sound recognition, metal detection and weight were used to detect glass and metal in trash bases based on spectrograms in highly controlled circumstances with an accuracy of 98 % (Funch et al., 2021); however, this is impractical for a large MRF.

5. Deep learning for visual CDW analysis

The application of DL technologies for CDW analysis in recycling facilities has evolved significantly over the past decade. Initial studies focused on traditional CV techniques, which often struggled with the complexities and variabilities inherent to CDW streams (K. Lin, K. et al., 2022). Early work highlighted the need for enhanced detection methodologies to manage the processing of diverse materials effectively (Lopes et al., 2024). As the field matured, researchers began to explore machine learning approaches that could automate the classification and sorting of CDW with improved accuracy, laying the groundwork for subsequent innovations in deep learning (Wang et al., 2020).

5.1. Object detection for visual CDW analysis

In a short space of time, driven by the increasing speed of object detection algorithms and the need for real-time sorting, classification has given way to object detection and segmentation algorithms (Diwan et al., 2023). MaskRCNN has been used with robot models for CDW detection. In one study, using an RGB-Depth camera, a real-time sorting robot was able to analyze CDW on a conveyor belt with a speed of 0.25 m per second and inference time not exceeding 1920 milliseconds (Li et al., 2022). Despite its accuracy, MaskRCNN's two stage architecture makes it less suitable for real-time applications compared with single stage algorithms (YOLO, SDD). Classifying materials from 15 to 25 m was also challenging for RGB cameras; this poses a problem for cameras placed

too far from the load drop off zone, or if those zones vary on the MRF floor (Wolff et al., 2021).

In the last few years, the research landscape has expanded beyond object detection to include comprehensive analyses of CDW composition. Demetriou et al. assessed 18 models, both single-stage (SSD, YOLO) and two-stage (Faster-RCNN) detector architectures coupled with various backbones feature extractors (ResNet, MobileNetV2, EfficientDet) on 6600 CDW samples belonging to brick, concrete and tile under working conditions with normal and heavily stacked adhered samples. YOLOv7 attained the best accuracy (mAP50:95 ~ 70 %) at the highest inference speed of less than 30ms (Demetriou et al., 2023).

5.2. Object segmentation for visual CDW analysis

Compared to object detection, segmentation can provide more precise information about mixed objects or regions within the image. Semantic segmentation performs classification at the pixel level, which may provide more granularity on waste composition compared with classification/object detection models. It is currently one of the most popular research fields in CV. It underlies technology in autonomous vehicles, medicine, image search engines, industrial inspection and augmented reality (Yu et al., 2023).

YOLOACT is a real-time instance segmentation algorithm that extends the YOLO framework by predicting object masks alongside bounding boxes and class labels. It has shown promising results in some studies, but its performance may be limited by the size and augmentation of the training dataset (Ko et al., 2022). Vision Transformers, a relatively new class of models, have demonstrated strong performance in various image recognition tasks, including object detection and semantic segmentation, and are considered a promising avenue for future research in construction and CDW applications (Dong et al., 2022).

The current leading segmentation algorithm for 'real-time' analysis of highly cluttered CDW, YOLOv9e-seg, achieves a mean average precision (mAP50:95) of 49.92 while processing at 125 frames per second. However, the mAP for small recyclable objects (<322 pixels) is significantly lower compared to medium (322–962 pixels) or large (>962 pixels) materials. The authors revealed that the examined models exhibited a bias towards objects situated within less visually complex environments. This preference could be attributed to the enhanced ease of object identification and segmentation within such contexts. The implementation of a patch-based inference strategy mitigated the detrimental impact of visual clutter on object detection performance, resulting in a mAP of 56.34. Notably, this improvement in detection accuracy was achieved without a significant compromise in classification accuracy or inference speed (Prasad and Arashpour, 2024).

6. Challenges and opportunities for DL in commercial CDW MRFs

6.1. Dataset challenge

A look at existing literature, as outlined in Table 1, shows numerous themes tied to the challenges of CDW dataset generation. Generating standardized datasets for this industry is difficult and expensive; the makeup of waste shifts over time and varies by location, indicating the need for flexible and thorough datasets that can accurately capture these changes. The challenges of annotating these datasets highlight not only the technical hurdles in identifying and categorizing materials but also the demanding and sometimes subjective nature of the labeling process (Demrozi et al., 2023). Despite advancements in automated data collection methods like image recognition and sensor fusion, there is still a notable lack of standardizing datasets for deep learning in CDW waste recycling.

The lack of a common standard results in datasets being developed inconsistently, which can negatively impact the performance and adaptability of deep learning models in various recycling settings (Liang

et al., 2020). Moreover, existing research often fails to address the differences found in real-world scenarios, where environmental factors and operational conditions can affect the accuracy and usefulness of the generated datasets. These issues stress the urgent need for more research to create best practices for dataset production, focusing on data quality, diversity and integration of both multi-modal and emerging sensor technologies with real-world recycling operations. Increased collaboration among academics, industry players, and policymakers could support the creation of standardized data collection protocols and promote the sharing of datasets across research and practical environments.

6.2. Generative AI and domain specific prompt engineering opportunities

Annotating complex CDW images is challenging due to the presence of small, innumerable, and difficult-to-identify materials. Stable Diffusion offers solutions by generating synthetic CDW images with controlled composition, augmenting existing datasets with label-preserving techniques, and even generating initial labels for human refinement (Valvano et al., 2024). Moreover, the rise of multi-modal AI (Barua et al., 2023), including the potential for CDW-specific models, combined with the increasing integration of IoT devices, promises to revolutionize CDW management by providing more comprehensive and accurate insights into recovery processes.

The advent of prompt engineering enables the use of pre-trained models for tasks by using customized prompts, which helps the models understand context better without needing a large amount of labeled data or complete retraining. Combining techniques like the YOLO system with infrared imaging shows how prompt engineering can improve detection accuracy in certain conditions by using different data enhancement methods (Yang et al., 2024). Overall, these strategies not only make the annotation process more efficient and cheaper but also set the stage for scalable solutions in many deep learning applications.

6.3. Business case challenges and opportunities

A business case supporting the purchase, development, installation, maintenance and use of a DL system should include an analysis of the expected costs and benefits. AI projects can be expensive and may not provide an immediate return on investment (ROI). The business case should start with a clearly defined problem that the model will solve and how it will improve efficiency, reduce costs, or open new opportunities. It should also quantify the problem, providing data demonstrating the extent and impact of the problem e.g., losses due to errors, time wasted on manual tasks, potential gains from new services or efficiency in current ones. The proposed solution should also outline the model functionality (classification, detection and segmentation) and the hardware or software services required (Enholm et al., 2021).

There are some spectacular examples of AI implementation failure (Olavsrud, 2024). Hence, knowing what an AI investment is worth and how to measure that value is a prerequisite for intelligent decision-making. An estimated 87 % of data science projects fail to make it into production (VentureBeat, 2019). Simply investing money and expecting a high-tech solution at the end of the project does not happen without the proper leadership support and conditions for success. The costs for development can be considerable and vary significantly; small to medium-scale projects can cost from \$10,000 to \$500,000 (CHI Software, 2024). Development considerations include data collection and annotation, model development (engineers, data scientists), infrastructure, and services. Deployment costs may include integrating models into existing systems and other expenses, for example cloud computing. Maintenance considerations include model updates, retraining, monitoring performance drift, ongoing technical support, and hardware replacement (edge devices, routers, power supplies) (Smith, 2023).

Several frameworks (Bevilacqua et al., 2023) have been published to help estimate ROI, but what seems clear is that the productivity dividend

of AI does not materialize immediately (Pandey et al., 2021). Specifically, for CDW processing, tangible benefits include reducing labor costs through task automation and reducing errors, increasing revenue through less plant downtime, speeding up processing times, and producing a cleaner product for marketing. Intangible benefits include improved safety or compliance, potential competitive advantage in the CDW recycling market and enhanced customer experience by providing more detailed and accurate information for billing and reporting purposes.

7. Discussion and conclusion

The study of CV and DL technologies in mixed CDW analysis at MRFs shows new innovations but also points out major challenges that limit their effective use. Key results from this review show that while DL models, especially convolutional neural networks, have improved material classification accuracy, the diverse nature of mixed CDW waste creates serious problems. Differences in material types and insufficient training datasets reduce the effectiveness and widespread use of these technologies in practical commercial situations such as MRFs. More large volume, high-quality, and varied datasets that represent the complexities of mixed waste streams to enable strong algorithm training are needed.

Reinforcing the main point, it is evident from literature that combining computer vision, sensors, and DL in CDW analysis is a critical area in waste management technology. Additionally, it highlights the importance of interdisciplinary approaches that mix engineering, environmental science, and waste management to develop better and more sustainable solutions. Developing an international collaborative data repository for CDW images may help to address issues around dataset standardization, sensors modality and model performance (Kras et al., 2020), if commercial barriers can be overcome. The environmental benefits of recycling efforts can also be better assessed using the database. By studying trends and patterns in CDW disposal, regions could pinpoint their most pressing waste streams and customize strategies accordingly. For instance, areas with high amounts of materials, like concrete or metal, can fine-tune their recycling facilities to focus more on these materials.

The synthesis of this literature also highlights future directions for research. The current focus is largely on the recyclable waste stream; however, components such as fire extinguishers, gas cylinders, lithium batteries, for example, are highly relevant, but largely absent from the current literature. In 2024, the annual report from the recycling industry fire protection firm found that there were 373 fires at MRFs and transfer stations across all recycling facility types, including CDW (Staub, 2024). The number of MRF fires is becoming increasingly more common. In the US and Canada, the number of major fires had increased by more than one third since 2017, numbering 390 in 2022. It is estimated that this figure is likely to be much higher as smaller fires often go unreported. These fires are due to several factors including the growing number of plants to deal with demand, major new hazards in the waste streams, and global shifts in policy and attitudes towards CDW management. Finally, a hotter and dryer climate will make it easier for fires to start and spread, destroying millions of dollars in infrastructure (Nugent, 2023).

Despite the extensive information available on CDW waste and its recycling potential, there remain gaps, especially regarding practical applications and stakeholder involvement. Much of the research has focused on measuring material types and volumes, often ignoring the social and economic factors that shape recycling practices. There is a deficit of comprehensive studies that assess how consumer knowledge and habits impact recycling outcomes. Future research could look at pile composition, as it is delivered to the MRF, to feed back to customers in real-time. This is essential for developing effective strategies to divert CDW waste, enhance recovery rates, reduce costs and improve overall environmental results.

The wider implications of these findings suggest that solving the identified issues could improve CDW waste management, enhancing material recovery efficiency and supporting sustainable practices that align with current environmental objectives. Applying advanced CV and DL technologies could automate complex sorting tasks, cut operational costs, and improve the quality of recovered materials, contributing to the circular economy. Successful use of these technologies can also lead to a better understanding of waste composition, guiding future regulations and aligning industry practices with sustainability standards.

The costs involved in generating large-scale datasets are not insignificant (ul Hassan et al., 2022). The literature demonstrates that a wide range of sensors and resolutions have been used previously (see Table 1). However, maintenance schedules are usually not listed. For example, cameras installed in the roof of MRFs may require special access equipment, such as a scissor lift, and trained staff for access and cleaning. Access may be difficult during business hours if there are technical issues. The video quality may degrade during the week if weekend cleaning schedules, during downtime, are adopted. This may affect the outcome or efficiency of the DL models. Furthermore, there would need to be a clear cost benefit for adopting these models as part of the deployment of new digital technologies for assessing and processing CDW.

To tackle these challenges, several areas of exploration are recommended, such as hybrid models that combine different data sources, like sensor data and CV results, to improve classification accuracy. Furthermore, collaborative frameworks involving industry, academia, and tech developers can drive innovation by creating feedback loops that enhance machine learning algorithms based on real-world experiences. There is also a need for studies on the ethical implications of AI in sorting tasks, as this technology is adopted, to ensure that automation does not reduce human oversight and accountability (Joseph et al., 2024).

In conclusion, the literature review gives a thorough overview of the critical issues facing CV and DL in analyzing mixed CDW in MRFs. By detailing both the progress made and the challenges that still exist, it underscores the pressing need for concerted research to realize the full potential of these technologies in achieving sustainable waste management practices.

CRediT authorship contribution statement

Adrian Langley: Writing – original draft, Visualization, Software, Methodology, Investigation, Conceptualization. **Matthew Lonergan:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Conceptualization. **Tao Huang:** Writing – review & editing, Supervision. **Mostafa Rahimi Azghadi:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Mostafa Rahimi Azghadi reports financial support was provided by Bingo Industries. Matthew Lonergan reports a relationship with Bingo Industries that includes: consulting or advisory. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by an Australian Research Training Program Scholarship.

Data availability

No data was used for the research described in the article.

References

Abeid Neto, J., Ardit, D., Evens, M.W., 2002. Using colors to detect structural components in digital pictures. *Comput.-Aided Civil Infrastruct. Eng.* 17 (1), 61–67.

Ahmadi, M., Ahmad Gholizadeh, L., Sharifi, A., Beris, A.T., Nouri, M., Amir Sharifzadeh, J., 2023. Application of segment anything model for civil infrastructure defect assessment. *arXiv.org*.

Ali, M., Courtenay, P., 2014. Evaluating the progress of the UK's material recycling facilities: a mini review. *Waste Manag. Res.* 32 (12), 1149–1157.

Aliew, F., 2022. An approach for precise distance measuring using ultrasonic sensors. *iecmca 2022*.

Anding, K., Linß, E., Träger, H., Rückwardt, M., Göpfert, A., 2011. Optical identification of construction and demolition waste by using image processing and machine learning methods.

Baaz, A., Yonan, Y., Hernandez-Diaz, K., Alonso-Fernandez, F., Nilsson, F., 2022. Synthetic data for object classification in industrial applications.

Barua, A., Ahmed, M.U., Begum, S., 2023. A systematic literature review on Multimodal machine learning: applications, challenges, gaps and future directions. *IEEE Access.* 11, 14804–14831.

Bashkirova, D., Abdelfattah, M., Zhu, Z., Akl, J., Alladkani, F., Hu, P., Ablavsky, V., Calli, B., Bargal, S.A., Saenko, K., 2021. ZeroWaste dataset: towards deformable object segmentation in cluttered scenes.

Bevilacqua, M., Berente, N., Domin, H., Goehring, B., Rossi, F., 2023. The return on investment in AI ethics: a holistic framework. *arXiv.org*.

Cao, B., Yuan, J., Liu, Y., Li, J., Sun, S., Liu, J., Zhao, B., 2024. SynArtifact: classifying and alleviating artifacts in synthetic images via vision-language model. *arXiv.org*.

Castro-Díaz, M., Osmani, M., Cavalaro, S., Cacho, I., Urias, I., Needham, P., Thompson, J., Parker, B., Lovato, T., 2023. Hyperspectral imaging sorting of refurbishment plasterboard waste. *Appl. Sci.* 13 (4).

Chen, J., Lu, W., Xue, F., 2021. Looking beneath the surface": a visual-physical feature hybrid approach for unattended gauging of construction waste composition. *J. Environ. Manage.* 286, 112233.

Chen, L., Xia, C., Zhao, Z., Fu, H., Chen, Y., 2024a. AI-driven sensing. *Technology: Rev. Sensor (Basel)* 24 (10), 2958.

Chen, X., Deng, L., Hu, C., Xie, T., Wang, C., 2024b. Dense small object detection based on an improved YOLOv7 model. *Appl. Sci.* 14 (17), 7665.

Chen, X., Huang, H., Liu, Y., Li, J., Liu, M., 2022. Robot for automatic waste sorting on construction sites. *Autom. Constr.* 141.

CHI Software, 2024. AI app development cost: calculations and factors to consider. <https://chisoftware.medium.com/ai-app-development-cost-calculations-and-factors-to-consider-8244acb17eef>. (Accessed 1 May 2024).

Choi, J., Lim, B., Yoo, Y., 2023. Advancing plastic waste classification and recycling efficiency: integrating image sensors and deep learning algorithms. *Appl. Sci.* 13 (18), 10224.

Davis, P., Aziz, F., Newaz, M.T., Sher, W., Simon, L., 2021. The classification of construction waste material using a deep convolutional neural network. *Automat. Construct.* 122.

Demetriou, D., 2022. Construction and demolition waste object detection dataset. In: Data, M. (Ed.), *Data Archiving and Networked Services (DANS)*, V1 ed.

Demetriou, D., Mavromatidis, P., Petrou, M.F., Nicolaides, D., 2024. CODD: a benchmark dataset for the automated sorting of construction and demolition waste. *Waste Manage* 178, 35–45.

Demetriou, D., Mavromatidis, P., Robert, P.M., Papadopoulos, H., Petrou, M.F., Nicolaides, D., 2023. Real-time construction demolition waste detection using state-of-the-art deep learning methods single-stage vs two-stage detectors. *Waste Manage* 167, 194–203.

Demrozi, F., Turetta, C., Machot, F.A., Pravadelli, G., Kindt, P.H., 2023. A Comprehensive Review of Automated Data Annotation Techniques in Human Activity Recognition. *Cornell University arXiv*.

Dimitrov, A., Golparvar-Fard, M., 2014. Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Adv. Eng. Inf.* 28 (1), 37–49.

Diwan, T., Anirudh, G., Tembhurne, J.V., 2023. Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimed. Tools. Appl.* 82 (6), 9243–9275.

Dong, Z., Chen, J., Lu, W., 2022. Computer vision to recognize construction waste compositions: a novel boundary-aware transformer (BAT) model. *J. Environ. Manage.* 305, 114405.

Driouache, Y., Milipied, J., Motamedi, A., 2024. Vision-based method to identify materials transported by dump trucks. *Eng. Appl. Artif. Intell.* 135, 108768.

Emsley, N.E.M., Holden, C.A., Guo, S., Bevan, R.S., Rees, C., McAinch, M.R., Martin, F.L., Morais, C.L.M., 2022. Machine learning approach using a handheld near-infrared (NIR) device to predict the effect of storage conditions on tomato biomarkers. *ACS. Food Sci. Technol.* 2 (1), 187–194.

Enholm, I.M., Papagiannidis, E., Mikalef, P., Krogstie, J., 2021. Artificial Intelligence and business value: a literature review. *Inf. Syst. Front.* 24 (5), 1709–1734.

Ferreira, D., Arnaout, R., 2023. Are foundation models efficient for medical image segmentation? *arXiv.org*.

Feščenko, V., Arents, J., Kadikis, R., 2023. Synthetic data generation for visual detection of flattened PET bottles, machine learning and knowledge extraction. pp. 14–28.

Frei, J., Kramer, F., 2023. Annotated dataset creation through large language models for non-english medical NLP. *J. Biomed. Inform.* 145, 104478.

Funch, O.I., Marhaug, R., Kohtala, S., Steinert, M., 2021. Detecting glass and metal in consumer trash bags during waste collection using convolutional neural networks. *Waste Manage* 119, 30–38.

Gao, X., Wang, T., Yang, J., Tao, J., Qiu, Y., Meng, Y., Mao, B., Zhou, P., Li, Y., 2021. Deep-learning-based hyperspectral imaging through a RGB camera. *J. Electron. ImAGING* 30 (5), 053014–053014.

Jagan, G.C., Jayarin, P.J., 2022. 2022 International conference on communication, computing and internet of things (IC3IoT). IEEE 1–4.

Joseph, S.A., Kolade, T.M., Val, O.O., Adebiyi, O.O., Ogungbemi, O.S., Olaniyi, O.O., 2024. AI-powered information governance: balancing automation and Human oversight for optimal organization productivity. *Asian J. Res. Comput. Sci.* 17 (10), 110–131.

Kirillov, A., Mintun, E., Nikhila, R., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Wan-Yen, L., Dollár, P., Girshick, R., 2023. Segment anything. *arXiv.org*.

Ko, P., Gopee, M., Yu, L., Garg, Y., Mansour, M., García de Soto, B., 2022. Development of an instance segmentation model for the automatic identification of recyclable objects in construction sites.

Kras, A., Celi, L.A., Miller, J.B., 2020. Accelerating ophthalmic artificial intelligence research: the role of an open access data repository. *Curr. Opin. Ophthalmol.* 31 (5), 337–350.

Kroell, N., Chen, X., Greiff, K., Feil, A., 2022. Optical sensors and machine learning algorithms in sensor-based material flow characterization for mechanical recycling processes: A systematic literature review. *Waste Manag.* 149, 259–290.

Kronenwett, F., Maier, G., Leiss, N., Gruna, R., Thome, V., Langle, T., 2024. Sensor-based characterization of construction and demolition waste at high occupancy densities using synthetic training data and deep learning. *Waste Manag. Res.* 734242X241231410.

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. *J. Mater. Cycles. Waste Manage* 23 (1), 84–95.

Lau Hiu Hoong, J.D., Lux, J., Mahieux, P.-Y., Turcay, P., Ait-Mokhtar, A., 2020. Determination of the composition of recycled aggregates using a deep learning-based image analysis. *Autom. Constr.* 116, 103204.

Li, J., Fang, H., Fan, L., Yang, J., Ji, T., Chen, Q., 2022. RGB-D fusion models for construction and demolition waste detection. *Waste Manag.* 139, 96–104.

Liang, X., Nguyen, D., Jiang, S.B., 2020. Generalizability issues with deep learning models in medicine and their potential solutions: illustrated with cone-beam computed tomography (CBCT) to computed tomography (CT) image conversion. *Mach. Learn. Sci. Technol.* 2 (1).

Lin, K., Zhao, Y., Zhou, T., Gao, X., Zhang, C., Huang, B., Shi, Q., 2022a. Applying machine learning to fine classify construction and demolition waste based on deep residual network and knowledge transfer. *Environ. Dev. Sustain.*

Lin, K., Zhou, T., Gao, X., Li, Z., Duan, H., Wu, H., Lu, G., Zhao, Y., 2022b. Deep convolutional neural networks for construction and demolition waste classification: VGGNet structures, cyclical learning rate, and knowledge transfer. *J. Environ. Manage.* 318, 115501.

Lopes, C.d.M.N., Cury, A.A., Mendes, J.C., 2024. Construction and demolition Waste management and Artificial intelligence – a systematic review. *RGSA: Revista de Gestão Social e Ambiental* 18 (10), e08810.

Lu, H., Liu, W., Zhang, B., Wang, B., Dong, K., Liu, B., Sun, J., Ren, T., Li, Z., Yang, H., Sun, Y., Deng, C., Xu, H., Xie, Z., Ruan, C., 2024. DeepSeek-VL: Towards Real-World Vision-Language Understanding. *Cornell University arXiv*.

Lu, W., Chen, J., 2022. Computer vision for solid waste sorting: a critical review of academic research. *Waste Manag.* 142, 29–43.

Lu, W., Chen, J., Xue, F., 2022. Using computer vision to recognize composition of construction waste mixtures: a semantic segmentation approach. *Resour. Conserv. Recycl.* 178.

Lux, J., Lau Hiu Hoong, J.D., Mahieux, P.-Y., Turcay, P., 2023. Classification and estimation of the mass composition of recycled aggregates by deep neural networks. *Comput. Ind.* 148, 103889.

Majchrowska, S., Mikolajczyk, A., Ferlini, M., Klawikowska, Z., Plantykon, M.A., Kwasigroch, A., Majek, K., 2022. Deep learning-based waste detection in natural and urban environments. *Waste Manag.* 138, 274–284.

Mookhaiah, S.S., Thangavelu, G., Hebbar, R., Haldar, N., Singh, H., 2022. Design and development of smart Internet of Things-based solid waste management system using computer vision. *Environ. Sci. Pollut. Res. Int.* 29 (43), 64871–64885.

Munappy, A.R., Bosch, J., Olsson, H.H., Arpteg, A., Brinne, B., 2022. Data management for production quality deep learning models: challenges and solutions. *J. Syst. Softw.* 191.

Na, S., Heo, S., Han, S., Shin, Y., Lee, M., 2022. Development of an artificial intelligence model to recognise construction waste by applying image data augmentation and transfer learning. *Buildings* 12 (2).

Nikhila, R., Gabeur, V., Yuan-Ting, H., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Kalyan Vasudev, A., Carion, N., Chao-Yuan, W., Girshick, R., Dollár, P., Feichtenhofer, C., 2024. SAM 2: segment anything in images and videos. *arXiv.org*.

Nikolenko, S.I., 2021. *Synthetic Data for Deep Learning*. Springer, Cham, Switzerland.

Nugent, C., 2023. Why recycling plants keep catching on fire. <https://time.com/6271576/recycling-plant-fire-indiana/>. (Accessed 12th January 2024).

Olavsrød, T., 2024. 10 famous AI disasters. <https://www.cio.com/article/1908885-famous-analytics-and-ai-disasters.html>. (Accessed 30 April 2024).

Oquab, M., Darcret, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Assran, M., Ballas, N., Galuba, W.,

Howes, R., Po-Yao, H., Shang-Wen, L., Misra, I., Rabbat, M., Sharma, V., Synnaeve, G., Hu, X., Jegou, H., Mairal, J., Labatut, P., Joulin, A., Bojanowski, P., 2023. DINOv2: Learning robust visual features without supervision. *Trans. Mach. Learn. Res.* **J.** 1–31.

Pal, A., Hsieh, S.-H., 2021. Deep-learning-based visual data analytics for smart construction management. *Autom. Constr.* **131**, 103892.

Pandey, S., Gupta, S., Chhajed, S., 2021. ROI of AI: effectiveness and measurement. *Int. J. Eng. Res. Technol. (IJERT)* **10** (5).

Panizza, R., Allam, A., Kasliwal, A., Nik-Bakht, M., 2024. Labeling Construction, Renovation, and Demolition Waste through Segment Anything Model (SAM). In: *Construction Research Congress 2024*, pp. 279–288.

Pickin, J., Wardle, C., O'Farrell, K., Stovell, L., Nyunt, P., Guazzo, S., Lin, Y., Caggiati-Shortell, G., Chakma, P., Edwards, C., Lindley, B., Latimer, G., Downes, J., Axiö, I., 2022. National waste report 2022.

Prasad, V., Arashpour, M., 2024. Optimally leveraging depth features to enhance segmentation of recyclables from cluttered construction and demolition waste streams. *J. Environ. Manage.* **354**, 120313.

Pučnik, R., Dokl, M., Fan, Y.V., Vučanović, A., Novak Pintarić, Z., Aviso, K.B., Tan, R.R., Pahor, B., Kravanja, Z., Ćućek, L., 2024. A waste separation system based on sensor technology and deep learning: a simple approach applied to a case study of plastic packaging waste. *J. Clean. Prod.* **450**, 141762.

Qiao, X., Poggi, M., Deng, P., Wei, H., Ge, C., Mattoccia, S., 2024. RGB guided ToF imaging system: a survey of deep learning-based methods. *Int. J. Comput. Vis.* **132** (11), 4954–4991.

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I., 2021. Learning transferable visual models from natural language supervision. arXiv.org.

Rajotte, J.-F., Bergen, R., Buckeridge, D.L., El Emam, K., Ng, R., Strome, E., 2022. Synthetic data as an enabler for machine learning applications in medicine. *iScience* **25** (11), 105331.

Rasmussen, I., Kvalsvik, S., Andersen, P.-A., Aune, T.N., Hagen, D., 2022. Development of a novel object detection system based on synthetic data generated from Unreal Game Engine. *Appl. Sci.*

Saka, A., Taiwo, R., Saka, N., Salami, B.A., Ajayi, S., Akande, K., Kazemi, H., 2024. GPT models in construction industry: opportunities, limitations, and a use case validation. *Dev. Built Environ.* **17**, 100300.

Schalkwyk, J., Silver, D., Pitler, E., Molloy, J., Isard, M., Moreira, E., Tucker, G., Krikun, M., Kazemi, M., Culp, L., Luan, Y., Lu, H., Garrette, D., Ale Jakse, H., Thanumalayan Sankaranarayana, P., Tao, C., Chen, M., Rao, A., Olszewska, K., Attaluri, N., Kannan, A., Bradbury, J., Godbole, V., Damoc, B., Agrawal, S., Ozturel, A., Swanson, C., Petrova, D., Brahma, S., Wang, L., Mincu, D., Cankara, Z., Neyshabur, B., Yang, F., Soparkar, K., Jhakra, P., Aroyo, L., Yang, Z., Besley, J., Dozat, T., Hu, H., Miao, Y., Elhawaty, M., Mu, J., Muir, C., Sholto, D., Baumli, K., Pavetic, F., Campos, V., Thacker, P., Ünlü, Ç., Joost van, A., Mandhane, A., Tung, M., Huot, F., Wu, S., Mohiuddin, A., Vieillard, N., Potluri, S., Prendki, J., Vijayakumar, A., Lam Nguyen, T., Paduraru, C., Kurylowicz, P., Choo, K., Feng, Z., Latkar, T., Reid, D., Finchelstein, D., Filonov, E., Chung, S., Fischer, F., Chu-Cheng, L., Liu, J., Xia, J., Polozov, A., Zheng, Z., Rosen, E., Elkind, C., Kemp, R., Ko, Y., Niu, N., Strudel, R., Tafti, P., Li, T., Parisi, A., Shroff, P., Bridson, D., Odom, S., Loher, L., Yadlowsky, S., Qureshi, H., Ainslie, J., Koukoumidis, E., Parashar, S., Mehta, H., Ruder, S., Nasr, M., Dean, J., Vinyals, O., 2023. Gemini: a Family of highly capable multimodal models. arXiv.org.

Shahab, S., Anjum, M., Umar, M.S., 2022. Deep Learning applications in solid waste management: a deep literature review. *Int. J. Adv. Comput. Sci. Appl.* **13** (3).

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *J. Big. Data* **6** (1), 1–48.

Shutterstock, 2024. Turbosquid. <https://www.turbosquid.com/>. (Accessed 25th March 2024).

Sirimewan, D., Harandi, M., Peiris, H., Arashpour, M., 2024. Semi-supervised segmentation for construction and demolition waste recognition in-the-wild: adversarial dual-view networks. *Resour. Conserv. Recycl.* **202**.

Sketchfab Inc, 2024. Sketchfab. <https://sketchfab.com/>. (Accessed 25th March 2024).

Smith, A., 2023. How much does it cost to build artificial intelligence software In 2024? <https://web-and-mobile-development.medium.com/how-much-does-it-cost-to-build-d-artificial-intelligence-software-in-2024-fcc2fe11fde5>. (Accessed 1 May 2024).

Song, L., Zhao, H., Ma, Z., Song, Q., 2022. A new method of construction waste classification based on two-level fusion. *PLoS. One* **17** (12) e0279472-e0279472.

Staub, C., 2024. Recycling facilities reported few fires in 2023. <https://resource-recycling.com/recycling/2024/03/19/recycling-facilities-reported-fewer-fires-in-2023/>. (Accessed 12th January 2025).

Sunwoo, H., Choi, W., Na, S., Kim, C., Heo, S., 2022. Comparison of the performance of artificial intelligence models depending on the labelled image by different user levels. *Appl. Sci.* **12** (6).

Tahmasbian, I., Morgan, N.K., Hosseini Bai, S., Dunlop, M.W., Moss, A.F., 2021. Comparison of hyperspectral imaging and near-infrared spectroscopy to determine nitrogen and carbon concentrations in wheat. *Remote Sens. (Basel)*.

ul Hassan, C.A., Khan, M.S., Irfan, R., Iqbal, J., Hussain, S., Sajid Ullah, S., Alroobaee, R., Umar, F., Abdelaziz, M., Mohamed, A., 2022. Optimizing deep learning model for software cost estimation using hybrid meta-heuristic algorithmic approach. *Comput. Intell. Neurosci.* **2022**, 1–20.

Valvano, G., Agostino, A., De Magistris, G., Graziano, A., Veneri, G., 2024. Controllable Image Synthesis of Industrial Data Using Stable Diffusion. Cornell University arXiv.

VentureBeat, 2019. Why do 87 % of data science projects never make it into production? <https://venturebeat.com/ai/why-do-87-of-data-science-projects-never-make-it-into-production/>. (Accessed 1 May 2024).

Wada, K., 2018. Labelme: image polygonal annotation with python. GitHub, GitHub repository.

Wang, X., Han, W., Mo, S., Cai, T., Gong, Y., Li, Y., Zhu, Z., 2023. Transformer-based automated segmentation of recycling materials for semantic understanding in construction. *Autom. Constr.* **154**, 104983.

Wang, Z., Li, H., Yang, X., 2020. Vision-based robotic system for on-site construction and demolition waste sorting and recycling. *J. Build. Eng.* **32**.

Wang, Z., Li, H., Zhang, X., 2019. Construction waste recycling robot for nails and screws: computer vision technology and neural network approach. *Autom. Constr.* **97**, 220–228.

Wolff, S., Wittmaier, M., Vögele, T., Runge, T., Keppner, M., Tiedemann, T., 2021. Concept of a robotic system for autonomous coarse waste recycling. In: *Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics*, pp. 493–500.

Xiao, W., Yang, J., Fang, H., Zhuang, J., Ku, Y., 2019a. Development of online classification system for construction waste based on industrial camera and hyperspectral camera. *PLoS. One* **14** (1) e0208706-e0208706.

Xiao, W., Yang, J., Fang, H., Zhuang, J., Ku, Y., 2019b. A robust classification algorithm for separation of construction waste using NIR hyperspectral system. *Waste Manage* **90**, 1–9.

Xiao, W., Yang, J., Fang, H., Zhuang, J., Ku, Y., 2020. Classifying construction and demolition waste by combining spatial and spectral features. *Proc. Inst. Civil Eng. - Waste Resour. Manag.* **173** (3), 79–90.

Xie, Z., Guan, B., Jiang, W., Yi, M., Ding, Y., Lu, H., Zhang, L., 2024. PA-SAM: prompt adapter SAM for high-quality image segmentation. arXiv.org.

Yang, M., Thung, G., 2016. Classification of trash for recyclability status. pp. 1–6.

Yang, X., Li, Y., Chen, Q., 2024. Automated image-based fire detection and alarm system using edge computing and cloud-based platform. *Internet of things (Amsterdam, Online)* **28**, 101402.

Yu, Y., Wang, C., Fu, Q., Kou, R., Huang, F., Yang, B., Yang, T., Gao, M., 2023. Techniques and challenges of image segmentation: a review. *Electronics. (Basel)* **12** (5), 1199.

Zhang, C., Liu, L., Cui, Y., Huang, G., Lin, W., Yang, Y., Hu, Y., 2023. A comprehensive survey on segment anything model for vision and beyond. arXiv.org.

Zhou, Q., Liu, H., Qiu, Y., Zheng, W., 2022. Object detection for construction waste based on an improved YOLOv5 model. *Sustainability* **15** (1).

Zhu, Z., Brilakis, I., 2010. Parameter optimization for automated concrete detection in image data. *Autom. Constr.* **19** (7), 944–953.