ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/384876120

Construction and Demolition Waste Management and Artificial Intelligence – A Systematic Review

Article *in* Revista de Gestão Social e Ambiental · October 2024

DOI: 10.24857/rgsa.v18n10-197

CITATION
1

READS
238

3 authors:



Carolina de Melo Nunes Lopes
Federal University of Ouro Preto
13 PUBLICATIONS **16** CITATIONS
SEE PROFILE



Alexandre Cury
Universidade Federal de Juiz de Fora
99 PUBLICATIONS **1,128** CITATIONS
SEE PROFILE



Julia Castro Mendes
Universidade Federal de Juiz de Fora
107 PUBLICATIONS **1,270** CITATIONS
SEE PROFILE

All content following this page was uploaded by Carolina de Melo Nunes Lopes on 14 October 2024.

The user has requested enhancement of the downloaded file.

RGSA – Revista de Gestão Social e Ambiental
ISSN: 1981-982X
Submission date: 07/05/2024
Acceptance date: 09/06/2024
DOI: https://doi.org/10.24857/rgsa.v18n10-197
Organization: Interinstitutional Scientific Committee
Chief Editor: Ana Carolina Messias de Souza Ferreira da Costa
Assessment: Double Blind Review pelo SEER/OJS

CONSTRUCTION AND DEMOLITION WASTE MANAGEMENT AND ARTIFICIAL INTELLIGENCE – A SYSTEMATIC REVIEW

Carolina de Melo Nunes Lopes [1]
Alexandre Abrahão Cury [2]
Júlia Castro Mendes [3]

ABSTRACT

Objective: The present study aims to identify how Artificial Intelligence (AI) techniques are being employed in Construction and Demolition Waste (CDW) management through a systematic review.

Method: The methodology adopted for this research comprises a systematic review of 37 papers focusing on AI applications in CDW management. The search for scientific papers was conducted using the Scopus and Web of Science databases. We extracted data on the year of publication, country of the first author, main models used, performance metrics, dataset size, dataset source, and key gaps and limitations identified by the authors.

Results: The incorporation of AI in CDW management is relatively recent, with a significant rise in publications noted from 2021. Most papers have a Chinese first author, and the most frequently used ML algorithms were the RF, GBM, CNN, and ANN. Three primary research topics were identified: CDW Generation Prediction, CDW Classification and Identification, and Decision Support Systems and Logistics. Despite the advancements, challenges such as insufficient data for model training and limited interpretability of AI models persist.

Research Implications: The research provides insights into how AI can enhance recycling processes and supports the integration of AI models with Decision Support Systems and logistics to promote sustainability and circularity in construction practices.

Originality/Value: The research underscores the potential for AI to transform CDW management practices, offering valuable perspectives for both academic research and practical applications.

Keywords: CDW, Construction, Artificial Intelligence, Machine Learning, Management.

GESTÃO DE RESÍDUOS DE CONSTRUÇÃO E DEMOLIÇÃO E INTELIGÊNCIA ARTIFICIAL – UMA REVISÃO SISTEMÁTICA

RESUMO

Objetivo: O presente estudo objetiva identificar como as técnicas de Inteligência Artificial (IA) estão sendo empregadas na gestão de resíduos de construção e demolição (RCD) por meio de uma revisão sistemática.

Método: A metodologia adotada para esta pesquisa compreende uma revisão sistemática de 37 artigos focados nas aplicações de IA na gestão de RCD. A busca por artigos científicos foi realizada nas bases de dados Scopus e Web of Science. Foram extraídos dados sobre o ano de publicação, país de origem do primeiro autor, principais modelos utilizados, métricas de desempenho, tamanho do conjunto de dados, fonte dos dados e principais lacunas e limitações identificadas pelos autores.

[1] Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais, Brazil.
E-mail: carolina.nunes@aluno.ufop.edu.br Orcid: https://orcid.org/0000-0003-2634-5536
[2] Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brazil. E-mail: alexandre.cury@ufjf.br
Orcid: https://orcid.org/0000-0002-8860-1286
[3] Universidade Federal de Juiz de Fora, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais, Brazil.
E-mail: juliacastro.mendes@ufjf.br Orcid: https://orcid.org/0000-0002-6323-5355

Resultados: A incorporação da IA na gestão de RCD é relativamente recente, com um aumento significativo nas publicações a partir de 2021. A maioria dos artigos tem um primeiro autor chinês, e os algoritmos de aprendizado de máquina mais utilizados foram RF, GBM, CNN e ANN. Três tópicos principais de pesquisa foram identificados: Predição de Geração de RCD, Classificação e Identificação de RCD e Sistemas de Suporte à Decisão e Logística. Apesar dos avanços, persistem desafios como dados insuficientes para treinamento dos modelos e a limitada interpretabilidade dos modelos de IA.

Implicações da Pesquisa: A pesquisa revela como a IA pode aprimorar os processos de reciclagem e apoia a integração dos modelos de IA com Sistemas de Suporte à Decisão e logística para promover a sustentabilidade e circularidade na construção civil.

Originalidade/Valor: A pesquisa ressalta o potencial da IA para transformar as práticas de gerenciamento de RCD, oferecendo perspectivas valiosas tanto para pesquisas acadêmicas quanto para aplicações práticas.

Palavras-chave: RCD, Construção, Inteligência Artificial, Aprendizado de Máquina, Gestão.

GESTIÓN DE RESIDUOS DE CONSTRUCCIÓN Y DEMOLICIÓN E INTELIGENCIA ARTIFICIAL – UNA REVISIÓN SISTEMÁTICA

RESUMEN

Objetivo: El presente estudio tiene como objetivo identificar cómo se están empleando las técnicas de Inteligencia Artificial (IA) en la gestión de residuos de construcción y demolición (RCD) a través de una revisión sistemática.

Método: La metodología adoptada para esta investigación comprende una revisión sistemática de 37 artículos centrados en las aplicaciones de la IA en la gestión de RCD. La búsqueda de artículos científicos se realizó en las bases de datos Scopus y Web of Science. Se extrajeron datos sobre el año de publicación, el país del primer autor, los principales modelos utilizados, las métricas de rendimiento, el tamaño del conjunto de datos, la fuente del conjunto de datos y las principales lagunas y limitaciones identificadas por los autores.

Resultados: La incorporación de la IA en la gestión de RCD es relativamente reciente, con un aumento significativo en las publicaciones a partir de 2021. La mayoría de los artículos tienen un primer autor chino, y los algoritmos de aprendizaje automático más utilizados fueron RF, GBM, CNN y ANN. Se identificaron tres temas principales de investigación: Predicción de Generación de RCD, Clasificación e Identificación de RCD y Sistemas de Apoyo a la Decisión y Logística. A pesar de los avances, persisten desafíos como la falta de datos para el entrenamiento de modelos y la limitada interpretabilidad de los modelos de IA.

Implicaciones de la Investigación: La investigación revela cómo la IA puede mejorar los procesos de reciclaje y apoya la integración de modelos de IA con Sistemas de Apoyo a la Decisión y logística para promover la sostenibilidad y circularidad en las prácticas de construcción.

Originalidad/Valor: La investigación destaca el potencial de la IA para transformar las prácticas de gestión de RCD, ofreciendo perspectivas valiosas tanto para la investigación académica como para aplicaciones prácticas.

Palabras clave: RCD, Construcción, Inteligencia Artificial, Aprendizaje Automático, Gestión.

RGSA adota a Licença de Atribuição CC BY do Creative Commons (https://creativecommons.org/licenses/by/4.0/).

1 INTRODUCTION

The construction industry plays an important role in the economic and social development of a community. At the same time, this sector is responsible for significant

environmental impacts. Around 30 to 40% of the total solid waste globally generated originates from the construction and demolition of buildings . Construction and demolition waste (CDW) occupy significant landfill space and, if not disposed of properly, contribute to soil sedimentation and erosion (Farinha *et al.*, 2019). Moreover, the construction industry is a major source of natural resource extraction and greenhouse gas emissions (Carvalho *et al.*, 2023).

Recent years saw an increase in research related to the management and recycling of CDW with the aim of mitigating its environmental impacts (Li *et al.*, 2020). However, although several studies demonstrate the potential use of CDW, particularly as recycled aggregate, its utilization rate remains low in many regions (Oluleye *et al.*, 2022; Souza *et al.*, 2022). This phenomenon can be attributed to multiple factors, including variations in waste composition, limited knowledge of its performance (due to the high variability), inadequate recycling infrastructure, lack of tax incentives, high recycling costs, and absence of a well-established market (Oluleye *et al.*, 2022).

Concrete produced with recycled aggregate has, in general, some disadvantages compared to those produced with conventional aggregates, such as higher porosity and lower compressive strength (Kannan *et al.*, 2021). Recent studies indicate that these disadvantages can be mitigated using chemical and mineral admixtures (Kannan *et al.*, 2021). Still, the transition of the construction industry to a circular supply chain faces challenges, especially behavioral and fiscal challenges. There is still limited societal acceptance regarding recycled construction materials, and there is a lack of investment and incentives favoring its use (Cosenza *et al.*, 2020).

Recently, Artificial Intelligence (AI) has significantly contributed to the optimization and automation of various sectors in society (Abioye *et al.*, 2021). AI can be defined as computational systems capable of performing tasks that typically require human intelligence, such as interpreting events, analyzing trends, and supporting and automating decisions (Abioye *et al.*, 2021). Machine Learning (ML) is a subset of AI that uses computational methods to recognize patterns in a dataset and then uses these discovered patterns to predict future outcomes (Abioye *et al.*, 2021). ML algorithms can be supervised or unsupervised. In supervised learning, the algorithm is trained using a labeled dataset (Mittelstadt *et al.*, 2016). That is, the algorithm is fed with input and output data to learn the relationship between them. In unsupervised learning, the algorithm deals with unlabeled data with the goal of finding patterns among the data and grouping them in a meaningful way in an automated manner (Mittelstadt *et al.*, 2016).

Deep learning (DL) is a subfield of ML that focuses on training artificial neural networks to perform tasks and make decisions in a similar manner to the human brain . It uses multiple layers of interconnected artificial neurons, which allow it to automatically learn and represent complex patterns and features from large datasets . DL has gained prominence due to its ability to handle unstructured data such as images, text, and speech, making it an important technology for image recognition and natural language processing tasks (Na *et al.*, 2022; Davis *et al.*, 2021).

In the construction industry, AI holds the potential to enhance productivity, safety, and quality (Abioye *et al.*, 2021). AI can automate the supervision of activities and assist on risk management, automatically checking the use of protective equipment, identify errors and opportunities in project planning, contribute to structure modeling and monitoring, and predict the performance of construction materials, such as concrete, optimizing resources, minimizing waste, and improving quality (Abioye *et al.*, 2021; Ensafi *et al.*, 2022; Penido *et al.*, 2022).

Regarding CDW management, AI can analyze large volumes of data concerning construction materials, their characteristics, and recycling potential (Lu *et al.*, 2021; Davis *et al.*, 2021). Davis *et al.* and Na *et al.* , for example, successfully developed CDW image recognition systems with the potential to streamline the waste separation process and reduce recycling costs. AI can also be used to predict the amount of waste generated in building demolition, contributing to decision-making regarding waste recycling and disposal (Cha *et al.*, 2022). And when combined with tools like GIS (Geographic Information System), AI can determine optimal landfill areas, identify illegal disposal sites, and optimize waste logistics (Mao *et al.*, 2022; Bi *et al.*, 2022; Paz *et al.*, 2018).

Important reviews regarding AI applications in the construction industry have been conduct in recent years. Abioye *et al.* reviewed AI applications in the construction industry from 1960 to 2020. Oluleye *et al.* reviewed papers related to circular economy practices in CDW management and observed a lack of AI-related studies in this field. On a further study, Oluleye *et al.* conducted a systematic review of advancements in AI and circular economy within the construction industry. These studies have provided valuable insights to the field, however, none of them specifically addressed the use of AI techniques in CDW management. Despite the recent publishing of these papers, the rapid advancement of AI in recent years suggests that research on this topic should be continuously updated. For instance, the review of Oluleye *et al.* have a cutoff date of 2022, whereas in our study, 9 out of the 37 considered articles (24%) are from 2023.

In this context, this study aims to systematically review how AI techniques are employed in CDW management, guided by three research questions: (1) What are the main

current research topics on CDW management with AI? (2) Which are the main algorithms used in CDW management models? (3) What are the main challenges and limitations in developing CDW management models?

We did not find any review addressing AI techniques specifically for CDW management. Thus, this study aims to enhance the understanding of the current state of this field, discuss challenges in model development, and encourage researchers and professionals to develop AI models for CDW management. Additionally, our findings provide a knowledge base for researchers, policymakers, and professionals, enabling them to direct resources to the identified gaps.

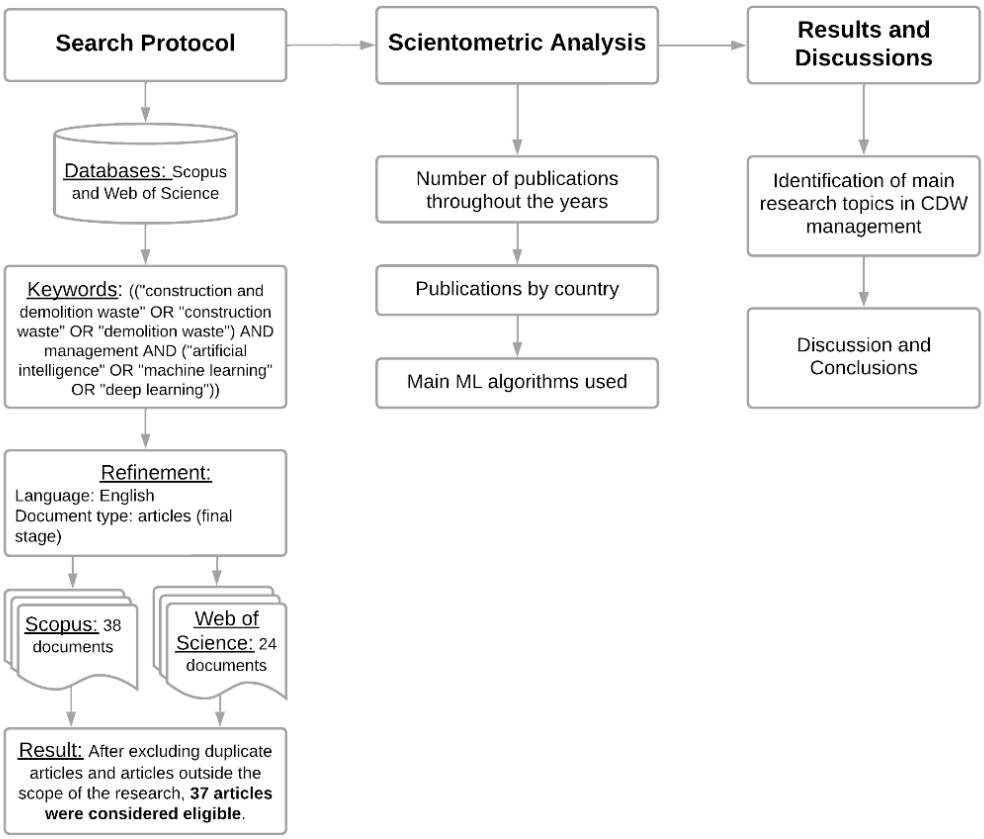
2 METHODS

This study comprises a systematic literature review of AI techniques that are being employed in CDW management. Figure 1 presents an overview of the research method. The search for scientific papers was conducted using the Scopus and Web of Science databases. Web of Science and Scopus are suitable platforms for bibliographic reviews (Wuni *et al.*, 2019) and are the most popular among the existing databases (Oluleye *et al.*, 2022). The search on Scopus was conducted on August 17, 2023, and on the Web of Science on September 26, 2023.

The search was carried out using the functionality "title/abstract/keyword" (Scopus) and “topic” (Web of Science), and the string search used was: [("construction and demolition waste" OR "construction waste" OR "demolition waste") AND management AND ("artificial intelligence" OR "machine learning" OR "deep learning")]. The search was conducted in English, as most journals indexed in these platforms are in English, and our study aimed for global coverage. Regarding the "document type", only articles in the final publication stage were considered. Review articles were not included because we sought to analyze original publications on the topic. Conference papers were also not considered since most conference proceedings are already written as journal articles (Abioye *et al.*, 2021).

Figure 1

Overview of the research method



There was no limitation on the publication year to analyze how research on the topic has evolved over the years. Nevertheless, the resulting articles are recent and date from 2011 to 2023. In total, 38 studies met all the restrictions in the Scopus database and 24 in the Web of Science. Only one paper among those resulting from the Web of Science search was not among the Scopus results. Therefore, 39 papers resulted from the search in both platforms.

The term "management" was used in the string search to obtain studies that specifically employed AI in CDW management. However, after fully reading the 39 articles, we excluded two that focused on analyzing the mechanical properties of construction materials with CDW, not management. Therefore, 37 articles were considered for this project. This quantity is in line with other valuable review papers, such as Oluleye *et al.* . To mitigate bias in the selection of sources, we limited the analyzed papers to the string search and the adopted criteria. The scientific merit or citation metrics of the selected articles were not considered.

To address the research questions of this review, we analyzed the 37 eligible papers, extracting data on the year, country of the first author, main models used, performance metrics, dataset size, dataset source, and key gaps and limitations identified by the authors. This information enabled us to perform a scientometric analysis, examining publication trends over

the years, leading countries, and commonly used ML algorithms. We then discussed the primary research topics utilizing AI in CDW management and main challenges addressed by the authors in model development. Finally, we identified critical gaps that require further attention.

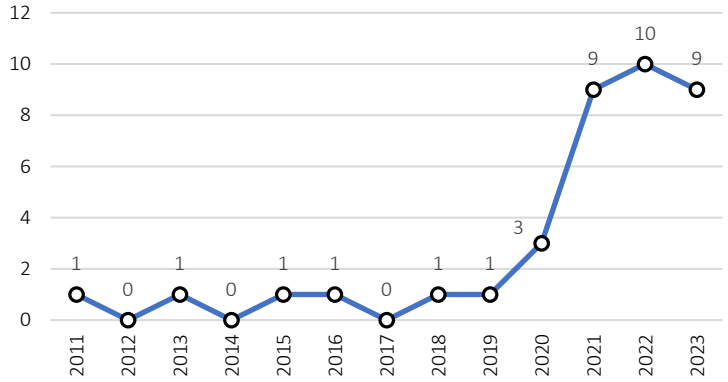
3 RESULTS

3.1 SCIENTOMETRIC ANALYSIS

Initially, we conducted a scientometric analysis, a valuable tool for researchers, institutions, and policymakers to gain insights into the structure, growth, impact, and trends within the field. Figure 2 shows the publication count over these years, revealing that the use of AI in CDW management is relatively recent, and a significant increase in publications occurred from 2021. Notably, the articles were collected in August and September 2023, therefore, it is likely that the number of 2023 publications will surpass those of 2022.

Figure 2

Number of publications throughout the years



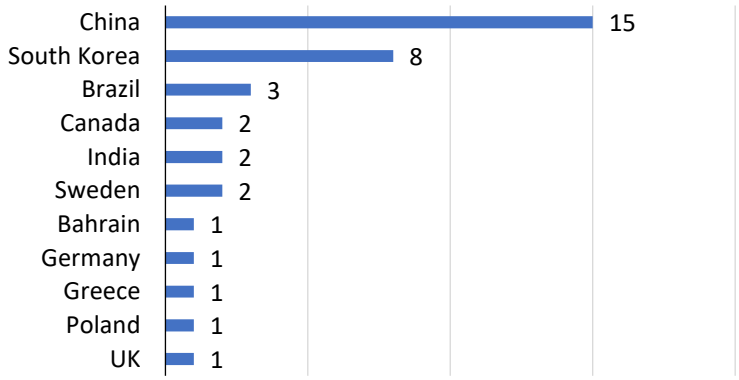
The recent increase in publications related to AI applications was also observed by other authors. Abioye *et al.* found that over 60% of papers on AI in the construction industry were published between 2010 and 2020. Meanwhile, Oluleye *et al.* noted that most articles published on circular economy related to CDW are still very recent, with a significant increase starting from 2020. These finds highlight the recent emergence of both AI-based tools and circular economy strategies in the construction industry.

Regarding the origin of the publications, 40.5% of the articles have a Chinese first author, as shown in Figure 3. This aligns with previous research highlighting China's prominent

role in AI research (Abdallah *et al.*, 2020; Tapeh & Naser, 2023). Abdallah *et al.* noted that China leads in publishing papers on AI applications in solid waste management. According to Tapeh and Naser , China is one of the main countries that publishes papers on AI techniques employed in Structural Engineering, alongside the U.S., Iran, and the U.K.

Figure 3

Publications by country

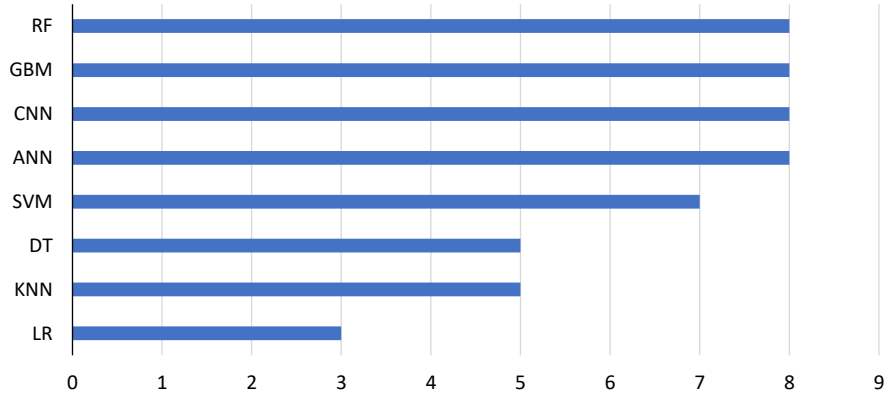


Many Chinese articles in this study utilized data from the Hong Kong Environmental Protection Department (HKEPD). According to Lu *et al.* , the HKEPD introduced the Construction Waste Disposal Charge Scheme in 2005, requiring projects exceeding HK$ 1 million to have a dedicated waste disposal account. Most Hong Kong renovation projects fall under this scheme, maintaining detailed records of waste disposal. These records facilitate the development of AI research in Hong Kong, as data scarcity is one of the primary challenges in advancing AI research, as will be discussed further on. Therefore, data availability, together with a strong research community in the field, are probably contributing factors to China's leadership position.

Figure 4 illustrates the main ML algorithms used in the articles - algorithms used only once were not considered in this analysis. The most frequently used ones were Random Forest – RF, Gradient Boosting Machine – GBM (other boost algorithms were considered in this class, such as XGBoost), Convolutional Neural Network – CNN, and Artificial Neural Networks – ANN. Figure 4 shows that researchers are exploring a wide variety of ML models to address CDW management issues, adapting their algorithm choices based on the specific needs and characteristics of each study.

Figure 4

Main ML algorithms used in the CDW management studies



Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree (DT), and Linear Regression (LR) are supervised learning algorithms that are used for classification and regression tasks. The SVM works by finding a hyperplane that maximally separates different classes of data points, making it a powerful tool for solving both linear and non-linear problems (Nguyen *et al.*, 2023). The KNN classifies a data point based on most classes among its "k" nearest neighbors in feature space . The DT is a tree-like structure with a root node at the top, internal nodes in the middle, and leaf nodes at the bottom (Gulghane *et al.*, 2023). Each node represents a decision based on features, and the tree divides data into subsets, making classification or regression decisions (Gulghane *et al.*, 2023). Finally, the LR is used for modeling the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data (Cha *et al.*, 2023a).

The Gradient Boosting Machine (GBM) and the Random Forest (RF) are considered ensemble learning algorithms, which combine multiple individual ML models to improve predictive performance and robustness. GBM is a boosting method that builds a predictive model by training a sequence of weak learners, typically DTs, where each new tree corrects the errors made by the previous ones (Cha *et al.*, 2023b). RF is a bagging method that constructs multiple independent decision trees by randomly selecting subsets of the data and features for each tree (Cha *et al.*, 2023b). The final prediction is determined by the average predictions from all the DTs (Cha *et al.*, 2023b).

Artificial Neural Network (ANN) can be applied either as ML or DL, while Convolutional Neural Network (CNN) are considered DL algorithms. ANNs are designed to simulate the interconnected neurons in the human brain, consisting of input, hidden, and output layers of nodes . CNNs are a specialized type of ANN tailored for image processing and

computer vision tasks (Davis *et al.*, 2021). The CNNs incorporate convolutional layers to automatically learn and detect patterns and features in images (Davis *et al.*, 2021).

The main algorithms used in CDW management models in this study are consistent with other reviews. Nguyen *et al.* , Abdallah *et al.* , and Tapeh and Naser , identified the ANN as the most common algorithm for predicting the mechanical properties of concrete with recycled aggregates, in solid waste management, and in structural engineering, respectively. In contrast, a recent review by Oluleye *et al.* on the advancements in AI and circular economy identified the CNN as the most used algorithm. ANNs' popularity may be related to their versatility and minimal data processing requirements, allowing for integration into hybrid models (Nguyen *et al.*, 2023). Meanwhile, CNNs have gained attention due to advancements in image classification and recognition systems (Davis *et al.*, 2021; Na *et al.*, 2022).

3.2 MACHINE LEARNING METRICS

To evaluate model performance, it is essential to understand key ML metrics. One commonly used metric in regression analysis is R^2, the coefficient of determination. R^2 represents the goodness of fit of a regression model (Lu *et al.*, 2023). It normally ranges from 0 to 1, with higher values indicating a better fit of the model to the data (Lu *et al.*, 2023). R, or Pearson's correlation coefficient, quantifies the linear relationship between two continuous variables. It ranges from -1 to 1, where -1 or 1 indicates a perfect linear correlation (negative or positive) and 0 indicates no linear correlation (Abdallah *et al.*, 2020).

The Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are metrics that lower values indicate better performance. MAE measures the average absolute difference between the predicted values and the actual values (Huang *et al.*, 2021). MSE measures the average squared difference between the predicted values and the actual values, while RMSE is the square root of the MSE and is often used in the same context . MAPE calculates the average percentage difference between the predicted and actual values (Huang *et al.*, 2021).

Accuracy is a metric used to measure the proportion of correctly classified instances in classification models, calculated as correct predictions divided by the total predictions. Recall evaluates the accuracy of identified positives, expressed as true positives divided by the sum of true positives and false negatives (Tapeh & Naser, 2023; Lin *et al.*, 2023). Precision quantifies the accuracy of positive predictions made by the model and is calculated as true positives divided by the sum of true positives and false positives (Lin *et al.*, 2023). The F1 score, a

harmonic mean of precision and recall, provides a single value that summarizes the model's ability to make accurate positive predictions while capturing most of the actual positives (Yang *et al.*, 2021). These metrics are often represented as percentages, with higher values indicating better performance .

The Intersection over Union (IoU) is a metric used to evaluate the performance of image segmentation algorithms, it measures the proportion of overlap between the predicted and true positive regions (Dong *et al.*, 2022). A higher IoU indicates a better agreement between the model's predictions and the actual ground truth. IoU values range from 0 to 1, where 0 represents no overlap, and 1 represents perfect overlap between the predicted and true positive regions.

Kappa coefficient measures the agreement between two raters (or models) beyond what would be expected by chance. It considers both the observed agreement and the expected agreement between the raters or models (Zhao *et al.*, 2022). Higher values usually indicate better agreement beyond what would be expected by chance. The area under the Receiver Operating Characteristic (ROC) curve – AUC, summarizes the performance of a binary classification model across all possible classification thresholds . AUC ranges from 0 to 1, where 0.5 indicates a model that performs no better than random chance, and 1 indicates a perfect model.

3.3 MAIN RESEARCH TOPICS

Throughout the reading of the 37 articles, we identified three primary research topics related to CDW management coupled with AI: CDW Generation Prediction (20 papers), CDW Classification and Identification (11 papers), and Decision Support Systems (DSS) and Logistics (6 papers). These research topics will be discussed individually in the following sections. They were categorized as follows:

a) CDW Generation Prediction: this area involves the development of models for predicting the generation of CDW;
b) CDW Classification and Identification: this area focuses on the use of AI for identifying, sorting, and classifying different types of CDW. This includes methods to distinguish hazardous waste and illegal disposal dumps;
c) Decision Support Systems and Logistics: papers in this category focus on the implementation of AI-driven decision support systems (DSS) and logistics solutions in the context of CDW management.

3.3.1 CDW generation prediction

The prediction of the volume and/or mass of CDW generated in a given construction site, city, or year is extremely useful for improving its management (Oluleye *et al.*, 2022). This knowledge enables the planning of waste management facilities, such as landfills or recycling plants, analysis of the feasibility of sharing landfills between neighboring municipalities, the quantification of carbon emissions, and implementation of other waste management strategies and policies (Lu *et al.*, 2021; Coskuner *et al.*, 2021; Cha *et al.*, 2023a).

According to Oluleye *et al.* , AI is a promising tool in forecasting CDW generation as it can analyze vast and complex datasets and provide reliable projections to assist in CDW management strategies (Oluleye *et al.*, 2022). Accordingly, 20 of the 37 papers (54%) addressed the CDW generation prediction. Abdallah *et al.* also observed in their review that most research regarding AI applications in solid waste management are about generation prediction. The papers regarding CDW Generation Prediction, and their main characteristics are shown in Table 1.

Most papers on predicting CDW generation utilized ML models, with only the three oldest relying on big data, statistical analysis, and web-based systems (Li & Zhang, 2013; Lu *et al.*, 2015; Paz & Lafayette, 2016). For instance, Li and Zhang developed a web-based system for CDW estimation requiring accurate data input for precise calculations. Lu *et al.* used big data and statical analysis to assess CDW generation rates in Hong Kong. Paz and Lafayette developed a software for estimating CDW based on project data, such as area, number of floors, and construction duration in Recife, Brazil.

Table 1

Publications on CDW Generation Prediction

Title	Source and Year	ML Model	R^2	MAE	RMSE	Others	Dataset size	Data source	Main Limitations
Quantification analysis and prediction model for residential building construction waste using machine learning technique	Gulghan*et al.* (2023)	KNN			0.49	Accuracy = 88.5%	134 projects	Collected by the authors	Small dataset and regional applicability
		DT			0.50	Accuracy = 88.3%			
Deep Learning-Based Swot Analysis in Construction and Demolition Waste Management	Rema & Nalanth (2023)	CNN	0.98	18.10			2280	India demolition sector	Not mentioned
Development of Machine Learning Model for Prediction of Demolition Waste Generation Rate of Buildings in Redevelopment Areas	Cha *et al.* (2023a)	ANN ReLu	0.90	71.93	100.80		160 buildings	Surveyed from buildings prior to demolition	There is room for improving results using hybrid ML models
		ANN Logistic	0.88	74.01					
		SVM polynomial	0.89	72.75	106.14				
		LR	> 0.80						
		RF	> 0.80						
		KNN	0.63 - 0.77						
Developing an Optimal Ensemble Model to Estimate Building Demolition Waste Generation Rate	Cha *et al.* (2023b)	RF	0.62		253.73	R = 0.79	782 buildings	Collected by the authors	Small dataset; low R^2 values
		ET	0.59		269.40	R = 0.77			
		GBM	0.61		253.09	R = 0.78			
		XGBoost	0.56		281.26	R = 0.75			
Performance Improvement of Machine Learning Model Using Autoencoder to Predict Demolition Waste Generation Rate	Cha *et al.* (2023c)	AE ANN	0.68	182.10	230.80	R = 0.83	782 buildings	Collected by the authors	Other algorithms and datasets should be investigated in the future
		AE SVM	0.11	302.20	384.30	R = 0.57			
		RF	0.62	197.04	249.60	R = 0.79			
Developing a Prediction Model of Demolition-Waste Generation-Rate via Principal Component Analysis	Cha *et al.* (2023d)	PCA LR	0.84	70.13	109.61		160 buildings	Surveyed from buildings prior to demolition	Small dataset; development of the model is costly; do not consider different types of CDW
		PCA KNN	0.90	69.01	102.12				
		DT	0.87	71.43	113.98				
A machine learning regression approach for pre-renovation construction waste auditing	Lu *et al.* (2023)	Tree regression ensemble	0.83	79.69	141.52		351 projects	HKEPD and Google Maps	Regional applicability; do not consider different types of CDW

Title	Source and Year	ML Model	R^2	MAE	RMSE	Others	Dataset size	Data source	Main Limitations
A hybrid machine-learning model for predicting the waste generation rate of building demolition projects	Cha *et al.* (2022)	ANN	0.46	356.70	316.19	R = 0.68	784 buildings	Collected by the authors	Small dataset; only compared two ML models
		CATPCA – ANN	0.46	236.02	300.31	R = 0.72			
		SVMR	0.01	1020.30	1116.26	R = 0.08			
		CATPCA – SVMR	0.59	202.23	261.45	R = 0.77			
Handling missing data for construction waste management: machine learning based on aggregated waste generation behaviors	Yang *et al.* (2021)	DT				F1 = 0.82	2,451 projects	HKEPD	Limited to new building and renovation projects; data imbalance; some models are not easily interpreted
		KNN				F1 = 0.81			
		SVM				F1 = 0.82			
		AdaBoost				F1 = 0.87			
Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables	Cha *et al.* (2021b)	RF	0.34 - 0.89	193.70		R = 0.73 - 0.95	690 buildings	Collected by the authors	Small dataset; there is room for improving feature selection
		GBM	0.22 - 0.84	206.20		R = 0.71 - 0.92			
Estimating construction waste generation in the Greater Bay Area, China using machine learning	Lu *et al.* (2021)	MLR	0.78				43 sets of socioeconomic, construction related and CDW generation data	National Bureau of Statistics of China	Small dataset
		DT	0.76						
		GM	0.98						
		ANN	0.91						
A method to improve the performance of support vector machine regression model for predicting demolition waste generation using categorical principal components analysis	Cha *et al.* (2021a)	SVMR - CATPCA	0.59	202.23	261.45	R = 0.77	690 buildings	Collected by the authors	Small dataset; the features employed had difficult in reflecting all the characteristics of the buildings
		SVMR	0.01	2020.30	1116.26	R = 0.08			
Application of artificial intelligence neural network modeling to predict the generation of domestic, commercial and construction wastes	Coskuner *et al.* (2021)	ANN	0.95; 0.99; 0.91 (domestic, commercial and CDW)			MSE = 0.44; 0.01; 0.09 (domestic, commercial and CDW)	20 years of waste generation rate	Official data from national authorities	Not mentioned

Title	Source and Year	ML Model	R^2	MAE	RMSE	Others	Dataset size	Data source	Main Limitations
Estimation of construction waste generation using machine learning	Nagalli (2021)	ANN	0.96			Accuracy = 43.3%	330 construction works	Collected by the authors	ANN's best configuration only predicts 43.3% of cases
LSTM-based forecasting for urban construction waste generation	Huang *et al.* (2021)	RR		A: 12.54; B: 36.7	A: 17.39; B: 37.84	MAPE = A: 6.83; B: 25.29	Dataset A: 32 years; Dataset B: 31 years	Shanghai statistical yearbook and Waste Reduction Office in the HKEPD	Only focus in 2 cities of China due to data limitation
		SVMR		A: 13.9; B: 6.13	A: 15.34; B: 6.73	MAPE = A: 8.20; B: 4.20			
		ANN BPNN		A: 10.18; B: 27.94	A: 13.84; B: 22.09	MAPE = A: 5.58; B: 15.57			
		RNN		A: 12.78; B: 12.16	A: 16.01; B: 14.19	MAPE = A: 7.90; B: 8.12			
		RNN LSTM		A: 6.79; B: 5.48	A: 10.14; B: 5.95	MAPE = A: 4.17; B: 3.98			
Development of a prediction model for demolition waste generation using a random forest algorithm based on small datasets	Cha *et al.* (2020)	RF	0.55 - 0.80			R = 0.69 - 0.87	691 buildings	Collected by the authors	Small dataset; do not compare the results with other ML models
Deep learning model for Demolition Waste Prediction in a circular economy	Akanbi *et al.* (2020)	DNN	0.97 (average)	17.93 - 19.04			2280 buildings	UK Demolition Industry	Dataset based in the UK - regional applicability
Forecasting of construction and demolition waste in Brazil	Paz & Lafayette (2016)	N/A					19 worksites	Collected by the authors	The method is limited to sites with large built areas
Benchmarking construction waste management performance using big data	Lu *et al.* (2015)	N/A					5764 projects	HKEPD	CDW generation rates per contract value vary from one place to another
Web-based construction waste estimation system for building construction projects	Li & Zhang (2013)	N/A					N/A	User provided data	Not mentioned

With recent advancements in AI, particularly in ML, several studies have emerged to predict CDW generation rates, primarily using ANNs and SVMs. Nguyen *et al.* noted that ANNs are popular for their accuracy but are often criticized as "black box" models, where the internal processes are not easily interpretable. Additionally, ANNs require substantial computational resources due to their hidden layers (Nguyen *et al.*, 2023). In contrast, SVMs are suitable for limited sample sizes and demand less computational power, but they generally have lower accuracy than ANNs (Nguyen *et al.*, 2023).

To address the "black box" issue, Lu *et al.* developed ML models to estimate CDW in China. They compared four models, including multiple linear regression (MLR), DT, Grey models (GM), and ANN. All four models performed well, with R^2 values ranging from 0.76 to 0.98. GM and ANN were highly accurate but less interpretable, while MLR and DT offered transparency at a slight cost in accuracy. The study highlights the importance of balancing simplicity and interpretability with complexity and accuracy, emphasizing ethical considerations in AI.

Regarding the performance metrics, most of the models presented high R^2 values, except for a few, such as Cha *et al.* (2023b) and Cha *et al.*, (2021a). However, it is important to understand that performance metrics are indicative of a model's performance on specific dataset used (Nguyen *et al.*, 2023), and some models addressed the limitation of imbalanced datasets (Yang *et al.*, 2021). According to Yang *et al.* , many CDW datasets have missing data, which makes model development difficult. In this sense, they employed various ML algorithms, including DT, KNN, SVM and AdaBoost to estimate missing values from a database of the Hong Kong Environmental Department's database.

A key limitation addressed by several authors is the scarcity of available data (Yang *et al.*, 2021; Coskuner *et al.*, 2021; Cha *et al.*, 2020). The dataset sizes in the reviewed papers varied significantly. This fact was also noted by Nguyen *et al.* in their review. Due to the lack of data, many authors opted to collect the data themselves, which is a very costly and time-consuming approach. Consequently, many models developed are only representative of a specific city or region. Authors such as Gulghane *et al.* (2023) and Lu *et al.* (2023) addressed the limitation of their model having a regional applicability.

Moreover, ML models primarily operate with numerical data, while CDW generation involves various categorical variables like construction type, project size, location, methods, and disposal regulations . Encoding these categorical variables is essential for compatibility with ML models. Cha *et al.* (2023b; 2023c; 2022; 2021a; 2021b; 2020) addressed this challenge developing ML models using small datasets comprised mainly of categorical variables.

Cha *et al.* employed the CATPCA, a method that transforms categorical variables into numerical ones. In both studies, the employment of CATPCA improved the performance of the models. Cha *et al.* assessed the performance of various ML algorithms, both individually and in combination with an autoencoder (AE) to predict CDW. AE are unsupervised neural networks used to compress and reconstruct input data that can convert categorical variables into numerical ones and extract relevant features from data (Cha *et al.*, 2023c). Similarly, Cha *et al.* combined ML models with Principal Component Analysis (PCA). PCA is a method for reducing dimensionality and feature extraction. It transforms high-dimensional datasets into new coordinate systems, capturing data's variability with fewer uncorrelated variables known as principal components (Cha *et al.*, 2023d).

Overfitting is a common challenge in ML when there is limited available data. It occurs when a model fits the training data too closely and loses its generalization ability (Santos *et al.*, 2018). Therefore, the model performs exceptionally well on the training data but fails to make accurate predictions on new data. Nagalli reports a possible situation of overfitting, when assessing the effectiveness of ANNs in forecasting CDW generation. Remarkable prediction outcomes, boasting an R^2 value of 0.96, were achieved. However, the best-configured ANN could accurately predict only 43.3% of cases. The author suggested that enhanced accuracy may be attainable by increasing the volume of simulated data.

Besides the importance of data, hyperparameters play a crucial role in improving ML model performance by adjusting the model configuration to optimize its effectiveness. Hyperparameters are settings of an ML model that are established (either manually or automatically) before the training process and are not learned from the data. By fine-tuning hyperparameters, the model's performance can be optimized by identifying the best combination of settings that minimize the discrepancy between predicted and actual values (Cha *et al.*, 2023a). Cha *et al.* focused on the significant influence of hyperparameters on ML development.

Cha *et al.* compared various ML algorithms (including ANN, KNN, LR, RF, and SVM) to predict CDW generation. The authors optimized hyperparameters to enhance the performance of the models. Similarly, Cha *et al.* created various ML models with different hyperparameters to achieve better performance. The most suitable models were RF ($R^2 = 0.62$) and GBM ($R^2 = 0.61$). However, in this last study, the authors addressed the low R^2 values as limitation.

3.3.2 CDW classification and identification

Classifying and identifying CDW plays a critical role in enhancing the management of this waste stream. Due to the waste's variability, manual separation is slow and costly, making material reuse difficult (Davis *et al.*, 2021). Effective classification allows for the efficient recycling, reuse, or proper disposal of CDW, as the heterogeneity of the waste is one of the main obstacles to its reuse (Hackenhaar *et al.*, 2019). In this regard, smart devices that automatically classify CDW can facilitate the waste separation process, making recycling a more attractive activity (Davis *et al.*, 2021).

Of the 37 articles, 11 focused on CDW Classification and Identification. These papers are relatively recent, being the oldest from 2019. Most studies focused on identifying, sorting, and classifying different types of CDW. Other papers focused on identifying CDW disposal sites and predicting the presence of hazardous waste. Image recognition algorithms were the basis for 73% of these papers. Recently, advancements in image classification and recognition systems have been noteworthy due to the successful application of the CNN (Na *et al.*, 2022; Davis *et al.*, 2021). A CNN can take an input image, assign weights (levels of importance), and differentiate the aspects present in it, recognizing patterns and detecting elements (Na *et al.*, 2022). Key characteristics of these studies are summarized in Table 2.

Despite advancements in image recognition, some authors (Dong *et al.*, 2022; Lu *et al.*, 2022) criticize that many studies are conducted in controlled environments, limiting their real-world applicability. Lu *et al.* argue that traditional computer vision methods oversimplify CDW classification by focusing on individual waste items against simple backgrounds. They proposed a semantic segmentation approach to handle CDW mixtures in complex, real-world environments. Using DeepLabv3+, a CNN model for semantic segmentation, they achieved an IoU of 0.56 and an F1 score of 69.2%. Limitations included imbalanced dataset and inconsistencies in predicted boundaries, suggesting room for boundary improvement.

Dong *et al.* introduced a deep neural network model with boundary-aware recognition (BAT) for refined recognition of CDW mixtures, using the same dataset as Lu *et al.* . According to the authors, the existing image recognition methods for CDW based on computer vision often suffer from low accuracy, struggling to distinguish boundaries between different waste. BAT employs a cascade decoder to enhance segmentation precision and a learning-enabled edge refinement scheme to improve boundary recognition (Dong *et al.*, 2022). The BAT model achieved 0.62 IoU, surpassing the results of Lu *et al.* .

Table 2

Publications focused on CDW Classification and Identification

Title	Source and Year	ML Model	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Others	Dataset size	Data source	Main Limitations
Applying machine learning to fine classify construction and demolition waste based on deep residual network and knowledge transfer	Lin *et al.* (2023)	CNN (ResNet) CDWNet-18 with knowledge transfer	73.3	73.7	73.3	73.1	Kappa = 69.8%	2836 images	Manually collected from Google or taken by authors cameras	Slowly training due to the large number of parameters; the model still has room for improvement
Vision-based automated waste audits: a use case from the window manufacturing industry	Martinez *et al.* (2022)	CNN (Xception) combined with vision-based algorithms		89.6	89.3	89.4		1752 images to train the CNN and 55 pieces of waste to test the hole model	Images captured by the authors in different lighting conditions and orientations	Waste profile measurement is limited to the field of view of the camera; cannot be generalized to other cutting stations
Computer vision to recognize construction waste compositions: A novel boundary-aware transformer (BAT) model	Dong *et al.* (2022)	DNN – BAT	72.8				IoU = 0.62	5366 photos of waste loads	HKEPD	Imbalanced dataset
Estimating construction waste truck payload volume using monocular vision	Chen *et al.* (2022)	CNN with semantic segmentation					Volume estimation error = 0.169	2,914 records from one sorting facility in Hong Kong	Provided by the sorting facility and collected by the authors.	Assumption of waste dump invariability (image from the top represents all the content)
Using computer vision to recognize composition of construction waste mixtures: A semantic segmentation approach	Lu *et al.* (2022)	CNN with semantic segmentation (DeepLabV3+)		70.6	68.7	69.2	IoU = 0.56	5366 photos of waste loads	HKEPD	Ambiguty among CDW types; imbalanced dataset; inconsistency of predicted boundaries
Identification of construction and demolition waste based on change detection and deep learning	Zhao *et al.* (2022)	CNN with semantic segmentation (DeepLabV3+)	91.7				Kappa = 86.4%	1053 patches	China Center for Resources Satellite Data and Application	Images compared were not taken on the same day, differing in spectral values

Title	Source and Year	ML Model	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Others	Dataset size	Data source	Main Limitations
Predicting the presence of hazardous materials in buildings using machine learning	Wu *et al.* (2022)	Logistic Regression	55 and 67 (asbestos and PCB)		62 and 68 (asbestos and PCB)			848 records - but only 188 were used for asbestos and 106 for PCB	Hazardous waste inventories of buildings from Gothenburg and Stockholm City Archives	Heterogeneous and unstandardized data; data extraction and compiling rely on manual operations
		SVM	50 and 67		58 and 69					
		KNN	74 and 81		78 and 81					
		RF	89 and 95		92 and 94					
		XGBoost	89 and 90		92 and 89					
		CatBoost	84 and 95		88 and 94					
Deep convolutional neural networks for construction and demolition waste classification	Lin *et al.* (2022)	CNN (VGGNet 16)	76.6	76.8	76.6	76.6	AUC = 0.87	2836 images	Images taken by authors' cameras or from Google search results	The models used still have room for improvement
Using computer vision to recognize construction material: A Trustworthy Dataset Perspective	Sun & Gu (2022)	CNN (DenseNet)	99.7	99.7	99.7	99.7		7000 images	Photographs taken by the authors	Similar framework is required to improve data quality for CDW object detection and semantic segmentation
		CNN (SqueezeNet)	99.6	99.6	99.6	99.6				
A data-driven approach to assess the risk of encountering hazardous materials in the building stock based on environmental inventories	Wu *et al.* (2021)	N/A						402 observations	Data collected from different authorities in Sweden	Data heterogeneity; high number of missing values
A robust classification algorithm for separation of construction waste using NIR hyperspectral system	Xiao *et al.* (2019)	RF	84.8 (characteristic reflectivity); 91.9 (first derivative)					1844 pieces of waste	Collected by the authors	Not mentioned
		ELM	95.6 (characteristic reflectivity); 86.7 (first derivative)							

Similar to the articles on CDW Generation Prediction, authors of CDW Classification and Identification studies also cite the lack of data or imbalanced datasets as a primary limitation. An imbalanced dataset poses a significant challenge in classification tasks because when working with such data, ML models may become biased towards the majority class, as global performance optimization often favors accuracy in the dominant class (Mittelstadt *et al.*, 2016). This can result in models performing well on the majority class but being ineffective in identifying the minority classes (Mittelstadt *et al.*, 2016).

Regarding the dataset source, many authors used images collected by themselves, highlighting the absence of a reliable database for such studies. In response to this, Sun and Gu focused on generating high-quality datasets for CDW management models, comprising 7,000 images. The authors presented a cross-review framework to create reliable construction material datasets and assess the performance of CNNs in material recognition. The characteristics of construction materials and construction techniques vary globally. Therefore, a high-quality dataset created for identification and classification studies in one region may not perform well in another region that employs different construction materials.

Similar to Sun and Gu , Wu et al. focused on developing high-quality datasets for ML models that assess the risk of encountering hazardous materials in existing buildings. In a further study, Wu *et al.* explored ML techniques to predict hazardous waste in buildings, focusing on asbestos pipe insulation in multifamily houses and PCB (printed circuit board) joints/sealants in school buildings.

Differently from the papers that focused on image recognition to classify CDW, Zhao *et al.* employed a CNN with multispectral images to identify the arrangement of CDW disposal in Beijing, China. Another different approach is the study of Xiao *et al.* that combined a hyperspectral camera and ML models to classify six types of CDW.

Regarding the dataset size, CDW Classification and Identification models employed a smaller variable dataset size when compared to the models developed to predict CDW generation. Models that relied on image recognition algorithms used at least 1,000 images in their datasets. Only studies on the prediction of hazardous materials in buildings used smaller datasets, around 400 – 800 observations (Wu *et al.*, 2021; 2022).

3.3.3 Decision Support Systems (DSS) and logistics

Decision Support Systems (DSS) combined with efficient logistics are crucial for CDW management. DSS offers a structured framework for data-driven decision-making, enabling

stakeholders to analyze and interpret information related to CDW generation, composition, and disposal . Efficient logistics involves meticulous planning, route optimization, and resource allocation, all of which are crucial for minimizing CDW transportation environmental impacts (Tsydenova *et al.*, 2021; Bi *et al.*, 2022). Of the 37 papers, 6 focused on DSS and Logistics in CDW management. These papers are detailded in Table 3.

Most of these studies (five out of six) did not use ML, but instead relied on mathematical simulations, optimization methods, and GIS. The oldest study, from 2011, by Banias *et al.* , introduced a software to propose end-of-life CDW management options. It estimates CDW quantities for 21 waste types in various building types and employs an optimization model to propose management options, considering economic and environmental factors. The model also proposes efficient disposal routes.

Recent studies have also developed DSS (Saeed *et al.*, 2023; Sobotka & Sagan, 2021). Saeed *et al.* developed a DSS to manage CDW, specifically concrete and glass. They quantified potential environmental benefits from various recovery methods (e.g., reuse, recycling, etc.) and employed multi-objective optimization to determine the ideal amount of material for each recovery scenario. It incorporates stakeholder preferences through parametric weights, balancing environmental and economic advantages.

Sobotka and Sagan developed a DSS for concrete waste management based on mathematical simulation. The system evaluates economic, environmental, and social factors to choose the best waste management option. The construction manager provides information about the construction, and the system considers three options: on-site recycling, landfill disposal, and transportation to another site where concrete waste will be recycled on-site.

Tydenova *et al.* introduced an optimized concrete recycling network. They employed a bi-objective optimization model to plan sorting screen locations and material flows within a regional recycling system in North Rhine-Westphalia, Germany. Similarly, Paz *et al.* designed a network for CDW in the city of Recife, Brazil using GIS. Their process involved mapping illegal disposal sites, categorizing recyclable waste, and identifying spots for voluntary drop-off points.

Bi *et al.* exclusively addressed the logistics of CDW management, employing ML. They focus on improving CDW transportation by utilizing electronic waybill technology. This technology streamlines waste generation, transportation, and disposal processes, enhancing management (Bi *et al.*, 2022). The study analyzes the transport trajectories of trucks to optimize routes, based on GPS-collected data points. Various ML methods were employed to predict waybill generation, with XGBoost achieving the best prediction accuracy at 90%.

Table 3

Publications which focused on Decision Support Systems (DSS) and Logistics

Title	Source and Year	ML Model	Accuracy (%)	Dataset size	Data source	Main Limitations
Environmental Impact and Cost Assessment for Reusing Waste during End-of-Life Activities on Building Projects	Saeed *et al.* (2023)	N/A		N/A	Case study in a university building in Canada	Only focused on concrete and glass waste
Identification of Working Trucks and Critical Path Nodes for Construction Waste Transportation Based on Electric Waybills: A Case Study of Shenzhen, China	Bi, *et al.* (2022)	DT	84.2	13,947 data of vehicles in waybill nodes	GPS devices	Path nodes that have high traffic accidents due to narrow roads and crowded people were not considered
		RF	87.7			
		GBDT	88.9			
		XGBoost	90.6			
Decision support system in management of concrete demolition waste	Sobotka & Sagan (2021)	N/A		N/A	User-provided information	Not clearly mentioned by the authors, but they only focus on concrete waste
Optimized design of concrete recycling networks: The case of North Rhine-Westphalia	Tsydenov *et al.* (2021)	N/A		N/A	N/A	Only focus on concrete waste; assumes a closed-loop cycle (do not consider the landfills as a final sink for CDW)
GIS-based planning system for managing the flow of construction and demolition waste in Brazil	Paz, *et al.* (2018)	N/A		N/A	Collect by the authors and official government data	It is necessary to evaluate the costs of implementation and operation of the proposed points
A web-based Decision Support System for the optimal management of construction and demolition waste	Banias *et al.* (2011)	N/A		565 irregular disposal points	User-provided information	Not mentioned

Banias *et al.* and Saeed *et al.* consider both economic and environmental factors in CDW management, while Sobotka and Sagan further incorporate social factors. A common limitation in three DSS and logistics studies is their narrow focus on a limited range of construction materials, primarily concrete. Additionally, Sobotka and Sagan and Banias *et al.* rely on user-provided data, making the model's success reliant on the manager's ability to supply accurate building information.

Bi *et al.* specifically explore logistical issues, although Banias *et al.* also propose efficient waste disposal routes, and Tsydenova *et al.* and Paz *et al.* propose CDW recycling networks. Remarkably, few studies address logistics, despite transportation and its environmental impact being a significant obstacle to CDW use (Tazi *et al.*, 2021). Therefore, a notable gap in addressing logistical challenges in CDW management remains. Oluleye *et al.* also noted a lack of AI studies on reverse logistics.

4 DISCUSSIONS

One of the main limitations reported by authors is the lack of quality data, including issues like unbalanced, faulty, or non-standardized datasets. The success of ML models depends heavily on data quality. Insufficient or poor-quality data can hinder their ability to generalize patterns, resulting in inaccurate or biased outcomes. Unbalanced datasets may skew models toward majority classes, ignoring important patterns in minority classes, while lack of standardization can introduce bias and reduce comparability across datasets.

The construction industry is one of the least digitized sectors when compared to manufacturing, resulting in a lack of data for developing robust AI algorithms (Abioye *et al.*, 2021). This slow technology adoption leads to project delays, poor performance, and uninformed decision-making. However, the rising use of Building Information Modeling (BIM) is likely to increase data availability for higher-quality models. In Brazil, for example, there is a decree mandating the use of BIM in public construction projects in the coming years . The dissemination of BIM will probably increase data availability in the construction sector, providing foundation for the application of AI techniques.

The lack of data also restricts model applicability on a broader scale. Many of the studies on CDW management developed so far are restricted to a specific city or region due to limited data availability. Moreover, authors often had to collect data themselves because of the lack of publicly accessible datasets. In this regard, Tapeh and Naser and Oluleye *et al.* highlight the need for open-access databases and code sharing to address these challenges. In our study, only

5 out of 37 papers provided public access to their datasets, while 9 offered data availability upon request. This lack of transparency creates barriers to scientific replicability and real-world model adoption.

Regarding code availability, 6 studies shared their code, and 1 paper (Banias *et al.*, 2011) even provided a link to a beta version of its software, indicating a commitment to community engagement. However, a significant majority (81%) did not mention code availability, highlighting a lack of transparency. This emphasizes the urgent need for standardized data and code-sharing practices to ensure credibility and collaborative progress in AI. While it may not always be feasible to generalize ML models globally due to varying construction characteristics and materials, valuable lessons from AI applications in different regions can still be applied if the credibility and soundness of the algorithms can be attested.

Beyond the data issue, many ML algorithms function as black boxes, making their decision-making processes difficult to understand. This complicates AI application in construction and raises ethical concerns. When algorithms make critical decisions, such as sizing civil structures or predicting concrete performance, clear explanations of their processes are essential to building trust (Abioye *et al.*, 2021). Thus, it is important to balance simplicity and interpretability with complexity and accuracy in ML model development (Lu *et al.*, 2021).

Similar to Abdallah *et al.* , we observed that while AI has advanced significantly in recent years, many models remain confined to academia with limited practical application. The number of DSS developed for CDW management so far is low – this was also noted by Oluleye *et al.* . Most papers that developed DSS rely on formulas and indices to estimate CDW generation rather than ML. Therefore, there is a need to integrate ML models for predicting and classifying CDW generation with DSS to enhance decision-making regarding optimal disposal options and routes.

Few studies have explored logistics in AI-driven CDW management, even though logistics are crucial for evaluating the viability of using CDW. Transportation has a significant impact on the feasibility of CDW use (Farinha *et al.*, 2019; Hackenhaar *et al.*, 2019). Therefore, ML models for CDW management still need to evolve to a level that integrates DSS and logistics to fully achieve circularity in the construction industry.

5 CONCLUSIONS

This study aimed to identify how AI techniques are being employed in CDW management through a systematic review of 37 research papers extracted from Scopus and Web

of Science. We observed that the incorporation of AI in CDW management is relatively recent, with a significant rise in publications noted from 2021. Most papers have a Chinese first author and the most frequently used ML algorithms were the RF, GBM, CNN, and ANN.

Three main research identified were CDW Generation Prediction, CDW Classification and Identification, and Decision Support Systems and Logistics. The main limitation addressed by the authors is the lack of data or unbalanced and non-standardized data. The construction industry's slow adoption of technology and digitization contributes to this data scarcity, hindering the development of robust AI algorithms.

AI research is progressing at a much faster pace than circularity in the construction industry. AI holds great potential to revolutionize CDW management, but without a proactive and ethical adoption of technology, the construction sector may face challenges in effectively reducing waste generation and contributing to the decarbonization efforts within the sector. Therefore, future endeavors should focus on overcoming data availability limitations, incorporating ethical considerations such as model interpretability, and integrating ML models with DSS and logistics to foster circularity in construction practices.

ACKNOWLEDGEMENTS

We acknowledge CAPES (Finance Code 001 - PhD scholarship for Carolina Lopes), FAPEMIG (grant APQ-01838-21 for Julia Mendes and PPM-00001-18 for Alexandre Cury), CNPq (grant 305818/2023-6 for Julia Mendes and grants 303982/2022-5 and 402533/2023-2 for Alexandre Cury) and PROPP/UFJF for the financial support. We are also grateful for the collaboration of the Research Group on Data Science in Engineering (CIDENG – CNPq).

REFERENCES

Abdallah, M., Talib, M. A., 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. doi:10.1016/j.wasman.2020.04.057

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Delgado, J. M., Bilal, M., & Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering, 44*, 103299.

Akhtar, A., & Sarmah, A. (2018). Construction and demolition waste generation and properties of recycled aggregate concrete: A global perspective. *Journal of Cleaner Production, 186*, 262-281. doi:https://doi.org/10.1016/j.jclepro.2018.03.085

Banias, G., Achillas, C., Vlachokostas, C., Moussiopoulos, N., & Papaioannou, I. (2011). A

web-based Decision Support System for the optimal management of construction and demolition waste. *Waste Management, 31*(12), 2497-2502. doi:10.1016/j.wasman.2011.07.018

Bi, J., Sai, Q., Wang, F., & Chen, Y. (2022). Identification of Working Trucks and Critical Path Nodes for Construction Waste Transportation Based on Electric Waybills: A Case Study of Shenzhen, China. *Journal of Advanced Transportation, 2022*, 7647121. doi:10.1155/2022/7647121

BRASIL. (2020). *Decreto Nº 10.306.* Brasília. Retrieved 11 14, 2023, from https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/D10306.htm

Carvalho, A. R., Calderón-Morales, B. R., Borba Júnior, J. C., Oliveira, T. M., & Silva, G. J. (2023). Proposition of geopolymers obtained through the acid activation of iron ore tailings with phosphoric acid. *Construction and Building Materials, 403*, 133078. doi:10.1016/j.conbuildmat.2023.133078

Cha, G. W., Choi, S.-H., Hong, W.-H., & Park, C.-W. (2023a). Development of machine learning model for prediction of demolition waste generation rate of buildings in redevelopment areas. *International Journal of Environmental Research and Public Health, 20*, 107. doi:10.3390/ijerph20010107

Cha, G. W., Choi, S.-H., Hong, W.-H., & Park, C.-W. (2023d). Developing a prediction model of demolition-waste generation-rate via principal component analysis. *International Journal of Environmental Research and Public Health, 20*(4), 3159. doi:10.3390/ijerph20043159

Cha, G. W., Hong, W.-H., & Kim, Y.-C. (2023c). Performance Improvement of Machine Learning Model Using Autoencoder to Predict Demolition Waste Generation Rate. *15*(4), 3691. doi:10.3390/su15043691

Cha, G. W., Hong, W.-H., Choi, S.-H., & Kim, Y.-C. (2023b). Developing an Optimal Ensemble Model to Estimate Building Demolition Waste Generation Rate. *15*, 10163. doi:10.3390/su151310163

Cha, G. W., Moon, H. J., Kim, Y.-M., Hong, W.-H., Hwang, J.-H., Park, W.-J., & Kim, Y.-C. (2020). Development of a prediction model for demolition waste generation using a random forest algorithm based on small datasets. *International Journal of Environmental Research and Public Health, 17*(19), 6997. doi:10.3390/ijerph17196997

Cha, G. W., Moon, H., & Kim, J. (2021a). A method to improve the performance of support vector machine regression model for predicting demolition waste generation using categorical principal components analysis. *International Journal of Sustainable Building Technology and Urban Development, 12*(3), 282-294. doi:10.22712/susb.20210023

Cha, G. W., Moon, H.-J., & Kim, Y.-C. (2021b). Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables. *International Journal of Environmental Research and Public Health, 18*(16), 8530. doi:10.3390/ijerph18168530

Cha, G. W., Moon, H.-J., & Kim, Y.-C. (2022). A hybrid machine-learning model for predicting the waste generation rate of building demolition projects. *Journal of Cleaner Production,*

375, 134096. doi:10.1016/j.jclepro.2022.134096

Cosenza, J. P., Andrade, E. M., & Assunção, G. M. (2020). Economia circular como alternativa para o crescimento sustentável brasileiro: análise da Política Nacional de Resíduos Sólidos. *Revista de Gestão Ambiental e Sustentabilidade, 9*(1), 16147. doi:10.5585/geas.v9i1.16147

Coskuner, G., Jassim, M. S., Zontul, M., & Karateke, S. (2021). Application of artificial intelligence neural network modeling to predict the generation of domestic, commercial and construction wastes. *Waste Management & Research, 39*(3), 499-507. doi:10.1177/0734242X20935181

Davis, P., Aziz, F., Newaz, M. T., Sher, W., & Simon, L. (2021). The classification of construction waste material using a deep convolutional network. *Automation in Construction, 122*, 103481. doi:10.1016/j.autcon.2020.103481

Dong, Z., Chen, J., & Lu, W. (2022). Computer vision to recognize construction waste compositions: A novel boundary-aware transformer (BAT) model. *Journal of environmental management, 305*, 114405.

Ekaba, B. (2018). What Is Deep Learning? In *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners* (pp. 327-329). Ottawa: Apress. doi:10.1007/978-1-4842-4470-8_27

Ensafi, M., Alimoradi, S., Gao, X., & Thabet, W. (2022). Machine Learning and Artificial Intelligence Applications in Building Construction: Present Status and Future Trends. Em *Construction Research Congress 2022 : Computer Applications, Automation, and Data Analytics* (pp. 116-124). Arlington: American Society of Civil Engineers. doi:10.1061/9780784483961.013

Farinha, C. B., Silvestre, J. D., Brito, J., & Veiga, M. d. (2019). Life Cycle Assessment of Mortars with Incorporation of Industrial Wastes. *Fibers, 7*(59). doi:10.3390/fib7070059

Gnat, S. (2021). Impact of Categorical Variables Encoding on Property Mass Valuation. *Procedia Computer Science, 192*, 3542-3550. doi:10.1016/j.procs.2021.09.127

Gulghane, A., Sharma, R. L., & Borkar, P. (2023). Quantification analysis and prediction model for residential building construction waste using machine learning technique. *Asian Journal of Civil Engineering, 24*, 1459–1473. doi:10.1007/s42107-023-00580-x

Hackenhaar, I. C., Waskow, R. P., Tubino, R., & Passuello, A. (2019). Life Cycle Assessment applied to construction and demolition waste treatment: proposal of a Brazilian scenario. *IOP Conference Series: Earth and Environmental Science, 323.*

Huang, L., Cai, T., Zhu, Y., Zhu, Y., Wang, W., & Sun, K. (2021). LSTM-based forecasting for urban construction waste generation. *Sustainability, 12*(20), 8555. doi:10.3390/su12208555

Kannan, S., Arunachalam, K., & Brindha, D. (2021). Performance analysis of recycled aggregate concrete with chemical admixture. *Structural Concrete, 22*, E8-E21. doi:10.1002/suco.201900380

Li, C. Z., Zhao, Y., Xiao, B., Yu, B., Tam, V. W., Chen, Z., & Ya, Y. (2020). Research trend

of the application of information technologies in construction and demolition waste management. *Journal of Cleaner Production, 263*, 121458.

Li, Y., & Zhang, X. (2013). Web-based construction waste estimation system for building construction projects. *Automation in Construction, 35*, 142–156. doi:10.1016/j.autcon.2013.05.002

Lin, K., Zhao, Y., Zhou, T., Gao, X., Zhang, C., Huang, B., & Shi, Q. (2023). Applying machine learning to fine classify construction and demolition waste based on deep residual network and knowledge transfer. *Environment, Development and Sustainability, 25*, 8819–8836. doi:10.1007/s10668-022-02740-6

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. doi:10.1016/j.resconrec.2021.106022

Lu, W., Chen, X., Peng, Y., & Shen, L. (2015). Benchmarking construction waste management performance using big data. *Resources, Conservation and Recycling*, *105*, 49–58. doi:10.1016/j.resconrec.2015.10.013

Lu, W., Long, W., & Yuan, L. (2023). A machine learning regression approach for pre-renovation construction waste auditing. *Journal of Cleaner Production, 397*, 136596. doi:10.1016/j.jclepro.2023.136596

Lu, W., Lou, J., Webster, C., Xue, F., Bao, Z., & Chi, B. (2021). Estimating construction waste generation in the Greater Bay Area, China using machine learning. *Waste Management, 134*, 78-88. doi:10.1016/j.wasman.2021.08.012

Mao, T., Liu, Y., Chen, W. Q., Li, N., Dong, N., & Shi, Y. (2022). Quantifying spatiotemporal dynamics of urban building and material metabolism by combining a random forest model and GIS-based material flow analysis. *Frontiers in Earth Science, 10*, 944865.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1-21. doi:10.1177/2053951716679679

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), 175. doi:10.3390/buildings12020175

Nagalli, A. (2021). Estimation of construction waste generation using machine learning. *Proceedings of Institution of Civil Engineers - Waste and Resource Management, 174*(1), 22-31. doi:10.1680/jwarm.20.00019

Nguyen, T. D., Cherif, R., Mahieux, P. Y., Lux, J., Aït-Mokhtar, A., & Bastidas-Arteaga, E. (2023). Artificial intelligence algorithms for prediction and sensitivity analysis of mechanical properties of recycled aggregate concrete: a review. *Journal of Building Engineering, 66*, 105929. doi:10.1016/j.jobe.2023.105929

Oluleye, B. I., Chan, D. W., & Antwi-Afari, P. (2023). Adopting Artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review. *Sustainable Production and Consumption, 35*, 509-524.

doi:10.1016/j.spc.2022.12.002

Oluleye, B. I., Chan, D. W., Saka, A. B., & Olawumi, T. O. (2022). Circular economy research on building construction and demolition waste: A review of current trends and future research directions. *Journal of Cleaner Production, 357*, 131927. doi:10.1016/j.jclepro.2022.131927

Paz, D. H., & Lafayette, K. P. (2016). Forecasting of construction and demolition waste in Brazil. *Waste Management & Research, 34*(8), 708-716. doi:10.1177/0734242X16644680

Paz, D. H., Lafayette, K. P., & Sobral, M. d. (2018). GIS-based planning system for managing the flow of construction and demolition waste in Brazil. *Waste Management and Research, 36*(6), 541-549. doi:10.1177/0734242X18772096

Penido, R. E., da Paixão, R. C., Costa, L. C., Peixoto, R. A., Cury, A. A., & Mendes, J. C. (2022). Predicting the compressive strength of steelmaking slag concrete with machine learning – Considerations on developing a mix design tool. *Construction and Building Materials, 341*, 127896. doi:10.1016/j.conbuildmat.2022.127896

Saeed, F., Mostafa, K., Rausch, C., & Hegazy, T. (2023). Environmental Impact and Cost Assessment for Reusing Waste during End-of-Life Activities on Building Projects. *Journal of Construction Engineering and Management, 149*(10), 04023099. doi:10.1061/JCEMD4.COENG-12943

Santos, M. S., Soares, J. P., Abreu, P. H., Araujo, H., & Santos, J. (2018). Cross-validation for imbalanced datasets: avoiding overoptimistic and overfitting approaches. *IEEE Computational Intelligence Magazine, 13*(4), 59-76. doi:10.1109/MCI.2018.2866730

Sobotka, A., & Sagan, J. (2021). Decision support system in management of concrete demolition waste. *Automation in Construction, 128*, 103734. doi:10.1016/j.autcon.2021.103734

Souza, F. d., Mendes, J. C., Morais, L. J., Silva, J. S., & Peixoto, R. A. (2022). Mapping and recycling proposal for the construction and demolition waste generated in the Brazilian Amazon. *Resources, Conservation and Recycling, 176*, 105896. doi:10.1016/j.resconrec.2021.105896

Sun, Y., & Gu, Z. (2022). Using computer vision to recognize construction material: A Trustworthy Dataset Perspective. *Resources, Conservation & Recycling, 183*, 106362. doi:10.1016/j.resconrec.2022.106362

Tapeh, A. T., & Naser, M. Z. (2023). Artificial Intelligence, Machine Learning, and Deep Learning in Structural Engineering: A Scientometrics Review of Trends and Best Practices. *Archives of Computational Methods in Engineering, 30*, 115-159. doi:10.1007/s11831-022-09793

Tazi, N., Idir, R., & Fraj, A. B. (2021). Towards achieving circularity in residential building materials: Potential stock, locks and opportunities. *Journal of Cleaner Production, 281*, 124489. doi:10.1016/j.jclepro.2020.124489

Tsydenova, N., Becker, T., & Walther, G. (2021). Optimised design of concrete recycling networks: The case of North Rhine-Westphalia. *Waste Management, 135*, 309-317.

doi:10.1016/j.wasman.2021.09.013

Wu, P.-Y., Mjörnell, K., Mangold, M., Sandels, C., & Johansson, T. (2021). A data-driven approach to assess the risk of encountering hazardous materials in the building stock based on environmental inventories. *Sustainability, 13*, 7836. doi:10.3390/su13147836

Wu, P.-Y., Sandels, C., Mjörnell, K., Mangold, M., & Johansson, T. (2022). Predicting the presence of hazardous materials in buildings using machine learning. *Building and Environment, 213*, 108894. doi:10.1016/j.buildenv.2022.108894

Wuni, I., Shen, G., & Osei-Kyei, R. (2019). Scientometric review of global research trends on green buildings in construction journals from 1992 to 2018. *Energy & Buildings*(190), pp. 69-85. Retrieved agosto 17, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0378778818334571

Xiao, W., Yang, J., Fang, H., Zhuang, J., & Ku, Y. (2019). A robust classification algorithm for separation of construction waste using NIR hyperspectral system. *Waste Management, 90*, 1-9. doi:10.1016/j.wasman.2019.04.036

Yang, Z., Xue, F., & Lu, W. (2021). Handling missing data for construction waste management: Machine learning based on aggregated waste generation behaviors. *Resources, Conservation and Recycling, 175*, Yang, Z., Xue, F., & Lu, W. (2021). Handling missing data for construction waste management: Machine learning105809. doi:10.1016/j.resconrec.2021.105809

Zhao, X., Yang, Y., Duan, F., Zhang, M., Jiang, G., Yan, X., . . . Zhao, W. (2022). Identification of construction and demolition waste based on change detection and deep learning. *International Journal of Remote Sensing, 43*(6), 2012-2028. doi:10.1080/01431161.2022.2054296

View publication stats