# Applying machine learning to fine classify construction and demolition waste based on deep residual network and knowledge transfer

**Abstract:** Few studies reported using the convolutional neural network with transfer learning to finely classify the construction and demolition waste. **Objectives:** This study aims to develop a highly efficient method to realize the finely sorting the construction and demolition waste, which is a key step for promoting the recycling system to realize carbon neutrality in the waste management sector. **Methodology:** C&DWNet models, ResNet structures based on knowledge transfer and cyclical learning rate, were proposed to classify ten types of construction and demolition waste. Indexes (confusion metric, accuracy, precision, recall, F1 score, sensitivity, specificity and kappa) were adopted to evaluate the performance of various C&DWNet models. **Results:** Knowledge transfer can reduce the training time and improve the performance of the C&DWNet model. The average training time is increased with the increase of the layer of C&DWNet architecture from C&DWNet-18 (946.7 s) to C&DWNet-152 (1186.6 s). The accuracy of various C&DWNet models is approximately 72~74%, the best accuracy is 73.6% in C&DWNet-152. C&DWNet-18 is more suitable for the classification of construction and demolition waste in terms of training time, accuracy, precision, and F1 score. Moreover, the t-distributed stochastic neighbor embedding can distinctly separate each type of construction and demolition waste. **Improvement:** The environmental applications and limitations of the C&DWNet module were also discussed, which could provide a reference for the intelligent management of construction and demolition waste and promote the development of the circular economy.

**Keywords:** Construction & demolition waste classification; Waste management; Machine learning; Deep residual network; Knowledge transfer

## 1 Introduction

The rapid global urbanization and population growth have brought considerable quantities of construction material consumption and produced a large amount of Construction and Demolition (C&D) waste (Huang et al., 2020). Unfortunately, less than 5% of C&D waste was reused or recycled in China (Duan et al., 2015). Most of them were randomly dumped or transported directly to landfills without distinction (Duan and Li, 2016), which would pose threat to air, water, soil, and limited landfills. On the other hand, about 40% of global greenhouse gas (GHG) emissions are caused by construction-related activities (Huang et al., 2020). Thus, promoting the level of C&D waste recycling will help to realize carbon neutrality in the waste management sector.

The traditional method of recycling C&D waste is very time-consuming and labor-intensive, leading to inefficient C&D waste management (Wang et al., 2020). In recent decades, smart classification, as means to improve the efficiency of waste classification, has become quite popular with the development of machine learning (Achu et al., 2021; Khosravi et al., 2019; Mao et al., 2021; Yan et al., 2021). Convolutional Neural Network (CNN), one of the state-of-the-art machine learning structures, was widely used for the task of computer vision like image classification, object detection, and semantic segmentation. This algorithm is being caught attention to classify the waste to improve waste management.

To be exact, Support Vector Machine (SVM) with scale-invariant feature transform and ResNet-50 with SVM was employed to classify the TrashNet, the dataset about the image of recyclable waste, and achieved an accuracy of 63% and 87%, respectively (Yang and Thung, 2016). ResNet-18 was integrated with a self-monitoring module for recyclable waste classification, which can recognize the six waste types in TrashNet with an accuracy of 95.87% (Zhang et al., 2021). A series of research like classifying wet waste (Yanai and Kawano, 2015), plastic waste (Bobulski and Kubanek, 2021; Sreelakshmi et al., 2019) and meal waste (Frost et al., 2019) have been carried out, while few studies have used the ResNet structure to classify the C&D waste. It is worthy to note that the essence of the CNN or ResNet model is a data-driven model and needs extensive information to achieve state-of-the-art performance (Lin et al., 2022).

Thus, measures of data augmentation and Knowledge Transfer (KT) may improve the performance of CNN algorithms. KT (so-called “transfer learning”) is a kind of machine learning technique, which can be applied to transfer knowledge in different physical scenarios (Sinno and Qiang, 2010). KT combined with CNN structures is transferrable and applicable to traffic object detection (Zhang et al., 2018) and recyclable garbage classification (Aral et al., 2018). However, to our best knowledge, limited literature has reported adopting KT integrated with ResNet structures for fine classification of C&D waste.

Notably, deep machine learning needs to identify the most suitable learning rate, which is one of the most hyper-parameters for the training process of CNN, as small or large learning rates will cause slow convergence or divergence in training algorithms (Y.Bengio, 2012). But to realize this, a large amount of computing sources is needed. Leslie (N.Smith, 2017) proposed the cyclical learning rate and proved its effectiveness in deep neural networks training.

This study aims to develop C&DWNet models based on ResNet structure and knowledge transfer to realize the fine classification of C&D waste. The technology of data augmentation and cyclical learning rate was applied to improve the training efficiency. Several evaluation metrics like confusion matrix, accuracy, precision, recall, F1 score, sensitivity, specificity, kappa and ROC also were used to evaluate the performance of C&DWNet models. In addition, algorithms of Principle Component Analysis (PCA) and t-Stochastic Neighbor Embedding (t-SNE) were applied to extract the representation of C&D waste images. The result of this study would provide the reference for the design of the C&DWNet models for the fine classification of C&D waste and promote the idea for the intelligence of C&D waste management.

## 2 Materials and Methods

### 2.1 Data Collection


Fig. 1 Example of C&D waste: concrete, brick, stone, ceramic tile, glass, metal scrap, gypsum board, wood, plastic, and paper

The C&D waste image dataset was manually collected from Google search or taken by authors' cameras. This dataset with a total of 2836 images was manually grouped into 10 categories: concrete, brick, stone, ceramic tile, glass, metal scrap, gypsum board, wood, plastic and paper (Fig. 1).

**2.2 Data augmentation**

Measures of color space transformation, flipping, rotation, and noise injection were taken to augment image samples, as shown in Fig. S1. After data augmentation, the number of concrete, brick, stone, ceramic tile, glass, metal scrap, gypsum board, wood, plastic and paper images was enlarged to 5526, 10908, 1116, 2430, 2376, 10224, 4500, 8568, 3402 and 1944, respectively. The size of C&D waste images was changed from

2836 to 50992. In addition, the size of the training dataset, validation dataset and test dataset was 36711, 4080 and 10203, respectively. The pixel size of all C&D waste images was reshaped as 224×224 (height)×(width) for training the neural network.

### 2.3 C&DWNet

C&DWNet models, five ResNet structures (ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152) based on knowledge transfer, were proposed to classify ten types of C&D waste images.

#### 2.3.1 Deep residual network (ResNet)

ResNet was proposed by He et al. (2016), it presented the best performance in ImageNet classification. ResNet is a network-in-network architecture with a large number of stacked residual units (Garcia-Garcia et al., 2018). The various ResNet structures were given in SI (Fig. S2), which include two deep building blocks: bottleneck 1 and bottleneck 2. Bottleneck 1 was applied to stack the structure of ResNet-18 and ResNet-34, while ResNet-50, ResNet-101 and ResNet-152 were stacked by bottleneck 2. In comparison to bottleneck 1, bottleneck 2 includes three layers 1×1, 3×3 and 1×1 convolutional layer, where the function of 1×1 layer is applied to reduce and increase the dimension of input, making the bottleneck of 3×3 layer with small input/output dimensions.

Identity mapping is a measure to address the degradation problem. The details of how it works were introduced as follows (Fig. S2 and equation (2-3)). As shown in bottleneck 1 in Fig. S2, x, $F(x,\{w_i\})+x$, and $F(x,\{w_i\})$ are represented the input, output vectors, and residual mapping for learning, respectively. The y in equation (2-1) is equal to $F(x,\{w_i\}) + x$, which operation is conducted by a shortcut connection (Fulkerson,

1996) and element-wise addition.

$$y=F(x,\{w_i\})+x \tag{2-1}$$

What is noteworthy is that average pooling was introduced and linked to the fully connected layer in Conv5_x, where the activation function of the rectified linear unit (ReLu) was used to predict classes based on the highest probability given the input data, which can be expressed in the mathematics form (2-2):

$$\Pr(Y=i|v, W, b)= \mathrm{Softmax}_i(Wv)+b=\frac{e^{w_i v+b_i}}{\sum_j e^{w_j v+b_j}} \tag{2-2}$$

Where elements of w and b refer to the weights and bias, respectively. Index j is applied to normalize the posterior distribution. The model prediction was the class with the highest probability, as presented in Equation (2-3):

$$y_{prediction}= \mathrm{argmax}_i \Pr(Y= i \mid v, W, b) \tag{2-3}$$

The elements of weights and bias in deep ResNet structures were also optimized by the error backpropagation algorithm, which adopted an error metric to calculate the distance between true class labels and the predicted class labels. The Cross-Entropy function (2-4) was chosen as the loss function to be minimized for dataset V.

$$L=\frac{1}{N}\sum_{I}\sum_{c=1}^{M} y_{ic}\log(p_{ic}) \tag{2-4}$$

Where L denotes the loss function; Here, $V=\{v^{(1)},v^{(2)},v^{(3)},\ldots,v^{(n)}\}$ refers to the set of input samples in the training dataset; $Y=\{y^{(1)},y^{(2)},\ldots,y^{(10)}\}$ presents the corresponding labels: brick, …, wood.

### 2.3.2 Knowledge transfer

Usually, knowledge transfer may significantly enhance the performance of learning by reducing the efforts of data labeling. Considering the explanation of concepts of domain and task in knowledge transfer, for instance, given a source domain, ImageNet ($X_s$) and a corresponding source task, Image label ($Y_s$), there will be marginal probability distribution $P_X(x,y)$ between $X_s$ and $Y_s$. For the task of C&D waste sorting, there is also a target domain, the C&D waste image dataset ($X_T$), as well as a target task, the corresponding image label ($Y_T$). The main purpose of knowledge transfer is to learn the target probability distribution $P_T(x,y)$ in $X_T$ with the knowledge gained from $X_s$ and $Y_s$.

The knowledge transfer was applied to obtain pre-trained ResNet models. The weights and biases from ImageNet were also adopted. Fine-tuning the ResNet model was taken by truncating the original softmax layer and changing 1000 categories to 10 categories. This method can obtain the pre-trained parameters and deal with the task of C&D waste classification.

**2.4 Cyclical learning rate**

The cyclical learning rate includes base learning rate and max learning rate, cycle, step size, batch size, batch, or iteration. Base learning rate and max learning rate define the boundaries of a range, where the learning rate will fluctuate (Vidyabharathi et al., 2021). The value of base learning rate and max learning rate was set as $10^{-8}$ and 10, respectively. The triangular policy was adopted for the variations in cyclical learning rate. In this research, the value of cycle, step size, batch size and iteration was 100, 50, 32 and 100, respectively. The learning rate either increases or decreases based on the outcome from the latest batch in the epoch (Samudre et al., 2022). After getting the learning rate, combined with the data distribution, the learning rate of each experiment was set one by one. As shown in Fig. S3, the value of the learning rate for C&DWNet-18, C&DWNet-34, C&DWNet-50, C&DWNet-101 and C&DWNet-152 was $5.0\times10^{-4}$, $7.5\times10^{-4}$, $7.0\times10^{-4}$, $8.0\times10^{-4}$ and $7.5\times10^{-4}$, respectively.

**2.5 Model evaluation metrics**

Confusion matrix, recall, precision, F1 score, accuracy, Receiver Operating Characteristic (ROC), and Area Under the Curve (AUC) were used to evaluate the performance of C&DWNet. Recall, precision, F1 score, and accuracy were defined as follows:

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}} \tag{2-5}$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}} \tag{2-6}$$

$$\text{F1 score}=\frac{2\times\text{TP}}{2\times\text{TP+FP+FN}} \tag{2-7}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FN+TN+FP}} \tag{2-8}$$

$$\text{Sensivity} = \frac{\text{TP}}{\text{TP+TN}} \tag{2-9}$$

$$\text{Specificity} = \frac{\text{FN}}{\text{FN+FP}} \tag{2-10}$$

$$\text{Kappa} = \frac{n\sum_{1}^{10}(\text{TN}_i\times a_i)\text{-}\sum_{1}^{10}(a_i\times b_i)}{n^2\text{-}\sum_{1}^{10}(a_i\times b_i)} \tag{2-11}$$

Where TP, TN, FN, FP, n, $a_i$ and $b_i$ present the numbers of true positives, true negatives, false negatives, false positives, the number of the tested sample, the number of true samples of each class and the number of predicted samples of each class, respectively.

## 2.6 Visualization

### 2.6.1 PCA

PCA was employed to represent data of waste sorting in a low-dimensional way (Thomaz and Giraldi, 2010). The matrix X of m × n training includes m input samples (C&D waste images) and n pixels (or variables). The covariance matrix $C$ of the data matrix can be obtained by Equation (2-12):

$$C = \frac{1}{\mathrm{m}} \mathrm{XX}^{\mathrm{T}} \tag{2-12}$$

The eigenvector P and eigenvalue Λ of covariance matrix $C$ can be calculated according to Equation (2-13):

$$\mathrm{P}^{\mathrm{T}} C \mathrm{P} = \Lambda \tag{2-13}$$

Therefore, such a set of eigenvectors P for the training set matrix X was regarded as the principal component.

### 2.6.2 t-SNE

t-SNE is a kind of technology that can reduce dimensionality tasks by minimizing the divergence between the pairwise similarity distribution of input points and low-dimensional embedding points (Maaten and Hinton, 2008; Retsinas et al., 2017). Considering the input points as $\{x_1, x_2, x_3, ..., x_n\}$ and their corresponding embedding points as $\{y_1, y_2, y_3, ..., y_n\}$, the pairwise similarity between points $x_i$ and $x_j$ can be obtained by using the joint probability, $p_{i,j}$, as presented in Equations (2-14) and (2-15):

$$p_{j|x} = \frac{\exp(-d(x_i, x_j)^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-d(x_i, x_j)^2 / 2\sigma_i^2)} \tag{2-14}$$

$$P_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \tag{2-15}$$

where $d(x_i,x_j)$ represents Euclidean distance function.

Using a similar method to get the pairwise similarity between points $y_i$ and $y_j$ in the embedding space (Equations (2-16)).

$$q_{j|x}=\frac{\exp(-\left\|y_i\text{-}y_j\right\|^2)}{\sum_{k\neq i}\exp(-\left\|y_i\text{-}y_j\right\|^2)} \tag{2-16}$$

The Kullback-Leibler divergence was minimized to calculate the embedding Y, which considers the pairwise similarity distribution for both initial and the embedding spaces as follows:

$$C(Y)=KL(P/\!/Q)=\sum_i\sum_j p_{ij}\log\frac{p_{ij}}{q_{ij}} \tag{2-17}$$

Gradient descent method was used to find an embedding Y to minimize the divergence, while the gradient of the divergence for each point of the embedding space was obtained according to Equation (2-18):

$$\frac{\delta C}{\delta y_i}=2\sum_j\left(p_{j|i}\text{-}q_{j|i}+p_{i|j}\text{-}q_{i|j}\right)\left(y_i\text{-}y_j\right) \tag{2-18}$$

**2.7 Research flow and Experimental platform**

The details of the research flow and the experimental platform were presented in Fig. S4 and Table 1, respectively.

Table1 Experimental platform for training C&DWNet models

| Item | Parameters |
|---|---|
| CPU | Intel (R) Core (TM) i9-10900K @ 3.70 GHz |
| Language | Python 3.8; Pytorch 1.7.1+cul10 |
| Hard drive | 2T |

| | |
|---|---|
| Operating system | Window 10 |
| Random-Access Memory (RAM) | 128 G |
| Graphic Processing Unit (GPU) | NVIDIA GeForce RTX 3080 |

# 3 Results and Discussion

## 3.1 Effect of knowledge transfer on the performance of C&DWNet-18 for the classification of C&D waste

Fig. S5 shows the effect of knowledge transfer (KT) on the performance of C&DWNet-18. As shown in Fig. S5 a) and b), the average training time in C&DWNet-18 without KT and C&DWNet-18 is 951.5 s and 946.7 s, respectively. It means that KT can slightly increase the efficiency of C&DWNet model training. As shown in Fig. S5 c) and d), the training loss and validation loss in C&DWNet-18 are much lower than in C&DWNet-18 without KT. They decrease with the increase of the epoch number in C&DWNet-18, which decreases from 1.48 (at epoch 1) to 0.12 (at epoch 10). While the training loss and validation loss in C&DWNet-18 increase from 0.54 (at epoch 1) to 0.65 (at epoch 3) and then decrease to 0.04 (at epoch 10).

The accuracy rates of the training dataset and validation dataset present an upward trend with the increasing epoch (Fig. S5 e) and f)). The accuracy from C&DWNet-18 in the training dataset and validation dataset is more than 80% since the $1^{st}$ epoch and reaches 99.99% in the $10^{th}$ epoch. While the accuracy of C&DWNet-18 without KT is less than 50% at the $1^{st}$ epoch and increases to 99.83% in the $10^{th}$ epoch. The result suggested that KT could promote efficiency and enhance the accuracy of C&D waste sorting.

In comparison to C&DWNet-18 without KT in Fig. S5 g), more C&D waste images like brick (1619), concrete (795), glass (409), gypsum board (676), metal scrap (1719), paper (266), plastic (308) and wood (1343), were found on the diagonal line in C&DWNet-18 in Fig. S5 h), indicating that KT can enhance the performance of C&DWNet-18 on the test dataset.

As shown in Fig. S5 i) and j), The area under the ROC curve (AUC) is regarded as an indicator for the effect of classification. The macro-average AUC of C&DWNet-18 (0.82) is higher than the C&DWNet-18 without KT (0.77), indicating that KT can improve the effect of C&D waste sorting.

The precision, recall, F1 score, accuracy, sensitivity, specificity and kappa were used to assess the performance of the C&DWNet-18 without KT and C&DWNet-18 model. The performance evaluation of C&DWNet-18 without KT and the C&DWNet-18 model was given in Table 2. The accuracy of C&DWNet-18 without KT on the C&D waste test dataset is 64.7%. C&DWNet-18 has a better performance on C&D waste sorting in terms of accuracy (73.3%), precision (73.7%), recall (73.3%), F1 score (73.1%), sensitivity (73.4%), specificity(4.9%) and kappa (69.8%).

In conclusion, the method of KT can shorten the training time and improve the performance of the C&DWNet-18 model on the classification of C&D waste.

Table 2 Performance evaluation of C&DWNet-18 without/with knowledge transfer on C&D waste test dataset (Note: FS, CT, GB, MS and WA represent F1 Score, Ceramic Tile, Gypsum Board, Metal Scrap and Weighted Average )

| C&DW Categories | C&DWNet-18 without knowledge transfer | | | | | | | C&DWNet-18 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Precision | Recall | FS | Accuracy | Sensitivity | Specificity | Kappa | Precision | Recall | FS | Accuracy | Sensitivity | Specificity | Kappa |
| Brick | 0.704 | 0.632 | 0.666 | 0.647 | 0.703 | 0.133 | 0.587 | 0.794 | 0.742 | 0.767 | 0.733 | 0.794 | 0.088 | 0.698 |
| CT | 0.568 | 0.570 | 0.569 | | 0.568 | 0.032 | | 0.636 | 0.558 | 0.594 | | 0.590 | 0.029 | |
| Concrete | 0.568 | 0.657 | 0.610 | | 0.568 | 0.061 | | 0.554 | 0.719 | 0.626 | | 0.554 | 0.044 | |
| Glass | 0.679 | 0.674 | 0.677 | | 0.678 | 0.024 | | 0.911 | 0.859 | 0.884 | | 0.911 | 0.008 | |
| GB | 0.649 | 0.671 | 0.660 | | 0.648 | 0.047 | | 0.830 | 0.751 | 0.789 | | 0.830 | 0.029 | |
| MS | 0.726 | 0.792 | 0.758 | | 0.726 | 0.079 | | 0.759 | 0.841 | 0.798 | | 0.759 | 0.042 | |
| Paper | 0.653 | 0.517 | 0.577 | | 0.653 | 0.029 | | 0.806 | 0.684 | 0.740 | | 0.806 | 0.016 | |
| Plastic | 0.514 | 0.278 | 0.360 | | 0.514 | 0.071 | | 0.653 | 0.452 | 0.534 | | 0.653 | 0.048 | |
| Stone | 0.259 | 0.321 | 0.287 | | 0.259 | 0.023 | | 0.365 | 0.308 | 0.334 | | 0.365 | 0.021 | |
| Wood | 0.645 | 0.709 | 0.675 | | 0.644 | 0.084 | | 0.753 | 0.784 | 0.768 | | 0.752 | 0.047 | |
| WA | 0.647 | 0.647 | 0.643 | - | 0.647 | 0.078 | - | 0.737 | 0.733 | 0.731 | - | 0.734 | 0.049 | - |

### 3.2 Comparison of various C&DWNet structure performance in training and validation dataset

Fig. S5 shows C&DWNet performances on various cases of C&D waste sorting. The average training time increases along with the increase of the layer of C&DWNet architecture. Namely, the training time in C&DWNet-18, C&DWNet-34, C&DWNet-50, C&DWNet-101 and C&DWNet-152 is 946.7, 951.6, 968.2, 1042 and1186.6 s, respectively. This occurrence is due to the fact that much more parameters are needed to be trained, as the layer increases.

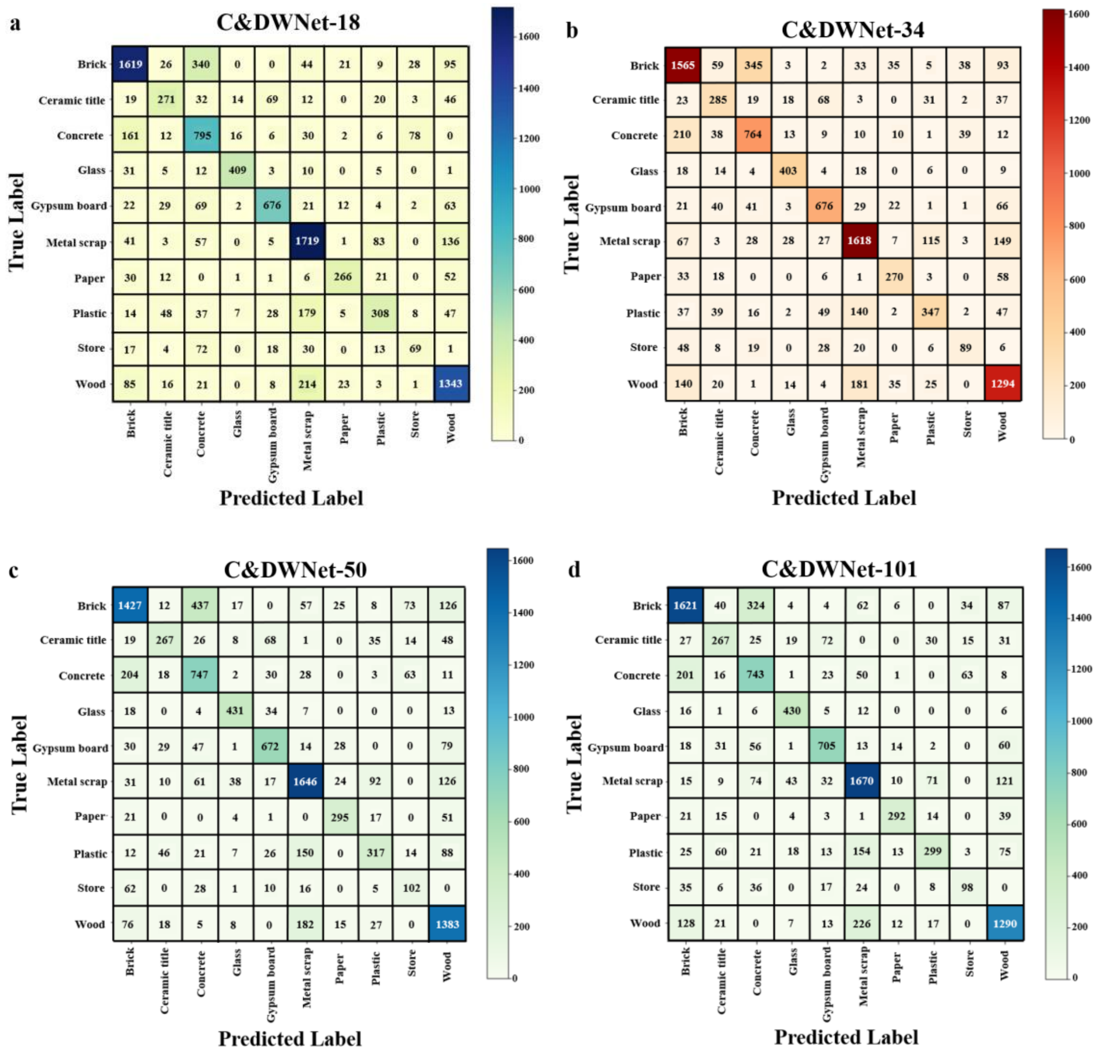
The loss in C&DWNet-34, C&DWNet-50 and C&DWNet-101 shows a similar trend, it decreases with the epoch number increase. By contrast, the loss in C&DWNet-18 and C&DWNet-152 increases at the early stage and then decreases. The trend of accuracy in various C&DWNet architectures has little difference with the iteration of the epoch.

### 3.3 Model performance in the test datasets

#### 3.3.1 Confusion matrix

The confusion matrix of the assessment used for C&DWNet models' (C&DWNet-18, C&DWNet-34, C&DWNet-50, C&DWNet-101 and C&DWNet-152) performance was shown in Fig. 2. The values of the C&D waste on the diagonal line represent correct classification, in contrast, the values outside the diagonal line represent unpaired labels and images. For example, the brick in C&DWNet-152, 1620 images were accurately sorted, on the contrary, 33 images of ceramic tile, 321 images of concrete, 2 images of gypsum board, 30 images of metal scrap, 10 images of paper, 3 images of plastic, 25 images of stone and 138 images of wood were wrongly identified as the brick (Fig. 2 e)).

Compared with other C&DWNet models in Fig. 2, most the C&D waste like brick (1620), ceramic tile (290), glass (439), plastic (346), stone (101) and wood (1410) are found on the diagonal line in C&DWNet-152, indicating that C&DWNet-152 can provide better performance on these C&D waste sorting. While most of the gypsum board (705) and paper (295) are found on the diagonal line in C&DWNet-101 and C&DWNet-50, respectively. As for concrete and metal scrap, C&DWNet-18 can provide better performance. Thus, different C&DWNet models are suitable for different kinds of C&D waste.



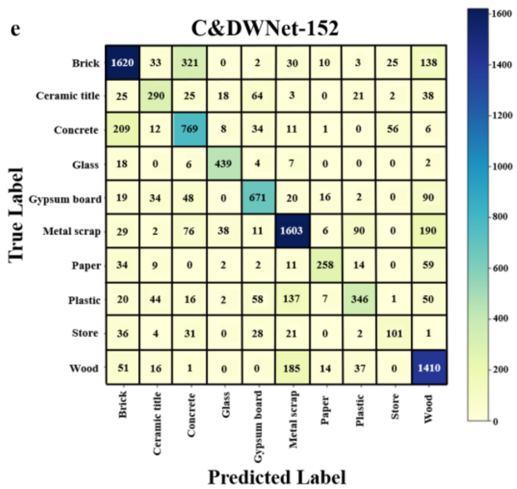



Fig. 2 Confusion matrices of C&DWNet model performance: a) C&DWNet-18; b) C&DWNet-34; c) C&DWNet-50; d) C&DWNet-101; e) C&DWNet-152

### 3.3.2 Accuracy, Precision, Recall, F1 score, Sensitivity, Specificity and Kappa

Several indexes were applied to quantitatively assess the performance of C&DWNet models for C&D waste sorting (Fig. 3). The accuracy of various C&DWNet models is approximately at 72~74%, the best accuracy is 73.6% in C&DWNet-152 in Fig. 3 a). However, there are limitations for the indicator of accuracy due to the unbalanced data (Dhillon and Verma, 2019). Thus, precision, recall, F1 score, sensitivity, specificity and kappa were also applied to assess the performance of C&DWNet models.

Precision represents the proportion of correctly predicted positive items to the total predicted items. As shown in Fig. 3 b), C&DWNet-18 and C&DWNet-152 show a better performance than other C&DWNet models in terms of weighted average at about (79%). Although both two models show a good performance, they still have room for improving their performance, especially for glass (36.5%) and wood (55.4%) in C&DWNet-18.

The recall value refers to the number of positive items correctly identified. As shown in Fig. 3 c), the weighted average of C&DWNet-18, C&DWNet-101 and C&DWNet-152 are at about 74%, these three models show a similar and a little better performance than other C&DWNet models.

F1 score represents a balance between recall value and precision. As shown in Fig. 5 d), the weighted average of F1 score is followed in this order: 76.7% (C&DWNet-18) > 76.4% (C&DWNet-152) > 75.6% (C&DWNet-101) > 72.1% (C&DWNet-34) > 69.9% (C&DWNet-50). This means that most of the C&D waste can be well classified by the model of C&DWNet-18 and C&DWNet-152. However, the F1 score of the gypsum board and glass in all the C&DWNet models are less than 60%, their classification performance needs to be improved.

The sensitivity of various C&DWNet models is almost at 71.6~73.6%, the best sensitivity is 73.6% in C&DWNet-152 in terms of weighted average in Fig. 3 e), while the specificity of various models follows this order in weighted average: C&DWNet-50 (5.6%) > C&DWNet-34 (5.5%) > C&DWNet-101 (5.1%) > C&DWNet-152 (5.0%) > C&DWNet-18 (4.9%). In addition, the index of kappa was also taken to evaluate the performance of various C&DWNet models, as shown in Fig. S6. The max value of kappa was 69.8% in C&DWNet-18, followed by 66.8% in C&DWNet-34, 66.6% in C&DWNet-101 and 66.5% in C&DWNet-18, and the mix value of kappa was 65.9% in C&DWNet-101, which indicated the result of various C&DWNet models showed well consistency (Hayden and Ghosh, 2014).

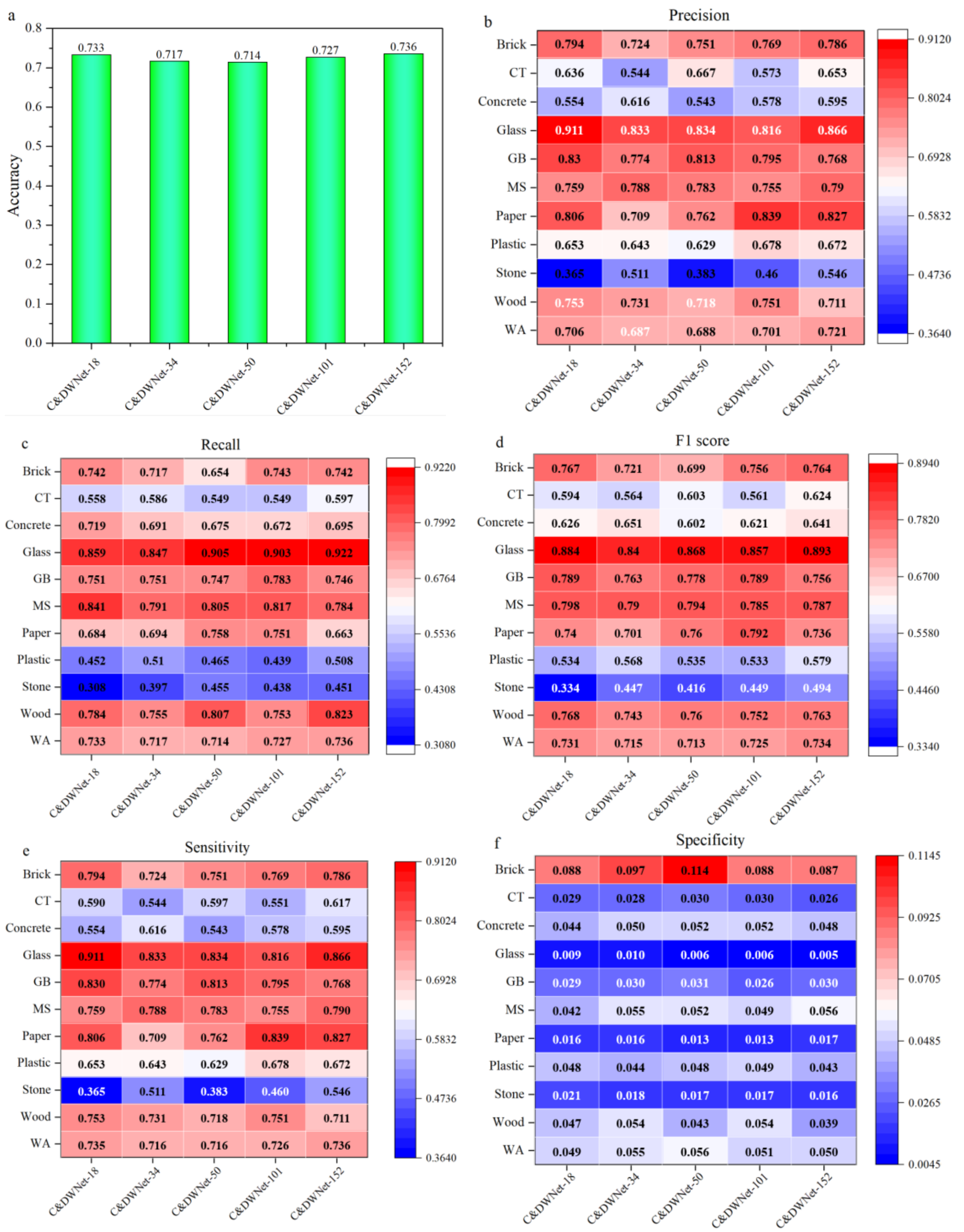



Fig. 3 Comparison of different C&DWNet performances in terms of accuracy, precision, recall, F1 score, sensitivity, specificity and F1 score: a) Accuracy; b) Precision; c) Recall; d) F1 score; e) Sensitivity; f) Specificity. Note: CT, GB, MS and WA represent ceramic tile, gypsum board, metal scrap and weighted average, respectively.

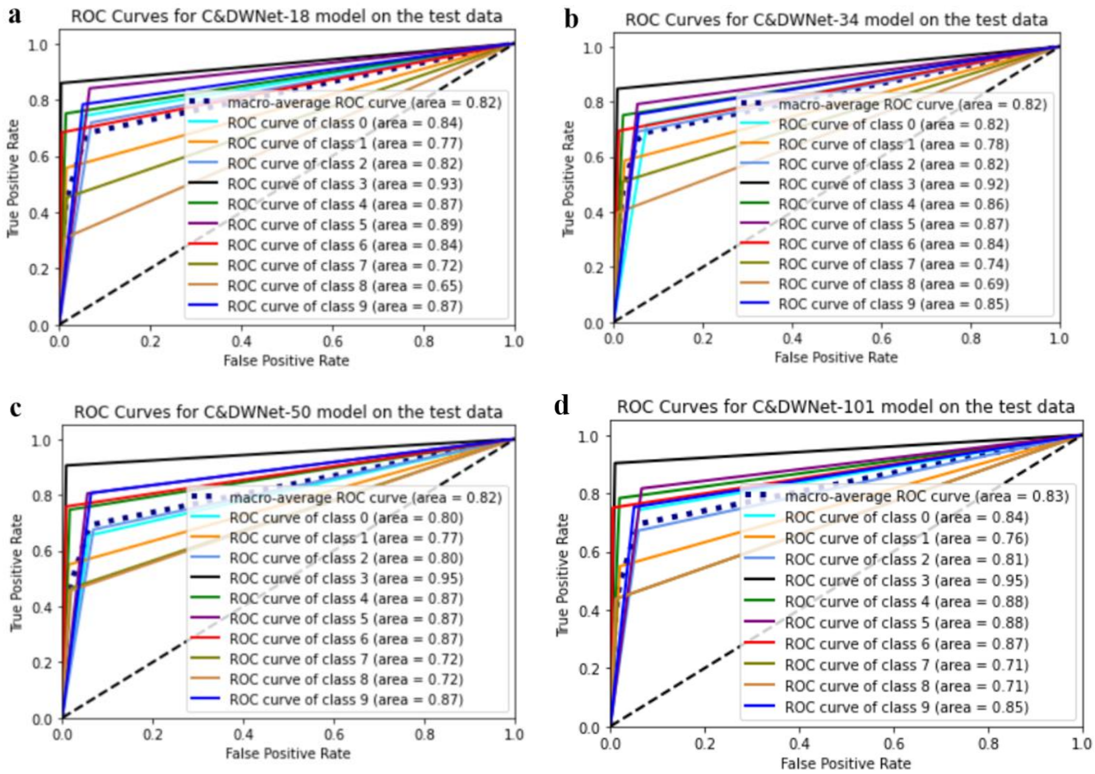
### 3.3.3 ROC

The area under the ROC curve (AUC) is an indicator for assessing the classification performance (Ahmad et al., 2020; Zhang et al., 2021). All the macro average AUC of C&DWNet models are 0.82 (Fig. 4). In Fig. 6, class 0, class 1, class 2, class 3, class 4, class 5, class 6, class 7, class 8, class 9 represent brick, ceramic tile, concrete, glass, gypsum board, metal scrap, paper, plastic, stone, wood, respectively.

As shown in Fig. 4, all the C&DWNet models show a similar performance. Almost all the AUC values are higher than 0.8, except for the AUC value of ceramic tile, plastic, and stone. The maximum AUC value of ceramic tile, plastic, and stone are 0.79, 0.75, and 0.72, respectively. The reason is that the ceramic tile, and stone waste images often were wrongly identified as gypsum board and concrete, respectively (Fig. S7). Therefore, the results demonstrate that most C&D waste can be well classified by the C&DWNet models, and the classification effect of ceramic tile, plastic and stone should be improved.

On the other hand, the result also suggested that different C&DWNet models show a good performance in different kinds of C&D waste. This is different from the conclusion of the ImageNet application, which indicated that the deeper ResNet structures, the better performance. This phenomenon can be ascribed to the dataset of C&D waste being not as complex as these from the ImageNet (Yang et al., 2021). Those results suggested that the structure and depth choice of C&DWNet models should be made according to the practical application.

Considering the index of accuracy, precision, recall, F1 score, sensitivity, specificity, kappa and AUC, the performance of C&DWNet models does not show obvious improvement in C&D waste sorting, but the training time would increase with the layer of increase. This would spend more and consume resources. Therefore, C&DWNet-18 is more suitable for the classification of C&D waste.



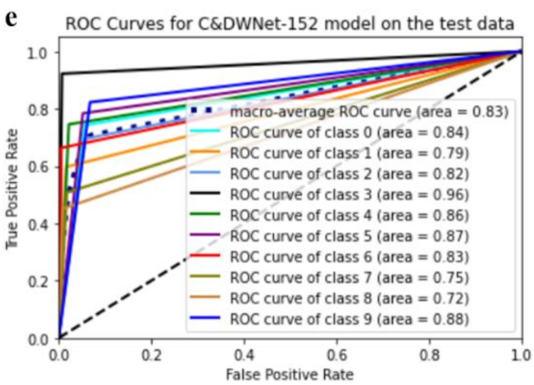



Fig. 4 ROC from C&DWNet models: a) C&DWNet-18; b) C&DWNet-34; c) C&DWNet-50; d) C&DWNet-101; e) C&DWNet-152. Note: Class 0, Class 1, Class 2, Class 3, Class 4, Class 5, Class 6, Class 7, Class 8, Class 9 represent brick, ceramic tile, concrete, glass, gypsum board, metal scrap, paper, plastic, stone, wood, respectively.

### 3.4 Visual explanation by using PCA and t-SNE

Methods of PCA and t-SNE were applied to present the distribution of the C&D waste image dataset in the C&DWNet-18 model. Fig. 5 a) shows 2-dimension extracted representations from the last layer of the C&DWNet-18 model obtained from the PCA algorithm. The features from the C&DWNet-18 model demonstrate an obscure semantic clustering. There is some overlap of C&D waste's features, which means there exists confusion.

Fig. 5 b) presents C&D waste representations that were better separated by t-SNE in comparison to PCA. The features of brick, ceramic tile, concrete, glass, gypsum board, metal scrap, paper, plastic, stone and wood were distinctly separated. Therefore, considering the advantages of t-SNE, it can be applied to accurately present the features of C&D waste sorting (Gisbrecht et al., 2015).

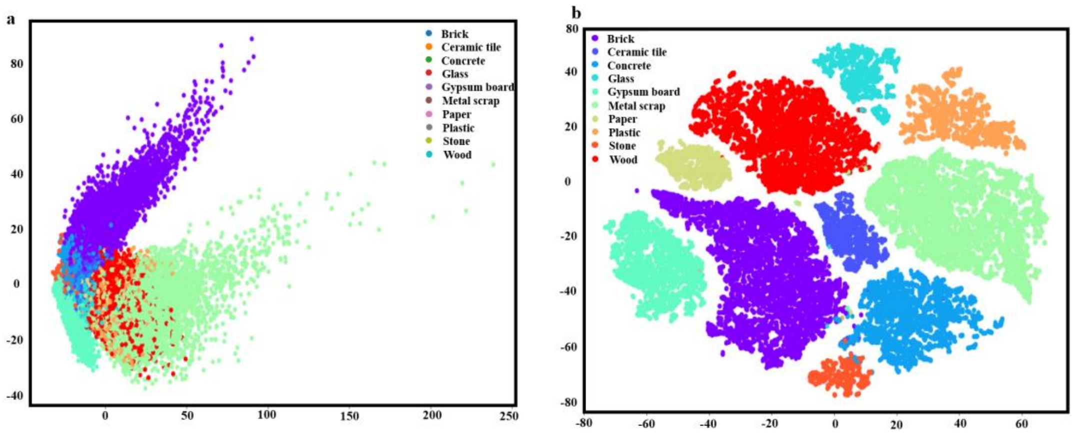



Fig. 5 A 2-D feature visualization of an image representation of waste images by the method of PCA and t-SNE for the last layer of C&DWNet-18: a) PCA; b) t-SNE. Note: Each color illustrates a different class in the dataset. PCA refers to Principal component analysis; t-SNE refers to t-distributed stochastic neighbor embedding

### 3.5 Environmental implications and future perspective

The study suggested that different C&DWNet models show a good performance in different kinds of C&D waste. Namely, the structure of C&DWNet can be designed according to the composition of C&D waste. For example, C&DWNet-152 may be a good choice for the classification of glass (F1 score was 0.893). C&DWNet module can be implemented with a mechanical arm to realize C&D waste sorting automatically. While this is related to C&D waste classification, object detection, and semantic segmentation, they would be explored in the future. In addition, the C&DWNet module can be also integrated with GIS or Drone to detect the behavior of dumping to strengthen the enforcement efforts.

However, the C&DWNet model has some limitations, which can be improved in future work. The training time of C&DWNet is slower due to the millions of parameters that are needed to be trained, it can be adjusted according to the structure of the ResNet model, aiming to reduce the number of parameters. In addition, the performance of the C&DWNet model still has great room for improvement, which can be combined with other algorithms like genetic algorithms, DenseNet and VGGNet to enhance accuracy, precision and recall.

## 4 Conclusion

C&DWNet models, five ResNet structures (ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152) based on knowledge transfer, were proposed to classify ten types of C&D waste. The cyclical learning rate was applied to quickly find the best global learning rate. The results showed that KT can reduce the training time and improve the performance of the C&DWNet model. The average training time increases with the increase of the layer of C&DWNet architecture from C&DWNet-18 (946.7 s) to C&DWNet-152 (1186.6 s). The accuracy of various C&DWNet models is approximately 72~74%, the best accuracy is 73.6% in C&DWNet-152. C&DWNet-18 is more suitable for the classification of C&D waste. The structure and depth choice of C&DWNet models should be made according to the practical application. Moreover, in comparison to PCA, the algorithm of t-SNE can distinctly separate each type of C&D waste. In addition, the C&DWNet module can be also integrated with

GIS or Drone to detect the behavior of dumping to strengthen the enforcement efforts. Through the C&D waste classification, it helps to promote the development of the circular economy. The code is available on: (Annyulin/C-D-waste-classification-by-ResNet (github.com)).

## Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2019YFC1904001), the National Natural Science Foundation of China (No. 52000143; No. 51878470), the Shanghai Municipal Science and Technology Major Project (2021SHZDZX0100) and the Fundamental Research Funds for the Central Universities, and the International Postdoctoral Exchange Fellowship Program (YJ20200280).

## References

Achu, A.L., Thomas, J., Aju, C.D., Gopinath, G., Kumar, S., Reghunath, R., 2021. Machine-learning modelling of fire susceptibility in a forest-agriculture mosaic landscape of southern India. Ecological Informatics 64.

Ahmad, K., Khan, K., Al-Fuqaha, A., 2020. Intelligent fusion of deep features for improved waste classification. IEEE Access 8, 96495-96504.

Aral, R.A., Keskin, Ş., Kaya, M., Hacıömeroğlu, M., 2018. Classification of trashnet Dataset based on deep learning models, IEEE International Conference On Big Data. IEEE, Seattle, WA, USA.

Bobulski, J., Kubanek, M., 2021. Deep learning for plastic waste classification system. Applied Computational Intelligence and Soft Computing 2021, 7.

Dhillon, A., Verma, G.K., 2019. Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial

Intelligence 9(2), 85-112.

Duan, H., Li, J., 2016. Construction and demolition waste management: China's lessons. Waste Manag Res 34(5), 397-398.

Duan, H., Wang, J., Huang, Q., 2015. Encouraging the environmentally sound management of C&D waste in China: An integrative review and research agenda. Renewable and Sustainable Energy Reviews 43, 611-620.

Frost, S., Tor, B., Agrawal, R., G.Forbes, A., 2019. CompostNet: An image classifier for meal waste, IEEE Global Humanitarian Technology Conference (GHTC). pp. 1-4.

Fulkerson, B., 1996. Pattern Recognition and Neural Networks. Cambridge university press.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing 70, 41-65.

Gisbrecht, A., Schulz, A., Hammer, B., 2015. Parametric nonlinear dimensionality reduction using kernel t-SNE. Neurocomputing 147, 71-82.

Hayden, M.S., Ghosh, S., 2014. Regulation of NF-kappaB by TNF family cytokines. Semin Immunol 26(3), 253-266.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA pp. 770-778.

Huang, B., Gao, X., Xu, X., Song, J., Geng, Y., Sarkis, J., Fishman, T., Kua, H., Nakatani, J., 2020. A Life Cycle Thinking Framework to Mitigate the Environmental Impact of Building Materials. One Earth 3(5), 564-573.

Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. Journal of Hydrology 573, 311-323.

Lin, K., Zhao, Y., Kuo, J.-H., Deng, H., Cui, F., Zhang, Z., Zhang, M., Zhao, C., Gao, X., Zhou, T., Wang, T., 2022. Toward smarter management and recovery of municipal solid waste: A critical review on deep learning approaches. Journal of Cleaner Production 346.

Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-SNE. Journal of Machine Learning Research 9, 2579-2605.

Mao, W.-L., Chen, W.-C., Wang, C.-T., Lin, Y.-H., 2021. Recycling waste classification using optimized convolutional neural network. Resources, Conservation and Recycling 164.

N.Smith, L., 2017. Cyclical learning rates for training neural networks, IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, Santa Rosa, CA, USA

Retsinas, G., Stamatopoulos, N., Louloudis, G., Sfikas, G., Gatos, B., 2017. Nonlinear manifold embedding on keyword spotting using t-SNE, International Conference on Document Analysis and Recognition (ICDAR). pp. 487-492.

Samudre, A., George, L.T., Bansal, M., Wadadekar, Y., 2022. Data-efficient classification of radio galaxies. Monthly Notices of The Royal Astronomical Society 509(2), 2269-2280.

Sinno, J.P., Qiang, Y., 2010. A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering 22(10), 1345-1359.

Sreelakshmi, K., Vinayakumar, R., Soman, K.P., 2019. Deep segregation of plastic (DSP): Segregation of plastic and nonplastic using deep learning, Big Data Recommender Systems - Volume 1: Algorithms, Architectures, Big Data, Security and Trust. pp. 169-191.

Thomaz, C.E., Giraldi, G.A., 2010. A new ranking method for principal components analysis and its application to face image analysis. Image and Vision Computing 28(6), 902-913.

Vidyabharathi, D., Mohanraj, V., Kumar, J.S., Suresh, Y., 2021. Achieving generalization of deep learning models in a quick way by adapting T-HTR learning rate scheduler. Personal and Ubiquitous Computing.

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

Y.Bengio, 2012. Neural networks: Tricks of the trade, chapter practical recommendations for gradient-based training of deep architectures, Springer Berlin Heidelberg.

Yan, B., Liang, R., Li, B., Tao, J., Chen, G., Cheng, Z., Zhu, Z., Li, X., 2021. Fast identification and characterization of residual wastes via laser-induced breakdown spectroscopy and machine learning. Resources, Conservation and Recycling 174.

Yanai, K., Kawano, Y., 2015. Food image recognition using deep convolutional network with pre-training and fine-tuning, IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, Turin, Italy, pp. 1-6.

Yang, K., Yang, T., Yao, Y., Fan, S.-d., 2021. A transfer learning-based convolutional neural network and its novel application in ship spare-parts classification. Ocean & Coastal Management 215.

Yang, M., Thung, G., 2016. Classification of trash for recyclability status, CS229 Projection report 2016. pp. 940-945.

Zhang, H., Wang, K., Tian, Y., Gou, C., Wang, F.-Y., 2018. MFR-CNN: Incorporating multi-scale features and global information for traffic object detection. IEEE Transactions on Vehicular Technology 67(9), 8019-8030.

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