Applying a deep residual network coupling with transfer learning for recyclable waste sorting

Kunsen Lin1 · Youcai Zhao1,2 · Xiaofeng Gao3 · Meilan Zhang1 · Chunlong Zhao1 · Lu Peng1 · Qian Zhang4 · Tao Zhou1,2

Received: 16 March 2022 / Accepted: 19 July 2022 / Published online: 26 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Recyclable waste sorting has become a key step for promoting the development of a circular economy with the gradual realization of carbon neutrality around the world. This study aims to develop an intelligent and efficient method for recyclable waste sorting by the method of deep learning. Thus, RWNet models, which refers to various ResNet structures (ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152) based on transfer learning, were proposed to classify different types of recyclable waste. Cyclical learning rate and data augmentation were taken to improve the performance of RWNet models. In addition, accuracy, precision, recall, F1 score, and ROC were taken to evaluate the performance of RWNet models. Results showed that the accuracy of various RWNet models is almost at 88%, and the best accuracy is 88.8% in RWNet-152. The highest precision, recall, and F1 score in terms of weighted average value appeared in RWNet-101 (89.9%), RWNet-152 (88.8%), and RWNet-152 (88.9%), respectively. The area under the ROC curve (AUC) is higher than 0.9, except for the AUC value of plastic (0.85), which indicated that most of the recyclable waste can be well sorted by RWNet models. This study demonstrates the good performance of RWNet models that can be used to automatically sort most of the recyclable waste, which paves the way for better recyclable waste management.

Keywords Recyclable waste · Waste classification · Waste management · Deep learning · Transfer learning

Abbreviations
AUC Area under the curve
CNN Convolutional neural network

Responsible editor: Guilherme L. Dotto

Xiaofeng Gao
gaoxiaofeng@cqu.edu.cn
Kunsen Lin
kslin@tongji.edu.cn

1 The State Key Laboratory of Pollution Control and Resource Reuse, College of Environmental Science and Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China
2 Shanghai Institute of Pollution Control and Ecological Security, 1515 North Zhongshan Rd. (No. 2), Shanghai 200092, People’s Republic of China
3 Key Laboratory of the Three Gorges Reservoir Region’s Eco-Environment, Ministry of Education, College of Environment and Ecology, Chongqing University, Chongqing 400045, China
4 Robert M. Buchan Department of Mining, Queen’s University, Kingston K7L 3N6, Canada

FN False negative
FP False positive
ISWA International Solid Waste Association
MSW Municipal solid waste
PCA Principal component analysis
ResNet Residual network
ROC Receiver operating characteristic
RWNet ResNet based on transfer learning for recyclable waste sorting
SDG Sustainable development goals
SVM Support vector machine
TN True negative
TP True positive
TL Transfer learning
t-SNE T-Distributed stochastic neighbor embedding
UNEP United Nations Environment Programme
USEPA US Environmental Protection Agency
Introduction

The rapid development of industrialization and urbanization has brought considerable quantities of municipal solid waste (MSW). The amount of global MSW generation already reached about 2.01 billion tons in 2020, and it is expected to reach 2.59 billion tons in 2030 and 3.40 billion tons in 2050 (Ram et al. 2021). In developed countries, 16% of the total population generates about 34% of the global MSW, and the MSW per capita is close to 2.1 kg/day, which is higher than the average of the world (0.74 kg/day) (Jaunich et al. 2019). Even for low- and middle-income countries, such as Africa and parts of Asia, it will also grow fast and account for about 35% of the global MSW (Kaza and Bhada-Tata 2018). However, the current MSW treatment capacity is insufficient. United Nations Environment Programme (UNEP) pointed out that this has become a significant problem affecting economic development, human settlements, and public health (Harris et al. 2021). To improve the level of global governance, UNEP and the International Solid Waste Association (ISWA) jointly released the “Global Waste Management Outlook,” trying to propose global solutions for solid waste management. Moreover, the United Nations released the 2030 Global Agenda for Sustainable Development, which proposed 17 Sustainable Development Goals (SDGs) and 169 specific goals (Schmidt-Traub et al. 2017). Especially, MSW management is the key goal target of SDG11, and waste reduction and recycling were also pointed out in SDG12 (Brunner and Rechberger 2015). MSW management would have directly impact on human life, economic, and business communication (Abbas et al. 2019a; Aman et al. 2021; Hussain et al. 2021; Yu et al. 2022). Therefore, it is urgent to enhance the level of MSW management.

It is worth noting that about 43–59% of MSW in developed countries and 26–47% of MSW in developing countries is relatively high-value components (so-called recyclable waste). Recyclable waste management is also emphasized by the circular economy, which is low consumption and low emissions and is also in line with the SDGs (Schmidt-Traub et al. 2017). The US Environmental Protection Agency (USEPA) already released the “National Recycling Strategy,” which pointed out that recyclable waste sorting should be rapidly improved (Regan 2021). Thus, recyclable waste sorting has become a key step for promoting the development of the circular economy with the gradual realization of carbon neutrality around the world.

In the “National Recycling Strategy” of USEPA, it indicated that efficiencies of recyclable waste sorting should be enhanced, including reducing the cost of sorting, decreasing the number of residual waste, and improving the quality and quantity of recyclables (Regan 2021). The classification of recyclable waste is related to the social responsibility, the propagation of social media and internet, especially technology innovation (Abbas et al. 2021; Hussain et al. 2019; Li et al. 2021b; Mubeen et al. 2021a). Recently, great attention has been caught on smart classification with the development of the Internet of Things, electronic technology, and control technology. Image recognition plays a crucial role in intelligent classification and automation. The method of traditional machine learning like support vector machine (SVM) (Wang et al. 2019), laser-induced breakdown spectroscopy and stoichiometric tools (Costa et al. 2017), and Fourier transform infrared spectroscopy coupled with principal component analysis (PCA) (Chophi et al. 2021) to realize image/data process. However, these methods are needed to extract hand-designed features, which show poor discrimination ability and weak generalization for high-level features (Zhang et al. 2021a).

Recently, deep learning has emerged as a powerful method for high-level features, which is being widely used in image classification, object detection, and semantic segmentation (Alom et al. 2019). In comparison to the traditional image classification method (Alom et al. 2019), by automating the extraction of representations, deep learning algorithms have strong robustness to stand up against noise, which is more robust, scalable, and generalized (Yan et al. 2021). Deep learning like a convolutional neural network (CNN) has already been applied in recyclable waste sorting (Adedeji and Wang 2019) and food waste sorting (Frost et al. 2019). Although the deep residual network (ResNet) is being widely used in image classification in other fields (Cheng et al. 2021, Ghosh and Ghosh 2022), with the capacity to address the problem of gradient degradation in network performance by identity mapping (He et al. 2016), limited literature has reported applying the residual network (ResNet) model to waste classification. Notably, the ResNet model is a data-driving model, which required a large number of data. Fortunately, transfer learning can be applied to solve this bottleneck (Pan and Yang 2009). In addition, the selection of hyper-parameters like learning rate is challenging and cumbersome work, as small or large learning rates will cause slow convergence or divergence in training algorithms (Bengio 2012). This problem could be overcome by the method of cyclical learning rate (Smith 2017). Deep learning is considered the black box, but integrating the t-distributed stochastic neighbor embedding (t-SNE) (Retsinas et al. 2017) and PCA (Serranti et al. 2012) could make the deep learning model visible.

Few studies used the method of transfer learning and cyclical learning rate with deep learning model to realize the finely classified the recyclable waste. This study aims to develop an intelligent and efficient method for recyclable
waste sorting through deep learning. Thus, RWNet models, various ResNet structures (ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152) based on transfer learning, combined with the technology of data augmentation and cyclical learning rate, was proposed to classify different types of recyclable waste images in this paper. Considering the shortcomings of the accuracy assessment method due to the unbalanced data, precision, and recall, F1 score and receiver operating characteristic (ROC) also were applied to evaluate the performance of RWNet models. Algorithms of t-SNE and PCA were adopted to extract the representations of recyclable waste from high dimension to low dimension. The main contributions of this article are summarized as follows: A recyclable waste image dataset with the 45,486 images was collected by the method of data augmentation. The experimental results indicated that the method of cyclical learning rate and transfer learning to improve the performance of RWNet is practical and feasible.

Materials and methods

Data collection

The datasets were collected from open source (https://github.com/garythung/trashnet), which contained cardboard, glass, metal, paper, plastic, and trash. The number of these categories was 403, 501, 410, 594, 482, and 137, respectively (Fig. 1).

Data augmentation

Considering recyclable waste in the environment is random and irregular, measures of color space transformation, rotation, flipping, and noise injection have been taken to augment the image sample (Abbas et al. 2019b; Abbasi et al. 2021; Mubeen et al. 2021b; Zhang et al. 2022), as shown in Fig. S1. After data augmentation, the numbers for cardboard, glass, metal, paper, plastic, and trash were enlarged to 7254, 9018, 7380, 10,692, 8676, and 2466, respectively. The pixel size of images was reshaped as 224×224 (height)×(width) for efficient training of the deep learning neural network. In addition, the number of the training dataset, validation dataset, and test dataset was 32,747, 3639, and 9100, respectively.

The experimental platform is given in SI (Table S1).

Convolutional neural network (CNN)

Process of convolution layer and pooling layer

The CNN is a kind of feedforward neural network, and it automatically extracts representations or abstractions from data by using its convolutional layer and pooling layer (Li et al. 2021a). The convolutional layer can be computed as in Eq. (2–1):

\[ conv_f^l[j] = f * \left( \sum_{k=0}^{K-1} w_{k}^l S_j + b_j^l \right) \]  

(2-1)

where \( conv_f^l[j] \) is the output of the \( j \)th neuron at layer \( l \); \( f \) is a nonlinear function; \( w_{k}^l \) and \( b_j^l \) are the scaler weight and bias of the \( j \)th neuron at layer \( l \), respectively; \( S_j \) represents the selection of the input maps.

For the pooling layer calculation, it can be obtained according to Eq. (2–2):

\[ Poolinglayer = \begin{cases} \text{Max}(conv_f^l) \\ \text{Average}(conv_f^l) \end{cases} \]  

(2-2)

where the Max\((conv_f^l)\) and Average\((conv_f^l)\) depend on the pooling layer type.

Residual network (ResNet)

ResNet won the 1st place in ImageNet classification, which was first introduced by He et al. (2016). ResNet is a network-in-network architecture that relies on many stacked residual units (Garcia-Garcia et al. 2018). The structure of ResNet was shown in Fig. 2, which was consisted of two deep building blocks: bottleneck 1 and bottleneck 2. ResNet-18 and ResNet-34 were stacked by bottleneck 1, while bottleneck 2 was used to stack the structure of ResNet-50, ResNet-101, and ResNet-152. The number of ResNet like 18 and 152 was represented the 18 and 152 layers, respectively. Compared with bottleneck 1, the three layers 1×1, 3×3, and 1×1 convolutions in bottleneck 2, where the function of 1×1 layer is reducing and increasing the dimension of input, making the 3×3 layer a bottleneck with small input/output dimensions (He et al. 2016). Identity mapping is a key measure for addressing the degradation problem, how it works introduce in below:

As shown in bottleneck 1 in Fig. 2, \( x, y, \) and \( F(x, \{ w_{i} \}) \) was represented the input, output vectors, and residual mapping for learning, respectively. The operation \( F(x, \{ w_{i} \}) + x \) is conducted by a shortcut connection (Fulkerson 1996) and element-wise addition, which could effectively avoid the degradation problem.

\[ y = F(x, \{ w_{i} \}) + x \]  

(2-3)

It is worth noting 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 as follows (2–4):

\[ Pr(Y = i|v, W, b) = \text{Softmax}_i(Wv) + b = \frac{e^{w_i v + b_i}}{\sum_j e^{w_j v + b_j}} \]  

(2-4)
where elements of w and b denote the weights and bias, respectively. Index j was used to normalize the posterior distribution.

The model prediction is the class with the highest probability:

$$y_{\text{prediction}} = \arg \max_i Pr(Y = i | v, W, b)$$  \hspace{1cm} (2-5)

The elements of weights and bias in deep ResNet structure are also optimized by the error backpropagation algorithm, which uses an error metric to calculate the distance between true class labels and the predicted class labels.

Cross-entropy function (2–6) was chosen as the loss function to be minimized for dataset V.

$$L(V, Y) = -\frac{1}{n} \sum_{i=1}^{n} y_i \ln a(v_i^0) + (1 - y_i) \ln (1 - a(v_i^0))$$  \hspace{1cm} (2-6)

where L denotes the loss function; here, $V = \{v^{(1)}, v^{(2)}, v^{(3)}, ..., v^{(n)}\}$ is the set of input samples in the training dataset; $Y = \{y^{(1)}, y^{(2)}, ..., y^{(n)}\}$ is the corresponding labels: cardboard, glass, metal, paper, plastic, and trash; the $a(v)$ represents the output of the ResNet given input v.

Fig. 1 Gallery of recyclable waste: cardboard, glass, metal, paper, plastic, and trash
Transfer learning

Transfer learning is a machine learning technique to migrate knowledge between different physical scenarios, involving the concepts of domain and task. Given a source domain, ImageNet ($X_s$), and a corresponding source task, Image label ($Y_s$), there is a marginal probability distribution $P(x,y)$ between $X_s$ and $Y_s$. For the task of recyclable waste classification, there is also a target domain, trash dataset ($X_T$) as well as a target task, corresponding label ($Y_T$). The objective of transfer learning is to learn the target probability distribution $P_T(x,y)$ in $X_T$ with the knowledge gained from $X_s$ and $Y_s$. 

Fig. 2 The structure of the ResNet: ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152
Comparison of the performance of waste sorting by ResNet-18 without/with transfer learning (TL): a training and valid loss in ResNet-18 without TL; b training and valid loss in RWN-18; c training and valid accuracy in ResNet-18 without TL; d training and valid accuracy in RWN-Net; e training time in ResNet-18 without TL; f training time in RWN-18 with TL. Note: TL was represented to transfer learning Y s. Noticeably, ImageNet consists of 12 million images and 1000 categories. Therefore, $X_i \neq Y_s$ or $X_i \neq Y_T$.

The measure of fine-tuning was taken to adjust the final learnable layers of the pretrained ResNet models, changing neurons in the output layer from 1000 to 6. This method can share the pre-trained parameters and handle the task of recyclable waste classification.

**Model evaluation metrics**

As for the definition of true positive (TP), true negative (TN), false positive (FP), and false negative (FN), taking the cardboard as example. The recyclable wastes which are cardboard and detected as TP; the recyclable waste that are not cardboard and actually predicted not-cardboard are denoted as TN; while, if recyclable waste that were actually cardboard but misjudged as other types of recyclable waste, which denoted as FN; as for FP, it can be defined that if recyclable waste were not misjudged as other types of recyclable waste, which denoted as TN; while, if recyclable waste that were actually cardboard but classified as cardboard in non-cardboard waste.

Confusion matrix, recall, precision, F1 score, accuracy, ROC, and area under the curve (AUC) were used to evaluate the performance of RWNet models. Recall, precision, F1 score, and accuracy were defined as follows:

\[
Recall = \frac{TP}{TP + FN} \tag{2-7}
\]

\[
Precision = \frac{TP}{TP + FP} \tag{2-8}
\]

\[
F1score = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{2-9}
\]

\[
Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2-10}
\]

\[
AUC = \frac{1}{2} \times \left( \frac{TP}{TP + FN} + \frac{TN}{TN + TP} \right) \tag{2-11}
\]

in which, TP, TN, FN, and FP are the numbers of true positives, true negatives, false negatives, and false positives, respectively.

**Visualization**

**PCA**

PCA was employed to create low-dimensional data representing and interpreting waste sorting (Thomaz and Giraldi 2010). First, the matrix X of m \times n training was set to include m input samples and n pixels (or variables). Secondly, the covariance matrix C of the data matrix was obtained according to Eq. (2–12):

\[
C = \frac{1}{m}XX^T \tag{2-12}
\]

Thirdly, the eigenvector P and eigenvalue $\Lambda$ of this covariance matrix $C$ were calculated according to Eq. (2–13):

\[
P^T C P = \Lambda \tag{2-13}
\]

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

**t-SNE**

The t-SNE is capable of reducing dimensionality tasks via minimizing the divergence between the pairwise similarity distribution of input points and low-dimensional embedding points (Maaten and Hinton 2008, Retsinas et al. 2017). The input points were set as $\{x_1, x_2, x_3, ..., x_n\}$ and their corresponding embedding points were $\{y_1, y_2, y_3, ..., y_n\}$. The pairwise similarity between points $x_i$ and $x_j$ can be measured using the joint probability, $p_{ij}$, as given by Eqs. (2–14) and (2–15):

\[
p_{ij} = \frac{\exp(-d(x_i, x_j)^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-d(x_i, x_j)^2/2\sigma_i^2)} \tag{2-14}
\]

\[
P_{ij} = \frac{p_{ij} + p_{ji}}{2N} \tag{2-15}
\]

where $d(x_i, x_j)$ represented distance function that is considered to be Euclidean.

A similar method was applied to calculate the pairwise similarity between points $y_i$ and $y_j$ in the embedding space, that is:

\[
q_{ij} = \frac{\exp\left(-\|y_i - y_j\|^2\right)}{\sum_{k \neq i} \exp\left(-\|y_i - y_j\|^2\right)} \tag{2-16}
\]

The Kullback–Leibler divergence was minimized to calculate the embedding Y that consider the pairwise similarity distribution for both initial and the embedding spaces as follows:

\[
C(Y) = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{2-17}
\]
Algorithm RWNet

1. **Training phase**: Input: \(X_{\text{train}}, Y_{\text{train}}\); \(X\): RGB ImageNet image, \(y\): class labels.
2. Setting up ResNet architecture and selecting hyper-parameters like batch size, epoch number, and work numbers.
3. Learning parameters of ResNet-ImageNet model with cyclical learning rate using \(X_{\text{train}}, Y_{\text{train}}\).
4. Using transfer learning and fine-tuning trained ResNet-ImageNet
5. Input: \(X_{\text{train}}, Y_{\text{train}}, X_{\text{valid}}, Y_{\text{valid}}\); \(X\): RGB waste image, \(y\): class labels.
6. Learning parameters of RWNet model using \((X_{\text{train}}, Y_{\text{train}})\) and \((X_{\text{valid}}, Y_{\text{valid}})\)
7. **Testing phase**: Input RWNet model, \(X_{\text{test}}\)
8. Compute RWNet inference for testing data \(\rightarrow\) recyclable waste classification
9. **Explanation phase**: Generating output features via PCA and t-SNE, respectively.

**Results and discussion**

**Learning rate**

Figure S2 shows the relationship between the learning rate and loss in various RWNet models based on the ImageNet dataset processed with cyclical learning rate. Instead of a fixed or exponentially decreasing value, the value of learning rate in the method of cyclical learning rate is varied between bounds. The optimum learning rate will be between the bounds and the near-optimum learning rate can be found throughout the training. The gradient between loss and learn rate reach maximum in RWNet-18, RWNet-34, RWNet-50, RWNet-101, and RWNet-152 when the value of the learning rate is \(1.8 \times 10^{-4}, 5.0 \times 10^{-4}, 4.8 \times 10^{-4}, 3.2 \times 10^{-4},\) and \(2.1 \times 10^{-4}\), respectively. These points are the optimum learning rate for various ResNet architectures. Therefore, \(1.8 \times 10^{-4}, 5.0 \times 10^{-4}, 4.8 \times 10^{-4}, 3.2 \times 10^{-4},\) and \(2.1 \times 10^{-4}\) of the learning rate from ImageNet dataset training were separately applied in RWNet-18, RWNet-34, RWNet-50, RWNet-101, and RWNet-152 to train the recyclable waste training and validation dataset.

**Effect of the transfer learning on RWNet-18 performance for recyclable waste classification**

Figure 3 shows the effect of transfer learning on ResNet-18 performance in terms of loss, accuracy and training time. Figure 3a and b separately plot train loss and valid loss with various epoch applied ResNet-18 without/with transfer learning (TL) to sort the recyclable waste. The tendency of loss from ResNet-18 without TL and ResNet-18 with TL was decreased with the increase of the epoch number in both the training dataset and validation dataset. At the 10th epoch, the loss in the training dataset from ResNet-18 without/with TL was 0.043 and 0.006, respectively.

As for accuracy, its tendency showed an upward trend with the increase of the epoch number both ResNet-18 without/with TL (Fig. 3c and d), and reached 98.59% and 99.78% in the training dataset at the 10th epoch. The accuracy from ResNet-18 with TL in the training dataset and validation dataset reached approximately maximum (98.18% and 98.18%, respectively) at the 5th epoch. Transfer learning could promote ResNet-18 to reach a better accuracy of recyclable waste sorting. As shown in Fig. 3e and f, the average training time in ResNet-18 without TL and ResNet-18 with TL was approximately 277.1 s and 231.3 s, respectively.

**Performance of various RWNet models in training and validation dataset**

Figure S3 shows various RWNet model performance on recyclable waste sorting in terms of training time, loss, and accuracy. The average training time was increased generally with the increase of the layer of ResNet architecture. It can be ascribed to much more parameters needed to be trained with the layer increased (Table S2). Specifically, the average training time was 231.3 s in RWNet-18, 242.9 s in RWNet-34, 272.8 s in RWNet-50, 350.3 s in RWNet-101, and 738.4 s in RWNet-152 (Fig. S1).

As shown in Fig. S3, the loss in various RWNet architectures showed a similar trend. The loss decreased with the epoch number increase in both the training dataset and the validation dataset. On the contrary, the trend of accuracy in various RWNet architectures increased with the iteration of the epoch.
Model performance in the test datasets

Confusion matrix

As shown in Fig. 4, confusion matrices of the assessment in model performance for RWNet models ran on the recyclable waste testing dataset. The number of the recyclable waste image on the diagonal line means correct classifications. And the values outside the diagonal line represent unpaired labels and images (Aliper et al. 2016). As for cardboard in RWNet-18, 1382 cardboard images were accurately sorted, in comparison to 0 images of glass, 4 images of metal, 4 images of paper, 3 images of plastic, and 0 images of trash were wrongly identified as the cardboard (Fig. 4b). It is noted that much more correct classifications in RWNet-18 compared with that in ResNet-18-without TL (Fig. 4a), indicating that RWNet models provided a better performance on the TrashNet test on the testing dataset.

In addition, most recyclable wastes like cardboard (1382), glass (1727), metal (1363), paper (1960), plastic (1194), and trash (418) were found on the diagonal line in Fig. 6b, and almost at the same level with other RWNet models. It indicated that each RWNet model can perform well on recyclable waste classification.

Accuracy, precision, recall, and F1 score

The meaning of accuracy, precision, recall, and F1-score was the proportion of the waste classification model correctly classified, the ratio of real correctly predicted positive items to all samples predicted to be positive, the ratio of correctly identified positive items to all real positive samples, and harmonic mean of recall and precision, respectively (Munir et al. 2019). The accuracy, precision, recall, and F1-score were adopted to quantitatively evaluate the performance of different RWNet models for recyclable waste sorting. The details of model evaluation metrics were shown in Fig. 5 and Table S3.
accuracy of various RWNet models is almost at 88%, and the best accuracy is 88.8% in RWNet-152 in Fig. 6a, which is much higher than the accuracy for recyclable waste in previous studies, 22% (Thung and Yang 2016), 63% (Thung and Yang 2016), and 87% (Adeyemi and Wang 2019). However, due to the unbalanced data, there are shortcomings of the accuracy assessment method (Dhillon and Verma 2019).

For the precision value, as shown in Fig. 5b and Table S3, RWNet-101 shows better performance than other RWNet models in terms of a weighted average (89.9%). There is an excellent performance in all categories of recyclable waste except glass. The precision value of glass in various RWNet models followed this order: RWNet-50 (88.4%) > RWNet-18 (88.0%) > RWNet-152 (87.1%) > RWNet-101 (89.9%).
Certainly, RWNet-101 still has room for enhancing the performance, especially for trash (55.6%).

For the recall value, as shown in Fig. 5c and Table S3, RWNet-152 shows a little better performance than other RWNet models in terms of weighted average (88.8%). Specifically, the recall value of RWNet-152 for 9100 images classified into cardboard, glass, metal, paper, plastic, and trash was 95.2%, 93.3%, 94.9%, 90.3%, 72.1%, and 87.4%, respectively. As given in Table S3, cardboard (97.5%), glass (95.7%), metal (94.9%), paper (91.6%), plastic (72.1%), brick (79.3%), and trash (87.4%) can be well separated by RWNet-34, RWNet-18, RWNet-152, RWNet-101, RWNet-152, and RWNet-152, respectively.

As shown in Fig. 5d and Table S3, the weighted average of F1-score follows this order: RWNet-152 (88.9%) > RWNet-50 (88.7%) = RWNet-101 (88.7%) > RWNet-18 (88.5%) > RWNet-34 (88.3%). This suggested that most of recyclable waste can be well sorted by RWNet models. In addition, the sorting performance for plastic and trash has great potential to improve.

**ROC**

The area under the ROC curve (AUC) can be used to assess the classification performance (Ahmad et al. 2020; Zhang et al. 2021b). It means a better classification effect along with a larger AUC value. All the macro average AUC of RWNet models was 0.93 (Fig. 6). Class 0, class 1, class 2, class 3, class 4, and class 5 represent cardboard, glass, metal, paper, plastic, and trash, respectively.

As presented in Fig. 6, all the RWNet models show similar performance. All the AUC value is higher than 0.9, except for the AUC value of plastic, which indicated that most recyclable waste can be well sorted by the RWNet models. The maximum AUC value of the cardboard is 0.98, followed by metal (0.97), glass (0.96), paper (0.94), trash (0.92), and plastic (0.85). It means that the classification effect of cardboard is the best, in contrast, the classification effect of plastic is the lowest. This is because of the plastic waste images, which often were wrongly identified as glass or trash (Fig. S4). In short, RWNet models show excellent performance in recyclable waste. The performance of RWNet models in terms of accuracy, precision, F1 score, and AUC is different from the conclusion of the ImageNet application, which indicates that the deeper ResNet structures, the better performance. This phenomenon can be ascribed to the dataset of TrashNet that are not as complex as those from the ImageNet (Yang et al. 2021). Those results suggested that the structure and depth choice should be appropriate for the practical application in recyclable waste sorting.

**Visual explanation by using PCA and t-SNE**

PCA and t-SNE were adopted to present the distribution of the TrashNet test dataset in RWNet models. Figure 7a shows 2-dimension representations extracted from the last layer of the RWNet-152 model by the method of PCA algorithm. The features from RWNet-152 models demonstrate a relatively clear semantic clustering. However, all the features of recyclable waste overlapped a little bit, which indicates there exists confusion.

Figure 7b shows that the features of cardboard, glass, metal, paper, plastic, and trash were clearly separated, indicating that representations of recyclable waste were better separated by t-SNE from high dimension to low dimension compared to PCA. Therefore, combined with its advantages, the t-SNE method can be used to accurately describe the classification features of recyclable waste (Gisbrecht et al. 2015).
Environmental Science and Pollution Research (2022) 29:91081–91095

RWNNet model limitations

Developing an intelligent and efficient method for recyclable waste classification and recycling is imperative due to the continuous increase of global MSW, especially during the period time of the global COVID-19 pandemic (Ge et al. 2022; Liu et al. 2021; Mamirkulova et al. 2022; Raza Abbasi et al. 2021; Zhou et al. 2021). RWNNet model has some limitations, which can be improved in the future work. The training time of RWNNet-152 is slower (average time for one epoch was approximately 700 s) due to millions of parameters need to be trained, it can be adjusted the structure of ResNet model to reduce the number of parameters. In addition, the images used in this study are not sufficiently realistic, and the real image of waste can be considered to test the robustness of RWNNet.

Conclusion

The optimized learning rate for RWNNet-18, RWNNet-34, RWNNet-50, RWNNet-101, and RWNNet-152 to train the trash dataset is $1.8 \times 10^{-4}$, $5.0 \times 10^{-4}$, $4.8 \times 10^{-4}$, $3.2 \times 10^{-4}$, and $2.1 \times 10^{-4}$, respectively. Transfer learning can help RWNNet models reach a better accuracy of recyclable waste sorting. The average training time increased with the increase of the layer of ResNet architecture from RWNNet-18 (231.3 s) to RWNNet-152 (738.4 s). The accuracy of various RWNNet models is almost at 88.5% in RWNNet-152. The highest precision, recall, and F1 score in terms of weighted average value appeared in RWNNet-101 (89.9%), RWNNet-152 (88.8%), and RWNNet-152 (88.9%), respectively. This suggested that most recyclable waste can be well sorted by RWNNet-152 models. Specifically, the recall value of RWNNet-152 for 9100 images classified into cardboard, glass, metal, paper, plastic, and trash was 95.2%, 93.3%, 94.9%, 90.3%, 72.1%, and 87.4%, respectively. Moreover, all the AUC value is higher than 0.9, except for the AUC value of plastic (0.85), which indicates that most recyclable waste can be well sorted by RWNNet models. The classification effect of cardboard (AUC value = 0.98) is the best, in contrast, the plastic is the lowest. In addition, representations of recyclable waste were better separated by t-SNE from high dimension to low dimension compared to PCA.

Policy recommendations

This study demonstrates the good performance of RWNNet models that can be used to sort most of the recyclable waste automatically, although the sorting performance for plastic has the potential to improve. Based on these findings, market-based environmental regulations should be introduced, which would encourage economic development as well as technology innovation.

Supplementary Information

The online version contains supplementary material at https://doi.org/10.1007/s11356-022-22167-w.

Author contribution

Kunsen Lin designed the research, performed the calculation and analysis, and drafted the manuscript. Meilan Zhang, Chunlong Zhao, Lu Peng, and Qian Zhang gave comments and revisions. Xiaofeng Gao, Tao Zhou, and Youcai Zhao carried out the main revisions of contents and figures.

Funding

This work was financially supported by the National Natural Science Foundation of China (no. 51878470; no. 52000143), the National Key R&D Program of China (no. 2018YFC1901400), and the International Postdoctoral Exchange Fellowship Program (YJ20200280).

Data availability

Enquiries about availability should be directed to the authors.

Declarations

Ethics approval

Not applicable.

Consent to participate

All the authors (Kunsen Lin, Youcai Zhao, Xiaofeng Gao, Meilan Zhang, Chunlong Zhao, Lu Peng, Qian Zhang, Tao Zhou) consent to participate.

Consent for publication

All the authors (Kunsen Lin, Youcai Zhao, Xiaofeng Gao, Meilan Zhang, Chunlong Zhao, Lu Peng, Qian Zhang, Tao Zhou) approved the version to be published.

Competing interests

The authors declare no competing interests.

References

Abbasi J, Mahmoud S, Ali H, Raza M, Ali G, Aman J, Bano S, Nurunnabi M (2019a) The effects of corporate social responsibility practices and environmental factors through a moderating role of social media marketing on sustainable performance of firms’ operating in Multan, Pakistan. Sustainability 11 https://doi.org/10.3390/su11123434

Abbasi J, Raza S, Nurunnabi M, Minai MS, Bano S (2019b) The impact of entrepreneurial business networks on firms’ performance through a mediating role of dynamic capabilities. Sustainability 11 https://doi.org/10.3390/su11113006

Abbasi J, Mubeen R, Iorember PT, Raza S, Mamirkulova G (2021) Exploring the impact of COVID-19 on tourism: transformational potential and implications for a sustainable recovery of the travel and leisure industry. Curr Res Behav Sci 2 https://doi.org/10.1016/j.crbeha.2021.100033

Adeleji O, Wang Z (2019) Intelligent waste classification system using deep learning convolutional neural network. Procedia...
