Proceedings of the Seventh International Conference on Research in Intelligent and Computing in Engineering pp. 255–260 ISSN 2300-5963 ACSIS, Vol. 33
DOI: 10.15439/2022R01

# Simple and Efficient Convolutional Neural Network for Trash Classification

1st Binh-Giang Tran
*Department of IT*
*Vinh University of Technology and Education*
Nghe An, Vietnam
binhgiangktv@gmail.com

2nd Duy-Linh Nguyen
*School of Engineering, Electronic and Computer Engineering*
*University of Ulsan*
Ulsan, Korea
ndlinh301@mail.ulsan.ac.kr

***Abstract*—Strong economic and city developments have given a great amount of trash. Trash is made continuously from families, public and commercial areas, construction places, hospitals, etc. The enlarging trash amount is a much more serious problem than climate change, and the cost of trash treatment will be a big burden to countries in the world. One of the effective trash treatment measures is to separate trash right from its source, especially domestic trash. The countries have applied many trash classification systems, but the requirements for infrastructure, implementation, and operation are quite complicated. In order to help people easily sort household trash at home, this paper proposes a simple convolutional neural network for trash classification. The network is trained and evaluated on the TrashNet dataset with an accuracy of 90.71%. In addition, this work also tests in real-time on low-computation devices such as CPU-based personal computer and Jetson Nano devices.**

***Index Terms*—Convolutional neural network, trash classification, trash recognition.**

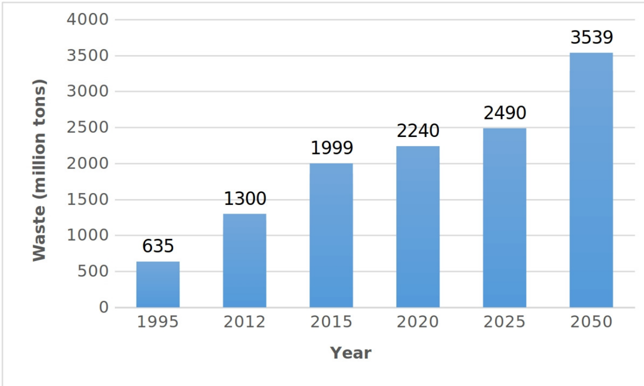


Fig. 1. The chart of the annual amount of trash on the globe.

## I. Introduction

Trash not only pollutes the environment but also directly affects people's health. Funding and ways to treat trash are also a dilemma for many countries around the world. More specifically, in 2012, humans released 1,300 tons of trash into the environment, double the amount of trash in 1995. In 2015, the total amount of trash worldwide was 1,999 tons [1]. The quantity of trash in 2020 is 1.1 times higher than in the previous five years, about 2,220 tons [2]. And with the current rapid trash generation rate, it is expected that by 2050, the weight of trash that humans generate every day could be 3,539 million tons [3]. Fig. 1 shows the chart of the annual trash on the globe. Trash is an inevitable consequence of human life and production. On average, a person generates 0.74 kg of trash every day. This amount of daily household trash includes leftovers, paper, old clothes, damaged electronics, old furniture, garbage from the yard, etc. The benefits of trash classification at the place of residence help to save resources, and bring a great economic source. From that experience, it contributes to reducing the total quantity of trash that needs to be destroyed, reducing the load on the environment, and saving costs of collection, transportation, and treatment. This also contributes to raising public awareness about the protection and rational use of natural resources and the environment. There are many ways to sort trash at home. The first way is based on chemical characteristic, we can separate organic trash and inorganic trash, then continue to classify organic trash for easy handling. The second way is to use the product standards listed on the packaging to separate them into recyclable trash and non-recyclable trash. The third way is optical sorting or eddy current-based sorting classification techniques [4]. However, to reduce the complexity for people, household trash can be classified into the following six groups: glass, paper, cardboard, plastic, metal, and trash. There are many trash sorting systems in the world. In 2010, Glave et al. announced an automated vacuum trash collection (AVAC) system, which transports trash at high speed through underground pneumatic tubes to a collection station, where they are compacted and sealed in containers. This system helps to classify and recycle garbage automatically but requires a large area of land, pipe system, big industrial fan system, sensors, and automated software [5]. In 2019, the identify system called Multi Re-use Facility (MRF) was used to recognize and classify trash on conveyor belt. Each stage is identified and the trash is gradually eliminated. The order of identification is hazardous trash such as batteries, aerosols; soft plastic materials; paper, bottles, cans and finally ferrous metals. This system also requires high installation costs, making it suitable for small trash treatment plants [6]. To ensure high accuracy, low installation cost, and suitability for each family's size, the paper proposes a simple convolutional neural network (CNN) to recognize trash in daily life according to the six groups mentioned above. This network exploits basic features such as convolutional layers, pooling layers, optimization method, and especially the global average pooling technique to replace fully connected layers to reduce computational cost. Because of the high accuracy and the low delay during real-time testing, the network is applied to classify the trash and help the person to put the correct trash in the designated bin.

The main contributions of this paper are shown as:
1 - Proposes a simple and efficient CNN with two new connectors to classify trash.
2 - Develops a trash classification system on low-computing devices such as personal computers and embedded devices. The rest of the paper is organized: Section II presents some traditional techniques and CNN in trash identification and analyzes their advantages and disadvantages. Section III describes the proposed network in detail. Section IV analyzes the experiments and evaluates the results. Finally, the conclusion and future development direction are presented in section V.

## II. Related Work

In this section, the paper introduces several traditional and CNN-based techniques which has been applied to identify trash and analyzes the advantages and disadvantages of those techniques.

### A. Traditional methods

The commonly used identification and classification algorithms include the K-Nearest Neighbor algorithm (KNN), Bayesian Classifier, and Support Vector Machine (SVM). The authors in [7] using the KNN algorithm in identifying and classifying trash with an accuracy of 93.8%. [8] presents the application of a Recurrent Neural Network (RNN) to monitor the amount of trash in the bin. The paper [9] developed a deep learning classifier to classify domestic solid trash into different categories. The classification results show that the trash classification accuracy is from 82% to 96%. In 2018, Yijian Liu et al. designed a trash classification system in which the software is based on the SURF-BoW algorithm and multi-class SVM classifier. This system focuses on sorting five types of trash: batteries, bottles, cans, paperboard, and paper boxes. Experimental results show that pin identification is the best [10]. However, the limitation of this algorithm is that it cannot recognize trash and broken bottles, which is one of the common types of household trash in every household.

### B. CNN-based methods

Currently, convolutional neural networks are one of the mainstream approaches to image recognition. To classify and identify trash, the following CNN networks can be used: Alexnet, VGGNet, GoogleNet, ResNets, DenseNet [11]. Faced with India's trash boom, Thanawala D. et al. approached a variety of techniques that use CNN to classify trash into three categories: recyclable, non-recyclable, and organic. Along with the proposed neural network, five standard CNN architectures including VGG-16, DenseNet, InceptionNet, MobileNet, and ResNet are also tested on the given dataset. The highest test accuracy is 92.65% and the lowest is 81.25% [12]. Gary has built an automatic trash classification technique based on CNN at the edge, allowing smart and quick decisions to be made without access to data from the cloud. With this technique, trash is classified into six categories: paper, cardboard, glass, metal, plastic, and others [13]. This method is suitable when applied to large sorting centers to automatically separate trash into different categories, making recycling easier. Hua Zheng and Yu Gu used CNN to identify domestic solid trash. First, three CNN networks including GoogLeNet, ResNet50, and MobileNetV2 are used as component classifiers for separate prediction.Next, UPMWS is used to determine the weight coefficients of the composite models. This technique separates trash into four groups: wet trash, recyclable trash, hazardous trash, and dry trash. This algorithm is difficult to install and achieves high efficiency with ensemble learning [14]. The authors in [15] proposed a CNN network to classify trash into three categories: recyclable, non-recyclable and organic trash, with an accuracy of 81.22%. The paper also mentions other structures such as VGG16, InceptionNet, DenseNet and MobileNet. Out of all these transfer learning models, MobileNet showed the highest accuracy of 92.65%. K.Ahmad et al. used the CNN network to classify trash into four groups: cardboard, paper, metal, and plastic. Through testing for 100 times, with images of size $50 \times 50$, the accuracy reaches 76% [16]. The work in [17] apply multilayer hybrid convolution neural network technique, consider TrashNet dataset to identify trash through images; classified into glass, metal, paper, plastic. This network has a similar structure to VGGNet but is simpler, with fewer parameters, faster processing time, and up to 92% accuracy. The disadvantage of this method is that it can't classify recyclable and non-recyclable trash.

## III. Proposed Methodology

The proposed trash classification network is described in Fig. 2. The network consists of feature extraction and classification modules.

### A. Feature extraction module

The feature extraction module plays an important role in the proposed CNN network. It performs feature extraction at different levels and is input to the processing in the next module. This module combines the advantages of convolution layer (Conv), depthwise separable convolution layers (DWConv), average pooling layer, batch normalization method (BN), and two new connectors to extract the best feature maps and optimize the network parameter. Its architecture consists of a stem and four connectors. The stem contains five main blocks of which three have exactly the same structure. The first block is designed based on two $7 \times 7$ Conv layers followed by a BN and a ReLU activation function. This block uses Conv layers with large kernel sizes to capture the basic information of the object in the input image. Therefore, the input image of size $224 \times 244 \times 3$ after passing through this block will reduce the

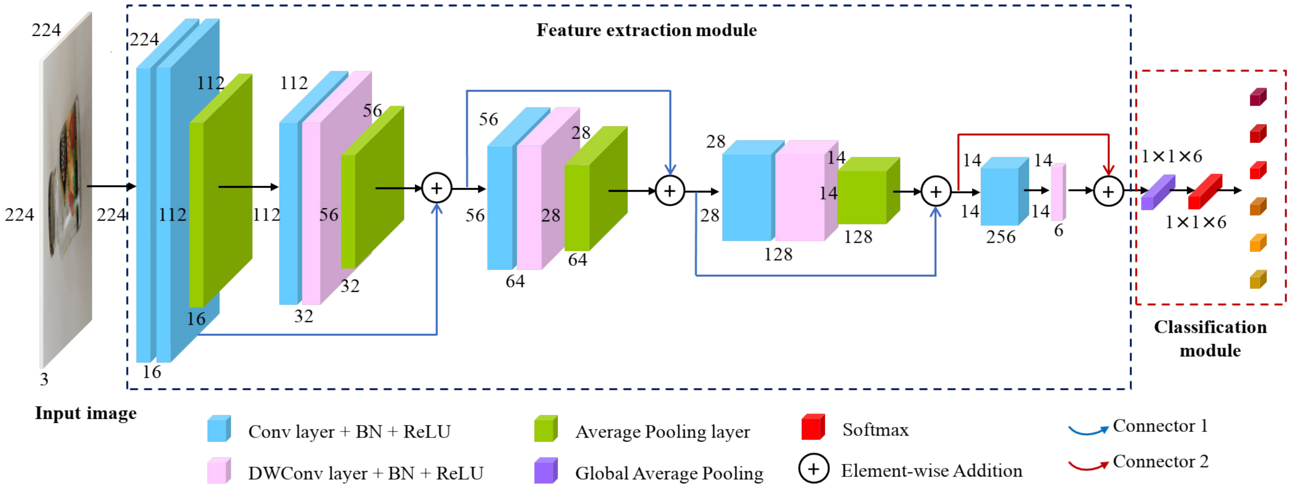


Fig. 2. The proposed trash classification network. This network consists of feature extraction and classification modules.

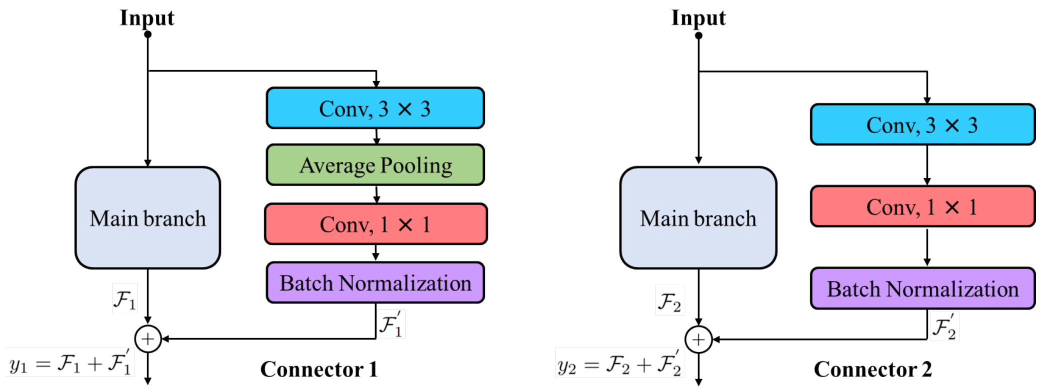


Fig. 3. The architecture of the connector 1 and connector 2.

dimension to $112 \times 122 \times 16$. At the intermediate level are three blocks with the same structure. Each block contains a Conv layer, a DWConv layer, an average pooling layer followed by a BN, and a ReLU activation function. Convolution layers use kernels with sizes reduced from $5 \times 5$ to $3 \times 3$. These blocks act as an intermediate feature extraction with more detailed information about the object. Through each block, the feature map has further reduced the dimension twice to form the output and link with the connectors. As a result, it generates three feature maps with dimensions of $56 \times 56 \times 32$, $28 \times 28 \times 64$, and $14 \times 14 \times 28$ respectively. The final block is made up of a Conv layer and a DWConv layer. This block acts as a bridge between the feature extraction module and the classification module. The feature map through this block will maintain the dimension 14x14 with the number of channels increasing to 256 and then decreasing to 6 (corresponding to the number of class in the dataset). The output from this block will also be connected to the previous feature map at the level through connector 2 using the element-wise addition operation.

As mentioned above, to combine feature maps at different levels this paper has proposed two types of connectors. Each connector consists of two branches, the main branch (stem) and the sub-branch. In the sub branch on connector 1 use a $3 \times 3$ Conv layer, an average pooling layer, and a $1 \times 1$ Conv layer followed by a BN. The two branches are then joined together via the element-wise addition operation. In connector 2, the sub-branch has the same architecture as the first connector, but it doesn't use average pooling in between the two Conv layers. With these connectors, the amount of information in the current feature maps is enriched with the information extracted from the previous level. The architecture of these connectors is shown in Fig. 3.

## B. *Classification module*

Typically, popular classification networks use fully connected layers in the classification module. However, this technique increases the network parameter significantly because it contains a lot of connections. To solve that, this work also proposes to replace all fully connected layers with just one global average pooling layer while still ensuring classification accuracy. Specifically, this module uses a global average pooling layer to extract a 1x1x6 feature map. It then applies a softmax function on this feature map to calculate the probability of each class (six classes).

### C. Loss function

The difference between the predicted value and the target value is calculated using the loss function during training. In this paper, the cross-entropy loss has been used and is defined as follows:

$$\mathcal{L}_{class} = -\sum_{i=0}^{5} \mathcal{P}_i^*.log(\mathcal{P}_i), \quad (1)$$

in which, $i$ is the index of a each class ($i$ from 0 to 5). $\mathcal{P}_i^*$ presents the target indicator (0 or 1) and $\mathcal{P}_i$ denotes the predicted probability from the proposed network. Function $log$ is a natural logarithm function.

## IV. EXPERIMENT

### A. Dataset preparation

The proposed network is trained and evaluated on the TrashNet dataset [18]. This dataset contains 2,527 images divided into six subclasses: glass (501 images), paper (594 images), cardboard (403 images), plastic (482 images), metal (410 images), and trash (137 images) ). The images were taken by placing the subject on a white posterboard using sunlight or room light. The devices used to take photos include the Apple iPhone 7 Plus, Apple iPhone 5S, and Apple iPhone SE. The original image size is $512 \times 384$. For a fair comparison with other research on this dataset, the experiment splits the dataset into 70% for training and 30% for evaluation.

### B. Experimental setup

The proposed trash classification network is built using the Python programming language based on the Keras framework. Training and evaluation were performed on a GeForce GTX 1080Ti GPU. In addition, this experiment also uses an Intel Core i7-4770 CPU 3.40 GH CPU (PC) and an Nvidia Maxwell GPU (Jetson Nano device) to evaluate the speed of the network in real-time. Training goes through 200 epochs with a batch size of 16 and an initial learning rate of $10^3$ (then reduced by a factor of 0.55 after 20 epochs if the accuracy is not improved). The Adam optimization method is used to update the weights during training. In addition, the network also applies some data augmentation methods to avoid overfittings such as random shift, random zoom, and random brightness.

### C. Experimental results

In order to evaluate the performance, this experiment conducts training and evaluation of the proposed network. Proceed to refine the common classifier networks by removing all fully connected layers and replacing with a global average pooling layer, and train and evaluate under the same conditions of the dataset. In addition, to ensure objectively the evaluation results of the proposed network are also compared with other researches using the same dataset. The results in Table 1 show that the proposed network achieves a quite high accuracy of 90.71% with only 653,888 network parameters. This result is superior to both the refined networks and the networks in other studies. When compared with the best-refined competitor (VGG19), the proposed network has 11.26% greater accuracy but its network parameters are 31.44 times less. For other best networks in other research, the proposed network is higher at 1.71% of accuracy. Several samples of the qualitative results of the garbage classification are shown in Fig. 4.

TABLE I
THE COMPARISON RESULT BETWEEN PROPOSED TRASH CLASSIFICATION NETWORK WITH REFINED CLASSIFICATION NETWORKS AND OTHER RESEARCH IN SAME DATASET. THE RED COLOR DENOTES THE BEST COMPETITOR. SYMBOL (∗) DENOTES THE REFINED NETWORKS

| Network | Parameters | Accuracy (%) |
|---|---|---|
| SqueezeNet∗ | 257,846 | 64.82 |
| **Proposed** | **653,888** | **90.71** |
| MobileNetV2∗ | 3,575,878 | 58.10 |
| MobileNetV1∗ | 4,284,614 | 75.69 |
| NASNetMobile∗ | 5,358,234 | 75.49 |
| DenseNet∗ | 8,093,254 | 64.63 |
| Xception∗ | 15,246,150 | 78.46 |
| VGG16∗ | 15,246,150 | 78.66 |
| VGG19∗ | 20,555,846 | **79.45** |
| InceptionV3∗ | 23,907,110 | 71.34 |
| ResNet50∗ | 25,692,038 | 64.70 |
| LeNet∗ | 78,430,076 | 64.63 |
| Method in [18] | - | 63.00 |
| CNN in [19] | - | 75.00 |
| InceptionV4 [20] | - | **89.00** |
| DenseNet169 [20] | - | 84.00 |
| MobileNet [20] | - | 84.00 |

The confusion matrix in Fig. 5 represents the prediction ratio of the proposed network per class in the dataset. Accordingly, the classes have equally balanced prediction rates from 91% to 94%, except for the two classes, plastic and metal, which have low prediction rates of 84% and 89%, respectively. This is easy to see because the structure and shape of several objects in the glass and trash classes cause confusion for the network during the learning process. Therefore, in the training process, data augmentation is essential to help the network learn a variety of shapes, structures, and contexts.

For evaluating the speed, the network was also tested with a simple real-time system based on a camera connected to a PC and Jetson Nano devices. With VGA camera resolution, the proposed network achieved 24.59 frames per second (FPS) and 15.89 FPS on PC and Jetson Nano devices, respectively. This result shows that the network can be implemented on low-computing devices for real-time trash classification with very low latency. However, the experimental process also presents that the accuracy of the system depends on several factors such as the distance from the camera to the object to be classified (trash), camera resolution, and light of the environment.

### D. Ablation study

Each component in the proposed network architecture has a specific role that affects the performance of the entire

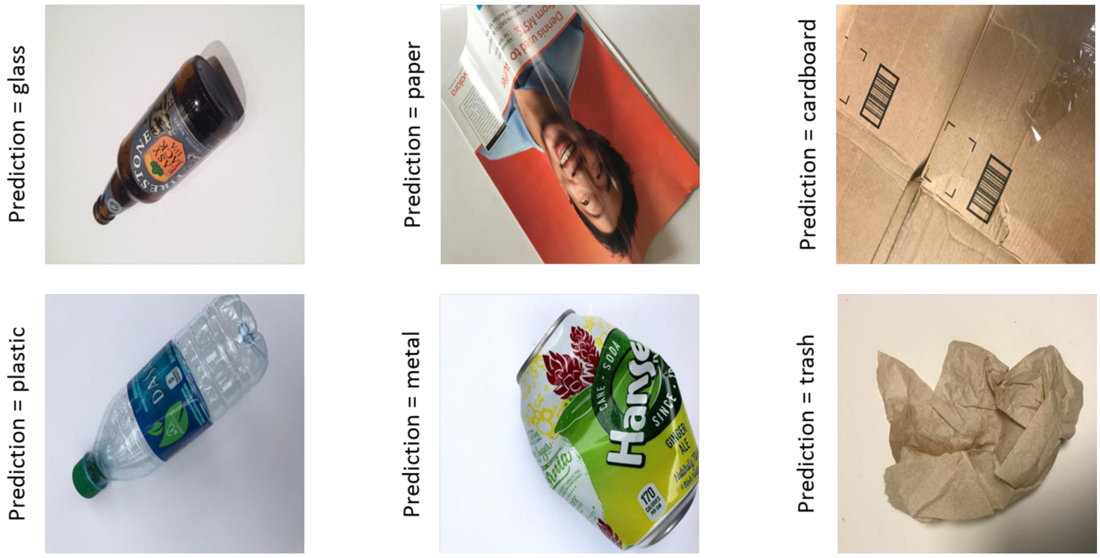


Fig. 4. Several qualitative results of trash classification on TrashNet dataset.

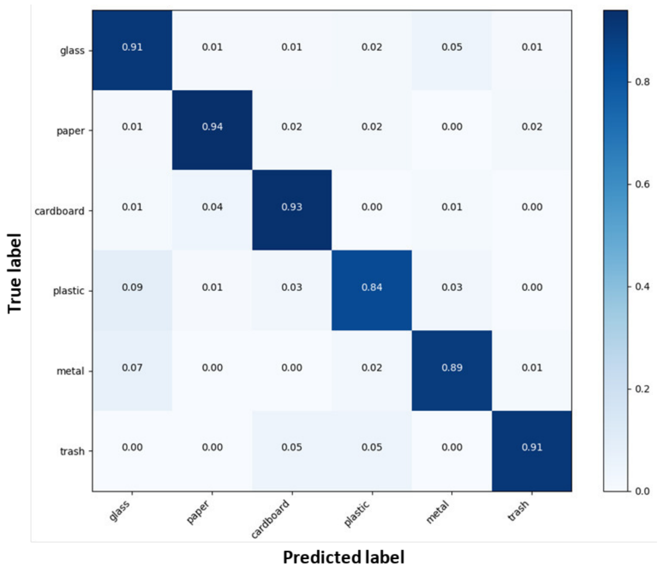


Fig. 5. The confusion matrix of trash classification network on TrashNet dataset.

network. This work conducted several ablation studies to evaluate such effects. First, the experiment conducted training and evaluation on the network without any connector (Stem). Then, gradually increase the connectors from one to four according to the depth of the network. The results in Table II show that when increasing the number of connectors from one to three (connector 1), the network parameter increases, but the accuracy tends to decrease or increase slightly. Until using the fourth connector (connector 2), the best accuracy is achieved.

TABLE II
ABLATION STUDY ON DIFFERENT NUMBER OF CONNECTOR. THE RED COLOR PRESENTS THE BEST NETWORK ARCHITECTURE.

| Connectors | Parameters | Accuracy (%) |
|---|---|---|
| 0 (just stem) | 445,608 | 87.88 |
| 1 (connector 1) | 448,552 | 87.62 |
| 2 (connector 1) | 460,072 | 88.01 |
| 3 (connector 1) | 505,640 | 87.75 |
| 4(Proposed) | **653,888** | **90.71** |

Next, the experiment investigates the influence of each type of convolution layer on the entire network parameter. The experiment uses all Conv layers instead of DWConv layers and then training and evaluation. The results in Table III demonstrate that the use of DWConv layers has resulted in significant savings in network parameters (195,872 parameters). While the combination of Conv layers and DWConv layers still maintains higher accuracy of the network.

TABLE III
ABLATION STUDY ON DIFFERENT CONVOLUTION TYPE. THE RED COLOR PRESENTS THE BEST NETWORK ARCHITECTURE.

| Convolution type | Parameters | Accuracy (%) |
|---|---|---|
| All Conv layer | 849,760 | 87.88 |
| Conv+DWConv layers | **653,888** | **90.71** |

## V. CONCLUSION AND FUTURE WORK

This paper has proposed a lightweight CNN for trash classification consisting of two main modules: feature extraction and classification. The network is built based on the features of Conv, DWConv, average pooling layers, and proposed connectors to extract feature maps. Then apply the global average pooling layer and softmax function to classify the objects. The network is trained and evaluated on the TrashNet dataset with high accuracy and negligible latency when tested in a real-time system. In the future, this trash classification network will be developed into an application that can be integrated into mobile devices.

## REFERENCES


[1] "The world's growing municipal solid waste: trends and impacts." Accessed: Jun. 24, 2022. [Online]. Available: https://iopscience.iop.org/article/10.1088/1748-9326/ab8659.

[2] "Solid waste management." Accessed: Jun. 24, 2022. [Online]. Available: https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management.

[3] "Humans on course to triple daily waste by 2100." Accessed: Jun. 24, 2022. [Online]. Available: https://www.nbcnews.com/sciencemain/humans-course-triple-daily-waste-2100-8c11497511.

[4] "Waste sorting." Accessed: Jun. 24, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Waste_sorting.

[5] G. Mangialardi, G. Trullo, F. Valerio, and A. Corallo, "Sustainability of a pneumatic refuse system in the metropolitan area: A case study in southern apulia region," *Procedia - Social and Behavioral Sciences*, vol. 223, pp. 799–804, 2016. 2nd International Symposium "NEW METROPOLITAN PERSPECTIVES" - Strategic planning, spatial planning, economic programs and decision support tools, through the implementation of Horizon/Europe2020. ISTH2020, Reggio Calabria (Italy), 18-20 May 2016.

[6] "Recycling facilities in the us." Accessed: Jun. 24, 2022. [Online]. Available: https://ibisworld.com.

[7] J. Yang, Z. Zeng, K. Wang, H. Zou, and L. Xie, "Garbagenet: A unified learning framework for robust garbage classification," *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 4, pp. 372–380, 2021.

[8] A. Camero, J. Toutouh, J. Ferrer, and E. Alba, "Waste generation prediction in smart cities through deep neuroevolution," in *Smart Cities* (S. Nesmachnow and L. Hernández Callejo, eds.), (Cham), pp. 192–204, Springer International Publishing, 2019.

[9] T. Sheng, M. Shahidul Islam, N. Misran, M. H. Baharuddin, H. Arshad, M. R. Islam, M. Chowdhury, H. Rmili, and M. Islam, "An internet of things based smart waste management system using lora and tensorflow deep learning model," *IEEE Access*, vol. PP, pp. 1–1, 08 2020.

[10] Y. Liu, K.-C. Fung, W. Ding, H. Guo, T. Qu, and C. Xiao, "Novel smart wgaste sorting system based on image processing algorithms: Surf-bow and multi-class svm," *Computer and Information Science*, vol. 11, p. 35, 06 2018.

[11] "Modern convolutional neural networks." Accessed: Jul. 10, 2022. [Online]. Available: https://d2l.ai/chapter_convolutional-modern/index.html.

[12] D. Thanawala, A. Sarin, and P. Verma, "An approach to waste segregation and management using convolutional neural networks," in *Advances in Computing and Data Sciences* (M. Singh, P. K. Gupta, V. Tyagi, J. Flusser, T. Ören, and G. Valentino, eds.), (Singapore), pp. 139–150, Springer Singapore, 2020.

[13] G. White, C. Cabrera, A. Palade, F. Li, and S. Clarke, "Wastenet: Waste classification at the edge for smart bins," 2020.

[14] H. Zheng and Y. Gu, "Encnn-upmws: Waste classification by a cnn ensemble using the upm weighting strategy," *Electronics*, vol. 10, no. 4, 2021.

[15] D. Ziouzios, D. Tsiktsiris, N. Baras, and M. Dasygenis, "A distributed architecture for smart recycling using machine learning," *Future Internet*, vol. 12, no. 9, 2020.

[16] K. Ahmad, K. Khan, and A. Al-Fuqaha, "Intelligent fusion of deep features for improved waste classification," *IEEE Access*, vol. 8, pp. 96495–96504, 2020.

[17] C. Shi, C. Tan, T. Wang, and L. Wang, "A waste classification method based on a multilayer hybrid convolution neural network," *Applied Sciences*, vol. 11, no. 18, 2021.

[18] M. Yang and G. Thung, "Classification of trash for recyclability status," *CS229 project report*, vol. 2016, no. 1, p. 3, 2016.

[19] "trashnet." Accessed: Jul. 24, 2022. [Online]. Available: https://github.com/garythung/trashnet.

[20] R. A. Aral, Ş. R. Keskin, M. Kaya, and M. Hacıömeroğlu, "Classification of trashnet dataset based on deep learning models," in *2018 IEEE International Conference on Big Data (Big Data)*, pp. 2058–2062, 2018.