

Research Article

Precision Measurement for Industry 4.0 Standards towards Solid Waste Classification through Enhanced Imaging Sensors and Deep Learning Model

Leow Wei Qin,¹ Muneer Ahmad ¹, Ihsan Ali ², Rafia Mumtaz ³,
Syed Mohammad Hassan Zaidi ³, Sultan S. Alshamrani ⁴, Muhammad Ahsan Raza ⁵,
and Muhammad Tahir⁶

¹Department of Information Systems, Faculty of Computer Science & Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

²Department of Computer System and Technology, Faculty of Computer Science & Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

³National University of Sciences and Technology (NUST), School of Electrical Engineering and Computer Science (SECS), Islamabad, Pakistan

⁴Department of Information Technology, College of Computer and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁵Department of Information Technology, Bahauddin Zakariya University, Multan 60000, Pakistan

⁶Department of Computer Science, Abdul Wali Khan University, Mardan, 23200, KPK, Pakistan

Correspondence should be addressed to Ihsan Ali; ihsanalichd@siswa.um.edu.my

Received 29 March 2021; Revised 20 April 2021; Accepted 27 April 2021; Published 17 May 2021

Academic Editor: Yuanpeng Zhang

Copyright © 2021 Leow Wei Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Achievement of precision measurement is highly desired in a current industrial revolution where a significant increase in living standards increased municipal solid waste. The current industry 4.0 standards require accurate and efficient edge computing sensors towards solid waste classification. Thus, if waste is not managed properly, it would bring about an adverse impact on health, the economy, and the global environment. All stakeholders need to realize their roles and responsibilities for solid waste generation and recycling. To ensure recycling can be successful, the waste should be correctly and efficiently separated. The performance of edge computing devices is directly proportional to computational complexity in the context of nonorganic waste classification. Existing research on waste classification was done using CNN architecture, e.g., AlexNet, which contains about 62,378,344 parameters, and over 729 million floating operations (FLOPs) are required to classify a single image. As a result, it is too heavy and not suitable for computing applications that require inexpensive computational complexities. This research proposes an enhanced lightweight deep learning model for solid waste classification developed using MobileNetV2, efficient for lightweight applications including edge computing devices and other mobile applications. The proposed model outperforms the existing similar models achieving an accuracy of 82.48% and 83.46% with Softmax and support vector machine (SVM) classifiers, respectively. Although MobileNetV2 may provide a lower accuracy if compared to CNN architecture which is larger and heavier, the accuracy is still comparable, and it is more practical for edge computing devices and mobile applications.

1. Introduction

Industry 4.0 standards require cutting-edge solutions to prevent, reduce, and even eradicate solid waste to ensure a pollution-free and sustainable environment [1]. As defined

by the Environmental Protection Agency (EPA), municipal solid waste (MSW) is a trash from sources that include residential, commercial, and institutional locations, such as businesses, schools, and hospitals [2]. However, the definition of MSW by EPA does not include industrial, hazardous,

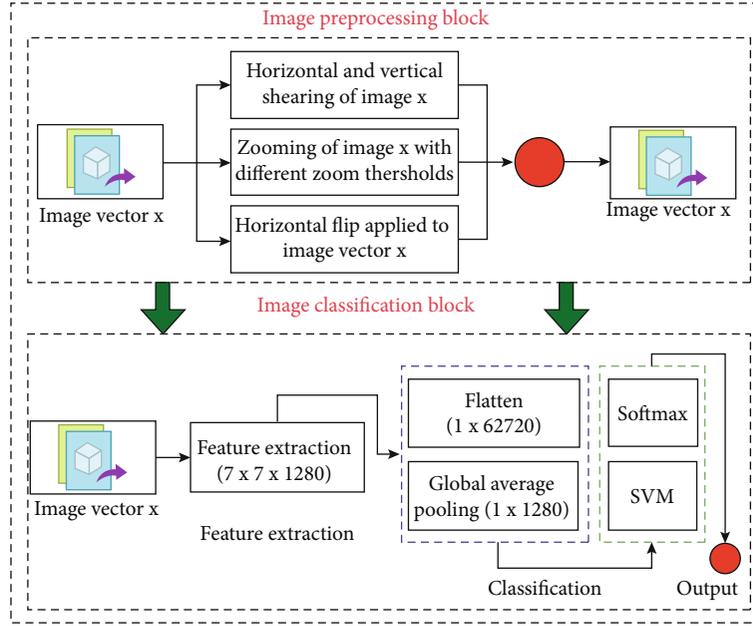


FIGURE 1: General architecture of the proposed classification model.

or construction and demolition (C&D) waste. MSW is locally handled by each municipality [3].

Municipal solid waste (MSW) can generally be classified into two categories: organic and inorganic categories [4]. To break it down further, organic waste consists of food waste and yard waste, while inorganic waste contains plastics, metal, cans, paper, glass, and others [5]. The waste composition differs from country to country as they are subjected to the influence of many factors, such as level of economic development, cultural norms, geographical location, energy sources, and climate. Inorganic waste such as plastics, paper, and aluminium increases while the relative organic waste decreases when a country urbanizes and populations become wealthier [6]. On the other hand, for low- and middle-income countries, organic waste would form the majority percentage in the urban waste stream, ranging from 40 to 85% of the total [7]. Paper, plastic, glass, and metal fractions increase in the waste stream of the middle- and high-income countries [8].

Significant advancement in the standard of living will inevitably increase MSW generation [9]. This is because when the population's spending power increases, consumption of goods and services increases as well [10]. When consumption increases, the population will demand more resources like water, energy, minerals, and land which will give rise to the amount of waste generated [11]. Along with a higher standard of living, the increasing population from rural areas to urban areas had also contributed to a higher MSW generation rate [12]. Based on an estimation by World Bank, currently, 1.3 billion tons of waste is generated annually over the world, and this amount will increase to 2.2 billion tons annually by 2025 [13]. Thus, if waste is not managed properly, it would bring about an adverse impact on health, the economy, and the global environment. It was

```

Loop (1): a in range (until the width of vector x)
Loop (2): b in range (until the height of vector x)
a' = width (vector x) - a - 1
End Loop (2):
End Loop (1):

```

CODE 1

also reported that improper management of waste will result in a cost higher than what it would have cost if waste was managed appropriately in the first place. As such, like it or not, all parties will need to assume more responsibility for waste generation and disposal, specifically, product design and waste separation [14].

Considering that waste is mainly a by-product of consumer-based lifestyles that is the main driver of the world's economies, in most cases, the fastest way to reduce waste generation is to reduce economic activity [15]. However, a reduction in economic activity would not be attractive [16]. With that, EPA has recognized MSW recycling as the second most environmentally sound strategy for dealing with urban waste [17].

The recent ICT-driven solutions have played a substantial role in MSW management. The state-of-the-art technologies, i.e., computer vision, edge computing, IoT, machine learning, and deep learning, have been proven to be an enormous support to achieve industry 4.0 standards [18]. The current industry employs the latest MSW management tools that are economically and computationally inexpensive and efficient [19].

The accurate classification of MSW employing machine learning models has considerably benefited the MSW

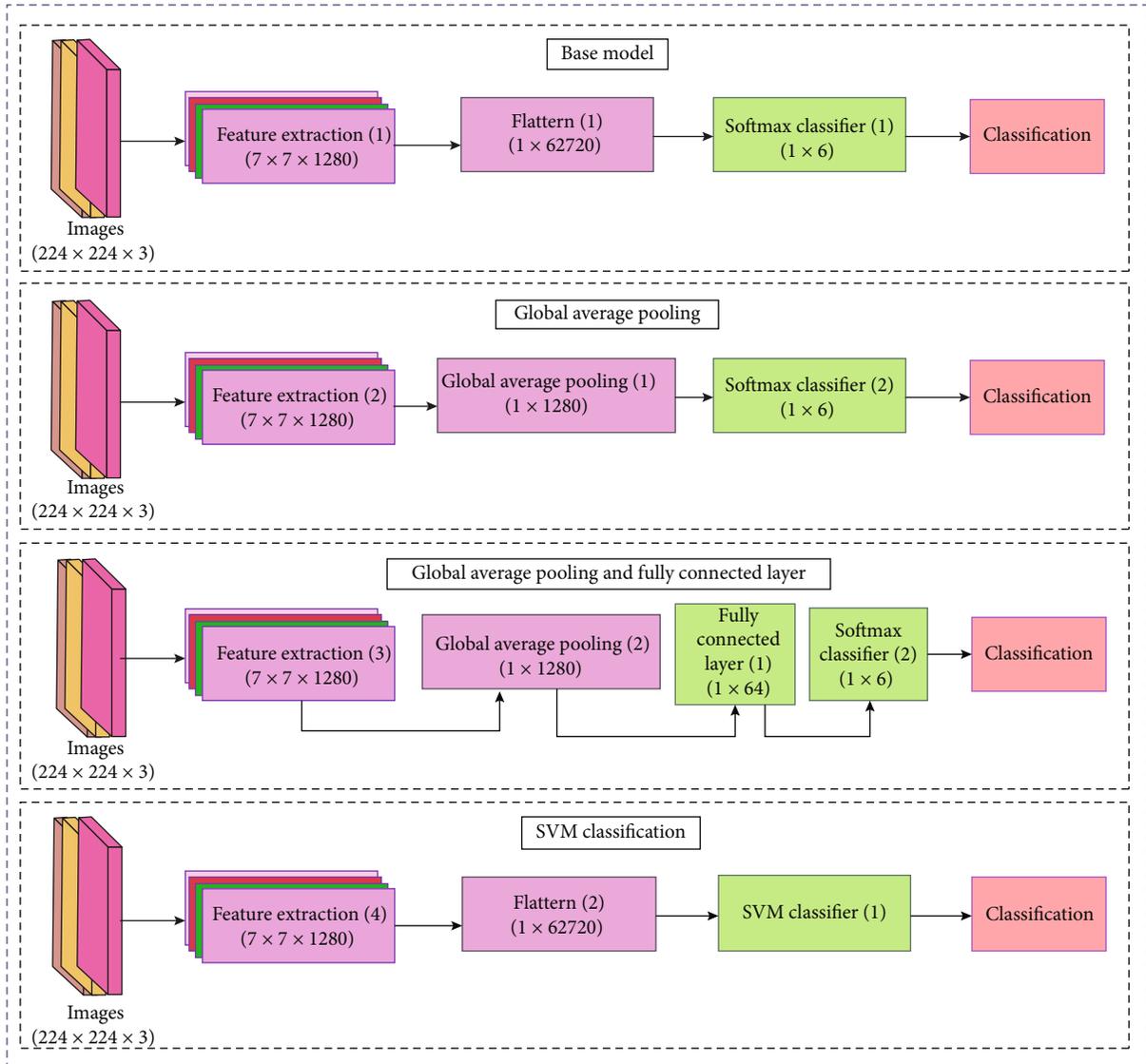


FIGURE 2: Component-based scenarios of the proposed classification model.

stakeholders. [20] proposed machine learning as a useful tool for accurate prediction and management of solid waste in Canada based on socio-economic and demographic variables. The study applied decision trees and neural networks to build the models. Ultimately, the study developed an integrated framework for accurate classification and management of municipal solid waste of 220 municipalities of Ontario. Similarly, [21] developed algorithms to understand the patterns of solid waste generations using machine learning and small area estimation techniques. The study incorporated the prediction of MSW to quantify the future estimates of waste generation to ease the recycling of waste materials for reusability. With these advancements, the ML algorithms, i.e., regression classification, naïve Bayes classification, support vector machine classification, decision trees, KNN, random forests, and CNN, have bestowed computationally inexpensive and fruitful solutions for MSW classifications. As the main advantages of these solutions, efficient resource utilisation has given economical solutions to local

and urban waste management [22]. Thus, significantly assisted in achieving a sustainable and pollution-free environment and enhanced the quality and standards of lives of common people [23].

2. Related Works

The study observes many citations related to forecasting future solid waste generation and its quantification, e.g., [24]. Similarly, another good number of citations can be observed for MSW using sensors and IoT devices, e.g., [25, 26]. On contrary, the study could find a little citable work related to optimization or enhancement of solid waste classification methods developed for edge computing devices or other lightweight applications. Among the recent citations, [27] developed an automated recognition system employing a deep learning algorithm for the classification of solid waste objects as biodegradable and nonbiodegradable.

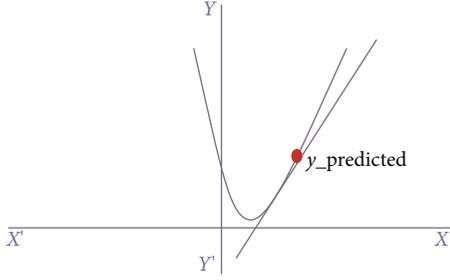


FIGURE 3: Gradient descent in the proposed approach.

```

Loop var i in range(10000):
y_predicted = m_curr * x + b_curr
md = -(2/n) * sum(x * (y - y_predicted))
yd = -(2/n) * sum(y - y_predicted)
m_curr = m_curr - rate * md
b_curr = b_curr - rate * yd
End Loop

```

CODE 2

Moving on, [28] used two models, namely, support vector machine (SVM) and CNN for waste image classification. SVM was used based on the principle of processing features in space higher than the original features space. Using a higher feature space, the data was divided into different categories using a hyperplane. SVM computed the selection of planes to develop the largest separation of different categories among the data. For CNN, the study had implemented an architecture that was similar to AlexNet but smaller due to computational constraints. At that point, there were no publicly available datasets, and the study initially found Flickr Material Database and Google Images. However, the images could not accurately represent the state of recycled goods. Such issues resulted in the creation of a new dataset named TrashNet that contained around 2400 images with six different classes. Data augmentation techniques such as random rotation, random brightness control, random translation, random scaling, and random shearing were applied to increase the dataset size. Surprisingly, SVM performed better than CNN even though SVM is a much simpler model. The authors believed that the poor performance of CNN was due to suboptimal hyperparameters.

Another work done by [29] proposed an image classification model related to MSW management. The study employed a camera to picture the solid waste and categorize it according to five defined waste categories trained on a dataset of 400-500 images. In a similar context, [30] employed SVM for the supervised classification of solid waste. The study segregated the solid waste images into recyclable or nonrecyclable.

The study done by [31] leveraged the transfer learning using a pretrained network utilizing AlexNet, GoogLeNet, VGG-16, and ResNet to classify waste images with different categories with the TrashNet dataset. In the final output layer

TABLE 1: Training result (10-fold crossvalidation) for MobileNetV2 using Softmax classifier.

Fold	Accuracy	Loss	Precision	Recall
Fold 1	91.832	1.605	0.915	0.914
Fold 2	92.574	1.112	0.911	0.910
Fold 3	87.129	1.704	0.912	0.910
Fold 4	85.396	3.124	0.915	0.913
Fold 5	91.584	1.323	0.917	0.916
Fold 6	91.337	1.506	0.915	0.914
Fold 7	89.851	1.831	0.915	0.914
Fold 8	84.901	2.714	0.913	0.912
Fold 9	85.856	2.013	0.906	0.904
Fold 10	88.089	2.052	0.923	0.922
Average	88.855	1.898	0.914	0.913

of the CNN, different classifiers were used, namely, Softmax and support vector machine (SVM). The study segregated training and testing data into equal halves without performing any further data augmentation.

[32] used VGG-19 for the waste image classification. The study used data augmentation such as shear, rotation, zooms, and shifts in addition to resizing images to $224 \times 224 \times 3$ dimensions before feeding the network. The classification layer of pretrained VGG-19 was removed and replaced by a fully connected layer of 256 neurons with ReLU activation function with a dropout of 0.6 and batch normalization. The model was trained using a batch size of 32, and the authors observed a learning rate between 0.001 and 0.0001 bestowed better results. The lower validation accuracy depicted the overfitting of the model even though [31] and [32] had managed to train the CNN model with high accuracy. The CNN architecture used, e.g., AlexNet, contained about 62,378,344 parameters, and over 729 million FLOPs required classifying a single image. A deep architecture as AlexNet or VGG increases the model complexity. As a result, the deployment of these huge CNN models is often unaffordable for common computers and mobile devices. Thus, this had given rise to the need for a lightweight CNN architecture that could still achieve good results when compared to huge CNN models such as AlexNet and VGG.

To create a lightweight deep neural network that could be used for embedded vision applications and mobile applications, [33] proposed a new architecture named MobileNet which is based on depthwise separable convolution. Depthwise separable convolution is a depthwise convolution followed by a pointwise convolution. Using a depthwise separable layer, the number of multiplication operations required is reduced significantly compared to normal convolution. Although being smaller in size, MobileNet has managed to achieve good results compared to other popular models such as VGG and GoogLeNet.

Following the success of MobileNet, [34] proposed an improved version of MobileNet named MobileNetV2. The study incorporated two changes to MobileNet. The first change is to incorporate an expansion layer before the depthwise convolution. This module takes a low-dimensional

Index	Cardboard	Glass	Metal	Paper	Plastic	Trash
Cardboard	67	0	1	12	0	1
Glass	0	77	7	3	13	1
Metal	0	13	64	1	3	1
Paper	2	0	2	107	1	7
Plastic	0	18	2	5	69	3
Trash	1	0	1	4	2	20

FIGURE 4: Confusion matrix (test set) using Softmax classifier.

TABLE 2: Precision and recall (test set) for MobileNetV2 using Softmax classifier.

Category	Precision	Recall
Cardboard	0.957	0.827
Glass	0.724	0.752
Metal	0.797	0.768
Paper	0.826	0.916
Plastic	0.769	0.722
Trash	0.613	0.679
Average	0.781	0.777

compressed representation which is first expanded to high dimensions and later on pass-through 1×1 convolution to project it back to a low-dimensional representation. Hence, the expansion layer will always have more output channels than the input. The second change proposed by [34] is to include a residual block similar to ResNet. Each layer has batch normalization, and the activation function used is ReLU6. Only the output of the projection layer does not have an activation function. This is because the projection layer produces a low-dimensional output, and applying nonlinearity on it will destroy useful information.

The magnitude of expansion is defined by the expansion factor, which is known as the width multiplier, α in MobileNet. For example, if there is an input with 24 channels and α is 6, the expansion layer will create a new feature with $24 \times 6 = 144$ channels. Subsequently, the depthwise convolution will be applied followed by the projection layer (1×1 convolution) which will project the 144 channels back into a smaller number, e.g., 24. The expansion and projection module is named as bottleneck layer [35]. Ideally, the number of channels should be large so that more information can be extracted. Combining both the expansion layer and projecting layer enables the model to extract more information while keeping the feature dimension rather small.

3. Methodology

In this research, we intend to develop an enhanced image classification model based on the lightweight feature extraction model MobileNetV2 to classify waste images into differ-

TABLE 3: Training result (10-fold crossvalidation) for MobileNetV2 using SVM classifier.

Fold	Accuracy	Precision	Recall
Fold 1	94.307	0.954	0.943
Fold 2	93.812	0.942	0.923
Fold 3	94.802	0.956	0.937
Fold 4	91.584	0.926	0.901
Fold 5	94.307	0.931	0.939
Fold 6	95.050	0.959	0.935
Fold 7	92.574	0.934	0.907
Fold 8	95.050	0.948	0.954
Fold 9	95.782	0.962	0.944
Fold 10	95.533	0.960	0.927
Average	94.280	0.947	0.931

ent categories according to their material. The prevailing citations of referred work in waste image classification employed convolution neural network models that were found to be larger and computationally expensive.

Let us consider a collection $S = \{x_1, x_2, x_3, x_4, \dots, x_n\}$, $n \leq N$ contains n images from N image space. Here, an image \mathbf{x} is a vector such that $\mathbf{x} \in R^D$ with D number of dimensions. Let us consider $\mathbf{x} \in R^{A \times B \times C}$ where each \mathbf{A} , \mathbf{B} , and \mathbf{C} represent the row vector, column vector, and color vector having specific indices i, j, k defined by $0 \leq i \leq D$, $0 \leq j \leq D$, and $0 < k \leq 3$.

Figure 1 depicts a general architecture of the proposed classification model. The input vector \mathbf{x} is passed through a series of preprocessing steps explained below. The feature extraction of output vector \mathbf{x}' assists in accurate classification.

We breakdown the preprocessing steps of image vector \mathbf{x} as follows:

- (1) We apply the horizontal and vertical shearing to image \mathbf{x} with coordinates a and b as described in Equations (1) and (2)

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a + mb \\ b \end{pmatrix} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \quad (1)$$

Index	Cardboard	Glass	Metal	Paper	Plastic	Trash
Cardboard	69	0	3	9	0	0
Glass	0	77	9	2	13	0
Metal	0	7	68	3	4	0
Paper	3	1	1	108	1	5
Plastic	2	6	2	2	82	3
Trash	0	0	2	4	3	19

FIGURE 5: Confusion matrix (test set) using SVM classifier.

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a \\ mx + b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \quad (2)$$

The horizontal and vertical shear displace the point of interest to adjust image \mathbf{x} for the desired shear.

- (2) We apply the zooming to image vector \mathbf{x} and achieve another image \mathbf{x}' such that \mathbf{x} is ($z * 10\%$) of vector \mathbf{x} . Here, z is a particular point in image \mathbf{x} having coordinates a and b . We can assume that zooming will give us zoomed image with points (za, zb) that specify zooming points with a displacement z . It is important to note down that various zooming can be applied to predict the network better
- (3) The horizontal flip applied to image vector \mathbf{x} with coordinates a and b produces an image \mathbf{x} with coordinates a' and b' such that $a' = \text{width}(\text{vector } \mathbf{x}) - a - 1$ while $b' = b$. We can describe this phenomenon as follows

Let us define the component-based scenario of our convolution neural network as

$$x^P \rightarrow w^P \rightarrow x^Q \rightarrow w^Q \rightarrow x^Z \rightarrow w^Z \rightarrow o^T. \quad (3)$$

Here, vector \mathbf{x} is an input image that is handled by the first layer of the network with weight \mathbf{w} (the vector that processes the input \mathbf{x} by applying weight \mathbf{w} , and the result of this computation is forwarded to the second layer). Next, the vector \mathbf{x}^Q , which was the output of the first layer, now serves as an input for the second layer managed by weight \mathbf{w}^Q . This process keeps on going for a defined number of layers until the outcome \mathbf{O}^T (where T represents the number of output channels, i.e., prediction of an image in two or more classifications) is reached. It is important to describe that \mathbf{o}^T is a trade-off between the last predicted value of vector \mathbf{x}^Z with the desired target (e.g., M) defined by Euclidean distance $d = 1/2 |\mathbf{M} - \mathbf{o}^T|^2$. The distance d can also be measured using other measures, e.g., Manhattan distance.

In an ideal scenario, the number of channels should be larger so to extract more information. Integration of two layers, i.e., expansion and projection layer, can significantly enable the model for extracting additional information. This

TABLE 4: Precision and recall (test set) for MobileNetV2 using SVM classifier.

Category	Precision	Recall
Cardboard	0.932	0.840
Glass	0.831	0.733
Metal	0.791	0.878
Paper	0.844	0.908
Plastic	0.794	0.835
Trash	0.840	0.750
Average	0.839	0.824

TABLE 5: Training result (10-fold crossvalidation) global average pooling using Softmax classifier.

Fold	Accuracy	Loss	Precision	Recall
Fold 1	88.861	0.323	0.909	0.820
Fold 2	87.624	0.384	0.913	0.816
Fold 3	85.891	0.437	0.914	0.819
Fold 4	82.426	0.499	0.918	0.828
Fold 5	89.604	0.302	0.911	0.813
Fold 6	89.109	0.373	0.913	0.821
Fold 7	88.614	0.301	0.912	0.819
Fold 8	85.891	0.349	0.913	0.819
Fold 9	89.578	0.336	0.911	0.817
Fold 10	88.337	0.372	0.914	0.820
Average	87.594	0.368	0.913	0.819

holds an added value of keeping the feature dimension even small in size. The activation function ReLU6 can also be significant in this context.

Figure 2 showed the summarized components based on various scenarios of the proposed classification model. The adjustment of weights is essentially important to reduce the error; a gradient descent can be adopted as described in Figure 3 and below algorithm.

Let us consider slope and intercept to be defined by m_{curr} and b_{curr} , respectively, and both initialized to 0. Let us assume the initial learning rate as 0.01. Let n represent the length of vector \mathbf{x} defined by $n = \text{len}(\mathbf{x})$.

Index	Cardboard	Glass	Metal	Paper	Plastic	Trash
Cardboard	67	0	3	11	0	0
Glass	0	69	6	1	23	2
Metal	0	5	68	3	6	0
Paper	5	0	1	108	0	5
Plastic	2	4	2	3	85	1
Trash	1	0	3	4	4	16

FIGURE 6: Confusion matrix (test set) using Softmax classifier global average pooling layer.

The range specifies the number of iterations; we can change the range to ensure a considerable reduction in the corresponding error value.

Image preprocessing, simulation of model, and classification are an end-to-end process described as follows:

- (1) Split the entire dataset into a training set and test set whereby the training set can be denoted as set S of n labelled examples: $S = \{x_i, i = 1, 2, 3 \dots n\}$ where $n = 2,019$ or $\sim 80\%$ of the entire dataset, while the test set has 508 images
- (2) Perform data augmentation only on the training set to increase the number of training images. Types of augmentation performed: shearing, zooming, and horizontal flipping of images
- (3) Upon performing data augmentation, the number of the training set is doubled where $n = 4,038$
- (4) The training set is fed to a pretrained MobileNetV2 CNN, which acts as a feature extractor. Note that the final output layer of the pretrained MobileNetV2 is removed
- (5) The extracted features are used as input to two classifiers, i.e., Softmax and SVM, with 10-fold crossvalidation
- (6) Both models are evaluated using the test set

4. Results and Discussion

We have employed TrashNet [32] in the research. The dataset contains 2,527 images of six waste category, namely, cardboard, glass, metal, paper, plastic, and thrash. The images were resized to dimension $224 \times 224 \times 3$ as required by the MobileNetV2. Data Augmentation such as shear, rotation, zooms, and shifts were used. We split the dataset into training and test sets for comparative analysis.

We train the model using system hardware with 16 GB RAM equipped with NVIDIA RTX 2060 GPU (6GB). To utilize GPU via Python, NVIDIA CUDA, and cuDNN are installed on the system.

4.1. Performance of MobileNetV2 Using Softmax. The feature extracted by the pretrained MobileNetV2 obtained from Keras was flattened and input into the newly trained Softmax

TABLE 6: Precision and recall (test set) for MobileNetV2 using Softmax classifier global average pooling layer.

Category	Precision	Recall
Cardboard	0.907	0.840
Glass	0.873	0.683
Metal	0.827	0.817
Paper	0.831	0.908
Plastic	0.712	0.866
Trash	0.640	0.571
Average	0.798	0.781

TABLE 7: Training result (10-fold crossvalidation) for MobileNetV2 with an additional fully connected layer.

Fold	Accuracy	Loss	Precision	Recall
Fold 1	91.337	0.205	0.936	0.889
Fold 2	91.337	0.348	0.936	0.894
Fold 3	88.366	0.387	0.934	0.887
Fold 4	94.059	0.172	0.933	0.885
Fold 5	89.604	0.393	0.931	0.883
Fold 6	89.851	0.295	0.934	0.886
Fold 7	91.089	0.274	0.934	0.887
Fold 8	91.584	0.249	0.932	0.881
Fold 9	91.067	0.255	0.936	0.891
Fold 10	91.067	0.282	0.935	0.888
Average	90.936	0.286	0.934	0.887

classifier. As there are six classes for the waste image dataset, the Softmax classifier layer has 6 nodes. This model will be used as the base model which will be compared with MobileNetV2 using an SVM classifier. Table 1 shows the accuracy, loss, precision, and recall throughout the 10-fold crossvalidation. The 10-fold crossvalidation is a good approximation method that gives unbiased and equal chance to all data points to be used fairly in the training of the model, and it ultimately enhances the accuracy, precision, and recall.

In terms of accuracy, the average accuracy for the 10-fold validation is 88.85%, with an average precision of 0.91 and an average recall of 0.91 as well. Generally, precision and recall

Index	Cardboard	Glass	Metal	Paper	Plastic	Trash
Cardboard	69	0	1	10	0	1
Glass	0	86	5	2	5	3
Metal	0	6	68	1	1	6
Paper	6	0	1	103	0	8
Plastic	1	10	2	3	73	7
Trash	1	0	0	2	0	25

FIGURE 7: Confusion matrix (test set) using MobileNetV2 using Softmax classifier with an additional fully connected layer.

close to 1 indicate that the model performs well in predicting all 6 different categories contained in our dataset. The accuracy of the model on the test set is 79.53%.

Figure 4 presents the confusion matrix of different objects and their classification. The blue cells depict the correct identification of objects related to other objects. Further, the precision and recall are shown in Table 2.

4.2. Performance of MobileNetV2 Using SVM Classifier. Similar to the base model presented above, the MobileNetV2 SVM model received the same features extracted by the pre-trained MobileNetV2. Instead of using the Softmax function, SVM was used as the classifier. C parameter in SVM dictates the amount of regularization performed. Small C values led to the SVM classifier setting a larger margin when performing classification, while a larger C value led to a smaller margin. A larger margin (smaller C values) increased the misclassification rate on the training set but tended to perform better on the test set as it is more generalized. In this study, the default C value of 1.0 is used.

Table 3 shows the accuracy, loss, precision, and recall throughout the 10-fold crossvalidation for SVM. The SVM model resulted in an average accuracy of 94.28%, average precision of 0.947, and recall of 0.931 which are higher than the 10-fold crossvalidation of the base model (Softmax classifier).

Figure 5 presents the confusion matrix of different objects and their classification. The blue cells depict the correct identification of objects related to other objects. Further, the precision and recall are shown in Table 4.

The MobileNetV2 using the SVM model is then evaluated using the test set, and it provided an accuracy of 83.26%, which is approximately 4% higher than the base model. This is aligned with the result obtained by [31], which had also stated that using the SVM classifier is more accurate when paired together with AlexNet, GoogLeNet, ResNet, and VGG-16 if compared to the Softmax classifier.

4.3. Performance of MobileNetV2 with Global Average Pooling Using Softmax Classifier. Based on the results above, it was shown that the SVM model is more accurate than the base model. However, the base model could be further enhanced to increase the accuracy.

A global average pooling layer is added to the model. To illustrate how the global average pooling layer works, given

TABLE 8: Precision and recall (test set) for MobileNetV2 using Softmax classifier with an additional fully connected layer.

Category	Precision	Recall
Cardboard	0.897	0.864
Glass	0.844	0.802
Metal	0.821	0.841
Paper	0.86	0.874
Plastic	0.88	0.753
Trash	0.478	0.786
Average	0.797	0.820

the feature extracted by the pretrained MobileNetV2 model has a dimension of $7 \times 7 \times 1280$, after applying global average pooling, the output dimension will be 1×1280 .

Table 5 depicts the training result of 10-fold crossvalidation using global average pooling with Softmax classifier with the global average-pooling layer, the model provided an average accuracy of 87.59%, average precision of 0.913, and an average recall of 0.819 from 10-fold crossvalidation. The training accuracy is slightly lower than the base model.

However, the accuracy on the test set is 81.10%, which is higher than the base model. This could be due to the global average-pooling layer which may enable the model to generalize better on the test set.

Figure 6 presents the confusion matrix of different objects and their classification. The blue cells depict the correct identification of objects related to other objects. Further, the precision and recall are shown in Table 6.

4.4. Performance of MobileNetV2 with Additional Fully Connected Layer. Using the model with a global average pooling layer, an additional fully connected layer of 64 neurons was added. The new fully connected layer was added so that the model could pick up more complex features, which may be helpful for the waste image dataset.

Table 7 presents the training result of 10-fold crossvalidation for MobileNetV2 with an additional fully connected layer. With 10-fold crossvalidation, MobileNetV2 with an additional fully connected layer provided an average accuracy of 90.93%, average precision of 0.934, and an average recall of 0.887.

Figure 7 presents the confusion matrix of different objects and their classification. The blue cells depict the correct identification of objects related to other objects.

Table 8 describes the precision and recall (test set) for MobileNetV2 using Softmax classifier with an additional fully connected layer. With the newly added fully connected layer, the model achieved an accuracy of 83.46% on the test set, which was higher than the model with a global average pooling layer (81.10%).

5. Conclusion

With the significant increase of human living standards, solid waste management, especially recycling, must be emphasised to prevent the adverse environmental problem before it is too late. As the recycling effort requires proper waste segregation by the public, data science, specifically deep learning paired with smart mobile phones, could be a tool to aid the effort in recycling. The development and advancement of CNN and the ability to perform transfer learning are crucial for image classification and model development purposes. Architecture such as VGG16 performs very accurately for image classification. However, the sheer size of a large CNN model, i.e., the number of parameters and operations required, is not feasible for mobile applications.

Although the accuracy of the MobileNetV2 with SVM classifier for waste image classification (83.46%) is lower than the accuracy of the VGG16 model with SVM classifier (97%), using MobileNetV2, which is a more lightweight CNN architecture, is perhaps more practical if the image classification model is to be implemented on mobile devices. Moreover, MobileNetV2 is indeed specifically designed for mobile applications.

In the current research, the feature extracted from MobileNetV2 is directly fed into the SVM classifier. To further improve the MobileNetV2 model SVM classifier, the hyperparameters could be tuned, and perhaps the feature extracted could be fed into multiple layers of the neural network before using SVM so that more complex features or relationships could be extracted. On top of that, the Softmax classifier could be further enhanced by adding more layers as well. The scope of this study is limited to solid waste classification only. We have used cardboard, glass, metal, paper, plastic, and trash but it can be scaled to a large variety of solid waste classification.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors have no conflict of interest in this research.

Acknowledgments

The authors are grateful to the Taif University Researchers Supporting Project number (TURSP-2020/215), Taif Univer-

sity, Taif, Saudi Arabia. This work is also supported by the Faculty of Computer Science and Information Technology, University of Malaya, under Postgraduate Research Grant (PG035-2016A).

References

- [1] A. A. Olajire, "The brewing industry and environmental challenges," *Journal of Cleaner Production*, vol. 256, p. 102817, 2020.
- [2] United States Environmental Protection Agency, *Solid Waste and Emergency Response (5306P)*, Washington, DC 20460, 2010, https://archive.epa.gov/epawaste/nonhaz/municipal/web/pdf/msw_2010_factsheet.pdf.
- [3] U. EPA, *What is integrated solid waste management? Solid Waste and Emergency Response*, Environmental Protection Agency (EPA), United States, 2002.
- [4] US EPA, *Advancing Sustainable Materials Management: Facts and Figures 2013*, United States Environmental Protection Agency, 2015, EPA530-R-15-003.
- [5] H. I. Abdel-Shafy and M. S. M. Mansour, "Solid waste issue: sources, composition, disposal, recycling, and valorization," *Egyptian Journal of Petroleum*, vol. 27, no. 4, pp. 1275–1290, 2018.
- [6] M. A. Kamaruddin, M. S. Yusoff, L. M. Rui, A. M. Isa, M. H. Zawawi, and R. Alrozi, "An overview of municipal solid waste management and landfill leachate treatment: Malaysia and Asian perspectives," *Environmental Science and Pollution Research*, vol. 24, no. 35, pp. 26988–27020, 2017.
- [7] D. Hoornweg and P. Bhada-Tata, *What a Waste: A Global Review of Solid Waste Management*, World Bank, 2014.
- [8] J. R. Jambeck, R. Geyer, C. Wilcox et al., "Plastic waste inputs from land into the ocean," *Science*, vol. 347, no. 6223, pp. 768–771, 2015.
- [9] S. Kumar, J. K. Bhattacharyya, A. N. Vaidya, T. Chakrabarti, S. Devotta, and A. B. Akolkar, "Assessment of the status of municipal solid waste management in metro cities, state capitals, class I cities, and class II towns in India: an insight," *Waste management*, vol. 29, no. 2, pp. 883–895, 2009.
- [10] J. Malinauskaitė, H. Jouhara, D. Czajczyńska et al., "Municipal solid waste management and waste-to-energy in the context of a circular economy and energy recycling in Europe," *Energy*, vol. 141, pp. 2013–2044, 2017.
- [11] E. I. Ugwu, A. C. Ekeleme, P. O. Awoyera, H. O. Ozioko, and U. Osinachi, "Effect of municipal solid waste contamination on some geotechnical properties of soil," *Journal of Materials and Environmental Sciences*, vol. 9, no. 2, pp. 585–590, 2018.
- [12] R. G. Alfaia, A. M. Costa, and J. C. Campos, "Municipal solid waste in Brazil: a review," *Waste Management and Research*, vol. 35, no. 12, pp. 1195–1209, 2017.
- [13] R. M. Rodzi, Z. M. Nopiah, and N. E. A. Basri, "Analysis of solid waste generation and composition in Malaysia TVET campus," *International Journal of Integrated Engineering*, vol. 11, no. 2, 2019.
- [14] B. Esmaeilian, B. Wang, K. Lewis, F. Duarte, C. Ratti, and S. Behdad, "The future of waste management in smart and sustainable cities: a review and concept paper," *Waste Management*, vol. 81, pp. 177–195, 2018.
- [15] W. A. Qazi, M. F. M. Abushammala, and M. H. Azam, "Multi-criteria decision analysis of waste-to-energy technologies for municipal solid waste management in Sultanate of Oman,"

- Waste Management & Research*, vol. 36, no. 7, pp. 594–605, 2018.
- [16] N. Pour, P. A. Webley, and P. J. Cook, “Potential for using municipal solid waste as a resource for bioenergy with carbon capture and storage (BECCS),” *International Journal of Greenhouse Gas Control*, vol. 68, pp. 1–15, 2018.
- [17] I. R. Istrate, D. Iribarren, J. L. Gálvez-Martos, and J. Dufour, “Review of life-cycle environmental consequences of waste-to-energy solutions on the municipal solid waste management system,” *Resources, Conservation and Recycling*, vol. 157, p. 104778, 2020.
- [18] P. Yadav and S. R. Samadder, “A critical review of the life cycle assessment studies on solid waste management in Asian countries,” *Journal of Cleaner Production*, vol. 185, pp. 492–515, 2018.
- [19] G. Bertanza, E. Ziliani, and L. Menoni, “Techno-economic performance indicators of municipal solid waste collection strategies,” *Waste Management*, vol. 74, pp. 86–97, 2018.
- [20] M. Kannangara, R. Dua, L. Ahmadi, and F. Bensebaa, “Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches,” *Waste Management*, vol. 74, pp. 3–15, 2018.
- [21] C. E. Kontokosta, B. Hong, N. E. Johnson, and D. Starobin, “Using machine learning and small area estimation to predict building-level municipal solid waste generation in cities,” *Computers, Environment and Urban Systems*, vol. 70, pp. 151–162, 2018.
- [22] K. Pardini, J. J. Rodrigues, S. A. Kozlov, N. Kumar, and V. Furtado, “IoT-based solid waste management solutions: a survey,” *Journal of Sensor and Actuator Networks*, vol. 8, no. 1, p. 5, 2019.
- [23] O. Adedeji and Z. Wang, “Intelligent waste classification system using deep learning convolutional neural network,” in *Procedia Manufacturing*, vol. 35, pp. 607–612, 2019.
- [24] M. Abbasi and A. El Hanandeh, “Forecasting municipal solid waste generation using artificial intelligence modelling approaches,” *Waste Management*, vol. 56, pp. 13–22, 2016.
- [25] J. N. Kalshetty, A. Anil, H. Jain, S. Beli, and A. K. Singh, “Smart monitoring for waste management using internet of things,” in *International Conference on Intelligent Data Communication Technologies and Internet of Things*, pp. 96–103, Cham, 2020.
- [26] A. Anil, H. Jain, S. Beli, A. K. Singh, and J. N. Kalashetty, “Smart monitoring for waste management using IoT,” *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 7502–7505, 2019.
- [27] S. Sudha, M. Vidhyalakshmi, K. Pavithra, K. Sangeetha, and V. Swaathi, “An automatic classification method for environment: friendly waste segregation using deep learning,” in *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, Chennai, India, 2016.
- [28] M. Yang and G. Thung, *Classification of Trash for Recyclability Status*, 2016, CS229Project Rep..
- [29] M. Satvilkar and P. N. Cosgrave, *Image Based Trash Classification Using Machine Learning Algorithms for Recyclability Status*, Doctoral dissertation, National College of Ireland, Dublin, 2018.
- [30] A. Arayakandy, “Design and development of classification model for recyclability status of trash using SVM,” *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 3, pp. 2146–2150, 2019.
- [31] U. Özkaya and L. Seyfi, “Fine-tuning models comparisons on garbage classification for recyclability,” 2018, <https://arxiv.org/abs/1908.04393>.
- [32] T. Kennedy, “OscarNet: using transfer learning to classify disposable waste,” *CS230 Report: Deep Learning*, Stanford University, CA, Winter, 2018.
- [33] A. G. Howard, M. Zhu, B. Chen et al., “MobileNets,” 2017, <https://arxiv.org/abs/1704.04861>.
- [34] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, “MobileNetV2: inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Salt Lake City, Utah, state of Utah, USA, 2018, Calvin L. Rampton Salt Palace Convention Center the week of June 18-22.
- [35] M. Hollemans, “MobileNet version 2,” *machine think BLOG*, 2018.