Android Waste Classification

Sean Sutherland[1], Richard Robinson[2], Emily Mendelson[3], Tung Pham[4]

QMIND – Queen's AI Hub
Queen's University, Kingston, Ontario K7L 3N6, Canada. Queen's

1 e-mail: sean.sutherland@queensu.ca
2 e-mail: 16rer@queensu.ca
3 e-mail: 17ecm1@queensu.ca
4 e-mail: 18tp12@queensu.ca

Abstract: *The gradual growth of solid waste in the urban area has been and is becoming a great concern for human health, and could result in environmental pollution and may be hazardous to humanity if not properly disposed. With that being said, an advanced waste management system is necessary to manage a variety of waste materials. The most important step of waste management is the separation of the waste into their designated categories. This process is usually done by manually hand-picking and sorting them into their designated bins. In order to save time and simplify the process, Queen's Waste Wizard introduces a waste classification computer vision model, which is developed using a pre-trained residual net (EfficientNet-B3) Convolutional Neural Network model, a machine learning tool. The model is used to classify waste into different groups such as* ***Blue Recycling****,* ***Grey Recycling****,* ***Landfill*** *and* ***Organics****. Our first proposed system was able to achieve an accuracy of 98% on the validation dataset. An overall 80% accuracy was achieved on our testing dataset with more realistic examples with background noise. This method is expected to be faster, by implementing the proposed system without or little human involvement.*

1. INTRODUCTION

1.1 Motivation

Globally, annual solid waste is expected to reach 2.2 billion tons by 2025 [1]. Improper waste management may lead to huge economic, environmental, and public health issues. As a result, there is a clear need for proper waste management within public spaces where a large proportion of waste is improperly discarded.

In 2016, the waste diversion rate of Queen's University was measured to be 43.23%, however, over 85% of the current waste stream was composed of items that can be diverted from landfills [2]. While Queen's University has already deployed a waste lookup application, this tool requires a high amount of user input. In contrast, computer vision models only require the user to take a single image of the waste item for classification. This type of application may help increase the waste diversion rate on campus by reducing the amount of misclassified waste.

While there are existing computer vision models for waste classification, there is limited use of such models within public spaces. Many models including those using AlexNet [3] and Inception-ResNet [4] have been trained for waste classification, however, these models are large and require a large number of parameters to achieve high accuracy. Consequently, many of these models are unable to run on easily deployable machines such as tablets.

1.2 Related Works

According to a paper published in 2019 by Quoc V. Le and Mingxing Tan from Cornell University, EfficientNet is a continuous family of models created by scaling each dimension with a fixed set of scaling coefficients. As a result, the depth, width and resolution of each variant of the EfficientNet models should be hand-picked to determine the best accuracy.

For instance, in a model used to classify Stanford Dogs, the model EfficientNet B0 was used. It has been shown in the study that transfer learning result is better for increased resolution if input images remain small. However, when training EfficientNet on smaller datasets, the model faces a risk of overfitting its data. Hence, data augmentation and pre-processing are important for EfficientNet. In other words, a useful tip is that in some cases, it might be beneficial to unfreeze only a portion of the layers rather than all, as this makes fine-tuning much faster when using larger models like B7. Another aspect to keep in mind is that larger variants of EfficientNet do no guarantee improved performance, especially for tasks with little data and few classes. In such a case, the larger variant of EfficientNet chosen, the harder it is to tune hyperparameters. In conclusion, it's important that the developers take time and experiment with all variants and play around with the layers in order to receive the best accuracy.

1.3 Problem Definition

The use of EfficientNet for waste classification may be ideal in public spaces where models with larger and more complex architectures are unable to run on small devices. Models such as ResNet scale up Convolutional Neural Networks (ConvNets) by adding more layers, and by scaling by depth [5]. However, it is not known if this is the most efficient scaling algorithm as previously, the process of scaling up ConvNets was poorly understood. This creates a problem in situations where models require a smaller size, yet still require high accuracy.

EfficientNet uses a new method of scaling to achieve better accuracy and efficiency greater than most traditional ConvNets [5]. Unlike conventional approaches to model scaling, where network dimensions are arbitrarily scaled, EfficientNet scales each dimension with a fixed set of scaling coefficients [5]. This results in a higher level of accuracy and efficiency. Furthermore, EfficientNet has several different versions along with EfficientNetLite versions that are specifically designed to run on mobile devices [5].

For waste classification, it is necessary for a model to classify waste quickly and accurately. Furthermore, when deployed in a public space, it is also necessary for the model to run on devices with limited storage capacity. Due to the high efficiency and accuracy of the EfficientNet, it is an optimal computer vision model to retrain for the purpose of waste classification.

2. METHODOLOGY

2.1 Training Data

The goal of our model is to successfully classify common waste items into four different categories. These categories are blue recycling (glass, plastic, and metal), grey recycling (paper and cardboard), along with organics, and landfill. These were based off the sorting categories in Kingston Ontario, as this is the preliminary location the model will be deployed. Using numerous public databases online, a collection of 4637 images was established for training data, summarized in Table 1. The quantity of each category was modified over time to reflect the difficulty the model had of classifying that category.

Table 1: Summary of Training Data

Category	**Image Quantity**
Blue Recycling	889
Landfill	1046
Organic	2301
Grey Recycling	404
Total	**4637**

2.2 Model Framework

To maximize the accuracy of the model, extensive research was conducted to determine the most appropriate framework. EfficientNet was found to be the most appropriate for this application as it can achieve high accuracy on the ImageNet dataset, while minimizing the number of parameters. This is very important for this application as the model will be run on an android tablet with limited computing power and must be capable of classifying an image in under a second. EfficientNet models between B0 and B5 were tested on our data, and it was found that the increase in input size from B0 to B3 caused significant improvements in accuracy, but further scaling had limited improvements. It was therefore determined that using the EfficientNet B3 framework using pretrained weights from the ImageNet dataset was the most appropriate baseline model. To tailor the model to this application, an average pooling layer, along with batch normalization, dropout, and fully connected layers were added on top of EfficientNet. A summary of the current model can be seen in Figure 1.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
efficientnetb3 (Functional)  (None, 10, 10, 1536)      10783535
_________________________________________________________________
global_average_pooling2d (Gl (None, 1536)              0
_________________________________________________________________
batch_normalization (BatchNo (None, 1536)              6144
_________________________________________________________________
dropout (Dropout)            (None, 1536)              0
_________________________________________________________________
dense (Dense)                (None, 768)               1180416
_________________________________________________________________
dense_1 (Dense)              (None, 4)                 3076
=================================================================
Total params: 11,973,171
Trainable params: 1,186,564
Non-trainable params: 10,786,607
_________________________________________________________________
```

Figure 1: Framework of image classification model

2.3 Training

The model was then trained for 20 epochs, until the validation set accuracy plateaued. to increase the size of the training dataset, data augmentation was used including rotating, zooming, shifting and flipping the preliminary training images. The model was trained in mini batches of 128 images, and learning rate decay was used to maximize the validation set accuracy. Before training, 10% of the data was set aside as the validation set which was used to measure the progress of the training. The accuracy and loss function of the training and validation sets are seen below in Figure 2.

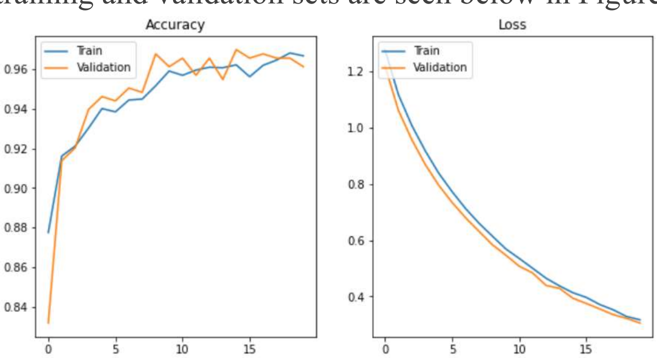

Figure 2 Accuracy and Loss function of training and validation data

As the training and validation set are not entirely representative of actual images seen by the model was deployed, a test set was developed to evaluate the model more accurately. Each of the four team members took approximately 100 images of common waste items around their home, in situations more representative of what the model will be expected to classify. These were then run through the model to predict how accurately the system would perform when deployed.

3. RESULTS AND DISCUSSION

By training the model shown above for 20 epochs, a peak validation set accuracy of 97% was achieved. The predicted categories compared to the true categories of the validation set for each of the categories is visualized in the confusion matrix in Figure 3.

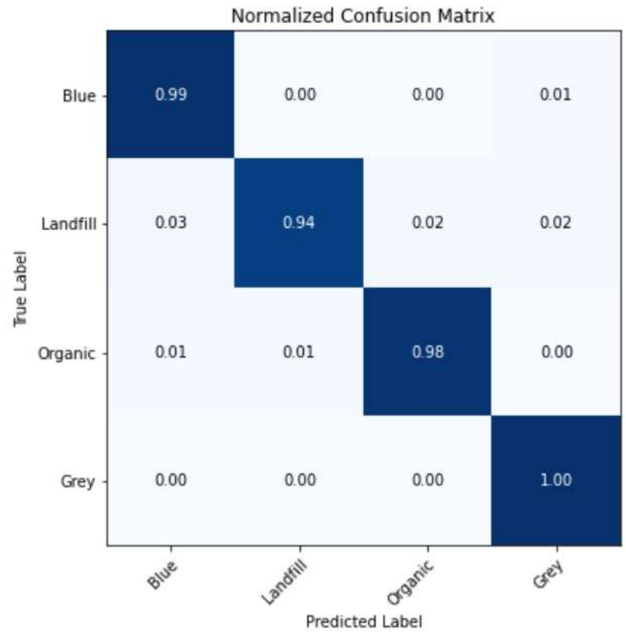

Figure 3: Normalized confusion matrix of validation set

The testing data collected by the team was then run through the model, producing an accuracy of 80%. This is likely representative of the accuracy the model achieves when deployed around campus, indicating that there is still further work to be done. The predicted categories compared to the true categories of the test set for each of the categories is visualized in the confusion matrix in Figure 4.

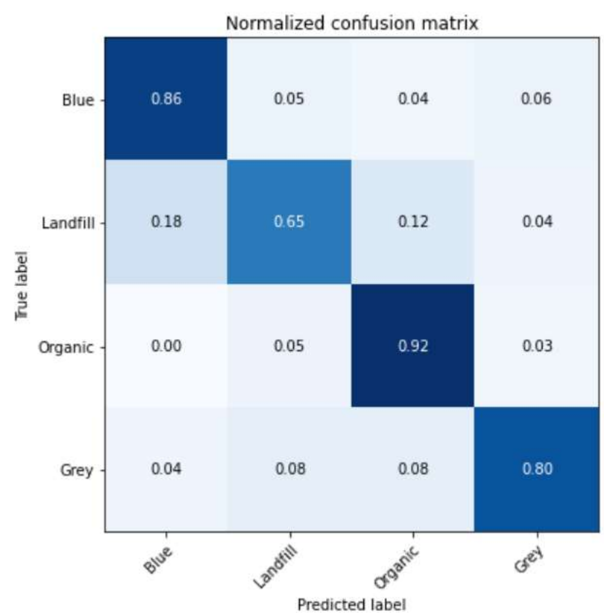

Figure 4: Normalized confusion matrix of test set

Testing revealed that there was a significant drop in accuracy between the validation set and test set. This means that the data used in the training and validation sets are not entirely representative of the testing set, which explains the drop in accuracy. When examining the training data, the majority of the photos are taken with white backgrounds, good lighting, and with the object taking up the majority of the frame. This is not the case with the testing data where the background is often a significant part of the image, taken in less than ideal lighting conditions. To improve the performance of the model on representative images, the training set would need to be further expanded, or the quality of the representative images would need to be improved by reducing background and lighting effects.

4. CONCLUSIONS AND FUTURE WORK

Recycling can be too easily contaminated when people do not ensure their waste is placed in the proper bin. This can be extremely damaging to the recycling initiatives and efficiencies. This project has accomplished the training and deployment of a convolutional neural network to properly classify waste items into their respective categories: blue recycling, grey recycling, landfill, and organic. This was done using transfer learning from the EfficientNet model which has been converted into a TensorflowLite model to be deployed locally on an Android tablet. This model will be used on Queen's campus to help Queen's students recycle more responsibly.

The model can properly classify common waste items with a 98% accuracy on the validation set. Even with background noise, activity in the background of images, that is commonly found in realistic deployment of software such as this one, the model is still able to perform with 80% accuracy.

Currently the development on the Android application is ongoing. Although the model is completely functional in the Android application the user experience is still being improved for ease of use to Queen's students. Along with improved UX additional resources are being implemented into the app so that in conjunction with the model the team can ensure students have assets easily accessible to completely responsibly dispose of their waste items. One of the major additions to the application is common exceptions with the waste disposal instructions. On campus certain products are specifically designed to be compostable even though visually entire plastic and similar cases. This is being done with location specific items so that managers of this software can easily set the location of the tablet to provide location specific instructions and suggestions.

Future steps also include the secure installation of Android tablets. To ensure the security of the tablets they are being installed with brackets at the most popular locations on campus. This way the model can be delivered with ease of use and peace of mind from any vandalism or theft.

REFERENCES

[1] D. Hoornweg, P. Bhada Tata, "What a Waste: A Global Review of Solid Waste Management, World Bank, 2012.

[2] GFL Environmental, "2016 Waste Audit Report", Queen's University, October 2016.

[3] Y. Chu, C. Huang, X. Xie, B. Tan, S. Kamal, X. Xiong, "Multilayer Hybrid Deep-Learning Method for Waste Classification and Recycling", Computational Intelligence and Neuroscience, 2018.

[4] V. Ruiz, A. Sanchez, J.F. Velez, B. Raducanu, "Automatic Image-Based Waste Classification", Bioinspired Systems and Biomedical Applications to Machine Learning, 2019.

[5] M. Tan, Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", International Conference on Machine Learning, 2019.