Waste Classification using Convolutional Neural Network on Edge Devices

Ashik Mohammed Sali, Harish Thampy, Immanuel S Vadakedam
Sunitha S Pillai, Assistant Professor
Department of Electronics and Communication
St. Joseph's College of Engineering and Technology, Palai, Kerala, India

Abstract:- Waste management is an important issue in the current scenario. Sorting of waste into different categorizes is one of the most important and tedious step in waste management.

Normally this is a manual (hand-picking) process which has its own cons and hence the need for an automated and efficient system to manage waste arises. Through this paper we propose an intelligent waste classification system, developed using a deep pre-trained (Xception) Convolutional Neural Network model that can classify solid wastes such as glass, metal, paper and plastic etc. on stand-alone edge devices. This system can be further deployed into a real time embedded system by adding a mechanism to physically separate wastes. The proposed system is trained on an open source dataset available online and is able to achieve a test accuracy of 92% on the dataset. Thus the system could make the separation process faster and intelligent without or reducing human involvement.

Keywords:- *Convolutional Neural Networks, Machine Learning, Deep Learning, Xception model, Dataset, Training, Sorting, Edge computing.*

I. INTRODUCTION

With waste generation increasing at an alarming rate, their disposal is really becoming a major problem. Currently the main method for dealing with waste is landfilling. However this method is inefficient, expensive and pollutes our natural environment. These landfills are filling up fast and can affect the health of people living around them. Another common method used for disposal is burning of waste causing air pollution. Hence arise the need to recycle solid waste in order to protect our environment and its inhabitants.

Sorting of waste can be subdivided into 2 processes. 1) Identification of waste (category) and 2) physical separation. In this paper we will be focusing on the first process. Many systems are being developed which aims to resolve this issue. Computing on the cloud is an option but providing stable network is a power hungry process and may not always be the best solution in remote areas. Hence, we focused on a solution which will be able to stand alone even in remote locations with minimal maintenance.

II. RELATED WORKS

The best way to identify a polymer is to study its spectroscopy. Even though it's fast and accurate, the high cost and maintenance factors make it inappropriate for commercial use. Different sensors can be integrated into the identification of waste [1].These systems use multiple sensors to detect different types of materials at different stages either using conveyor belts [2] or ramps. The disadvantage of such a design is that it's more complicated. The probability of failure increases especially considering the non-uniform shape of waste being put into these machines.

Another way to identify waste is by using image processing and machine learning. The increase in computing power and sophistication of algorithms has improved the accuracy of such models. Mindy Yang and Gary Thung compare the use of Support Vector Machine and Convolutional Neural Networks [3] to solve waste classification while providing a dataset. The accuracy received was not enough but another attempt [4] using Convolutional neural network (CNN) on the same dataset received an accuracy of 87%.

The concept of transfer learning relaxes the hypothesis that the training data must be independent and identically distributed with the test data, which motivates us to use transfer learning to solve the problem of insufficient training data[5].Basically we can use networks trained on other datasets (mostly similar objects) and transfer that knowledge to our case. Availability of large models trained on ImageNet [6] and the high accuracy [7] it offers compared to stand-alone models is the major reason for its increased popularity.

After waste is been identified, the next is to segregate it. A project at Columbia University by Huafeng Shi,Saurabh Bondarde and Vishakh B.V [8] demonstrated the separation of waste into two compartments using a flap mechanism. A camera with an infra-red sensor detects the image and is processed in the cloud. The result is implemented with the help of high torque motors that control the flaps. This method was again prototyped using Convolutional Neural Network [9].

TABLE I. THIS TABLE PROVIDES STARTING POINT FOR THE ACCURACY AMONG VARIOUS MODELS

Model	*Test accuracy*	*Data Aug.*	*Epoch*
ResNet	75%	-	100
MobileNet	76%	-	500
Inception ResNetV2	90%	+	200
DenseNet 121	85%	-	100
DenseNet 169	82%	+	100
DenseNet 201	85%	-	200
Xception	85%	+	100

III. DATASET

Our basic objective is to classify solid wastes for recyclable materials. So we recognized some of the most common recyclables. For this work, we are using a TrashNet dataset which was created by Gary Thung and Mindy Yang [3]. This dataset originally consists of 2527 images divided into six different classes, but we are only considering 5 classes among them for our application- glass, paper, plastic, metal and trash with each image consisting of different materials in different poses and lighting. Each image is resized down to to 512 x 384 pixel and size of original dataset is almost 3.5GB [3].Table I shows the number of images in each class and Fig 1 shows sample images from the dataset. We are considering only 5 classes.

TABLE II. DATASET INFORMATION

TrashNet Dataset		
SI.No	***Material Type***	***No.of each material***
1	Plastic	482
2	Glass	501
3	Paper	594
4	Trash	134
5	Metal	410
TOTAL		**2124**

Fig. 1. Sample images of dataset - Plastic,Glass,Paper,Trash,Metal

IV. METHODOLOGY

A. CNN-Convolutional Neural Network

Convolutional Neural Network (CNN) is a special type of deep neural network [10] that performs impressively in computer vision problems such as image classification, object detection [11], etc. As shown in Fig 2, CNN basically consists of two parts[12][13], a **convolutional base** composed of convolutional and pooling layers stacked together which are basically hidden and a **classifier** section consisting of fully connected layer. The major goal of convolutional base is feature extraction and classifier section is classification based on extracted features. Feature extraction part consist of a combination of linear and nonlinear operations i.e. convolutional operations and activation functions. These activation functions may be ReLU, sigmoid, tanh...etc. may be all in the same network based on the problem and type of data used.

ISSN No:-2456-2165

B. Transfer learning

Training an entire model from its scratch is a very tedious process due to lack of availability such a large dataset and large computational cost. In transfer learning we first train a base network on a base dataset and task, and then we repurpose the learned features or transfer them, to a second target network to be trained on a target dataset and task. This process will tend to work if the features are general, that is, suitable to both base and target tasks, instead of being specific to the base task. So it is preferred to pre-train the model on a huge dataset like ImageNet containing millions of images and then use it as a feature extractor or initialization. This method of using the knowledge gained in one or more source tasks and using it to improve learning in a related target task is called transfer learning [14].

C. Repurposing the pre-trained model

Transfer learning is achieved through pre-trained model. So when we have to reuse [13] a pre-trained model, firstly we load layers from the pre-trained model and then freeze them so as to avoid destroying any of the information they have during any of future rounds. Now we remove the original classifier section and add a new classifier section that fits our purpose. Finally the model is trained on our dataset and then fine tuning of the model is carried out according to one of the three following strategies.

- Train the entire model
- Train some layers and leave the other layers frozen
- Freeze the convolutional base and tune the classifier section

The type of strategy adopted depends upon the size of dataset used and the similarity of our dataset with that dataset used for the pre-trained model.

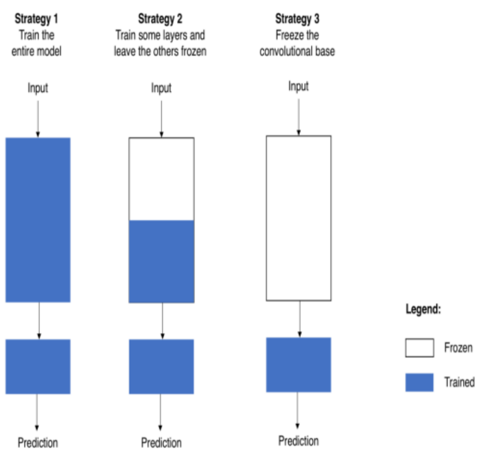


Fig. 2. Three different fine tuning strategies

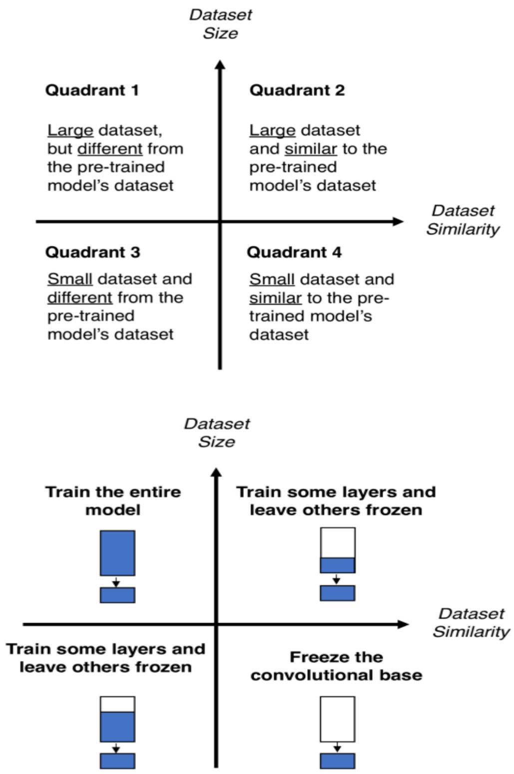


Fig. 3. Size-similarity matrix (top) and decision map for fine tuning pre-trained models (bottom)

This can be easily decided based on the size-similarity matrix and decision map given in Fig 3.So from practical perspective the entire transfer learning process can be summarized as

1. Selecting a pre-trained model
2. Classifying our problem according to size-similarity matrix
3. Fine tune the model

D. Xception model

We are using Xception model as our base model for transfer learning. Extreme inception or in other words Xception model uses depth-wise convolution model [15] instead of inception modules. It can be described as a stack of linear depthwise separable convolution layers with residual connections. Fig 4 shows the architecture of Xception model which has an input size of 299x299.The efficient architecture relies on two main points: depthwise separable convolution and shortcuts between convolution blocks. Model has about 36 convolutional layers forming the feature extraction base of the network. Data entering first goes through the entry flow, then through the middle flow which is repeated eight times and finally through the exit flow. Note that all Convolution and Separable Convolution layers are followed by batch normalization (not included in the diagram).All separable Convolution layers use a depth multiplier of 1 (no depth expansion).

ISSN No:-2456-2165

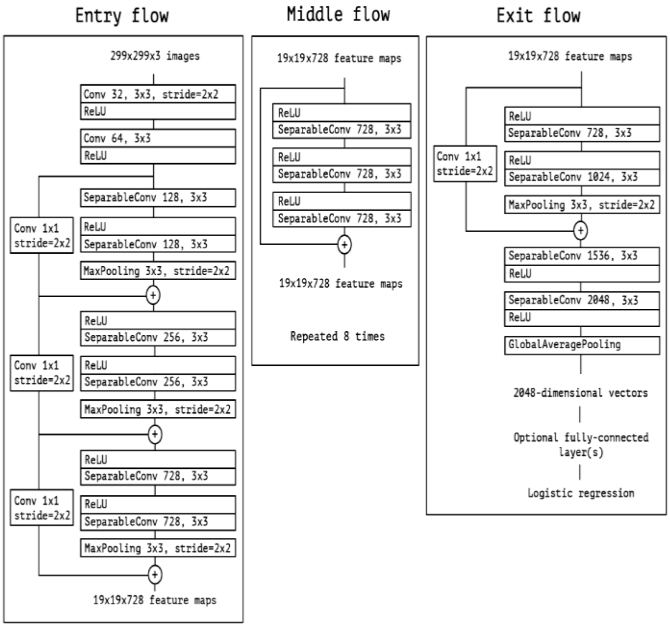


Fig. 4. Xception model architecture

V. IMPLEMENTATION AND MODEL TRAINING

Implementation of the research was performed using keras library as API with TensorFlow backend on Google colaboratory platform. Google colab provides a runtime fully configured for deep learning and free-of-charge access to a robust GPU.We loaded the Xception model pre-trained on Imagenet excluding its classifier section. This acts as the convolutional base for feature extraction. On top of this base model, a global average pooling layer is added which will be acting as our classifier section and fed its output directly into softmax activated layer. Dataset was split into a test-train split ratio of 30:70 i.e., 1486 training images and 638 testing images. In order to take advantage of model capacity, fine-tuning of the weights is performed on the selected best performed model. We initially ran the network through 15 epochs and then later 50 epochs for fine tuning. Layers of model up to 120 are freezed and then fine tuning was carried out from layer 120 onward. Adam was used as an optimizer with its learning rate kept at a default of 0.001.For all the training experiments, batch size selected was 32.As a performance metrics train and validation accuracies, their respective losses as well as a confusion matrix for the network is measured and plotted.

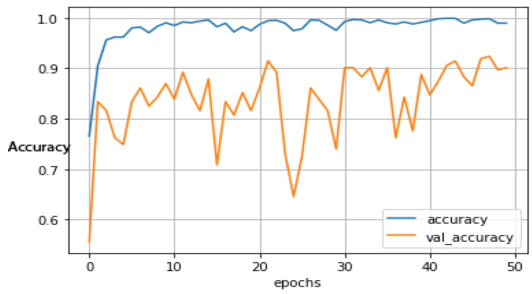
ALGORITHM

1. Downloading the libraries
2. Loading input images from dataset
3. Splitting into train and test set
4. Batch size=32,No.of classes=5
5. Resizing images into 299x299
6. Converting into array
7. Loading the transfer learning model
8. Building the classifier layer
9. Compiling the model
10. Training and fine tuning the model

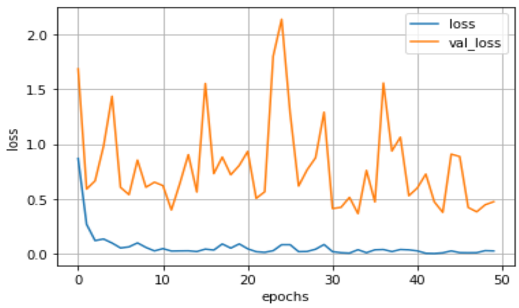
VI. RESULTS

On initial training of the model we were only able to achieve a validation accuracy of about 85% and a train accuracy of about 99% with batch size of 32. In order to make better performance, we fine-tuned the model. By fine tuning weights of the previous model, we were able to increase the model's validation accuracy up to 92%.The initial training was about 15 epochs and 50 epochs for fine tuning due to hardware limitation. It takes around 20 minutes for the entire training with virtual GPU on Google Colab Notebooks. Adam optimizer was utilized for training process with a learning rate of 0.001.

Respective plots are shown in fig 6.We found that, when applying feature extraction with pre-trained deep model on a very small dataset, the model could achieve higher accuracy than the baseline model and faster even on CPU mode.



(a)



(b)

Fig. 5. (a) Train-validation accuracy and (b) train-validation loss via fine tuning

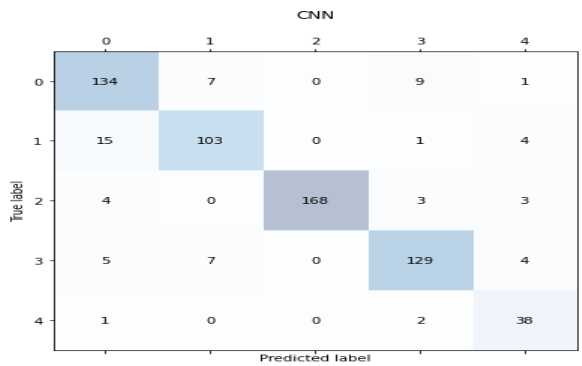


Fig. 6. Confusion matrix illustrating the errors in prediction on test set.[0,1,2,3,4][‘Glass’ ‘Metal’ ‘Paper’ ‘Plastic’]

This shows that when high accuracy is necessary and time is pressed, very deep models are better option than the shallow one. The maximum and minimum training loss is 8.0% and 1.0% respectively. Minimum loss is desired and a decreasing value means the model is learning during the training. Training can be stopped anytime if the loss is not decreasing anymore.

Fig 6 shows the confusion matrix resulting from evaluating the test set of dataset using the best performing model. It can be observed that metal often interestingly gets misclassified as glass, which can be justified given how metal images and glass images in dataset often produce similar features.

The trained model was then converted to TensorFlow Lite . The model was quantized by converting the weights from floating point to integers. This reduces the latency while having minimal effect on accuracy. The model is optimized for both size and latency.

The inference was run on a Raspberry Pi 4 Model B on TF lite interpreter. The inference results using single images was satisfactory but when tested in a real time environment using a camera module , the lack of computing power was evident even after reducing the frame rate.

VII . CONCLUSION

We proposed a waste classification system capable of identifying waste into different categories . We obtained 92% accuracy on the trained model with a minimal dip in accuracy when converted to use on an edge device. Even though the result when tested in a real time environment was not up to the mark, the results were still promising. Further optimization of the model will lead to a further decrease in latency. The implementation of computer vision on edge devices will increase in the future considering the improving accuracy of deep learning models and the increasing computational powers of microprocessors. In conclusion we believe this is a step in the right direction for using technology to help us safeguard our environment.

REFERENCES

[1]. A. Chandramohan, J. Mendonca, N. R. Shankar, N. U. Baheti, N. K. Krishnan and M. S. Suma,“Automated Waste Segregator," *2014 Texas Instruments India Educators' Conference (TIIEC)*, Bangalore, 2014, pp. 1-6.
[2]. Sofiya A,Edwina G Rodrigues,Nazim E,Alex P Johnson,“Integrated Smart Trash Bin,"*International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, vol. 1, no. 2, pp.236-242,March 2018.
[3]. M. Yang, G. Thung, “Classification of trash for recyclability status", 2016.
[4]. Olugboja, Adedeji & Wang, Zenghui,“Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network,"*Procedia Manufacturing*,Vol.35, pp.607-612, doi:10.1016/j.promfg.2019.05.086.
[5]. C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu,“A survey on deep transfer learning",*arXiv:1808.01974*,2018, [Online]
[6]. Available: https://arxiv.org/abs/1808.01974.
[7]. [Online]Available:http://image-net.org/,[Accessed:March 7, 2020].
[8]. S. Kornblith, J. Shlens and Q. V. Le,“Do Better ImageNet Models Transfer Better?,"*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR*), Long Beach, CA, USA, 2019, pp. 2656-2666.
[9]. Columbia University,“AutoRecycle,"*Columbia University EECS E4764 Fall '16 IoT*,[Online].
[10]. Available:http://icsl.ee.columbia.edu/iot-class/2016fall/group8/ [Accessed:Jan.6,2020].
[11]. Sachin Hulyalkar,Rajas Deshpande,Karan Makode,Siddhant Kajale, “Implementation of Smartbin using Convolutional Neural Networks,"*International Research Journal of Engineering and Technology (IRJET*), 2018.
[12]. G. Thung,“Trashnet,” GitHub repository, 2016.
[13]. A. Krizhevsky, I. Sutskever, and G. Hinton.“Imagenet classification with deep convolutional neural networks", *Proc. NIPS*,pp.1097-1105, 2012.
[14]. R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra and J. M. Z. Maningo, “Object Detection Using Convolutional Neural Networks," *TENCON 2018 - 2018 IEEE Region 10 Conference*, Jeju, Korea (South), 2018, pp. 2023-2027.
[15]. Pedro Marcelino,“Transfer learning from pre-trained models,"*Towards Data Science*,Oct. 23, 2018.[Online].
[16]. Available:https://towardsdatascience.com/ transfer-learning-from-pre-trained-models-f2393f124751. [Accessed:March 20, 2020].
[17]. Hussain, M., Jordan J., Bird, J. J., and Faria, D. R., “A Study on CNN Transfer Learning for Image Classification,”in *Proc. of 18th Annual UK Workshop on Computational Intelligence*,Nottingham,UK,2018,pp.191-202.
[18]. F. Chollet, “Xception: Deep learning with depthwise separable convolutions",*Proc. CVPR*,2017.