Analysis of Fine-Tuned Deep CNN Models for Waste Classification

Dipesh Gyawali[1], Alok Regmi[2], Aatish Shakya[3], Ashish Gautam[4], Surendra Shrestha[5]

Department of Electronics and Computer Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal

[1]072bex418@ioe.edu.np, [2]072bex403@ioe.edu.np, [3]072bex401@ioe.edu.np, [4]072bex406@ioe.edu.np, [5]surendra@ioe.edu.np

Abstract

Our research focuses on analyzing possibilities for automatic waste sorting and collecting in such a way that helps it for further recycling process. Various approaches are being practiced managing waste but not efficient and require human intervention. The automatic waste segregation would fit in to fill the gap. This research tested well known Deep Learning Network architectures for waste classification with waste image dataset combined from own endeavors and Trash Net. The convolutional neural network is used for image classification. The hardware built in the form of dustbin is used to segregate those wastes into different compartments. Without the human exercise in segregating those waste products, the study would save the precious time and would introduce the automation in the area of waste management. The situation is win-win for both government, society and industrialists. Because of fine-tuning of the ResNet18 Network, the best validation accuracy was found to be 87.8%.

Keywords: *Waste Classification, Transfer Learning, Fine Tuning, Confusion matrix, Recycling*

1. Introduction

With the tremendous increase in population worldwide, the solid waste production has increased abundantly. Improper waste management has an adverse effect in economic, public health and the environment. Deep learning is a class of machine learning algorithm that uses multiple layers of data representation and feature extraction. Using convolution neural network [1], a class of deep feed-forward artificial neural network has been applied successively to analyze the image. After the classification of image using CNN, the corresponding waste is automatically segregated using the smart dustbin after detecting using the camera.

2. Proposed System

A total of 1800 waste images are taken from the Stanford TrashNet Dataset[3] . We took 4 labels for classification i.e. Paper, Plastic, Metal, Glass. A total local waste images of 1302 are captured using the mobile phones for four classes of data. A total of 3102 images are used for analysis and classification of the waste. Deeper neural networks are more difficult to train due to necessity of high computation power. Due to the reason, we have used pre-trained model which was trained on ImageNet Dataset. Using transfer learning, the model seems to perform extremely well on our validation set after fine tuning the model.

Image Preprocessing techniques were performed to ensure best feature extraction across CNN layers. Initially the size of image was extremely large (2000-4000) pixels. Hence, the image was resized to 512 pixels from (2000-4000) pixels for the training process. Secondly, Data normalization is an important step, which ensures that each input parameter (pixel for images) has a similar data distribution. The distribution of such data resemble a Gaussian curve centered at zero. Also, a total number of 3102 images are not sufficient in order to train the neural network and get improved performance. We have done different types of image augmentation like horizontal flip, random crop, zooming to make variations inside the data so that it can generalize the unseen data accurately.

3. Experiments

We have implemented our method using Pytorch framework. Training was done on an NVIDIA GTX 1050 with 768 CUDA cores. We trained across various pre-trained ImageNet Models and compare the metrics across each model. So, we used convolutional layer of pre-trained network, ran the new data through it, and then trained a new classifier on top of the model. In the feature extraction process, we extract features by unfreezing the last three convolutional layers of pre-trained network then added our fully connected layers and trained on our dataset. After performing the feature extraction, another model reuse technique which is identical to feature extraction is used called fine tuning. All the layers except last layers are frozen and a custom layer was built in order to perform the classification. In the last layer, two linear layers were added with ReLU activation function and finally the Softmax activation function. Then the convolutional layer and newly added classifier are jointly trained which improved model performance after fine tuning.

4. Results and Analysis

Comparison between several model's strengths and weaknesses were clearly observed before finalizing which model should be used as to applicable standards. The final chosen model was ResNet18[4] to be deployed into the hardware to make some real time predictions in the system. Fig. 9. shows a consistent difference between the training and validation loss curve across the model ResNet50. The training error was decreased consistently but the model's generalization capacity was not very good. VGG16[5] model was observed characteristically similar to the ResNet50 as it also had capacity of representing more complex model than the prepared dataset required. The loss curve was observed as overfitting since the training loss was observed to be around 0.2 and still decreasing. However, the model's validation did not improve with reduction in training loss. The ResNet18 model seem to be good after visualizing the accuracy of both training and validation. The condition of overfitting and under fitting did not occur significantly in the model as shown in Figure 1.

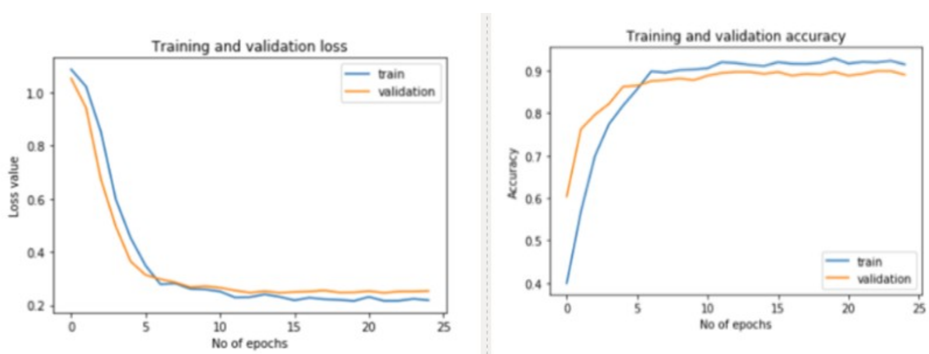

Figure 1. Loss and Accuracy for ResNet18 architecture

The confusion matrix across ResNet18 in Table 1. shows that paper is mostly classified correctly more than any other type of waste. The higher diagonal values of the confusion matrix indicate more correct predictions. The features used by the model seem to be similar to human comparison between wastes since the confusion matrix shows confusion in those places that confuse humans as well.

Table 1. Confusion Matrix for ResNet18 architecture

	Glass	**Metal**	**Paper**	**Plastic**
Glass	**83.8**	3.7	3.7	8.8
Metal	1.7	**89.0**	2.5	6.8
Paper	0.1	1.1	**90.2**	8.6
Plastic	7.5	4.6	5.2	**82.8**

5. Conclusion

Through this research, we obtained the accuracy above 87% and analyzed different models we adapted to train our data. For this, we have been putting a continuous effort for getting better result for each evaluation metrics. Also, the performance of model can be improved further by increasing the number of images and taking steps for preprocessing.

Acknowledgement

We would also like to thank the Department of Electronics and Computer Engineering IOE for giving us an amazing opportunity to carry out the research successfully.

References

[1] Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proceedings of Neural Information Processing System Conference, Lake Tahoe, CA, USA, December 2012.

[2] O.Awe, R.Mengistu, and V.Sreedhar:“Smart trashnet: Waste localization and classification,” CS229 Project Report2017, 2017.

[3] Thung, Gary and M. Yang. “Classification of Trash for Recyclability Status.” (2016).

[4] K. He, X. Zhang, S. Re, J. Sun, “Deep Residual Learning for Image Recognition”, arXiv preprint 1512.03385 (2015).

[5] K. Simonyan, A. Zisserman , “Very Deep Convolutional Networks for Large-Scale Image Recognition”, arXiv preprint 1409.1556 (2014).