PRCV Final Project Report

Plastic Classification And Sign Language Prediction

Gopisainath Mamindlapalli
Dept. of Mechanical Engineering,
College of Engineering,
Northeastern University
Boston, USA
mamindlapalli.g@northheastern.edu

Vinesh Krishna Anne
Dept. of Mechanical Engineering,
College of Engineering,
Northeastern University
Boston, USA
anne.v@northeastern.edu

Krishna Prasath Senthil Kumaran
Dept. of Electrical and Computer Engineering, College of Engineering,
Northeastern University
Boston, USA
senthilkumaran.k@northeastern.edu

***Abstract* — This document provides a report on two projects: plastic classification on a pre-classified image set of plastic and sign language detection. The ResNet18 network is used for both projects and the MobileNet-v2 network is used for plastic classification only. We gave these networks the pre-classified image set of plastics to train and evaluate the overall network performance. We also used pre-classified photos of hand signs to train the network and enable real-time hand sign identification for sign language prediction.**

Keywords* — *networks, detection, training, hand signs.

I. Introduction

Plastic classification is one of the most prominent research tracks for computer vision. The work is first fed by the computer images of the same type of object, finds common identifying features, and uses these features to detect and identify the same objects in new images.

Especially it is for waste management specially to detect and classify plastic waste. As the use of plastic is prevalent, plastic waste becomes a complex problem because it has a substantial impact not only on land and ocean but also on humans and animals. If it is not handled suitably, this potentially causes an unprecedented environmental crisis in the long term. Recognizing various plastics waste by appropriate classes will make the sorting task more efficient both at home and recycling centre. However, the challenge could be the sufficiency of the training dataset that will be representative of the plastic waste model that will be applied. This is due to the very diversity of plastic waste in shapes, colours, and dimensions. Also, they have absorbance and reflectance characteristics in a certain spectrum [2].

For the plastic classification project, we used ResNet18 and MobileNet-v2 networks. These pre-trained models were utilized to explore transfer learning for our project. The rationale behind using pre-trained networks is that they are already trained on massive datasets. The weights of these networks were frozen, and we then trained them with a pre-classified image set that consisted of various types of plastics commonly found in grocery store products. We evaluated the performance of these networks by comparing their accuracy across different numbers of epochs and batch sizes and selected the network that exhibited the highest accuracy for our classification task. Additionally, we attempted to classify the images by clustering them.

Sign Language system is the predominant sign language in India that contains standard hand-based gestures which are used by speech-impaired people for communication purposes. The general population has no knowledge of sign language gestures as these gestures are very extensive and complex which in turn hampers the knowledge sharing between speech-impaired people and the general population[3].

For sign language prediction, we trained the ResNet18 network using a pre-classified image set that contains 29 classes of different characters and tested the network to detect hand signs in images and then carried out it in real time.

II. Background

A. Paper [1]

In this paper, the images are clustered in situ to classify objects. When the image is given as input from the situ, it goes through the clustered images and gives out all the similar objects as output.

We were inspired by this approach and worked on the image clustering for the plastic classification but the output seemed to be not as expected due to the low amount of trained data set of similar objects.

B. Paper [2]

This paper focuses on real-time plastic waste detection and classification using object recognition with YOLOv3. The algorithm achieved good confidence values for detecting and classifying six macro plastic waste classes, including plastic bags, bottles, and straws. However, further improvements such as increasing the iteration number during training and reviewing image variety, angle, lighting, and resolution are necessary to achieve higher confidence values.

The plastic classification by proposing a real-time plastic waste detection and classification method using object recognition and the method that aims to efficiently sort plastic waste, which can contribute to the overall plastic classification effort gave us the idea to classify the plastic using pre-trained networks like ResNet and MobileNet.

C. Paper [3]

Here, they used two approaches, one utilizing Microsoft Kinect, achieving a high classification accuracy of static and dynamic hand gestures, and another using semantic segmentation with U-Net and ResNet 101, which removes the need for a Kinect camera and achieves excellent performance in real-time.

The inspiration for image classification of plastics using pre-trained models such as ResNet18 was taken from this reference where they use a similar setup for a similar task. We have used the same ResNet18 pre-trained model for detecting hand signs of sign language using images and real-time.

III. Method

A. Problem

The main objective is to find the amount of plastic being generated through grocery store sales throughout the year.

And sign language is to enable communication and facilitate inclusion for individuals who use sign language as their primary mode of communication.

B. Approach and solution

We developed a deep-learning network to classify the 6k images according to the amount of plastic in them. We tried using ResNet18 and MobileNet-v2 pre-trained networks to classify the images and to find the optimal model with a different number of epochs and batch sizes of the networks.

We also developed a sign language identification algorithm in which we used pre-trained ResNet18 to identify and classify the images. We further developed this algorithm to work in real-time but the accuracy is low because of the intensity, brightness, and contrast differences.

IV. Experimentation

In the Plastic classification project, we conducted experiments with ResNet18 and MobileNet-v2 neural networks. Initially, we trained the ResNet18 network dataset for 5 epochs using a batch size of 64. The training accuracy and the testing accuracy (Figure 1) along with the predicted outputs (Figure 2) are shown below.

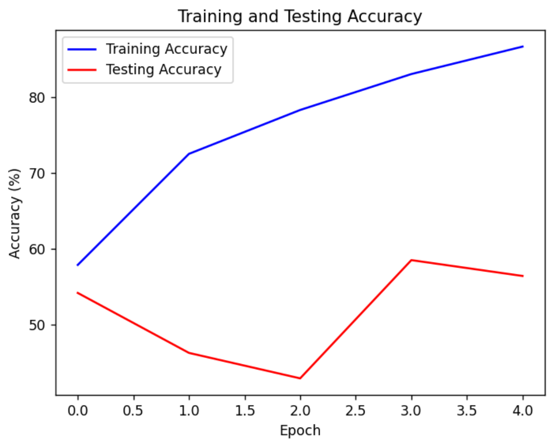

Figure 1: Accuracy plot for ResNet with 5 epochs

Figure 2: Prediction for ResNet with 5 epochs

The highest accuracy achieved for ResNet is around 60%. To improve the performance of the model, we increased the number of epochs to 10 and the batch size to 128 with the training and testing accuracy plots along with the predicted outputs.

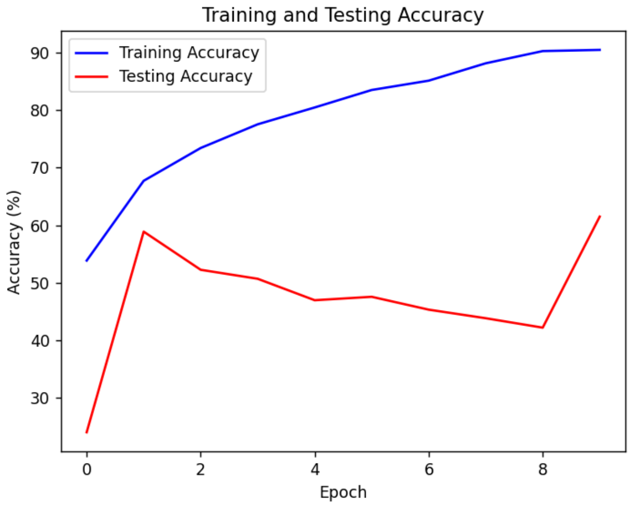

Figure 3: Accuracy plot for ResNet with 10 epochs

Figure 4: Prediction for ResNet with 10 epochs

We further trained the MobileNet-v2 network with 10 epochs and 64 batch size.

Figure 5: Accuracy plot for MobileNet with 10 epochs

Figure 6: Prediction for MobileNet with 10 epochs

The highest accuracy achieved is around 70% for MobileNet.

Although the accuracy of the networks is not remarkable, we can increase the accuracy of the models by collecting more data and training with a lot more images. We can further increase accuracy by using more sophisticated algorithms such as a clustering algorithm to cluster the images based on different objects and then classify them using the deep learning networks. By using both clustering and classification algorithms we can get more insights into the images and can easily classify them. We tried clustering the dataset and tried to find the optimal number of clusters and tried to visualize the clusters.

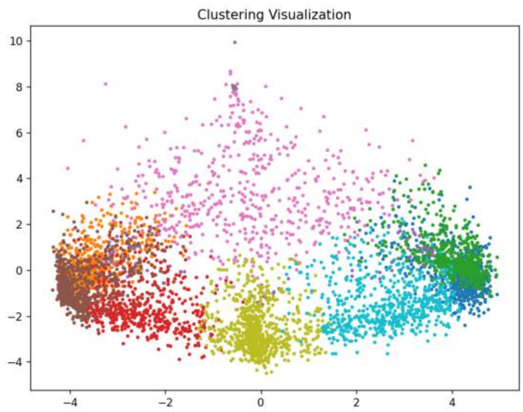

Figure 7: Visualization of the clustered images

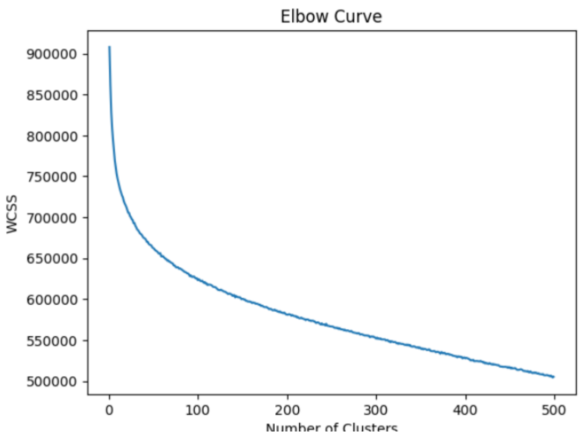

Figure 8: Optimal number of clusters

For the sign language prediction algorithm, we used pretrained ResNet18 to classify images into 29 classes. We used a dataset from the Kegel which contains 3k examples of each class. The algorithm is able to identify the sign language with an accuracy of 92%. These are some output predictions by the network for the provided image set.

Figure 9: Sign language prediction from image set

We tried implementing the sign language prediction in real-time and successfully implemented it, but the accuracy of the network is limited to 40%. There are many factors that affect the accuracy, such as differences in intensities, contrasts, and brightness of the frame.

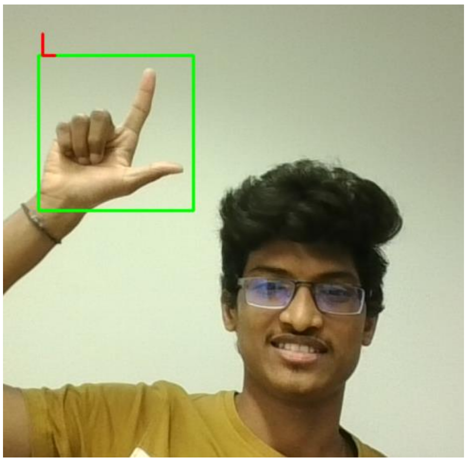

Figure 10: Sign language prediction in real-time

V. Conclusion

To summarize the findings of the two projects, it was anticipated that the MobileNet would outperform other models in the Plastic classification project, and incorporating clustering in addition to classification would enhance the accuracy of plastic classification. In the sign language prediction project, accurate sign predictions were achieved through image analysis, but real-time accuracy was affected by several factors in the environment, leading to decreased performance.

Acknowledgment

We express our sincere thanks and gratitude to Prof. Bruce Maxwell for giving us this opportunity. The success and outcome of this project required a lot of insights, feedback,

and guidance from the professor which have been invaluable in shaping and completing our project.

REFERENCES

[1] Bell, Sean & Bala, Kavita. (2015). Learning visual similarity for product design with convolutional neural networks. ACM Transactions on Graphics. 34. 98:1-98:10. 10.1145/2766959.

[2] I. W. Raka Ardana, I. Bagus Irawan Purnama, and I. M. Sumerta Yasa, "Application of Object Recognition for Plastic Waste Detection and Classification Using YOLOv3," 2020 International Conference on Applied Science and Technology (iCAST), Padang, Indonesia, 2020, pp. 652-656, doi: 10.1109/iCAST51016.2020.9557735.

[3] P. Likhar, N. K. Bhagat and R. G N, "Deep Learning Methods for Indian Sign Language Recognition," 2020 IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin), Berlin, Germany, 2020, pp. 1-6, doi: 10.1109/ICCE-Berlin50680.2020.9352194.

[4] Z. Tian, D. Sun, C. Yu, J. Li, G. Ma and H. Xia, "Garbage Classification Algorithm Based on Deep Learning," 2021 40th Chinese Control Conference (CCC), Shanghai, China, 2021, pp. 8199-8203, doi: 10.23919/CCC52363.2021.9549336.

[5] C. Chu, Q. Xiao, J. Xiao and C. Gao, "Sign Language Action Recognition System Based on Deep Learning," 2021 5th International Conference on Automation, Control and Robots (ICACR), Nanning, China, 2021, pp. 24-28, doi: 10.1109/ICACR53472.2021.9605168.

[6] A. Khalil et al., "Mobile Deep Classification of UAE Banknotes for the Visually Challenged," 2022 9th International Conference on Future Internet of Things and Cloud (FiCloud), Rome, Italy, 2022, pp. 321-325, doi: 10.1109/FiCloud57274.2022.00053.

[7] K. Hanapi, A. Setyanto and A. Dwi Hartanto, "Comparison of ResNet50 and SqueezeNet1.1 for Plastic Bottles and Cans Classification," 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 2022, pp. 525-529, doi: 10.1109/ICITISEE57756.2022.10057634.

[8] S. Hussain, R. Saxena, X. Han, J. A. Khan and H. Shin, "Hand gesture recognition using deep learning," 2017 International SoC Design Conference (ISOCC), Seoul, Korea (South), 2017, pp. 48-49, doi: 10.1109/ISOCC.2017.8368821.

[9] Agbehadji, Israel & Abayomi, Taofeek & Bui Khac Hoai, Nam & Millham, Richard & Freeman, Emmanuel. (2022). Nature-Inspired Search Method and Custom Waste Object Detection and Classification Model for Smart Waste Bin. Sensors. 22. 6176. 10.3390/s22166176.