Classification of Plastic Level in Grocery Store Products Employing Transfer Learning With ResNet-50

Bharat Chawla
Khoury College of Computer Sciences
Northeastern University
Boston, Massachusetts
chawla.bh@northeastern.edu

Himaja Reddy Ginkala
Khoury College of Computer Sciences
Northeastern University
Boston, Massachusetts
ginkala.h@northeastern.edu

Abstract—Recycling of plastic is an indispensable measure to protect the environment of our planet. Despite this reality, it continues to be one of the largest ecological problems the world faces today. There still exists a lack of knowledge of accurate classification of plastic levels in various products. The ability to effectively do so would minimize the detrimental impacts of plastic and aid in developing mitigation strategies. Artificial intelligence, particularly deep learning paired with computer vision, has resulted in powerful convolutional neural networks available to utilize on various classification tasks. This could be an effective approach to autonomously identify different levels of plastics among various sets of items. The main problem considered in this paper is the identification and application of a pre-trained neural network model, which can be employed via transfer learning, to differentiate plastic level in commonly seen grocery store products. For this classification task, a pre-classified dataset of images was categorized into 4 different categories: heavy-plastic, some-plastic, no-plastic and no-image. The overall training accuracy after 20 epoch cycles resulted in a 70.51% accuracy rate.

Keywords—computer vision, classification, deep network, convolutional neural network, transfer learning, ResNet, plastics

I. Introduction

The mass production and usage of plastic continues to be on a rising trend, securing its place as one of the core issues in waste generation today. Plastic pollution has a significant impact on life everywhere, from densely populated cities to sea life throughout the planet's oceans. Therefore plastic management has been and continues to be one of humanity's foremost concerns as it has a direct effect on the environment and the sustainability of the planet in the future. In a rapidly growing world, the identification and classification of plastics remains a challenge as the use of plastics and the number of types of the substance continue to grow as well.

One of the most common places an abundance of plastic can be seen is inside a general grocery store. From juices being packaged in clear plastic to candy being wrapped in a colorful plastic covering, the frequency is limitless. However, not all products are relatively the same level of plastic. While some can be considered *heavy-plastic*, such as bottles of spices, others can be classified as some-plastic, for example a box of cereal, and a third category consisting of products such as canned foods can be recognized as no-plastic.

Through the use of deep neural networks, advances in object recognition and image classification have seen incredible growth. Therefore the task of segregating various grocery store products into their level of plastic classifications was taken up with this approach. Instead of working with a detection network, a whole image network was determined to be the path to take for the classification task at hand. While there are many ways in which to configure a unique deep learning neural network, there are also numerous powerful pre-trained networks available for object classification purposes. Through the idea of transfer learning, these pre-trained neural networks can be utilized on new datasets, such as the grocery store products dataset being addressed in this task.

II. Related Work

Transfer learning is a deep learning technique in which the pre-trained convolutional neural network is retrained on another dataset by using the particular weights of the original network. [1] explores various convolutional neural networks to analyze the effect of transfer learning when taking up tasks of object classification. In their particular experiment, the CUB-200-2011 dataset was used to detect different types of birds. Five different pre-trained convolutional neural network models were used, and the results were then added with the majority voting scheme, leading to an accuracy rate of 97.45%. Transfer learning proved to be incredibly effective in providing likable accuracies for object classification tasks.

When looking into the various pre-trained convolutional neural networks available to train on the grocery store products dataset, multiple comparative analysis studies were encountered. One such study was [2] which saw that increasing the number of layers in a network and the number of the neurons in each layer, a positive growth was seen in accuracy rates. It was found however that as the network becomes larger and larger, the problems of gradient dispersion and gradient explosion arise. Both these issues are highly undesirable and were solved on [3, 4], however as the depth of the network is increased, situations were encountered in which the accuracy reached saturation and then declined rapidly during training. This occurrence is called degradation, and was solved through the use of a residual learning network.

As proposed in [5], residual networks, ResNets, are essentially based on shortcut connections. The issue of degradation of performance, connecting directly to the vanishing gradient problem, is solved through the skip connections in ResNet. These connections allow for an alternate shortcut path for the gradient to flow through and also allow the network model to learn the identity mappings in such a way that ensures that the higher layer will perform at least as good as the lower layer, and not worse. The described advantages of ResNet for object classification led to this group of networks being chosen to implement transfer learning for the plastic level classification task in common grocery store products.

III. Methodology

A. ResNet-50

The original ResNet architecture was the ResNet-34, which consisted of 34 weighted layers in the network. It pioneered a new way to add more convolutional layers to a CNN, without running into the vanishing gradient problem through the use of shortcut connections that skipped over some layers in the network, converting it to a residual network. The special feature about ResNet-50 is that it utilizes a bottleneck design for the building block, which is a residual block using 1x1 convolutions. This reduces the number of parameters and matrix multiplications and hence allows for much faster training for each layer in the network, a highly desirable feature for any deep learning neural network.

B. Dataset Modification

The provided dataset consisted of over 6,000 images of grocery store products pre-classified into four categories: *heavy-plastic*, *some-plastic*, *no-plastic*, and *no-image*. While the first three image categories consisted of actual images of commonly seen products, the fourth category was the exception as it was composed of "no image" images, essentially no product images. Examples of the four different classes can be seen in Figure 1. Each of the three datasets with some degree of the "plastic" classification contain images of products such that no one image is the same. There are sets of images within each classification in which there exist images very similar to one another, but never an exact match. There is always a variation in either color, angle, size, or combinations of the three. The dataset for the no-image class however contains multiple repeated images, and hence has limited variation in comparison to the datasets of the other classes.

Fig 1. Grocery store product examples of the 4 classes of classification: (a) heavy-plastic (b) some-plastic (c) no-plastic (d) no-image

Most of the pictures in the dataset were simply a capture of the object against a white background, but there were some images which showcased a real-life situation background or a dark background. For purposes of consistency and uniformity to help the network learn better, the dataset was filtered such that only images with white backgrounds were used. This was done by simply thresholding the image corners to check for the desired color value, white in this case. This process resulted in a dataset consisting of a total of 5,174 images: 1,640 of *heavy-plastic*, 2,199 of *some-plastic*, 1,068 of *no-plastic*, and 267 of *no-image*.

IV. Experiments and Results

The ResNet-50 network was employed via transfer learning for the plastic level classification task. It was trained on the filtered dataset of 5,174 images, which consisted of four subsets: (a) heavy-plastic, (b) no-image, (c) no-plastic, and (d) some-plastic. Compared to the classes of images which described an actual level of plastic, the distinction of the dataset of no-image class was hypothesized to be a good fit for stabilizing the loss rate.

A. Training Loss and Accuracy Rate

For neural network models, it is common to examine learning curve graphs to decide on model convergence. Generally, during the training process of a model, the training loss tends to diminish and accuracy rate tends to increase as the number of epochs cycle increases. However, after some point, it is expected that both training loss and training accuracy rate begin to stabilize.

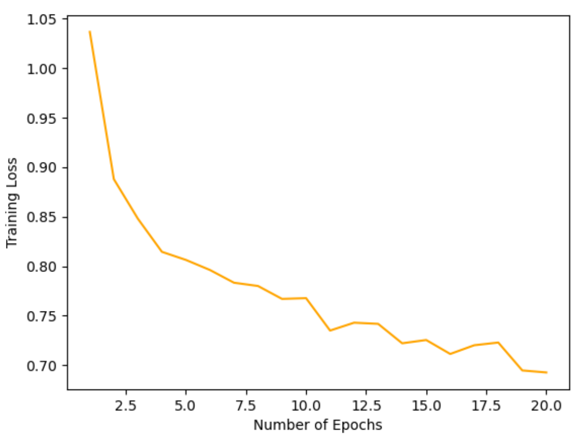

Fig 2. Average training loss over 20 epoch cycles.

In Figure 2, it can be seen that the training loss showcases a diminishing effect with the increase in number of epoch cycles. This results in better performance on accurately classifying plastic level in grocery store products as longer training leads to better convergence.

Usually, neural networks are expected to converge after training for a number of epochs, and the weights are updated iteratively as they consist of a gradient based algorithm. Therefore for sophisticated models such as ResNet-50, a single epoch cycle of training is not enough and leads to underfitting of the data. In Figure 3, it can be observed that the learning curve graphs get better and better until a point of convergence appears. However, one caveat is that it may lead to overfit, and hence validation errors begin to increase if the model is continuously training. For this experiment, the model was evaluated over 20 epoch cycles, and accuracy rate was observed to steadily increase over the execution of each epoch cycle. The overall training accuracy after 20 epoch cycles resulted in a 70.51% accuracy rate.

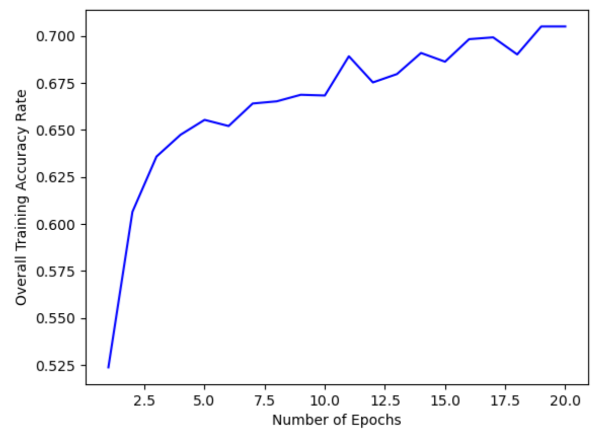

Fig 3. Overall training accuracies over 20 epoch cycles.

B. Class Accuracies

The comparison of the accuracy rates of each individual class resulted in the observation that the *no-plastic* class seemed to underperform in comparison to the other classes when classifying over 20 epoch cycles. At epoch cycle 20, the accuracy rates of the individual classes were as follows:

- heavy-plastic — 69.45%
- some-plastic — 76.63%
- no-plastic — 54.96%
- no-image — 88.76%

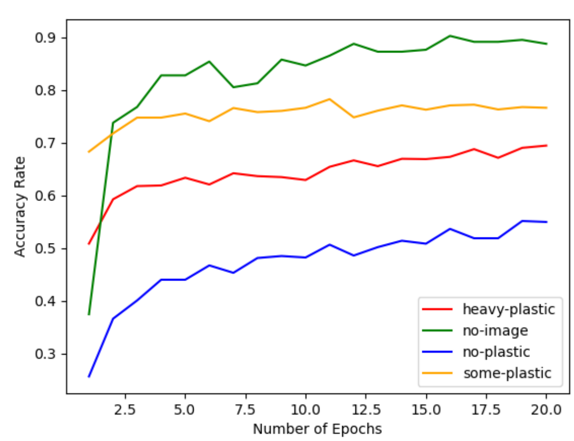

Fig 4. Prediction accuracies for each of the 4 classes over 20 epoch cycles.

As can be seen in Figure 4, transfer learning employed with the ResNet-50 model on a custom dataset is able to classify the *no-image* class at the highest rate. To increase the further performance of the model on classification of the actual plastics classes, an increase in the number of epoch cycles may lead to better results, but this must be done in accordance to diminish the factors that would result in overfitting the training model.

V. Discussion and Conclusion

From the results of the experiment, the *no-plastic* class resulted in the lowest accuracy rate, while the *no-image* class resulted in the highest accuracy rate. This can be deduced to be so mainly because the dataset for the *no-image* class contains multiple repeated images, as opposed to the other classes where there is much greater variation. Of the remaining two classes, the *some-plastic* class showed the second highest accuracy rate, followed by *heavy-plastic*. This is understandable because while an item of heavy plastic is clearly visible to be containing more of the element than an item of some plastic to the human eye, it is not necessarily as clear to the vision of a machine. The variation in color, designs of the plastics, and the lettering printed on a particular product as compared to another are all possible factors which make the classification task difficult. Additionally, the even more challenging task is that of knowing which products contain more plastic inside the packaging, such as individually wrapped snack packets, resulting in an even heavier amount of plastic contained.

There are a number of other pretrained models that can be experimented with this given custom dataset to observe their performances and training accuracy rates in comparison to the ResNet-50 network. Another whole image network that can be explored is DenseNet. This model connects each layer to every other layer, and layers are usually created using feature maps from all previous levels, and past ones are used in all future layers to create new layers. This process solves the vanishing-gradient problem, and improves feature propagation while significantly reducing the number of parameters. Another suggestion of neural network models to look into is the state-of-the-art You Only Look Once (YOLO) network. It is an object detection model which uses a single neural network to predict bounding boxes and class probabilities directly from full images in one evaluation. These various models could possibly provide more knowledge into the difficulty in classifying the *no-plastic* class, since it proved to be the toughest to differentiate.

In conclusion, classification of level of plastic in regular, familiar grocery store products proved to be an undeniably challenging task. However difficult, this is a task that must continue to be experimented on as the power of computer vision will play a large role in solving the various environmental problems our world faces today due to the abundance of plastic. It is hoped that by experimenting with various neural networks, different sets of models and abundant datasets, the knowledge gained will aid in minimizing the effects of this ecological problem on our planet.

Acknowledgement

We thank Professor Amanda Welsh and her team for providing us with the pre-classified image database consisting of over 6,000 images. We thank Professor Bruce Maxwell for bringing light to this particularly intriguing research opportunity.

References

[1] Bulbul Bamne, N. Shrivastava, L. Parashar, and U. N. Singh, “Transfer learning-based Object Detection by using Convolutional Neural Networks,” 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Jul. 2020, doi: https://doi.org/10.1109/icesc48915.2020.9156060.

[2] Y. Xiao et al., “A review of object detection based on deep learning,” Multimedia Tools and Applications, vol. 79, no. 33, pp. 23729–23791, Sep. 2020, doi: https://doi.org/10.1007/s11042-020-08976-6.

[3] Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: International conference on learning representations (ICLR), doi: https://doi.org/10.48550/arXiv.1409.1556

[4] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[5] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.