

One-Shot learning based classification for segregation of plastic waste

Shivaank Agarwal

Department of Computer Science
BITS Pilani
Hyderabad, India.
shivaank.agarwal@gmail.com

Ravindra Gudi

Department of Chemical Engineering
IIT Bombay
Mumbai, India.
ravigudi@iitb.ac.in

Paresh Saxena

Department of Computer Science
BITS Pilani
Hyderabad, India.
psaxena@hyderabad.bits-pilani.ac.in

Abstract—The problem of segregating recyclable waste is fairly daunting for many countries. This article presents an approach for image based classification of plastic waste using one-shot learning techniques. The proposed approach exploits discriminative features generated via the siamese and triplet loss convolutional neural networks to help differentiate between 5 types of plastic waste based on their resin codes. The approach achieves an accuracy of 99.74% on the WaDaBa Database [1].

Index Terms—Computer Vision, Deep Learning, Plastic Waste Segregation

I. INTRODUCTION

As stated by the Society of the Plastics Industry, Inc. (SPI), the purpose of plastic resin code is to provide a consistent system to facilitate the recycling of post-consumer plastic waste [2]. We propose the automatic classification of plastic waste based on its resin code using deep learning approaches. Deep Learning using convolutional neural networks has achieved significant breakthrough in tasks such as image classification and object detection [3], [4]. However, these algorithms do not perform well when less amount of data is available for training. In such circumstances, it is better to learn discriminative features to distinguish between classes. We propose using one-shot learning to generalise over the entire dataset, where the amount of data for certain classes is not sufficient to train traditional convolutional neural networks. One-shot learning methods directly learn the similarities between objects of the same class and also the differences between objects of different classes. In this article, we propose the use of siamese [5] and triplet loss [6] networks on the plastic waste images available in the WaDaBa database.



Fig. 1. 7 categories of Plastic Resin codes

II. RELATED WORK

WaDaBa is the only dataset available which classifies plastic waste based on its resin code. There have been two previous works on this database. In [1], the classification was attempted using histogram analysis, and an accuracy of 75.68% was achieved. In [7], the authors used image

augmentation (rotation) to increase the size of the database from 4000 to around 140,000 images and then used CNNs for classification. They achieved an accuracy of 99.92%. The authors classified images in 4 out of the 5 available categories in the database, due to a shortage of images in the category of other plastics (category 7 in Fig. 1).

While the above method already achieves high accuracy, it requires substantial rotation augmentation. This does not add any unseen instances in the test data since deep CNNs are powerful enough to learn rotational invariant filters on their own, given a sufficient number of training images. Hence the method may not work well on new instances of data that do not have a rotational variant in the training set. In this paper, we propose a model that does not require any augmentation to increase the size of the database, but achieves a comparable accuracy.

III. WADABA DATABASE

The WaDaBa database contains 4000 images belonging to the 5 categories - PET (01-polyethylene terephthalate), PE-HD (02-high-density polyethylene), PP (05-polypropylene), PS (06-polystyrene) and other (07). Table I shows the number of images per category in the database. Each image consists of a single object which is subjected to various degrees of deformations to simulate natural conditions.

TABLE I
IMAGES PER CATEGORY IN WADABA

Category	No. of Images
01 PET	2200
02 PE-HD	600
05 PP	640
06 PS	520
07 Other	40

IV. PROPOSED APPROACH

We learn image representations via a supervised metric-based approach with siamese and triplet loss neural networks, and then use the trained network weights for evaluation. For both the approaches, the CNN architecture used is inspired by [5] and shown in Fig. 2. The siamese and triplet loss networks are explained in the next subsections.

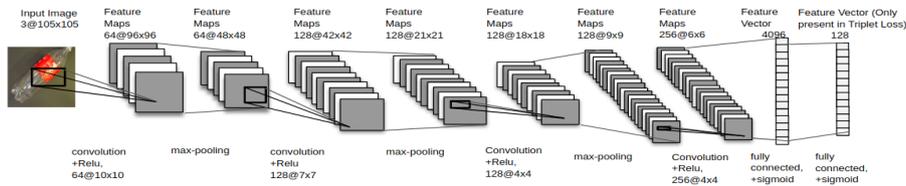


Fig. 2. Convolutional Neural Network Structure.

A. Siamese Network

Two images are taken as the input and the label is set to 0 if they belong to the same class and 1 otherwise. Both images are passed through two identical CNN networks which have shared weights as shown in Fig. 3. The resultant output is a 4096 feature vector for each image. Then the Euclidean distance between the two feature vectors is computed, passed through a sigmoid layer and the resultant cross-entropy loss is propagated backward to both the identical networks.

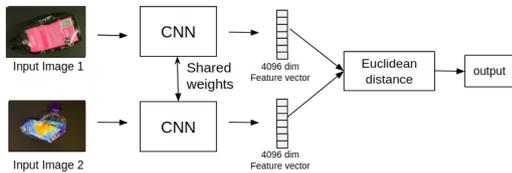


Fig. 3. Siamese Network Structure

B. Triplet Loss

The architecture consists of three identical networks with shared weights. One image at random is selected and it is referred to as the anchor image. The other two images are called as positive (image of the same class) and negative (image of another class) images. All the three images are passed into the three identical CNN networks. The CNN network used is identical to the siamese, but an additional fully-connected layer of 128-dim is added after the last fully-connected layer of the siamese network. The loss is computed in equation (1) and it is propagated backward into each of the three identical networks.

$$\|f(x^a) - f(x^p)\|_2^2 - \|f(x^a) - f(x^n)\|_2^2 + 0.4 \quad (1)$$

where $f(x^a)$, $f(x^p)$, $f(x^n)$ represent the anchor, positive and negative embedding respectively obtained after passing the corresponding images through the CNN.

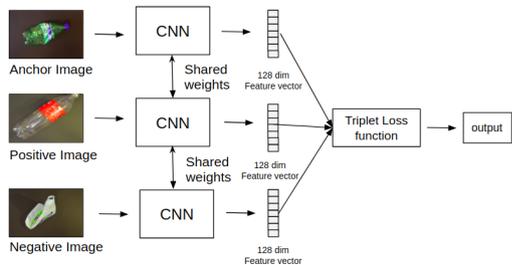


Fig. 4. Triplet Network Structure

V. EXPERIMENTS

A. Training

Siamese network is trained for 50 epochs, while the triplet loss network is trained for 100 epochs. In both the networks 5000 instances are generated per epoch. In both the cases Stochastic Gradient Descent optimizer is used with a learning rate of 0.001 and momentum 0.9.

B. Evaluation and Results

Two methods for evaluation are used. The first one is 1-shot 5-way accuracy in which 400 random sample images are evaluated. For each sample image, 5 test images are chosen randomly, one from each category. These 5 images are each paired with the sample image and sent as an input into the network. The sample image is assigned to the category of the test image for which the network's output is the minimum. Count of correctly classified images is incremented if the label of the sample image matches the assigned category. In the second method, the features from the last fully-connected layer are extracted and stored in a database. Each image in the database is compared to all other images using euclidean distance, and accuracy is measured using the K-nearest Neighbour algorithm. Table II and Table III show the accuracy obtained from the 1-shot 5-way and KNN evaluation methods respectively.

TABLE II
1-SHOT 5-WAY ACCURACY

Method	Accuracy
Siamese	99.50
Triplet Loss	99.0

TABLE III
KNN ACCURACY

Method	KNN	PET	PE-HD	PP	PS	Other	Average
Siamese	K = 3	99.59	99.66	99.35	100	100	99.72
Siamese	K = 5	99.4	99.83	99.00	100	100	99.64
Siamese	K = 7	99.59	99.66	99.20	100	100	99.69
TripletLoss	K = 3	99.68	100	99.30	100	100	99.79
TripletLoss	K = 5	99.50	100	98.5	98.5	100	99.56
TripletLoss	K = 7	99.40	100	98.40	99.8	100	99.52

VI. CONCLUSION AND FUTURE WORK

We achieved a maximum accuracy of 99.79 on the WaDaBa database across all the 5 plastic categories without image augmentation. Further work will focus on object localization followed by classification in case of multiple objects within a single image.

REFERENCES

- [1] Bobulski, Janusz, and Jacek Piatkowski. "PET waste classification method and plastic waste DataBase-WaDaBa." International Conference on Image Processing and Communications. Springer, Cham, 2017.
- [2] Wilhelm, Richard. "Resin Identification Codes—New ASTM Standard Based on Society of the Plastics Industry Code Will Facilitate Recycling". Standardization News (September/October 2008). ASTM International.
- [3] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [4] Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision. Springer, Cham, 2016.
- [5] Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." ICML deep learning workshop. Vol. 2. 2015.
- [6] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
- [7] Bobulski, Janusz, and Mariusz Kubanek. "Waste classification system using image processing and convolutional neural networks." International Work-Conference on Artificial Neural Networks. Springer, Cham, 2019.