Classification of images of recyclable trash into different categories

Mohammad Abbas Meghani, Qihua Zhong.

Introduction

- We will tackle the task of classifying images of different recyclable garbage into their corresponding categories (metal, plastic, glass etc.) with a special focus on CNNs.
- This project aims to develop different models for the task and compare their performance on different metric. We will also experiment with different hyperparameters and examine the effects on the final performance

Data

- We will collect publicly available dataset of recyclable garbage pictures for the classification task.[1]
- This data contains images of garbage classified into 6 categories – paper, plastic, glass, cardboard, metal and trash.
- There are 2527 sample images in total. We use a 70% train, 30% test split for training and validation purpose.
- The images are 512 x 384 pixels. We further rescale and shape the image for modelling depending on requirements.

Sample Images:

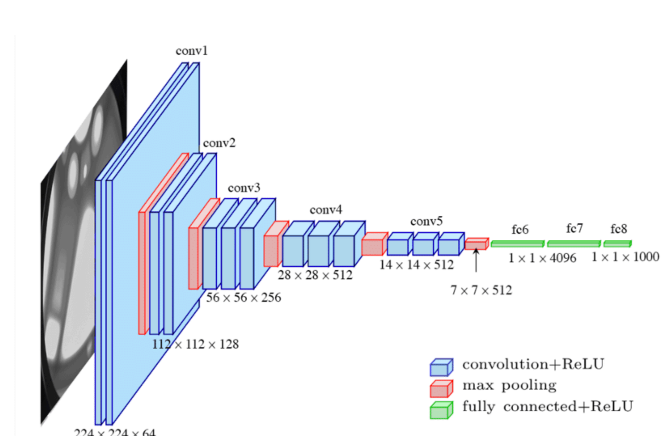
Models

SVM

- We use SVM as a baseline model for comparison with CNN.
- We create two models one with a linear kernel with default parameters and one using GridSearchCV for exhaustive search of different hyperparameters.

CNNs

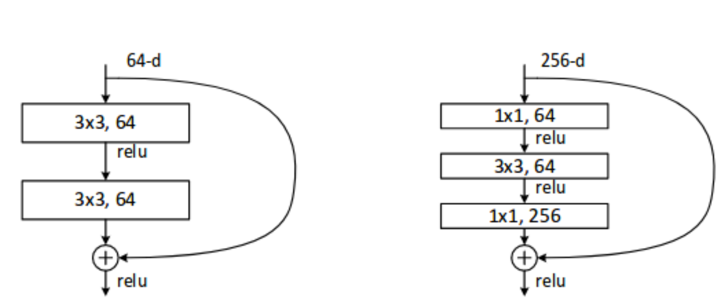
- Inspired by VGG16[3], we trained various CNNs similar to the VGG16. The models were trained from scratch with different hyperparameters:
- whether to use batch normalization: [True, False]
- pooling type: ['Max', 'Average']
- number of layer:[16, 13, 10]



Standard VGG16 structure[4]

Pre-trained Resnet34[2] :

- ResNets solve the problem of vanishing gradients by using residual blocks.
- Resnet34 is a 34-layers residual network with short-cut connections and bottleneck layers. We initialized the ResNet34 with pre-trained weights and continue to train the neural networks on our recyclable trash dataset.



Examples of Residual block[2]

.

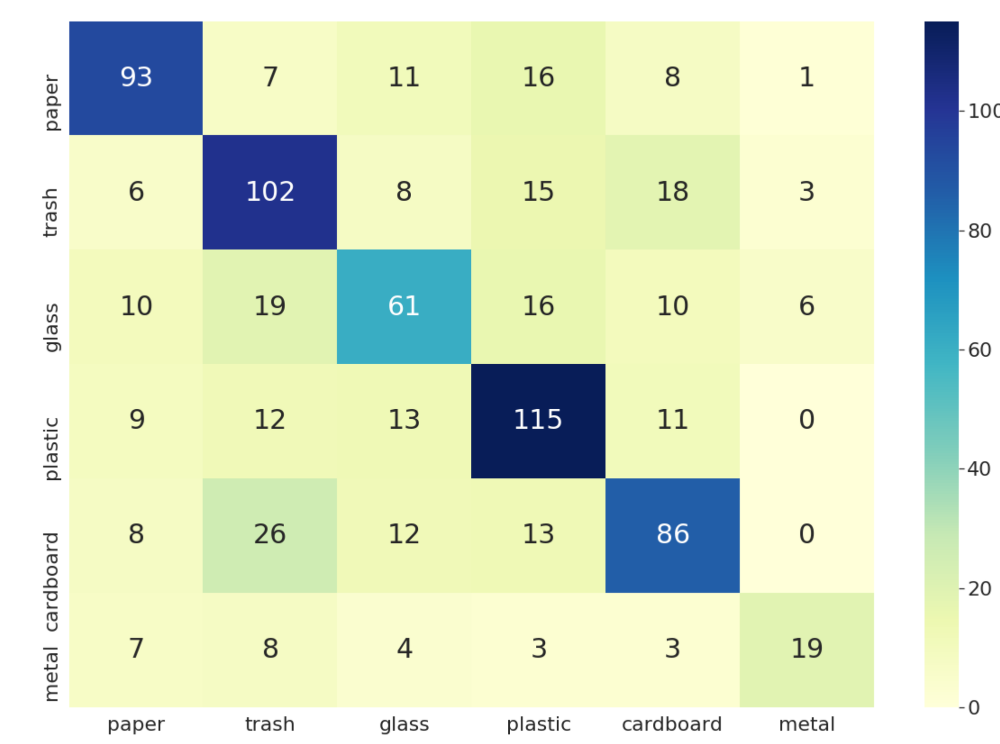
Results

SVM Linear

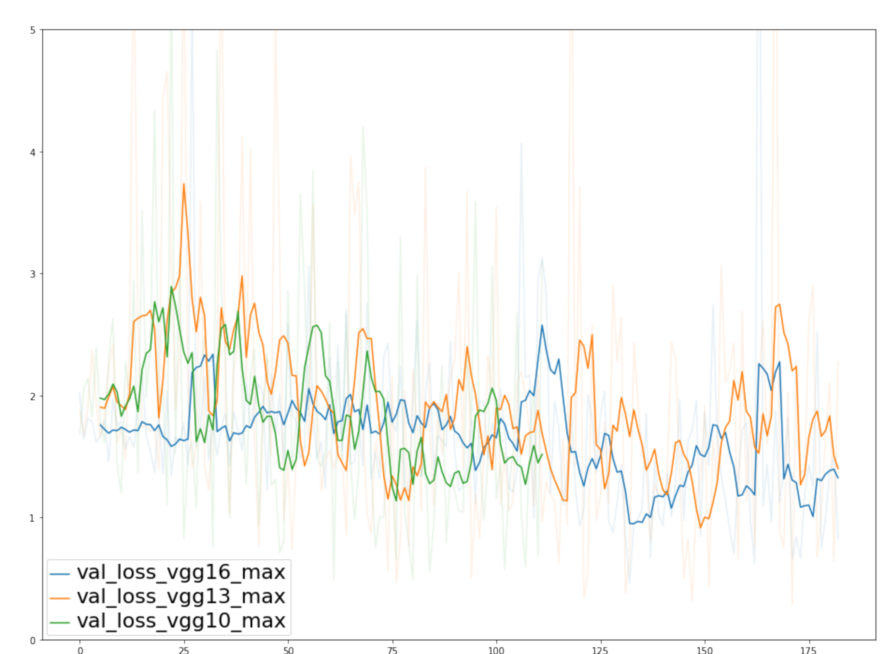
Classification report

	precision	recall	f1-score	support
Cardboard	0.46	0.49	0.47	136
Glass	0.36	0.43	0.39	152
Metal	0.36	0.31	0.33	122
Paper	0.60	0.59	0.59	160
Plastic	0.51	0.46	0.49	145
Trash	0.41	0.39	0.40	44

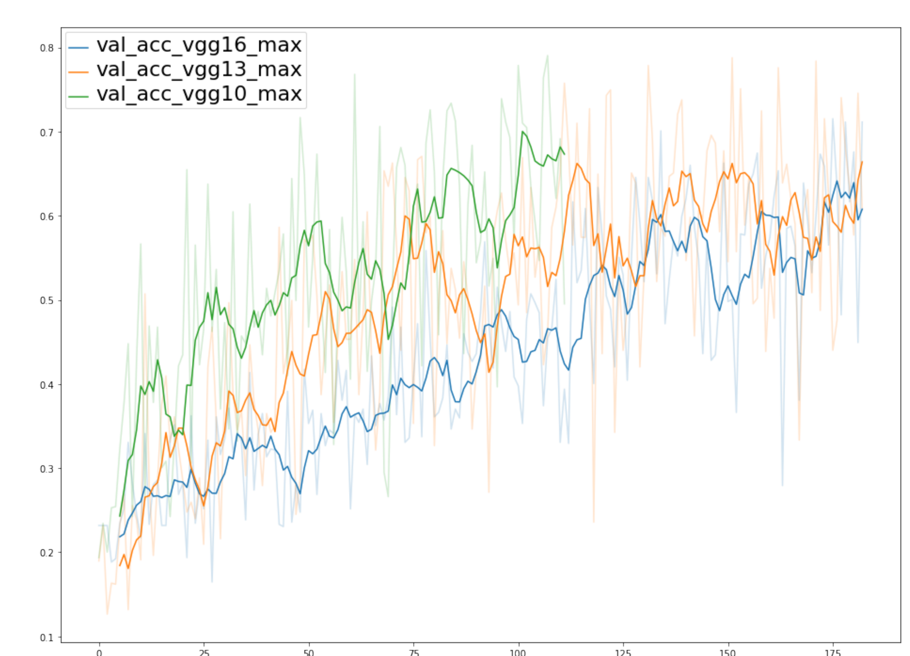
Confusion Matrix



CNNs



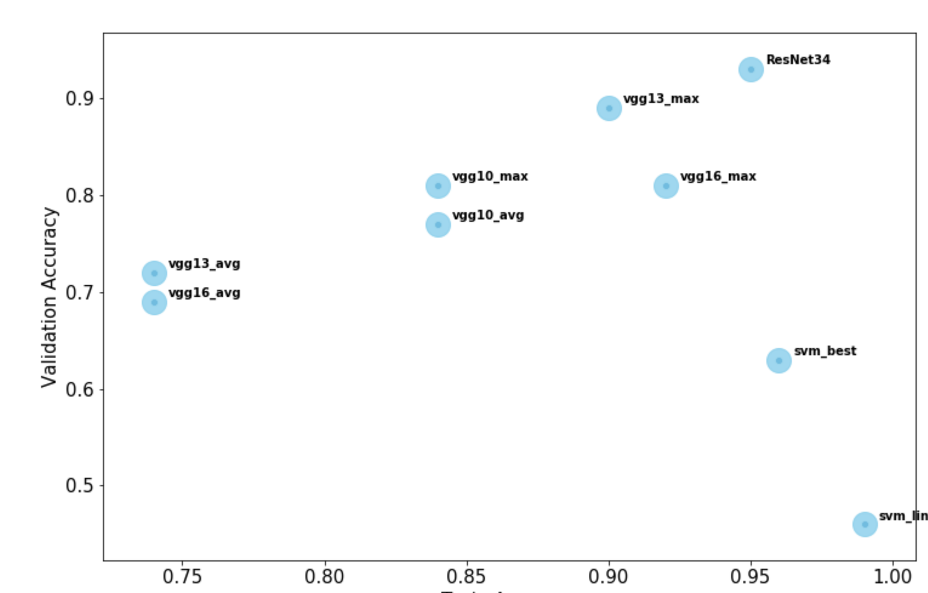
Validation Loss



Validation Accuracy

Performance Comparison

Model	Val Accuracy	Train Accuracy
vgg16_avg	0.69	0.74
vgg13_avg	0.72	0.74
vgg10_avg	0.77	0.84
vgg16_max	0.81	0.92
vgg13_max	0.89	0.90
vgg10_max	0.81	0.84
svm_lineaer	0.46	**0.99**
svm_best (rbf)	0.63	0.96
ResNet34	**0.93**	0.95



Findings

For this specific image classification task, we have the following findings:

- **Transfer Learning:** The pre-trained ResNet34 significantly outperforms all variations of SVMs and of CNNs that we experimented.
- **Batch Normalization** is essential for training deep networks.
- **Data Augmentation**: Models start to overfit quickly without data augmentation.
- **Number of layers:** Networks with less layers are easier to train compared to deeper networks
- **Pooling**: Max pooling is consistently better than average pooling in this study.

References

[1] https://github.com/garythung/trashnet

[2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *CVPR*, 2016.

[3] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[4] Ferguson, Max & ak, Ronay & Lee, Yung-Tsun & Law, Kincho. (2017). Automatic localization of casting defects with convolutional neural networks. 1726-1735. 10.1109/BigData.2017.8258115.