Recycling Machine

By Yeon Hwa Choi

The Problems

- The pollution
- 2.12 billion tons of waste are dumped/ year
- Lack of recycling

Question

- What can help with this pollution problem?
- How data science can be used in reducing waste?
- Can a machine do recycle for us?

Objective

To find the best **deep learning** model that will predict the content of image

Parameters to be compared:

1. Number of layers
2. Activation functions
3. Batch size
4. Learning rate

Dataset

Total 25,077 number of images

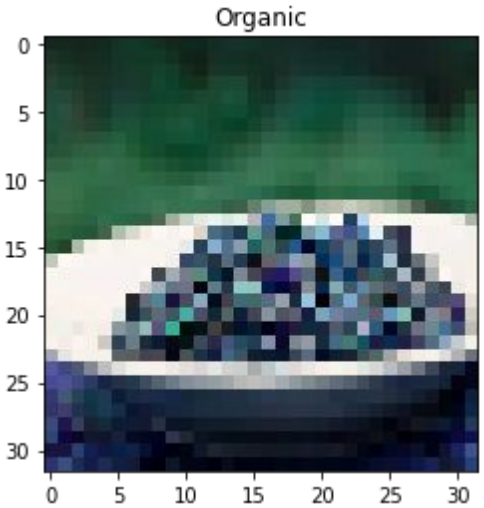
Training:

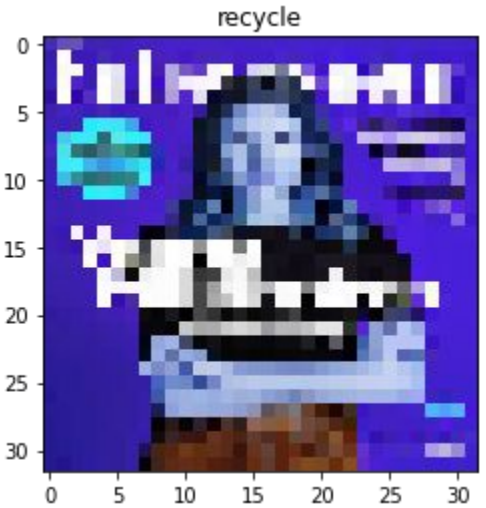
- 12,565 of organic waste images
- 9,999 of recycle waste images

Testing:

- 1,401 of organic waste images
- 1,112 of recycle waste images

Visualization





- Resized the data in 32 x 32.
- Not clear to see visually
- Unnecessary to resize them into higher resolution to built deep learning models

Objective 1 - Number of Layers

# of layers	3	5	7
Accuracy score	48.59%	62.36%	71.55%

Objective 2 - Activation Function

Activation Func.	ReLU	Tanh	Linear
Accuracy score	71.55%	74.61%	65.98%

Objective 3 - Batch Size

Batch Size	20	128	60,000
Accuracy score	75.25%	74.61%	69.76%

Objective 4 - Learning Rate

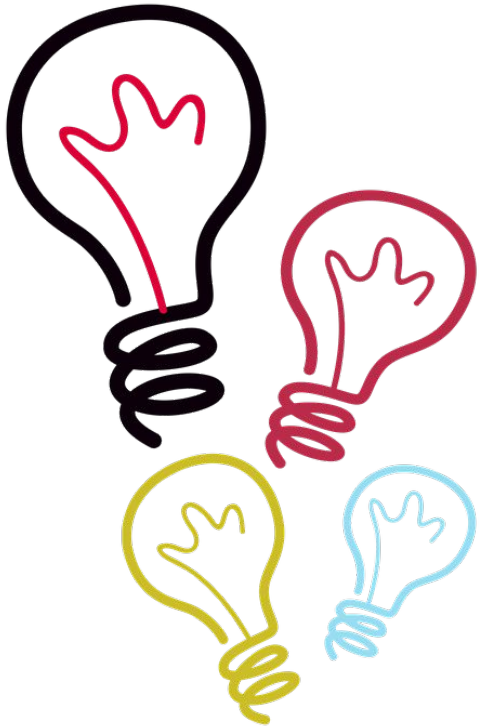
Learning Rate	Default	0.01	0.001
Accuracy score	75.25%	100%	98.65%

*Although the learning rate of 0.01 gives the accuracy of 1, it deemed to be overfitting. Thus, 0.001 was concluded to be the best model.

Conclusion - Best Model

Best Model To Recognize the Image of Waste:

Activation Function**: Tanh**
Number of Layers: **7 layers**
Batch Size: **20 batch**
Learning rate: **0.001**



Challenge

- Number of target variable: data is binary vs. real problem is multi-class
- Size of the data - real problem data could be too large to run on certain computers
- Further development of the model needed to be applied for real problems

Recommendation

This project was started from the idea of recycling machine which will scan the waste and recognize the type of the waste and allocate it by the type of waste. Although the model will need further development and deeper research, the deep learning model with activation functions of Tanh, 7 layers, batch size of 20, and learning rate of 0.001 is recommended to be used to start creating the recycling machine!