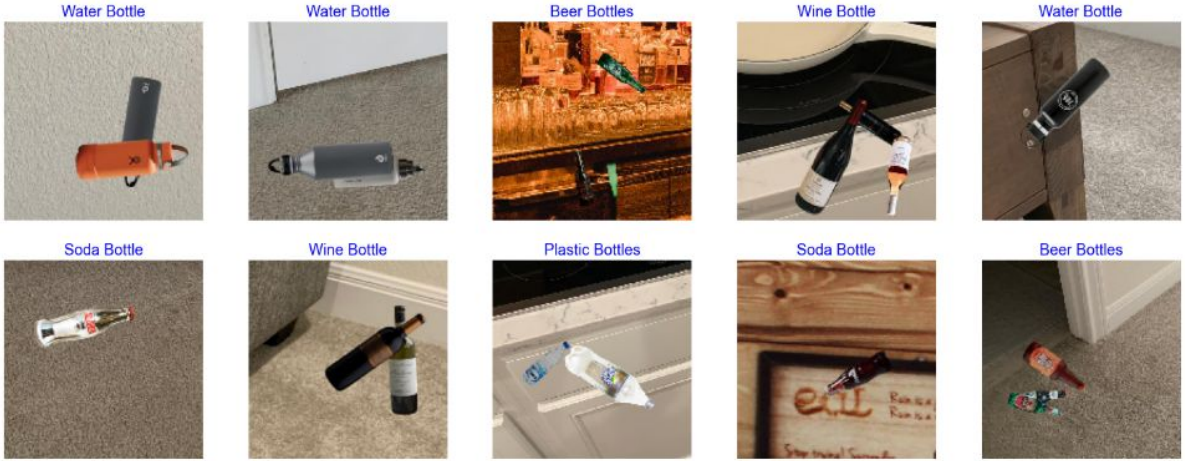
VANDERBILT UNIVERSITY

Bottle Image Classification Final Presentation

Team 5: Sydney Simmons, Anna Cameron, Cameron Rondeau

Problem

The problem we are solving through our project is to be able to **classify different types of bottles through image classification**. We think this could be useful in practical applications, like sorting for recycling. The below shows different types of bottles that we want to be able to label.



Initial Modeling

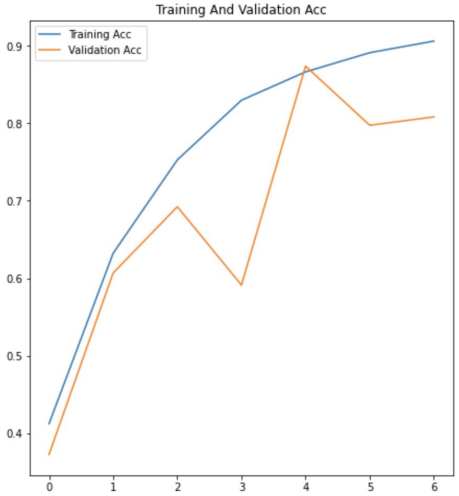
- In our last progress presentation, we presented the below results for our baseline model and a more advanced model architecture using ResNet
- Our takeaways from that presentation were to explore other model architectures and explore a realistic approach to reducing the long runtime of the ResNet model

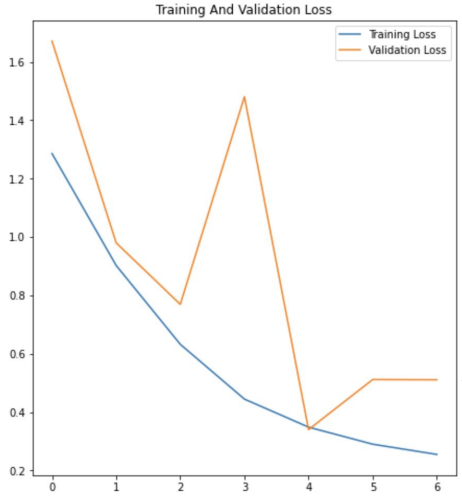
Model	Accuracy	Model Runtime
Baseline CNN Model	0.91 (10 epochs)	~ 1 hour
ResNet	0.96 (9 epochs)	~ 3 hours (timed out)

Increased accuracy using ResNet model architecture

ResMLP Model

- In exploring other architectures, we implemented another model architecture using ResMLP with a validation accuracy of ~80% but much faster training time
- In this code, we implemented early stopping based on the validation loss increasing for two epochs





Early stopping due to validation loss

```
early_stopping = tf.keras.callbacks.EarlyStopping(monitor='val_loss', patience=2, min_delta=1e-4)
```

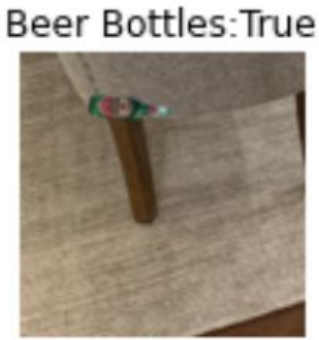
Overall Results

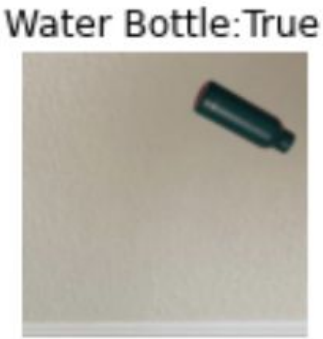
- Due to the lower accuracy of the ResMLP architecture, we chose the ResNet architecture as our final model
- To realistically deal with the runtime issue for ResNet, we reduced the epoch number for the ResNet model to 5 without sacrificing accuracy and drastically improving runtime

Model	Accuracy	Model Runtime
Baseline CNN Model	0.91 (10 epochs)	~ 1 hour
ResNet	0.95 (5 epochs)	~ 1 hour
ResMLP	0.80 (6 epochs)	~ 20 minutes

Model Testing (1/2)

- To confirm the model was predicting correctly, we tested the model on random images that we choose from the data and confirmed that the images matched the labels





Model Testing (2/2)

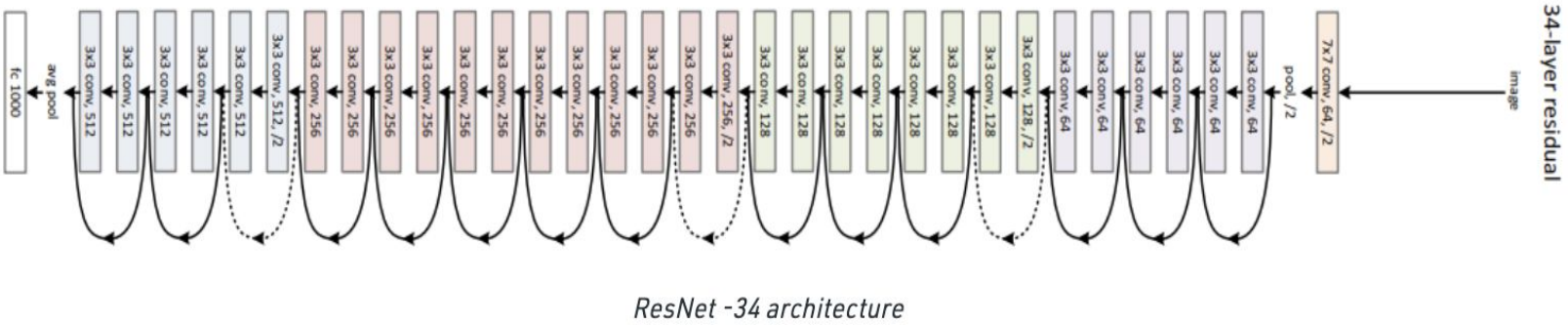
- We also tested the model on images that we took ourselves to confirm that the model could predict correctly outside of the dataset

Conclusions

VANDERBILT UNIVERSITY

- Image classification accuracy improved with advanced model architectures
- Having a large dataset is helpful to training, but can create an issue for training time
  - Did not do data augmentation as many images in the dataset were already shrunken, rotated or on different backgrounds
  - Would consider data augmentation for a different data set with less images
- Lengthy training time might be an issue, even with GPU resources
- Github Link: https://github.com/sydneysimmons/bottle-image-classification-project

Appendix: ResNet



ResNet -34 architecture

Source:https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/

Appendix: ResMLP

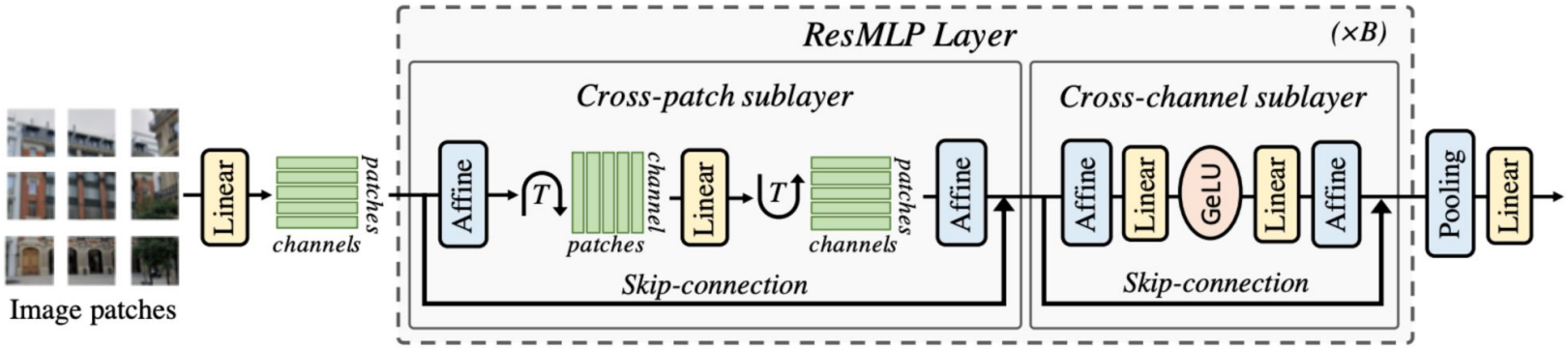


Figure 1: **The ResMLP architecture.** After linearly projecting the image patches into high dimensional embeddings, ResMLP sequentially processes them with (1) a cross-patch linear sublayer; (2) a cross-channel two-layer MLP. The MLP is the same as the FCN sublayer of a Transformer. Each sublayer has a residual connection and two Affine element-wise transformations.

Source: https://paperswithcode.com/method/resmlp