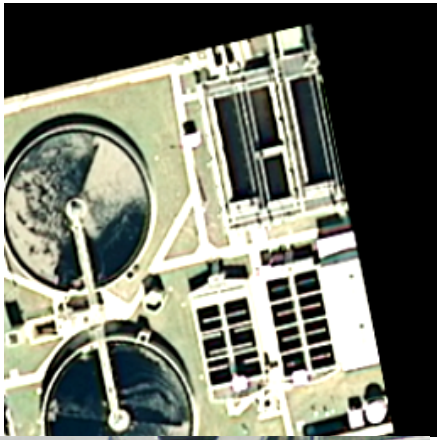
Land use classification

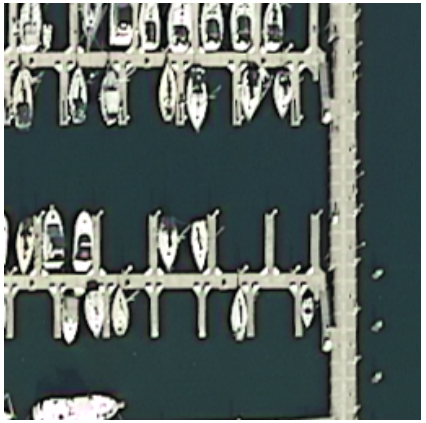
Kevin Fagan

Introduction

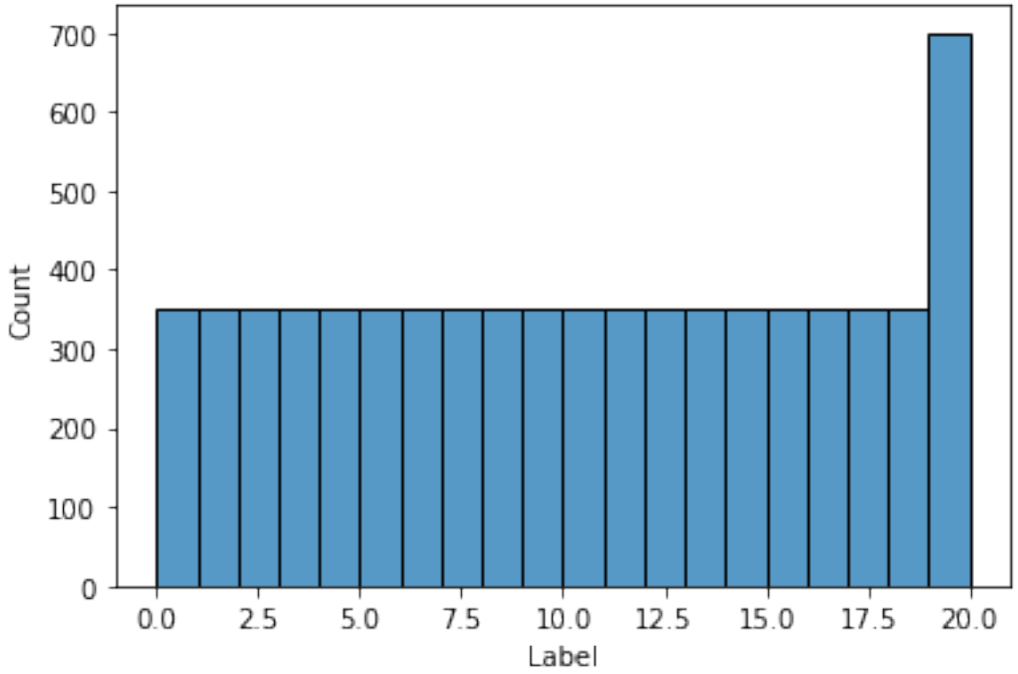
- Business problem.
  - The Landsat program is a joint program between NASA and USGS. It is the longest running enterprise for obtaining satellite imagery of earth.
  - Although it generates tons of data, it does not have a built in way to classify images.
- Goal
  - Use an image recognition model to classify images generated by the Landsat satellites.

The Images



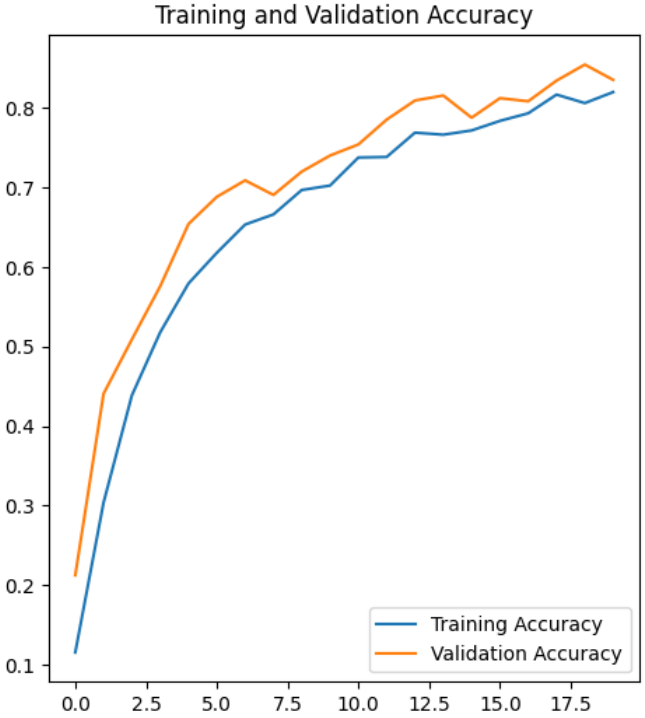


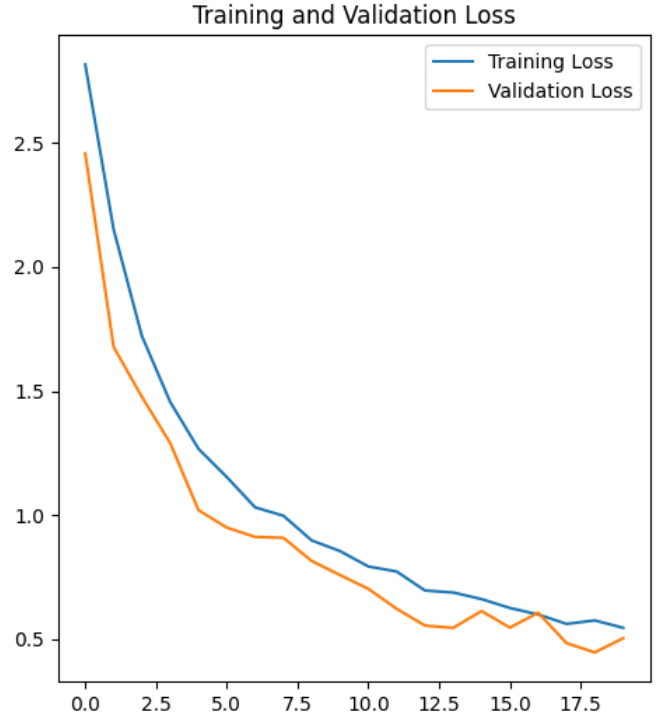
Distribution



CNN 1

First we created a model with 4 Conv2d layers and 1 fully connected dense layer of 512 nodes. It did very well right off the bat.





CNN 2

83% accuracy on test data right off the bat is pretty good.

Unfortunately, all subsequent models did significantly worse, even when they were much deeper or more complicated. Most shapes struggled to get about 40%

RESNET 50

ResNet50, developed in 2015, has a very good reputation for image classification, and was routinely able to get above 80% accuracy (with enough epochs), although it is much slower, and I don’t feel the extra processing time is worth it.

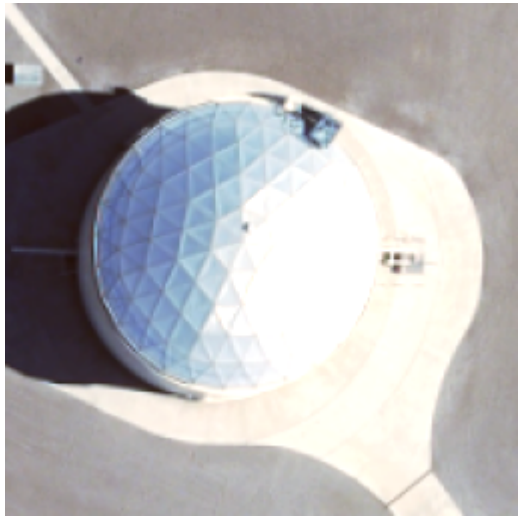
Classification

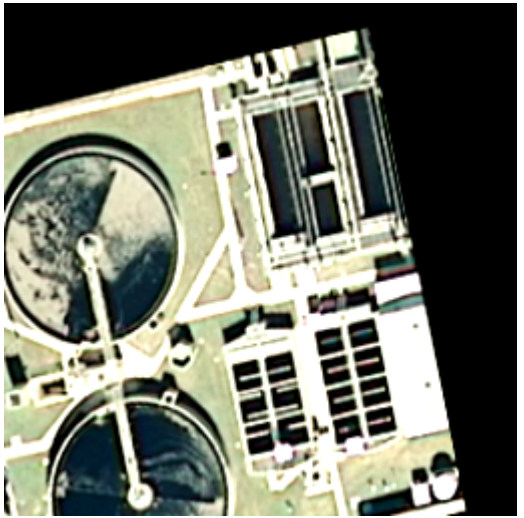
- We were able to do pretty well (83%) with our first CNN model, which was very wide, but not very deep.
- All attempts to improve it did poorer.
- RESNET 50 is widely considered one of the best image classification models out there.

Next steps.

- All images used to train and test the models are 256 * 256 pixels.
- Images from other earth sensing satellites could be added.
- More data should be obtained and from more categories to make the model more useful.
- The category labeling can be improved, but it would be a slow and tedious process.
- As the model is improved, it can be used for insdustrial and infrastructure planning.

Label problem





Questions

- Kevin Fagan

  kevinmfagan@gmail.com

  linkedin.com/kevin_m_fagan