ENEL 525 Fall 2024 – Final Project

Nicholas Lam – 30115728

Dec 17, 2024

University of Calgary

# Introduction

This project involves developing a convolutional neural network (CNN) to classify aerial satellite images based on their land use categories. The goal is to utilize deep learning techniques, extract relevant features, and achieve accurate classification of various land types. By going through the exercise of identification and recognition of different patterns in the dataset, I can learn about the effectiveness of CNNs in handling image classification and how they work.

# Methodology

I have been given a dataset containing 2100 aerial satellite images, consisting of 21 different land use categories. To help identify these land use categories, I require a neural network model which can identify new input images to a high degree of accuracy.   
  
I will be utilizing Python with the TensorFlow library to develop a model to identify the land use category of input images, training with 80% of the images and using the last 20% for validation/testing (2/3 for validation, 1/3 for testing).

# Parameters

**Learning Rate:** 0.0005

Learning rates from 0.001 to 0.0001 were tested, it was found that 0.001 led to poor training which was quick but very inconsistent in the resulting accuracy and loss. As the learning rate was tuned down, 0.0005 seemed to be a sweet spot where accuracy was consistently above 90% with a loss of less than 0.3 in training and 0.5 in validation. These results seemed to be the maximum performance from the model, being the same at rates lower than 0.0005, as a result, the learning rate was raised again to 0.0005.

**Batch Size:** 100

Batch size was fixed arbitrarily, other parameters were used to tune training stability and computation speed.

**Number of Epochs**: 25

The number of epochs was determined by when the learning rate plateaued. The model was left running with 50 epochs and an early stopping mechanism when the rate of loss stagnated for 3 epochs in a row. It was found that this model stagnated at around 23 epochs, as such, the number of epochs was set to 25.

**Optimizer:** Adam

The Adam optimizer was selected as it was recommended from class and the easiest to manipulate using the parameters. It seemed like a good default choice which would have good application to this model.

**Dropout:** 0.1

The value of the dropouts was determined through testing after the learning rate was set. Values from 0.4 to 0.01 were tested, with lower values in the range of 0.05 to 0.15 giving the smoothest and consistent accuracy and loss curves. High values often caused the model to take longer to train, while giving similar results, while values below 0.05 caused the models testing accuracy to be significantly lower than the training accuracy (overfitting). A value of 0.1 was settled on, as it consistently allowed the model to reach 90%+ accuracy in around 25 epochs.

**Kernel Size:** 3x3

The kernel size of 3x3 was first selected as it was used in the exercises prior to the project. 5x5 and 7x7 was tested after the learning rate and dropouts were finalized, however, they did not perform as well as 3x3, so 3x3 was left.

**Activation Function:** ReLU

Similar to the optimizer, the ReLU activation function as selected as it was used in class. Leaky ReLU was looked at after all testing was completed, but it’s primary benefit, the prevention of dead neurons, was mitigated with the use of dropout, making it unfavorable to attempt to change.

**Activation Function:** Softmax

Softmax was required to get an output for multi-class classification.

**Padding:** None

Padding is not used, as most of the dataset had features on the edges of the image. The airplane, sparse residential, intersection, and baseball diamond categories were the only categories which contained their distinctive features primarily in the center of the image.

# Training

The training dataset utilized 80% of the images and the last 20% was used for validation/testing (2/3 for validation, 1/3 for testing).