Forest Area Mapping using U-NET Segmentation Model

Eric Muthomi Gitonga

March 2025

1 Objectives

- Design and develop a U-Net model for accurate semantic segmentation of aerial images of forested regions.
- Successfully perform transfer learning by using a pre-trained ResNet model as the backbone encoder.
- Compute and analyse segmentation performance with Intersection over Union(IoU) as a performance metric.

2 Methodology

The model is a modified U-Net architecture using ResNet50 as the encoder with ImageNet features and an upsample decoder coupled with skip connections to preserve spatial resolution. The output layer uses a sigmoid activation.

2.1 Training Procedure

Data Preprocessing: I used a dataset of forest aerial images from Kaggle. I implemented $ImageDataGenerator$ for data augmentation and normalization [0,1] to even out the parameter magnitudes, improving training and generalization. I paired the image samples with their corresponding masks using a custom generator based on $zip()$ method. All images and masks were resized to 256×256 pixels for uniformity and computational efficiency. The (train-validation) and test split was a 0.85-0.15 split resulting in 3690 training samples, 651 validation samples and 767 test samples.

Hyperparameters: The model was compiled using $Binary\,Cross{-}Entropy$ as the loss function, sigmoid activation function the output layer for binary segmentation(forest vs non -forest) and Adam as the optimizer.

I iteratively fine tuned the epochs, batch sizes(BS), learning rate(LR) and EarlyStopping patience score on validation runs and recorded the results.

3 Results and Evaluation

I ran the model with varying hyperparameters and the below results were the best case scenarios:

1. **Round One**: BS=32, LR=0.0001, epochs=30, ES=5

Accuracy	Val Accuracy	IoU	Val IoU	Loss	Val Loss
0.9236	0.7428	0.7341	0.5343	0.1703	0.7674

2. **Round Two**: BS=32, LR= 0.0001, epochs= 30, ES= 10

Accuracy	Val Accuracy	IoU	Val IoU	Loss	Val Loss	Avg Test IoU
0.9759	0.8376	0.8778	0.7051	0.0396	0.6103	0.6999

After several training runs, I settled on round 2 metrics; a batch size(BS) of 32 coupled with a learning rate(LR) of 0.0001 and running for 30 epochs. I used $EarlyStopping(ES)$ patience to stop training which occurred after 22 of the possible 30 epochs. I then predicted the masks using the test data, obtaining a test IoU score. The predicted masks showed promising results, only slightly deviating from the actual mask on several occasions which reflected an acceptable IoU score.

4 Challenges and Future Works

- **Scarce and complex data**: Initially my project was to be based on Sentinel-2 satellite imagery of tea plantations. However, lack of open-source segmentation data posed a challenge. I attempted to annotate my own satellite images and create the masks, but the complex nature of the data which contains 13 spectral bands proved to be time consuming considering the specified coding time(8 hours). This forced me to obtain alternative data on forest aerial images

- **Computational Expenses**: A ResNet50 with ImageNet weights contains approximately 71 million trainable parameters. This greatly lengthened the training period. Limited T4 GPU runtime from Google Colab also hindered multiple training rounds. The memory allocation from Colab's GPU was also not sufficient for running an instance with a batch size of 64, returning an Out of Memory(OOM) error.

Future works

I aim to follow on my initial research scope, utilising sentinel-2 imagery to produce high resolution imagery from different spectral bands and the corresponding masks. With the MERIT program actively sourcing for a premium Colab subscription, I will also have more computational power at hand to train my model for more epochs and larger batch sizes.

A Segmentation Results

This appendix contains visual examples comparing the input image, the predicted segmentation mask, and the ground truth mask from the test set.

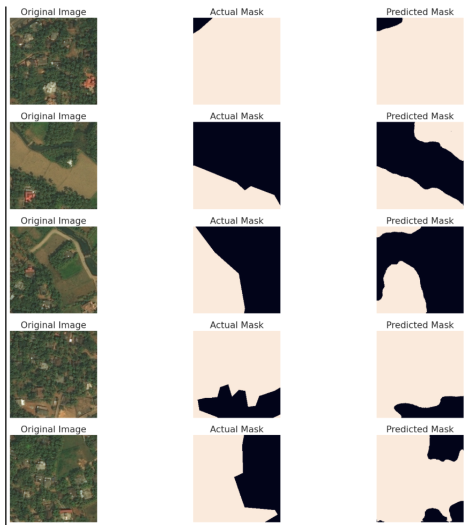

Figure 1: Predicted Masks on Test Samples

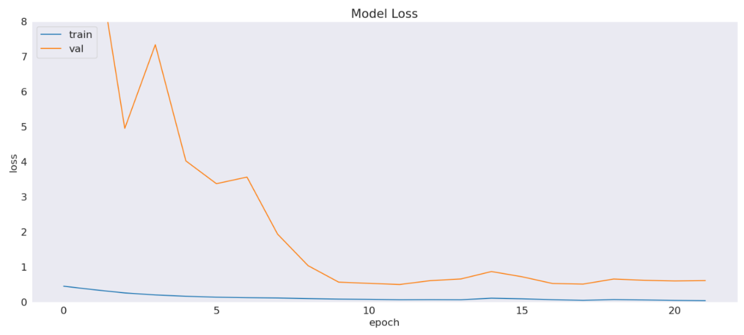

Figure 2: Training and Validation losses