Road Segmentation and Elevation Classification

Rik Das

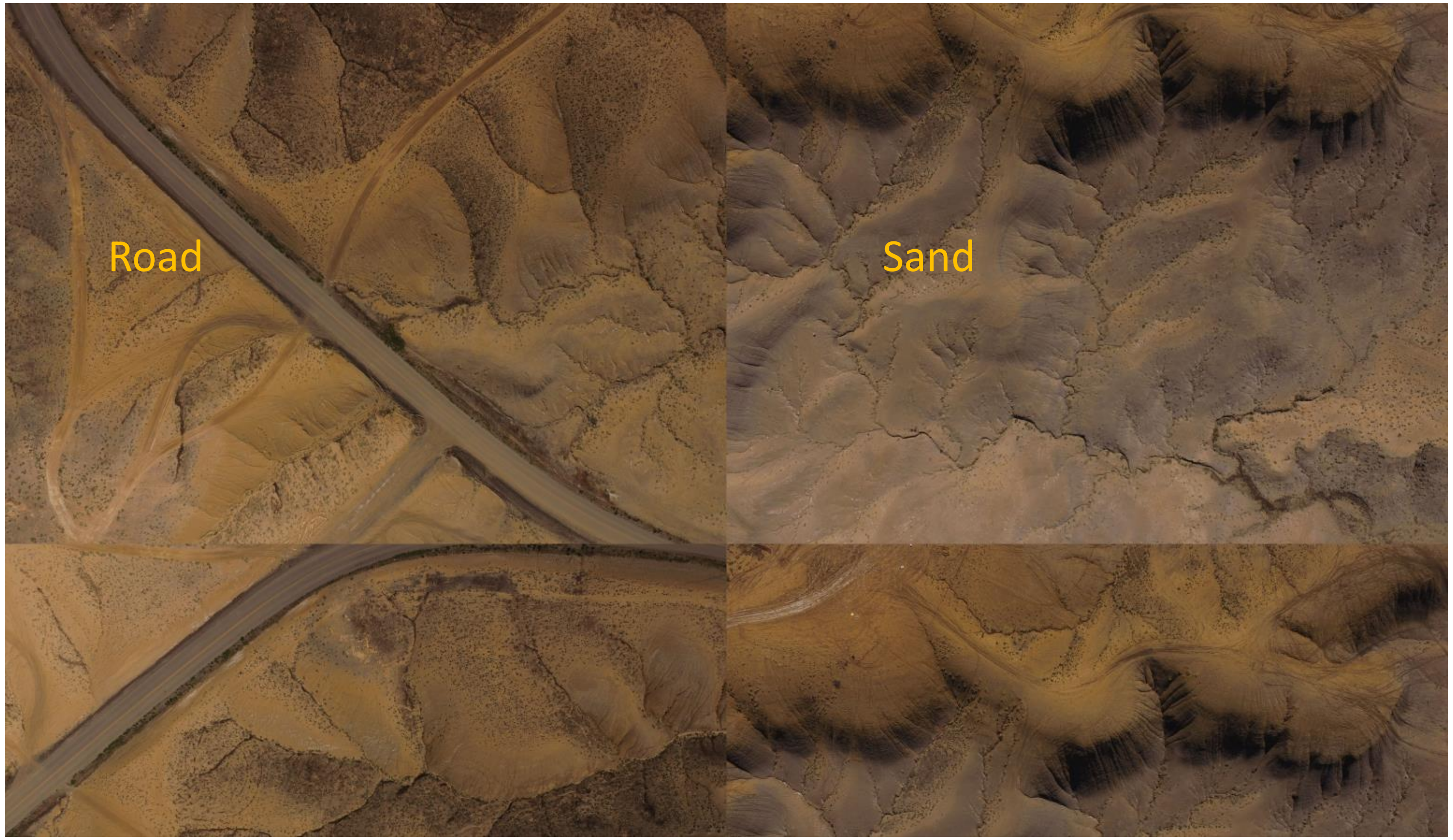
Agenda

- **Prelude**
- **The Problem Statement**
- **Proposed Solution (Road Segmentation)**
    - Preparation of Training Data for Segmentation
    - The U-net architecture
    - Training Accuracy and Loss
    - Testing with the saved weights
- **Proposed Solution (Elevation Classification)**
    - Some Initial Groundwork
    - Grondtruth Preparation
    - The Approach
    - Comparison with Correlation
- **Concluding Remarks**

Prelude

- Identifying a road with a machine used to be an impossible task till the mass acceptance of computer vision
- This has paved the way for multiple applications of Content Based Image Classification (CBIC) in real time.
- Image Segmentation is one of the significant applications which has exhibited comprehensive applications
- Identifying objects and boundaries with image segmentation has facilitated several contemporary automations

The Problem Statement

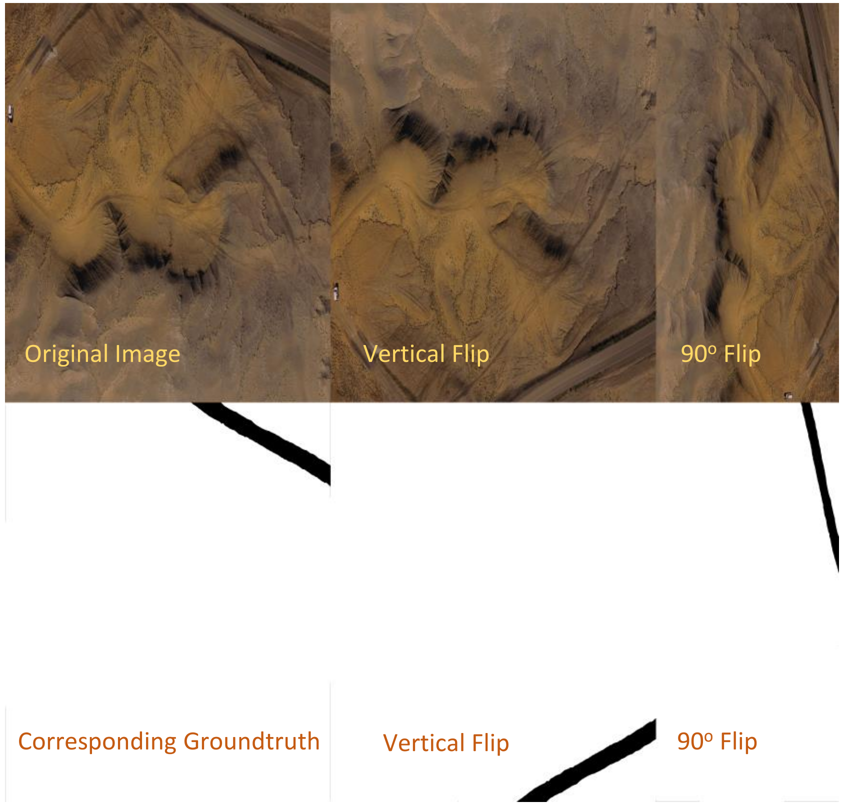


- Road segmentation and sand elevation classification have to be done from a dataset of 160 images captured with Drone Camera
- The dataset has no supporting groundtruth images or any other labels for training
- The images are taken from top view and the elevations are not visible in three dimensions
- The location at which the image is captured is also not disclosed

Proposed Solution(Road Segmentation)

- Primarily, a solution approach for road segmentation is undertaken
- Creation and labelling of groundturth is carried out using handcrafted technique
- The images having visible roads are initially cropped with only the object of interest
- The images without road are entirely cropped
- The images are further binarized using global thresholding technique to create the groundtruth

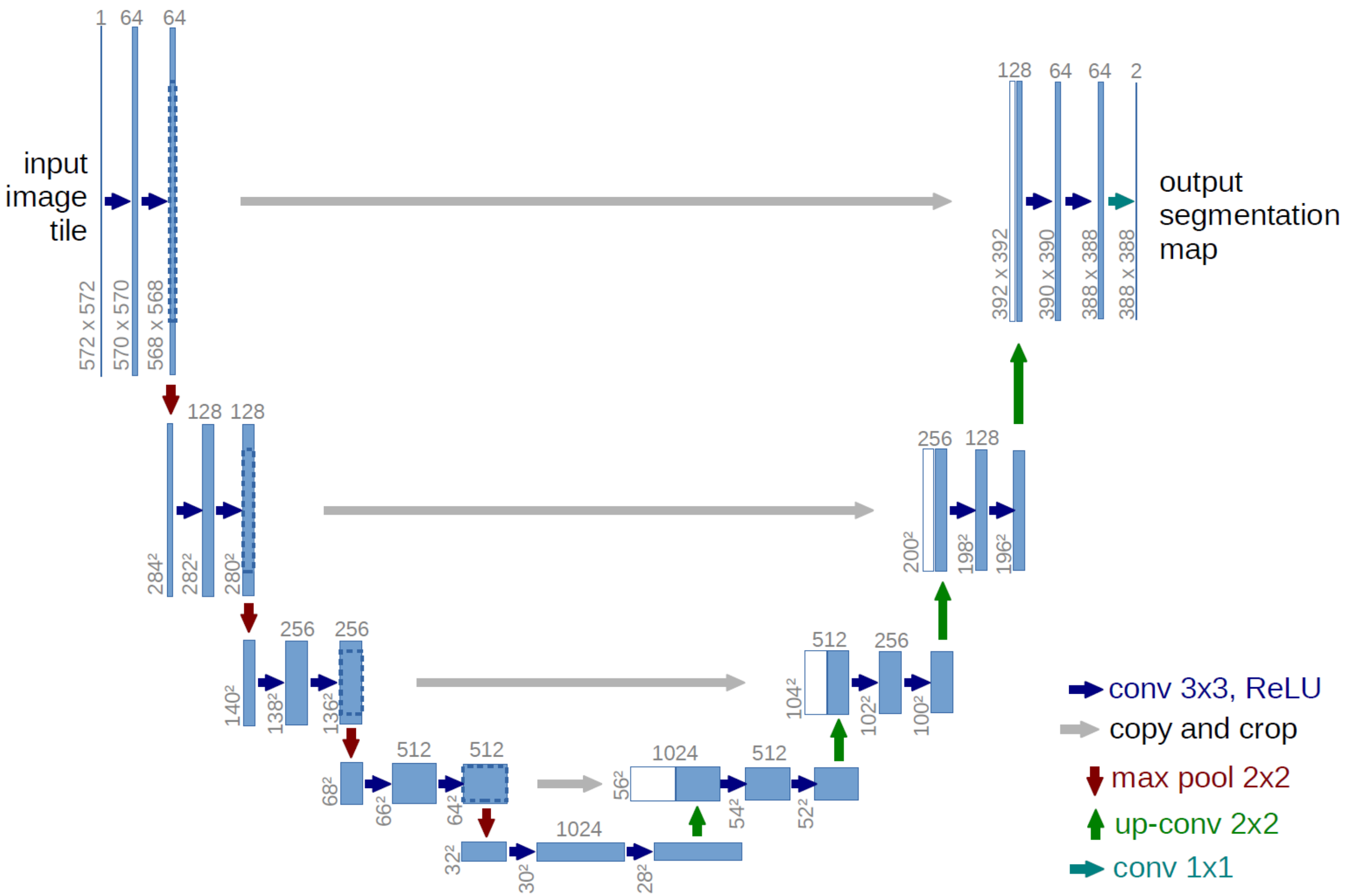
Preparation of Training Data for Segmentation



- The segmentation is carried out with an U-net architecture
- The entire dataset is divided into an 80/20 split with 80% of the images in training and 20% in testing
- Training is carried out primarily without data generation for training set augmentation
- Since the number of training records is only 128, another approach considering data augmentation is also conducted during the training with U-net
- It is carried out using Image Data Generator which has created several variations of the training images
- This has helped in leveraging two different approaches for generating segmented images providing option to choose the best

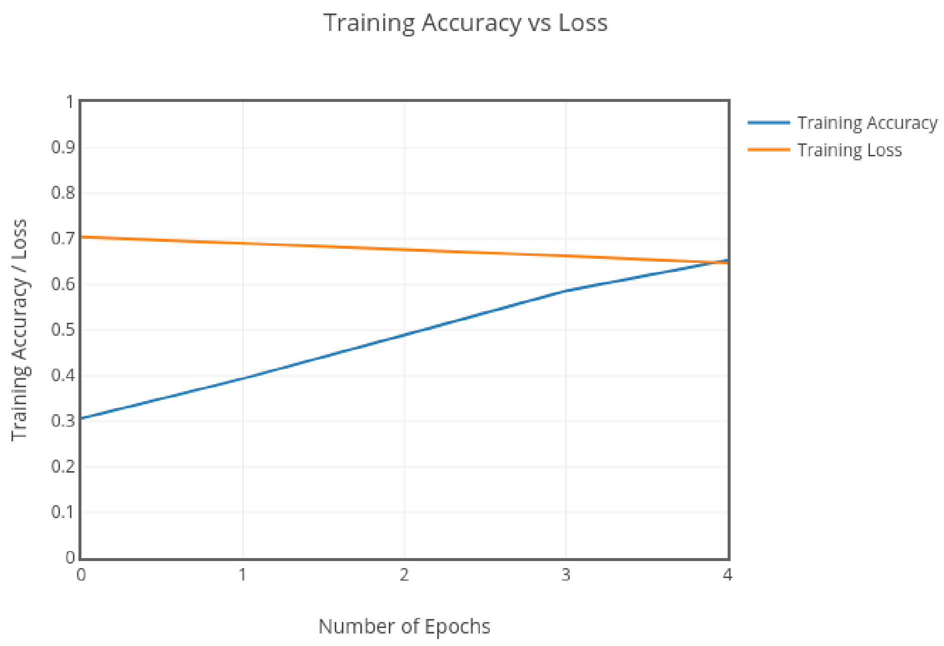
The U-net architecture

- U-net is chosen because of it's ability to carry out segmentation in automated manner
    - In contrast to conventional CNN, upsampling operators are used in U-net in place of pooling
    - Three different learning rates are tested to finalize with the learning rate of 0.000001
    - Binary cross entropy is chosen as the loss function for this binary classification problem
    - The U-net is 20 layer deep with convolution and max pooling function
    - A sigmoid activation function is chosen since the expected output is binary



Training Accuracy and Loss

- The training accuracy has increased upto 71% as shown in the graph above
- The training loss has shown inversely proportional relation to training accuracy
- Training is only possible with limited number of epochs since after that a " resource exhaust" error is declared everytime by Google Colab



Testing with the saved weights

- The image dataset has a test split of 20% of the entire image data.
- The data is used to generate the segmented images from the test set
- Samples are shown with images on the right
- Higher number of epochs in training with better computational resources could have resulted in further prominent results

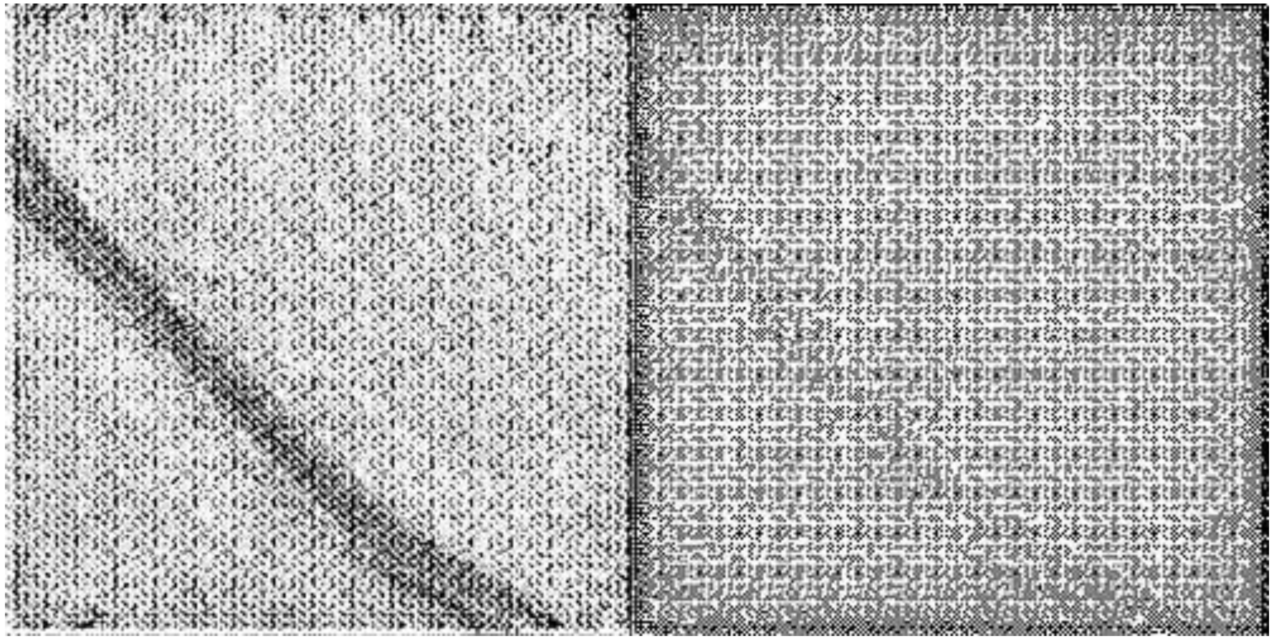


Image Generated without Data Augmentation

Image Generated with Data Augmentation

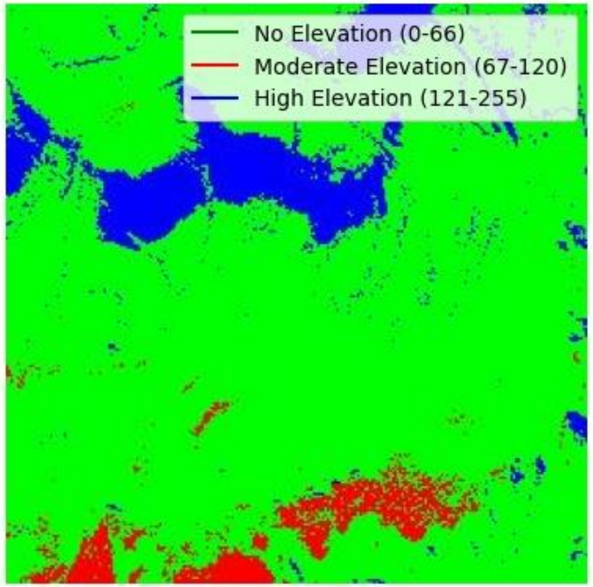
(only 5 epochs training)

Proposed Solution(Elevation Classification)

- The desert images are having sand elevation at different levels.

- The solution approach has considered three different categories of sand elevations
  - High Elevation (121-255)
  - Moderate Elevation (67-120)
  - No Elevation (0-66)

- The three levels are represented by three primary colors, namely, Red, Green and Blue as shown in the Figure



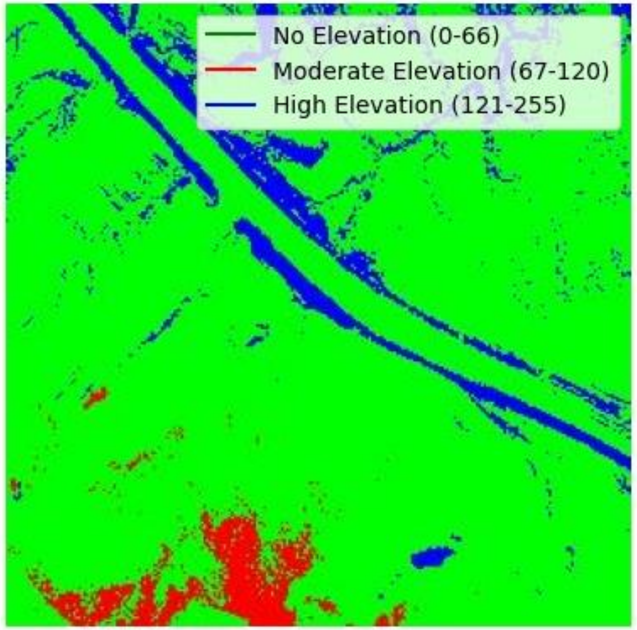
Some Initial Groundwork

- A PyTorch based implementation was primarily attempted which is for depth estimation of objects in images
- The model was not functioning properly due to resource constraints since the existing trained model is on a very different kind of data
- Hence, training from scratch is required for the model for which a “ resource exhaust” error is shown by Google Colab notebook just after 10 epochs since it works on shared resources
- Hence an unsupervised K means clustering based approach is chosen which is lightweight and can be implemented without exhausting much resources

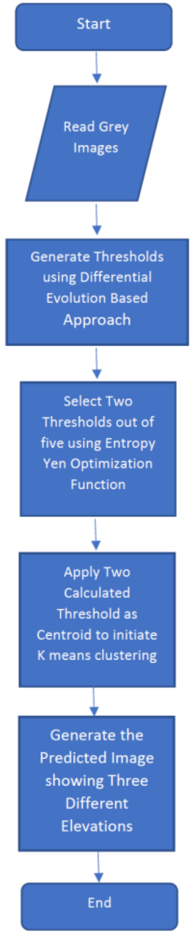
Groundtruth Preparation

- The groundtruth preparation is carried out by selecting two different thresholds dynamically

- The thresholds are selected to separate the images into three different labels, namely
  - High Elevation (121-255)
  - Moderate Elevation (67-120)
  - No Elevation (0-66)



Groundtruth Image



The Approach

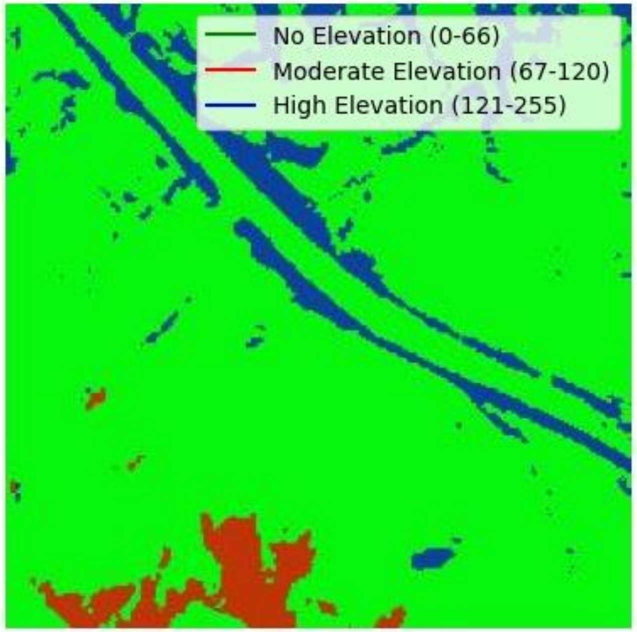
A differential evolution based k means clustering is designed to model the elevation classification approach

An optimization function named Entropy Yen is used to choosing the optimal threshold to initiate the contouring process

The process of contouring is followed by k means clustering for categorizing the different labels into 3 classes

Correlation is used as a metric of similarity matching since it yields the maximum result for locations where the generated image matches with corresponding groundtruth (pixel by pixel)

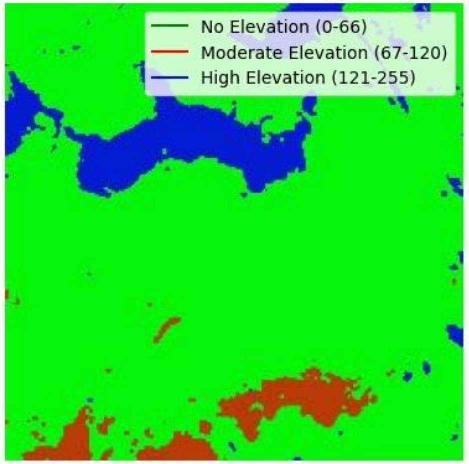
This helps to evaluate the relevance of the elevation classification carried out using the proposed algorithm



Generated Image

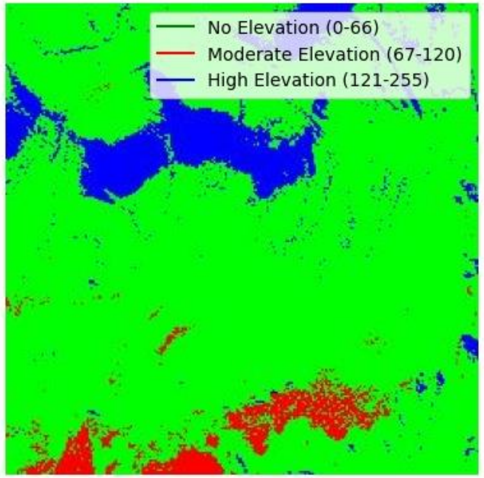
Comparison with Correlation

Original Image



Generated Image using Differential Evolution based k means Clustering

Correlation Coefficient 0.86



Groudtruth Image

Concluding Remarks

The solution approaches have proposed two different techniques for road segmentation and elevation classification respectively

The resultant outputs are shown under restricted environment of resources

Improved results can be achieved with better resources by enhancing the number of epochs for road segmentation

The elevation classification approach can be also attempted using pretrained neural network supported by sufficient resources

THANK YOU