PERFORMANCE OF DIFFERENT U-NET ARCHITECTURES FOR INVENTORY OF COCONUT PLANTATIONS USING CARTOSAT-2 MULTISPECTRAL DATA

Sujeeth A Vankudari, Navneet Raju, Anirudh Maiya, Uma D and Shylaja SS
Center for Data Science and Applied Machine Learning
PES University, Bengaluru,Karnataka India

and

Hebbar R and Ganesha Raj, K.
Regional Remote Sensing Centre - South
Indian Space Research Organization, Bengaluru,Karnataka,India

Problem statement - Overview

Overview

- Over the last decade, the horticulture sector has become one of the important driving forces for the rapid development of agriculture in India.
- We aim at detecting coconut plantations from multispectral satellite images where morphological features of coconut trees are not visible.

Multispectral Satellite Images

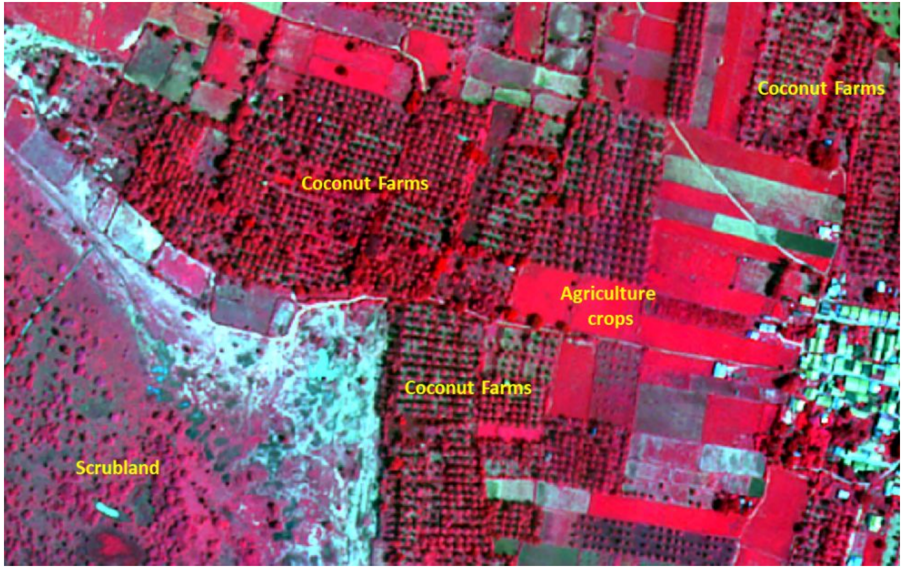
- We used multispectral images obtained from the Indian remote sensing satellite Cartosat-2 for this study.
- These multi-spectral images consist of four bands viz., Blue, Green, andRed and Near Infrared with spatial resolution.
- These multispectral image from CartoSat-2 has spectral resolution of 1.6 m

Benefits of Deep Learning to Coconut farmland detection

- Current status - Existing techniques account for spectral features and not the morphological characteristics of tree patterns.
- Although, OBIA has shown promising outcomes for classification of high-resolution satellite imagery, the post classification refinement is very tedious and time-consuming.
- The basic requirements for proper planning of these crops are the availability of reliable spatial information in terms of area and production at different spatial and administrative hierarchies

Classes Annotated -

- Coconut Farmland.
- Agricultural Crops.
- Other Scrublands.



Dataset Collection - Cartosat-2 Images

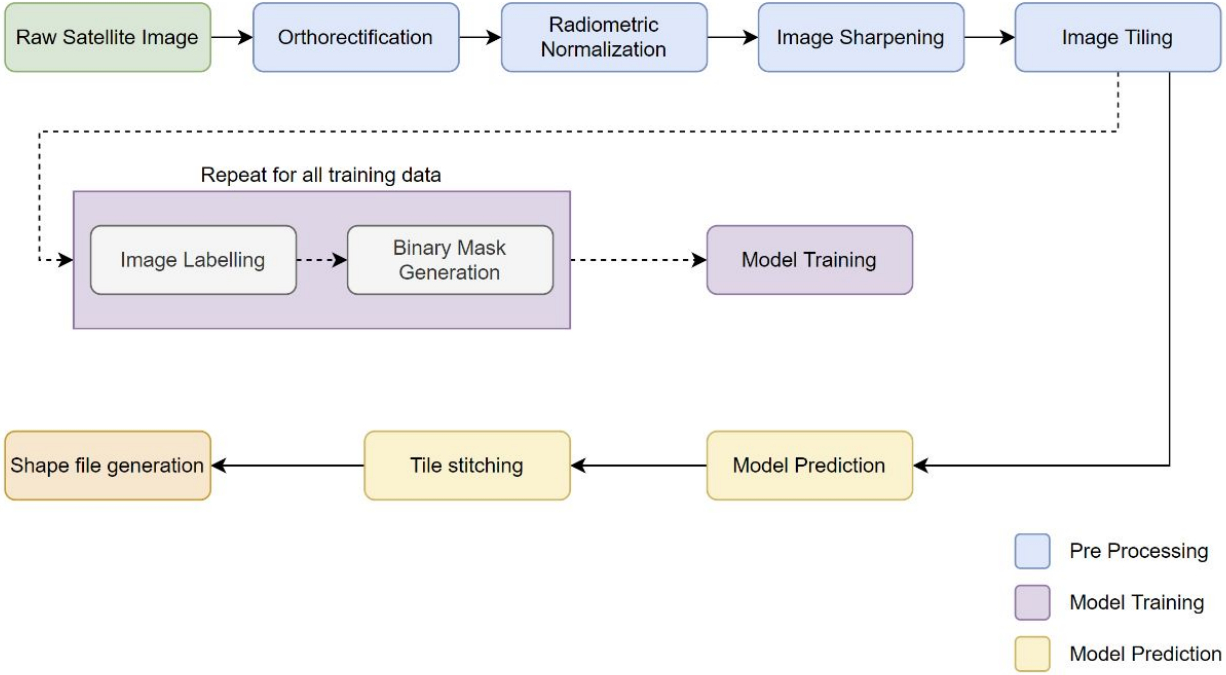
- Extracted 4/13 bands - Tumkur, Karnataka, India
- Bands : 2 - Blue, 3 - Green, 4 - Red, 8 - Near Infrared
- Source: Cartosat-2 Satellite.
- Image Dimension: 4* 9800*9800

Implementation

Pre-Processing

Deep Learning Model

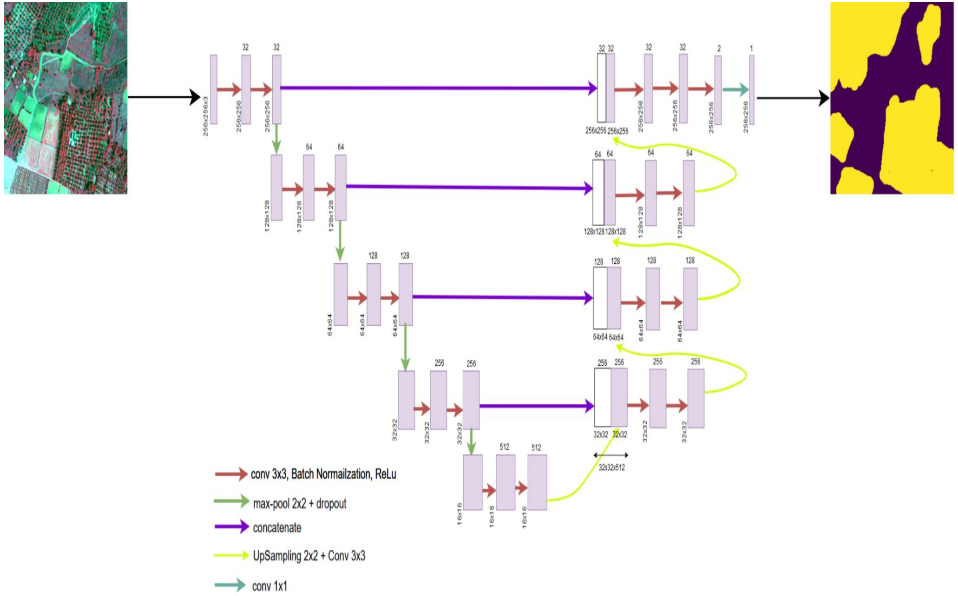
Overall flow of the system



Pre-Processing

- Orthorectification
- Radiometric normalization
- Image sharpening

Semantic Segmentation Using U-Net



What is Semantic Segmentation?

The goal of semantic image segmentation is to label each pixel of an image with a corresponding class of what is being represented.

Why U-Net?

It was invented to deal with biomedical images where the target is not only to classify whether there is an infection or not but also to identify the area of infection.

Hyperparameters

Hyperparameters	Values
Weight initialization	He initialization
Learning Rate (LR)	0.001
LR Decay Factor	2
Loss Function	Jaccard Distance
Batch Size	32
Activation Function	Relu

Custom U-Net

Model Details

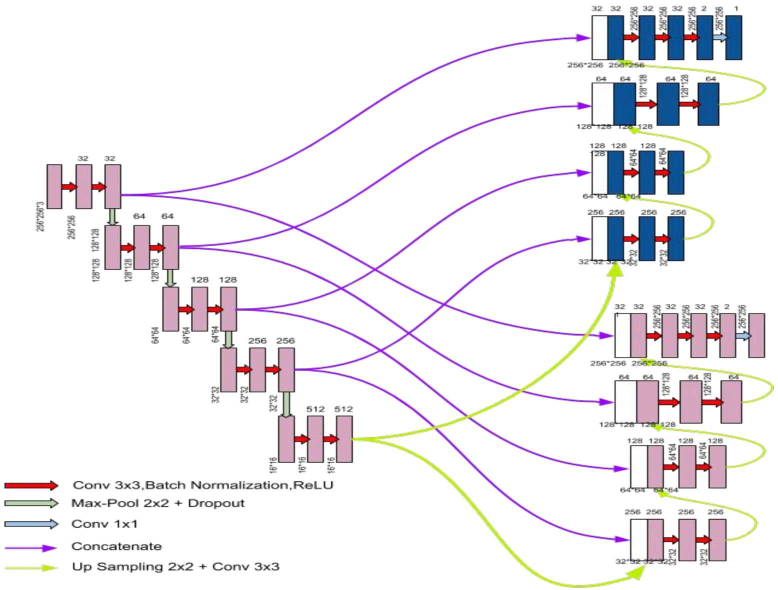
- This architecture consists of symmetrical contracting and expansive paths to downsample and upsample feature map respectively.
- Contracting Path: Convolutional Layers, ReLu Activation Function with max-pooling layer.
- Expansive path: Bilinear-upsampling along with Convolutional layers and ReLu Activation.

Model Configuration

- **Input Shape:** (4*256*256)-- 4 bands (Red, Green, Blue, NIR).
- **Loss Function:** Jaccard Distance
- **Dropout Rate:** 0.025
- **Output Shape:** (1*256*256)

Siamese U-Net

- We also introduce a new variant of U-Net inspired from Siamese Network.
- Unlike Custom U-Net, Hybrid Siamese Network has two expansive paths connected to a single contracting path.
- These expansive paths learns features contradictory to each other.



Siamese U-Net

Training Phase can be formulated as:

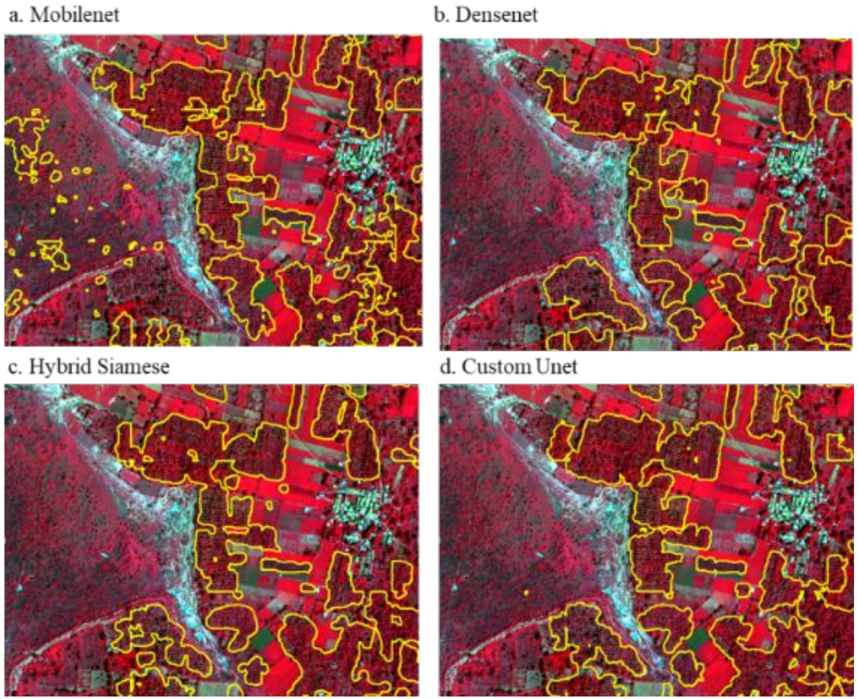
$$\begin{aligned} \hat{y_1} &= \mathcal{F}_{\theta_1}(x) \\ \hat{y_2} &= \mathcal{F}_{\theta_2}(x) \\ \min_{\theta_1, \theta_2} \mathcal{L}(\hat{y_1}, \hat{y_2}, y, \tilde{y}) &= \mathcal{L}(\hat{y_1}, y) + \mathcal{L}(\hat{y_2}, \tilde{y}) \end{aligned}$$

After the model has converged, let θ^*_1 and θ^*_2 be the parameters of the expansive paths. The final output y can be calculated as:

$$y^* = \mathcal{F}_{\theta_1^*}(x) \vee \neg \mathcal{F}_{\theta_2^*}(x)$$

Results

Spatial depiction of coconut plantations derived from four DL models



Metrics for comparison

1. <u>IoU Score (Intersection over Union)</u>

   Area of overlap between the predicted segmentation and the ground truth divided by the area of union between the predicted segmentation and the ground truth

2. <u>Accuracy</u>

   Classification accuracy was assessed using 224 independent samples collected from the field

IoU Score Comparison

Backbone of U-Net Architecture	IoU Score
MobileNet	0.6668
SE-ResNet-18	0.6668
ResNet-152	0.6668
VGG-19	0.6769
DenseNet-121	0.6861
Hybrid Siamese U-Net	**0.6921**
Custom U-Net	**0.7035**

Comparison of Model Accuracies

Backbone of U-Net Architecture	Accuracy %
MobileNet	83.51
SE-ResNet-18	85.83
ResNet-152	84.62
VGG-19	86.96
DenseNet-121	87.93
Hybrid Siamese U-Net	**89.17**
Custom U-Net	**91.88**

Note: Classification accuracy was assessed using 224 independent samples collected from the field

Future Work

- Despite the good results obtained with the proposed approach, there is further scope for improvements for large scale applications.
- Development of multi-class DL approaches is one of the priority areas of research in view of the diverse cropping patterns existing in India with multiple plantation crops.