Oil Palm Plantation Monitoring from Satellite Images using Deep Learning Approach

Mohd Asyraf Zulkifley

Universiti Kebangsaan Malaysia

Sustainable Plantation

Star ONLINE | News | Business | Sport | Metro | Lifestyle | Food | Tech | Education | Opinion | Videos | Property | Jo

Unfair to blame palm oil industry for haze, says Teresa Kok

NATION

Monday, 23 Sep 2019

12:08 PM MYT

By JO TIMBUONG

PUTRAJAYA: It is unfair to keep blaming the palm oil industry for the haze, as many companies adhere to sustainable agriculture standards, says Teresa Kok.

Introduction

- ★ Palm oil is one of the major commodities that contributes a lot to Malaysia economy.
- ★ In Jun 2019 alone, Malaysia exports a total of RM 5 billion worth of palm oil and its derivatives [Trade statistics].
- ★ It is the most popular vegetable oil with the benefit of Vitamin A and E.
- ★ Sustainable Oil Palm Plantation is the way moving forward.
- ★ Government role: Monitoring the land areas used for the plantation.

Remote Sensing

- ★ Traditional approach: Land survey - limited access, costly and tedious.
- ★ Drone: Limited coverage and issue with sampling repeatability.
- ★ Remote sensing: Full coverage, data coherency and acceptable sampling rate.
- ★ Challenges: Cloud and spatial resolution.

Plantation Size Prediction

- ★ Plantation areas will be segmented automatically.
- ★ Plantation size will be inferred according to the segmented pixels.
- ★ Size is a mapping of spatial resolution of a pixel and the total segmented areas.
- ★ Problem? It is hard to get a good segmentation map.

Sample Images

Deep Learning

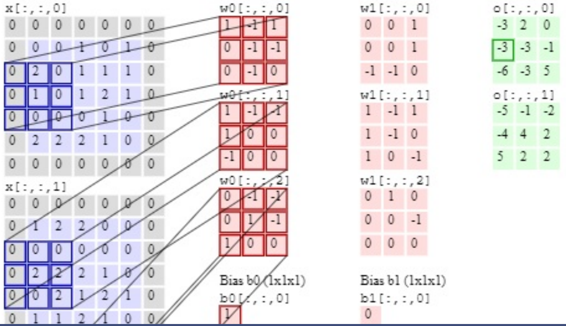
- ★ Traditional Machine Learning: Hand crafted features + trained classifier.
- ★ Deep learning: trained features + trained classifier.
- ★ It is an end to end training process. (input=original image, output=segmented image).
- ★ The features is trained by using convolutional neural networks (CNN).

Convolutional Neural Networks

★ The core is Convolutional layer: a set of learnable kernels/filters.

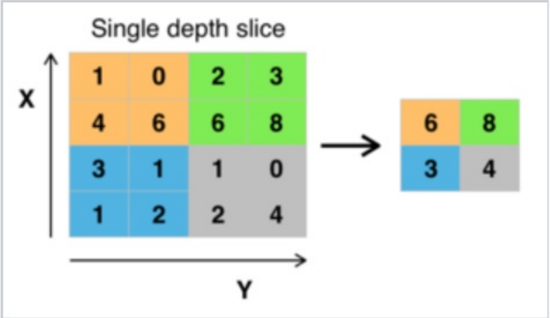
★ The learnable kernels will be trained for a specific task such as to classify the palm oil image.

★ Usual additional operations, 1) Pooling function and 2) Activation function.



Pooling Layer

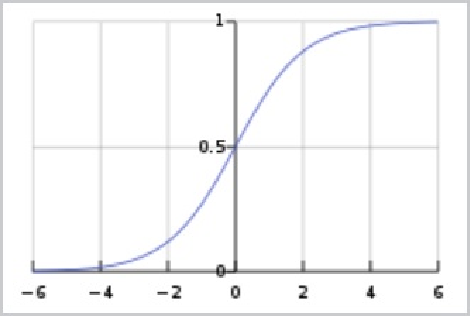
- ★ Pooling: to downsample the input so that the best response will be captured.
- ★ Pooling method: Max pooling, average pooling and min pooling.



Max pooling with a 2x2 filter and stride = 2

Activation Function

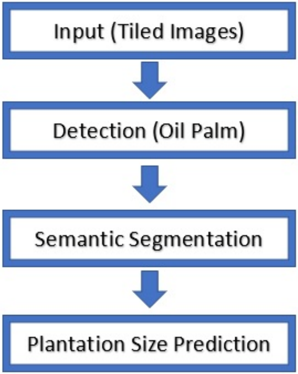
- ★ Activation function: to convert non linear separable input data to linear separable output.
- ★ Examples: Rectified Linear Unit function, Sigmoid function, Tanh function, Softmax function.



Logistic activation function

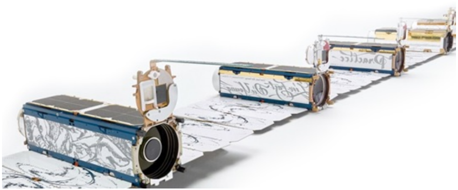
Monitoring System

- ★ Detection: Encoder + Dense.
- ★ Segmentation: Encoder + Decoder.



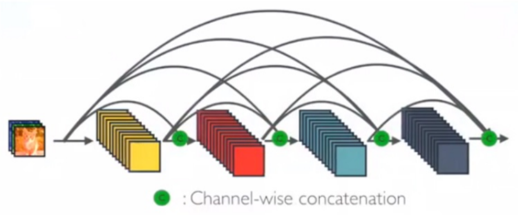
Data

- ★ Kaggle competition WiDS Datathon 2019.
- ★ PlanetScope: 120 unit sensors with four spectral bands of red, green, blue and near-infrared.
- ★ Total image: 15262 images.
- ★ 14320 negative images and 942 positive images.
- ★ 3 meters spatial resolution per pixel.



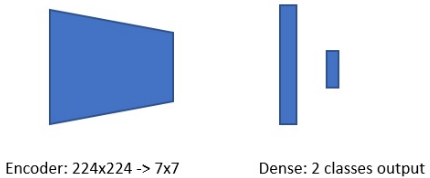
Detection

- ★ Detection is done to make sure that there is at least an oil palm plantation in the test image.
- ★ Without detection module, false negative of the segmentation module will be high.
- ★ To ensure that other images are excluded before segmentation process.
- ★ DenseNet architecture is chosen as the basis (Best paper for CVPR 2017).



DenseNet

- ★ Encoder: 4 Dense Blocks + 4 Transition Layer.
- ★ Dense: Global average pooling.
- ★ Dense Block layer concatenates the layers of the previous block.
- ★ Contrary to ResNet, where sum operation is used to combine the skip connection.
- ★ Three basic architectures are explored, which are DenseNet-121, DenseNet-201 and DenseNet-264.



Squeeze and Excitation Module

- ★ Recalibrates feature responses by explicitly modelling interdependencies between channels.
- ★ Improve generalization ability, which will help to distinguish between uniform forest and oil palm plantation.
- ★ SE module is added at the end of each dense block.

Global average pooling
Dense fully connected, ReLu
Dense fully connected, Sigmoid (Squeeze connection)
Scale (Skip connection x Squeeze connection)

Detection Architecture

Layers	Output size	DenseNet-121	DenseNet-201	DenseNet-264
Convolution	112x112	7 x 7 Conv, Stride 2		
Pooling	56x56	3 x 3 Max Pool, Stride 2		
Dense Block (1)	56x56	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 6$
	56x56	Squeeze and Excitation Module		
Transition Layer (1)	56x56	1 x 1 Conv		
	28x28	2 x 2 Average Pool, Stride 2		
Dense Block (2)	28x28	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 12$
	28x28	Squeeze and Excitation Module		
Transition Layer (2)	28x28	1 x 1 Conv		
	14x14	2 x 2 Average Pool, Stride 2		
Dense Block (3)	14x14	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 64$
	14x14	Squeeze and Excitation Module		
Transition Layer (3)	14x14	1 x 1 Conv		
	7x7	2 x 2 Average Pool, Stride 2		
Dense Block (4)	7x7	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ Conv} \\ 3 \times 3 \text{ Conv} \end{bmatrix} \times 48$
	7x7	Squeeze and Excitation Module		
Classification Layer	1X1	Global average pooling		
		Binary dense fully connected, SoftMax		

Training Setup

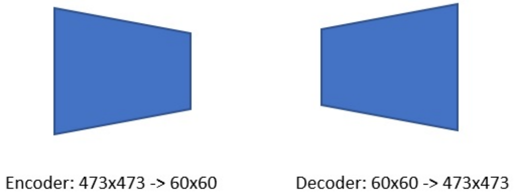
- ★ Loss function: categorical cross-entropy between classes.
- ★ Gradient optimizer: Adaptive Moment Estimation (Adam optimizer).
- ★ Platform: Python with Tensorflow GPU with Keras frontend(RTX 2080Ti).
- ★ Learning rate: 0.0001, no. of epoch=50, batch size=16.
- ★ Random initialization: No transfer learning.

Results

Architecture	Original (%)	DenseNet+SE (%)
DenseNet-121	96.194	97.179
DenseNet-201	96.784	98.294
DenseNet-264	93.830	97.835

Segmentation

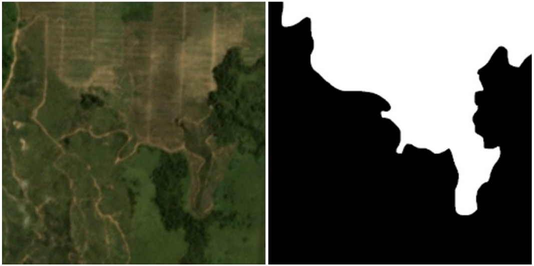
- ★ Only images that have been identified to consist oil palm plantation will be processed.
- ★ All segmentation networks follow the encoder-decoder configuration.
- ★ The goal is to identify either each pixel belong to the plantation or not.
- ★ All identified pixels will be analyzed to calculate the size of plantation areas.



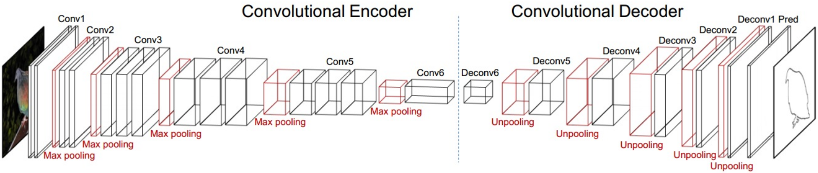
Methods

★ Four methods are explored:

1. FCN-8
2. SegNet
3. UNet
4. PSPNet



Example of CNN Encoder-Decoder



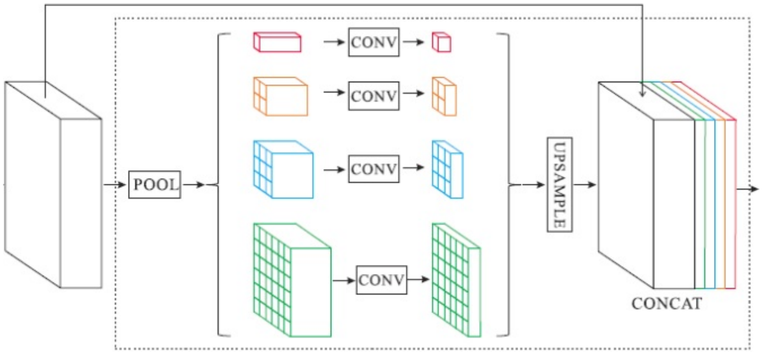
Method

- ★ FCN-8: 12 CNN layers for the encoder side and 5 CNN layers on the decoder side.
- ★ Utilizes max pooling for downsample and linear interpolation for upsample.
- ★ SegNet: 13 CNN layers for both encoder and decoder sides.
- ★ Utilizes similar downsample and upsample procedures as FCN-8 with the addition of fixed location for the upsample kernels.
- ★ Uses batch normalization and ReLu activation function.

Method

- ★ UNet: Uses the same architecture as SegNet but 11 CNN layers on both encoder and decoder sides.
- ★ Four skip connections are concatenated to the decoder side.
- ★ PSPNet: Have a complex decoder and simplified encoder.
- ★ Its encoder consists of ResNet-50 + Pyramid Pooling Module + 2 CNN layers
- ★ It does not have real encoder, the encoder is just bilinear interpolation to the original input size.

PSPNet



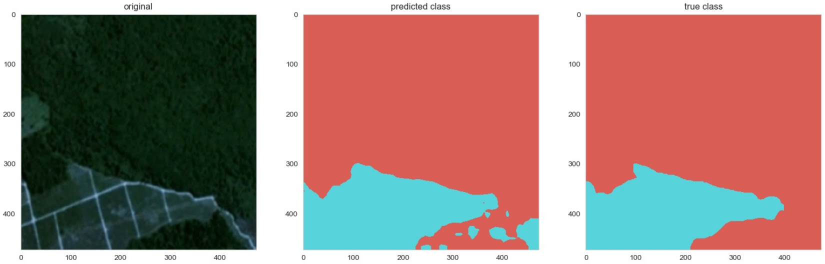
Training Setup

- ★ Loss function: categorical cross-entropy between classes (each class is represented by a channel).
- ★ Gradient optimizer: Stochastic Gradient Descent (SGD optimizer).
- ★ Platform: Python with Tensorflow GPU with Keras frontend(RTX 2080Ti).
- ★ Learning rate: 0.01, no. of epoch=300, batch size=4 (limited by GPU RAM).
- ★ ResNet-50 (PSPNet) and VGG-16 (FCN-8): Use pretrained model.
- ★ Other parameters: random initialization.

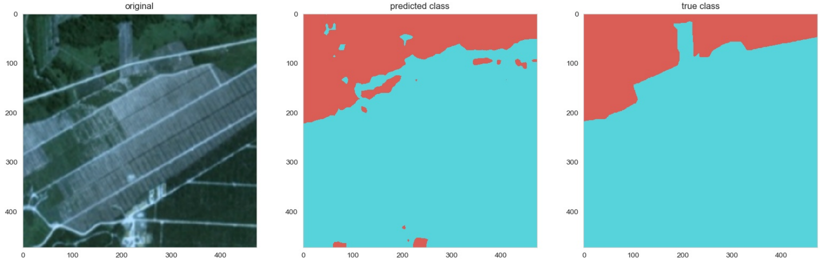
Results

Method	Accuracy (%)
FCN-8	68.495
SegNet	80.565
Unet	82.059
PSPNet	83.894

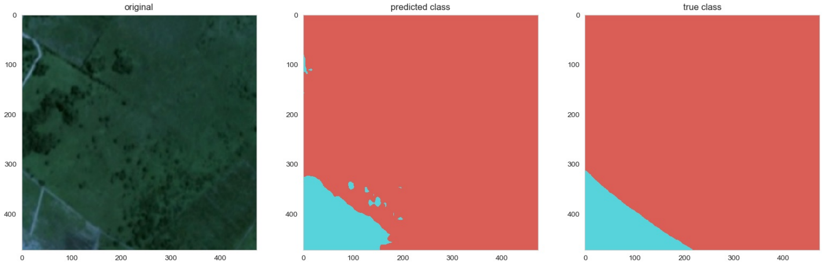
PSPNet



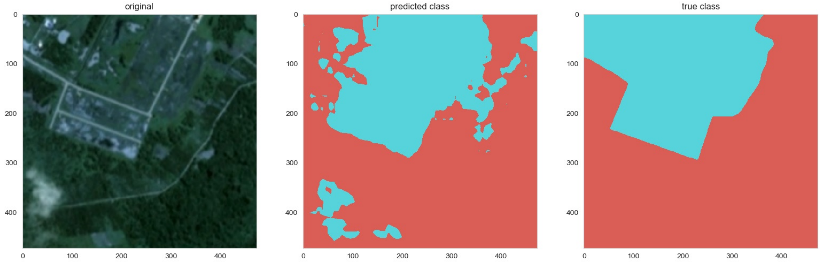
PSPNet


original
predicted class
true class
0
100
200
300
400
0
100
200
300
400

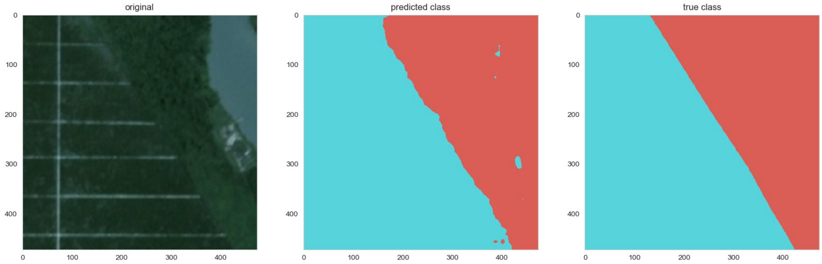
PSPNet



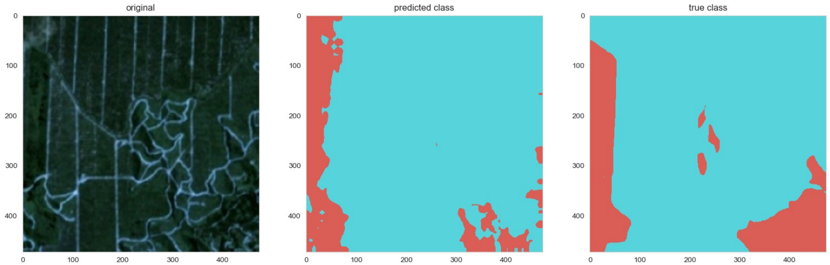
PSPNet



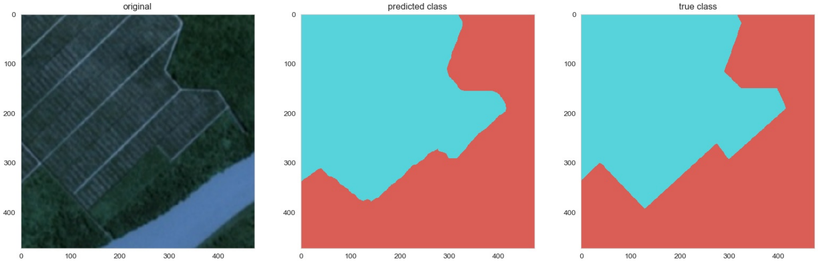
PSPNet



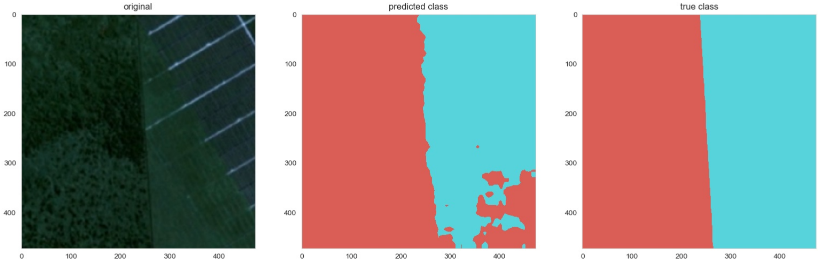
PSPNet



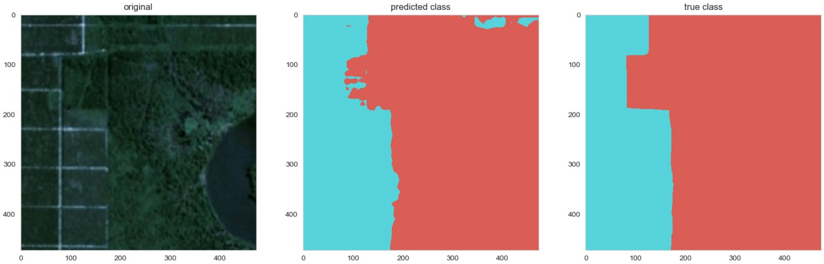
PSPNet



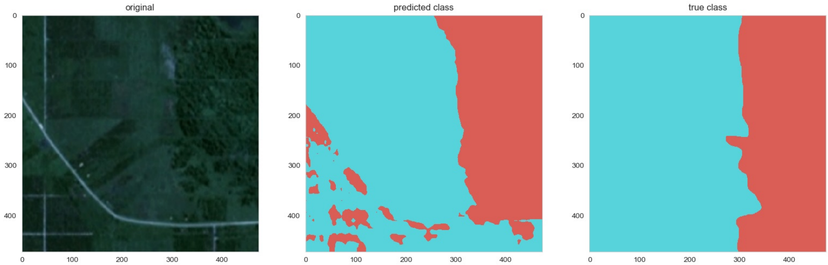
PSPNet



PSPNet


original
predicted class
true class
0
100
200
300
400
0
100
200
300
400

PSPNet

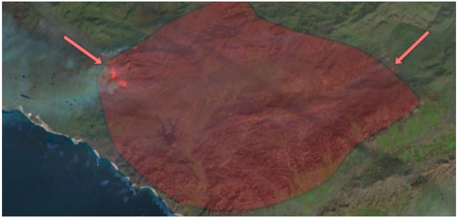


Results

- ★ PSPNet predicts that the plantation size is 10.426 KM^2.
- ★ The ground truth of the plantation size is 11.899 KM^2.

Conclusion

- ★ It is important for the government to monitor the growth of oil palm plantation.
- ★ Deep learning-based classifier produces a high detection accuracy for oil palm plantation detection.
- ★ Deep learning-based segmentation allows the monitoring of land usage for oil palm plantations.
- ★ Future work: deep learning-based hotspot detection.



THANK YOU
Question & Answer