The Eye in the Sky

Aaryaan Sharma, Anand Gokhale, Sumanth Hegde
IIT Madras

The Challenge

Using labelled data of only 14 satellite images, to implement a segmentation technique for 8 classes : Water, Swimming Pools, Trees, Grass, Buildings, Roads , Railways and Bare Soil.

Our Solution

- Followed a divide-and-conquer strategy.
- Domain specific solutions such as CCCI for certain classes.
- Mix of classical computer vision and deep learning.
- Hard mining for underrepresented classes

MOTIVATION

Motivation - No Deep learning

- Difficult to develop a robust deep learning solution that would not suffer from poor generalization.

- Approached the problem by first exploiting simple features in the satellite images using classical computer vision.

Motivation - No Deep learning ... Yet

- For classes such as *Buildings* and *Roads* - inter-class similarity and intra-class variance.
- Implemented individual segmentation models .

Why individual models ?

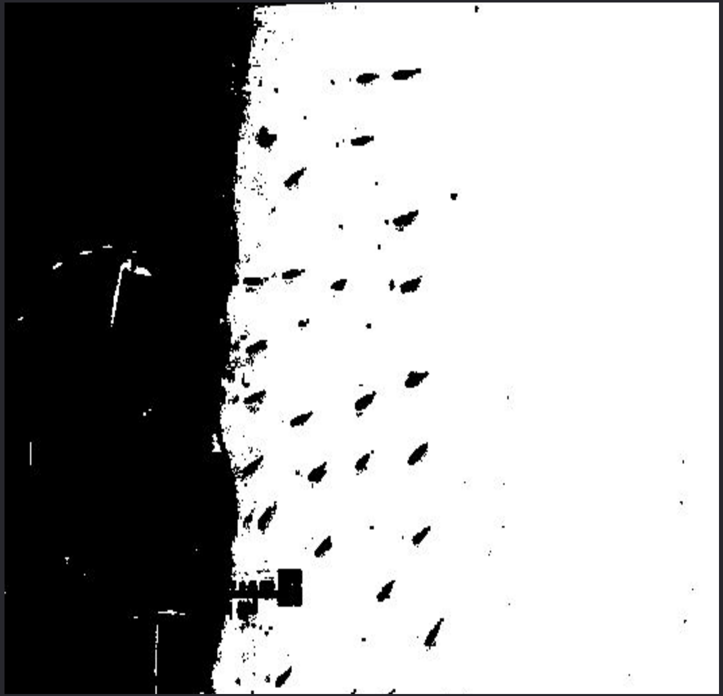
Simply because , joint classification approaches fail to provide reasonable accuracy and individual binary classification masks better exploit the unique features of each class.

Water Bodies - Water and Swimming Pools

- Water bodies absorb most of the NIR light that is incident upon them.

- Normalized from varying maxima and minima to a uniform scale for all images.(Why?)

- Thresholded the NIR band of the images to separate water bodies from rest of the image.



Swimming Pools

- Performed a pixel-wise binary AND operation on the NIR-based mask and a mask generated based on the typical colour of a swimming pool.
- Removed noise by removing pixels that have already been classified as other water bodies.
- Performed Morphological transforms to remove any remaining noise.

WATER

- Performed an intersection of the water mask and an RGB mask fine-tuned to minimise shadows.
- Eliminated noisy inner contours using morphological transformations
- To further eliminate noise, we then put a threshold on the contour areas of this new mask.

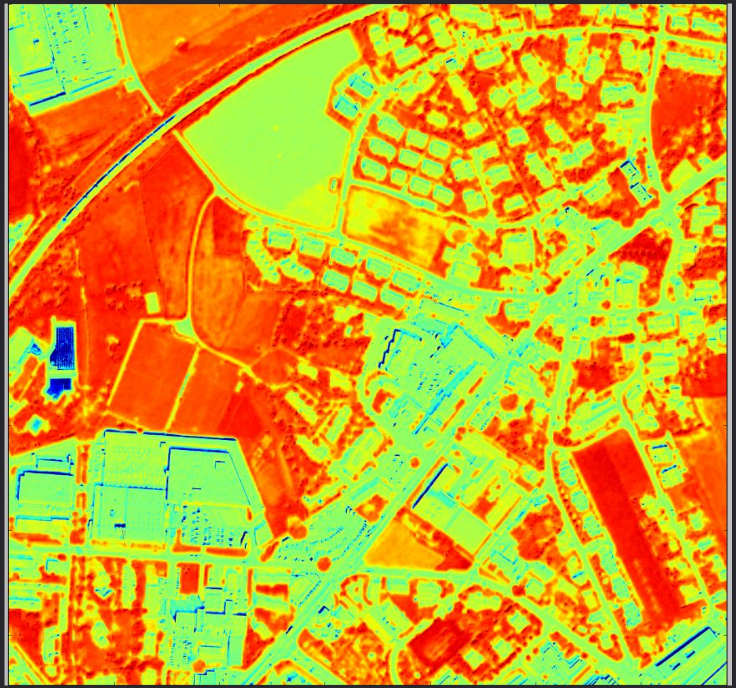
FORESTS AND GRASS - The CCCI Index

- Created a mask containing all trees and grass using the Canopy Chlorophyll Content Index(CCCI).

$$NDRE = \frac{NIR - RED}{NIR + RED}$$

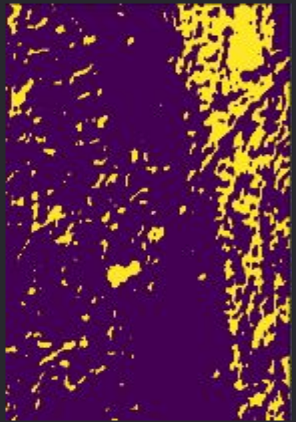
$$CCCI = \frac{NIR - NIR_{min}}{NIR_{max} - NIR_{min}}$$

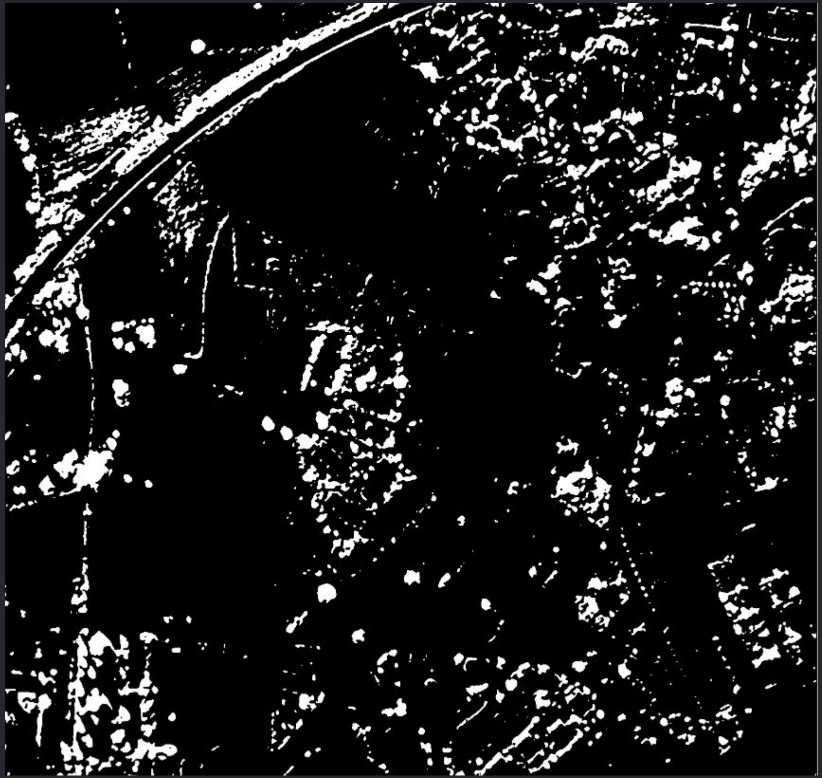
- Required mask is obtained by thresholding the image after the above transformations.

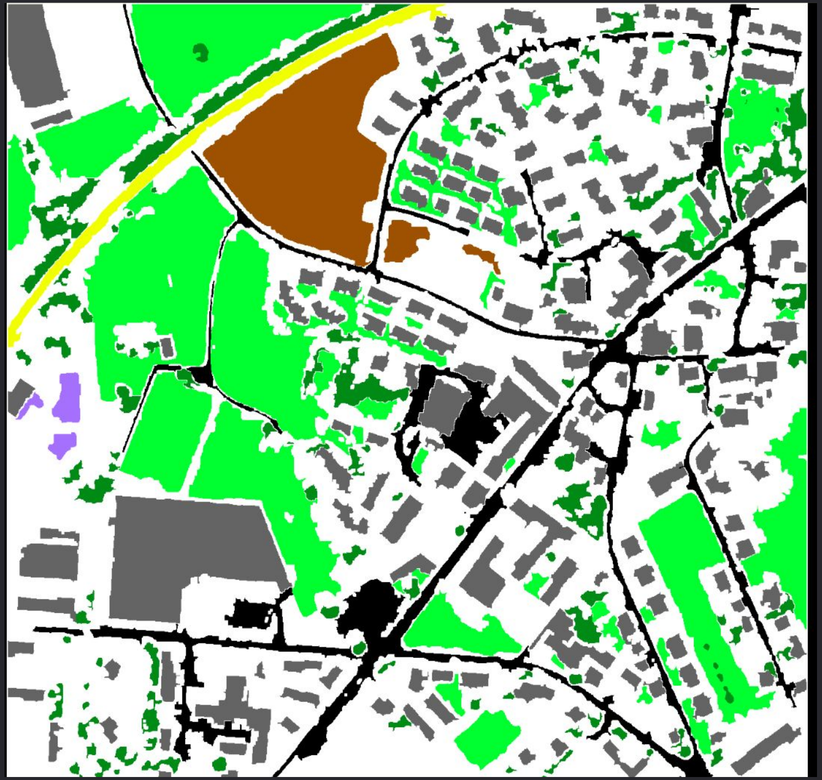


Separating Grass from Trees

- Shifted to the HSV Colorspace and then applied a bound based on the value parameter to separate grass from trees. obtained in the previous CCCI-based mask.
- We report marginally lower accuracy because of noisy ground truth.





BUILDINGS, ROADS, RAILWAYS AND BARE SOIL : THE GENERAL STRATEGY

PREPROCESSING

- Histogram-equalised for contrast enhancement.
- Broken down into patches of size 64x64.
- To incorporate boundary effects, added padding with reflections.
- Used data augmentation, employing rotations and shifts, incorporating rotational and translational invariance in the neural net.

MODEL ARCHITECTURE

- *U-net* - the standard architecture for semantic segmentation constrained by the size of the dataset.
- Our motivation - Deeper architectures such as *Deep U-net* have obtained better results than the U-net architecture, but are based on further downsampling of the input image.
- Used two variants of the *U-net* : *U-net(a)* and *U-net(b)*
- *U-net(a)* has more feature channels at each block when compared to the U-net.
- *U-net(b)* is similar to *U-net(a)*, but with more convolutional layers at each block.

HOW DO WE PREVENT OVERFITTING ??

- Batch normalization after every convolutional layer in our networks and kernel regularization as an additional measure to prevent overfitting.

- Batch normalization, along with a regularization effect on our model, makes the model more stable during training.

- Batch normalization replaces dropout regularization, which is rather ineffective in *FCNs* ,where spatial relationships in the feature maps make the activations highly correlated.

ROADS

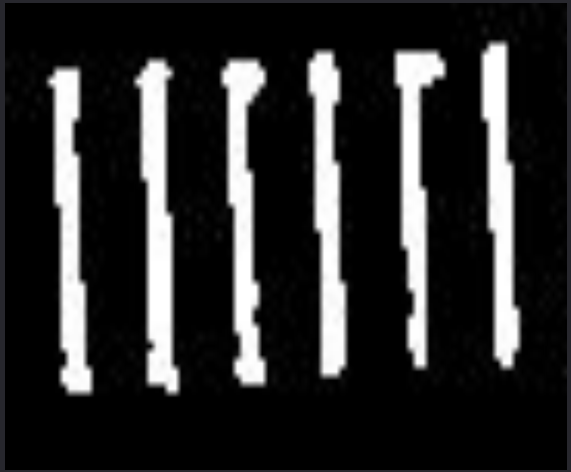
- The inputs of the size 64x64x4 were fed to a model of *U-net(b)* architecture.
- Used 4128 patches for training and 538 of validation.
- The loss function used was binary cross entropy loss.

BUILDINGS

- The inputs of size 64x64x4 - fed to a model based on the *U-net(b)* architecture.

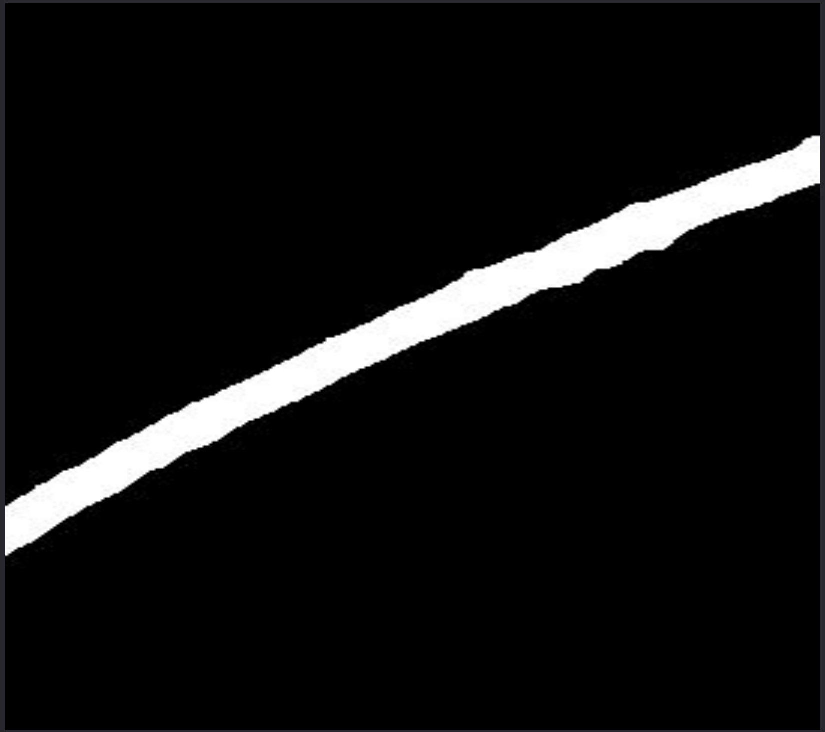
- Our train and validation split : 4128 patches for training and 538 for validation.

- Used a weighted binary cross entropy loss with the weight for the positive targets term to be 0.6.



RAILWAYS - THE HARD MINING APPROACH

- Huge class imbalance when it comes to the railways class.
- Model at a high risk of overfitting and learning a black mask for any input!
- Hard mining to tackle this problem : 50:50 ratio of “positive” and “negative” examples.
- *U-net(b)* architecture based on binary cross entropy loss.

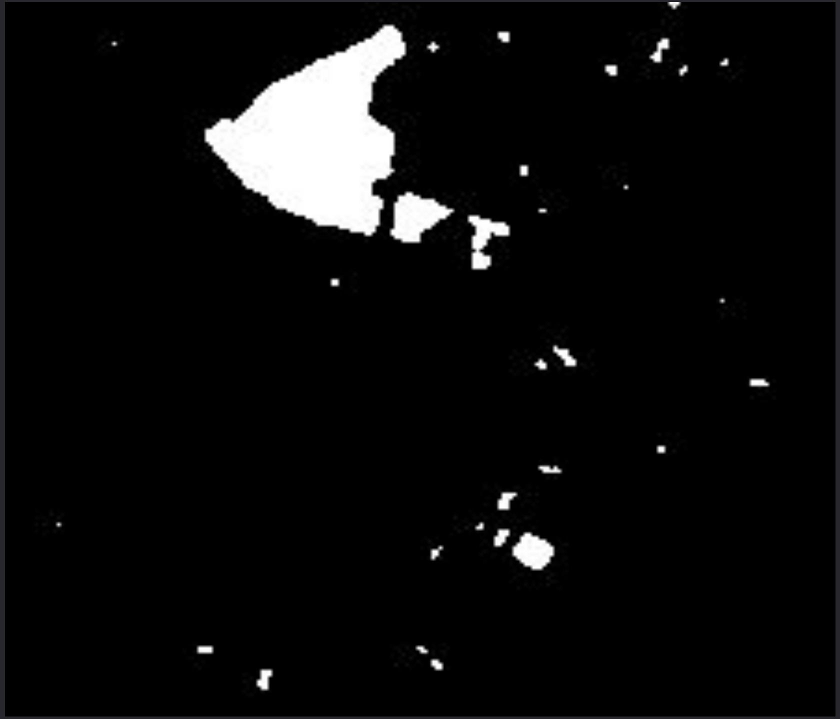


BARE SOIL

- With similar class imbalance with this class - use the same hard mining approach.

- Positive examples - patches containing at-least 15 pixels of this class.

- *U-net(a)* architecture - to have sufficient model complexity as well to prevent overfitting and binary cross entropy loss.



THE FINAL STEP!

- Different methods of blurring and thresholding specifically tuned for each class-wise mask.
- Combined all the 8 masks generated in the increasing order confidence to further eliminate errors.

References

CCCI:

Cammarano, Davide & Fitzgerald, Glenn & Basso, Bruno & O'Leary, Garry & Chen, Deli & Grace, Peter & Costanza, Fiorentino. (2011). Use of the Canopy Chlorophyl Content Index (CCCI) for Remote Estimation of Wheat Nitrogen Content in Rainfed Environments. Agronomy Journal. 103. 1597-1603. 10.2134/agronj2011.0124.

U-net :

O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv e-prints (2015) arXiv:1505.04

Deep U-net:

R. Li, W. Liu, L. Yang, S. Sun, W. Hu, F. Zhang, W. Li, DeepUNet : A Deep Fully Convolutional Network for Pixel-level Sea-Land Segmentation, arXiv e-prints (2017) arXiv:1709.0020

THANK YOU!

ANY QUESTIONS?