Identifying Nutrient Deficiency in Satellite Farmland Imagery

By Steven Tran

Photo link

Problem Statement

Considering global population growth and failure to adequately mitigate the disastrous effects of global warming, one area of civilised society we must be be vigilant over is our food production.

Producers in the agriculture industry all over the world (especially in less developed industries) need all the help they can get to protect agricultural yields to ensure a stable supply of food for the populace. Maximizing crop yields and minimizing losses requires a solution for staying informed about crop and soil health.

Proposed Solution & Task

Use labelled satellite imagery and computer vision to identify patterns of ailments which negatively impact field conditions and crop yields.

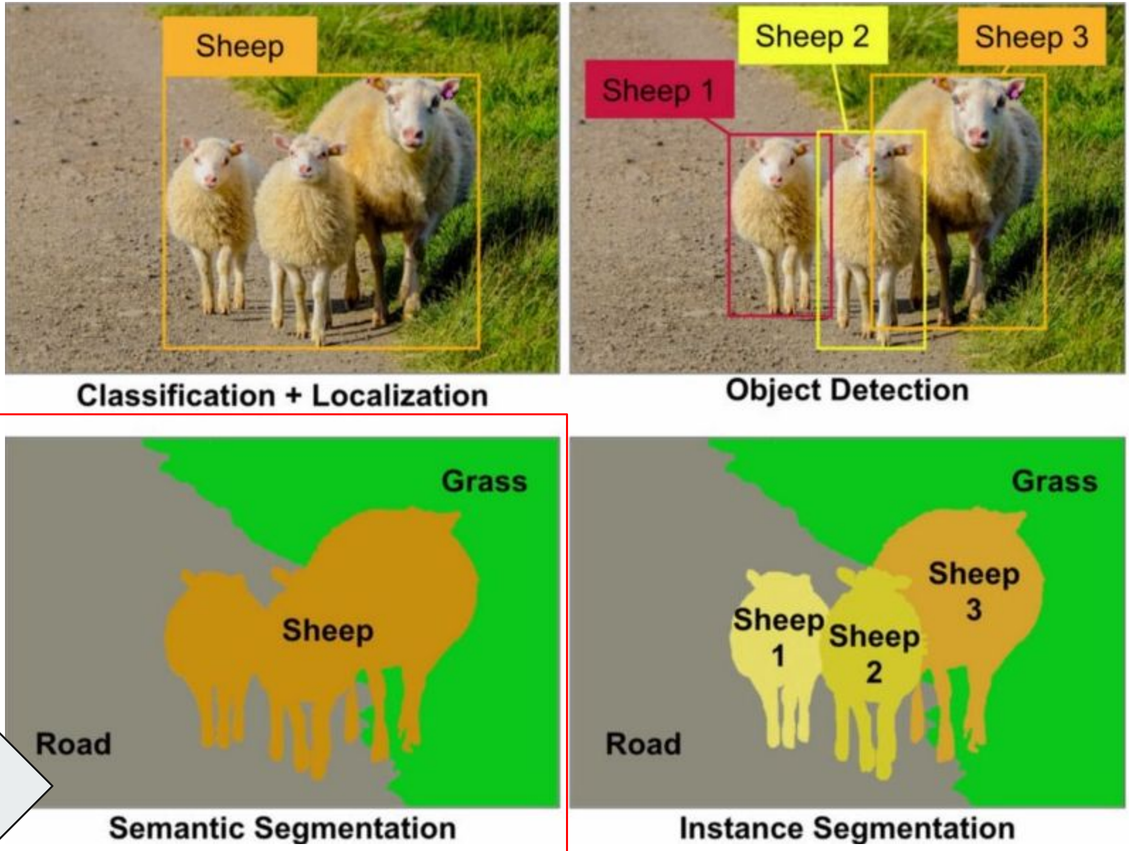
For this project, I attempted to identify **nutrient deficiency** in satellite photography of croplands.

Proposed Solution & Task

There are a variety of defined tasks in the computer vision domain.

This task is called **semantic segmentation.**

https://miro.medium.com/max/1400/1*rgliupBanbeMYW7xXYH5Lw.jpeg



I'm focusing on this one!

AgricultureVision was accessed on December 10, 2021 from https://registry.opendata.aws/intelinair_agriculture_vision.

The Data

- The AgricultureVision dataset published in 2020 consists of "**94,986** images sampled from 3,432 farmlands"
- The images were captured from 2017 through 2019
- The cameras used captured 4-channel field images consisting of **Near-Infrared, Red, Green,** and **Blue** color channels
- The publishers had five expert agronomists use software to **annotate image regions** with different crop ailments

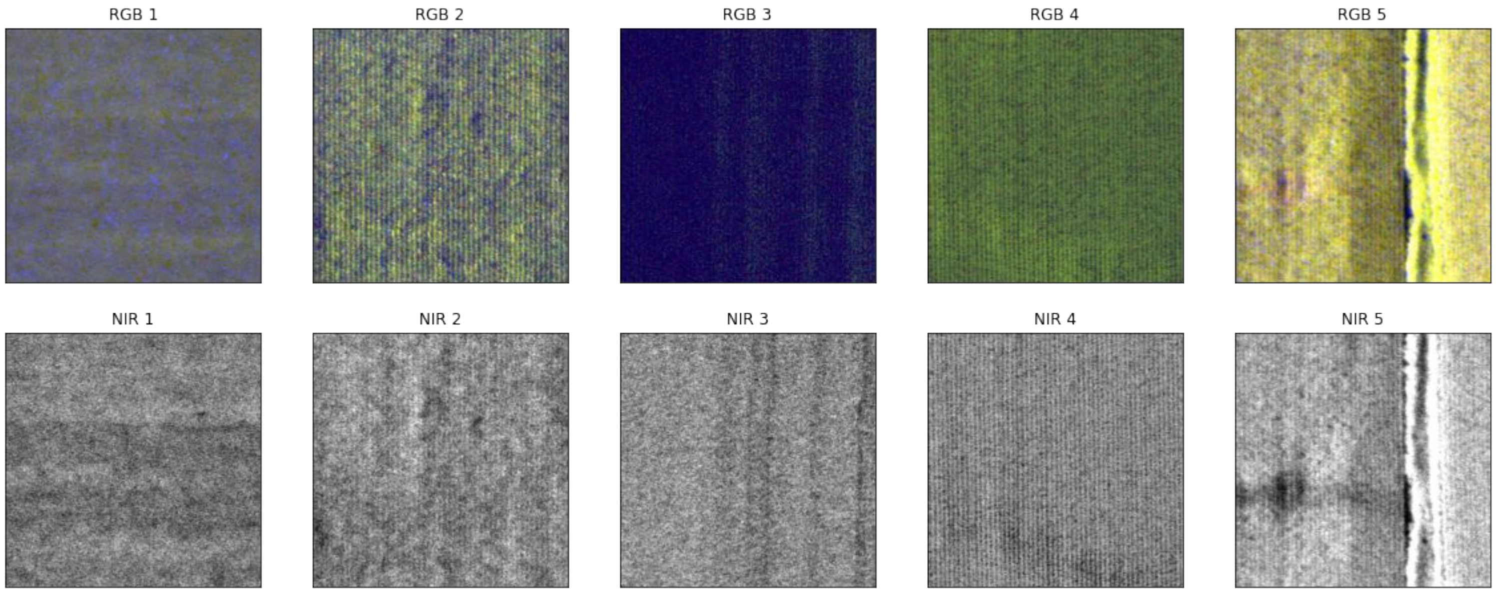
AgricultureVision was accessed on December 10, 2021 from https://registry.opendata.aws/intelinair_agriculture_vision.

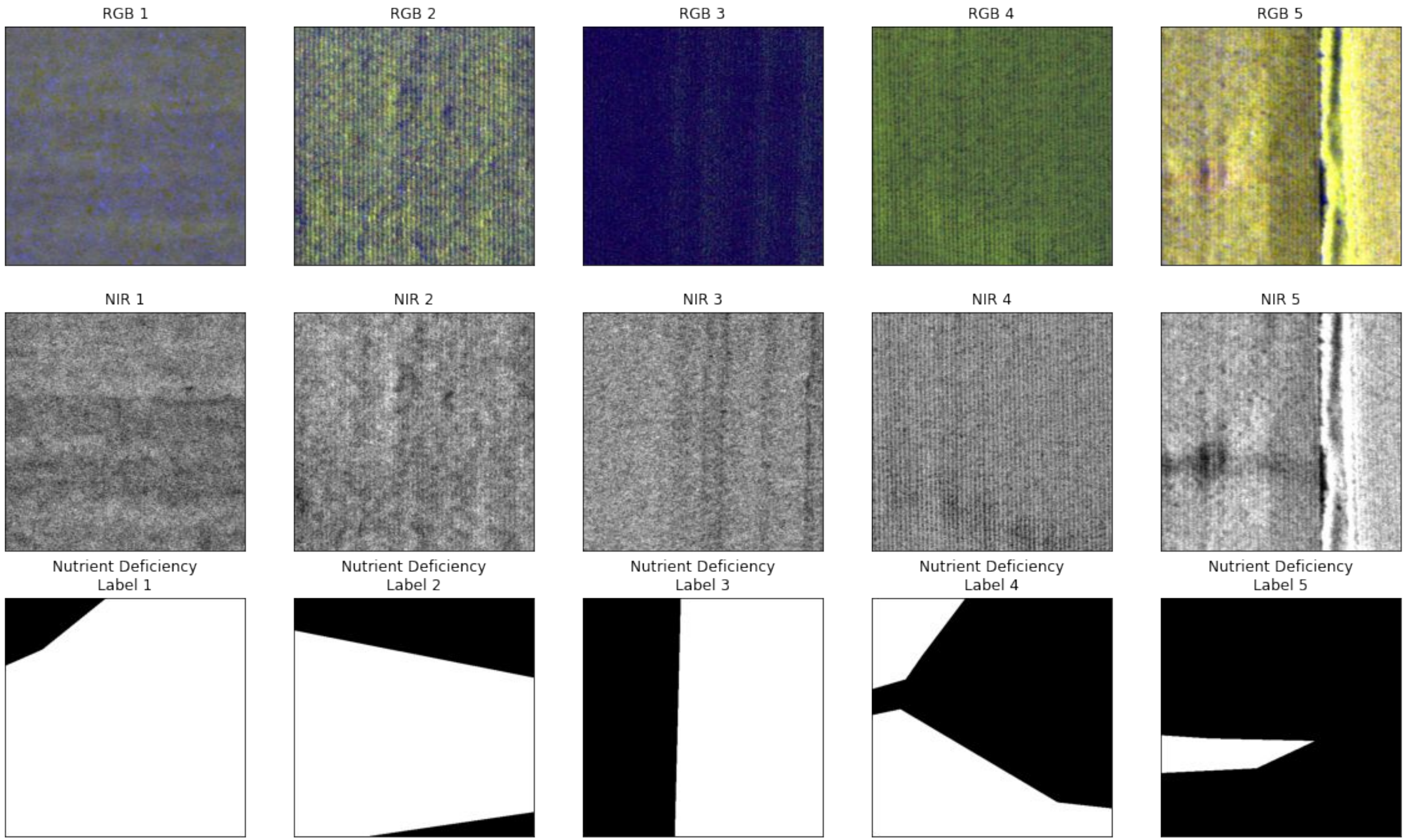
The Data

- For this analysis, I limited the dataset to only the 2018 photos which included **nutrient deficiency** annotations

- In total, I used 14,712 images (totaling ~3GB of disk space) for training (90%) and validation (10%)

Let's take a look at a few examples!

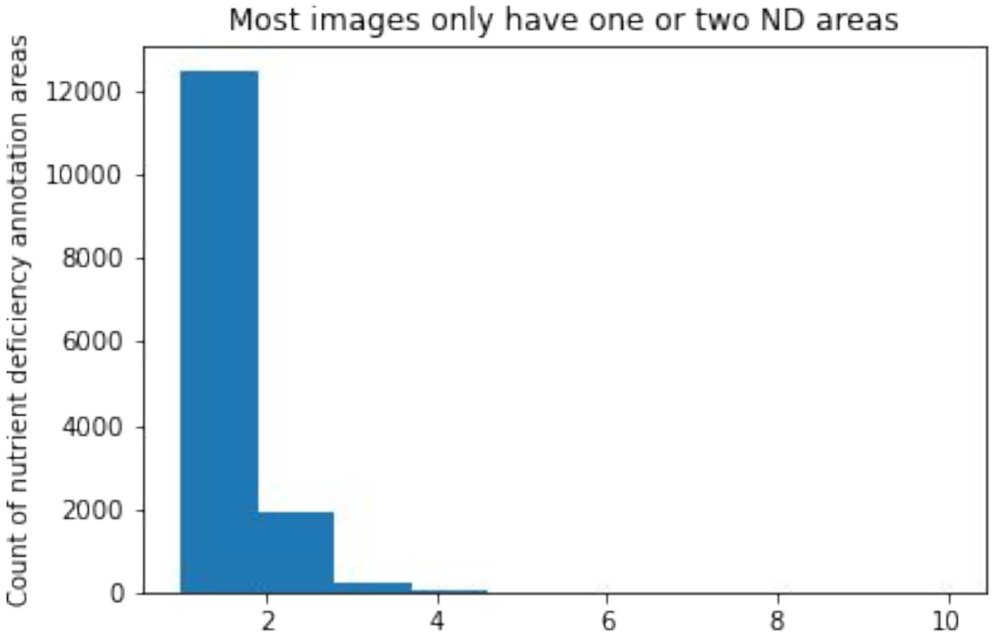

RGB 1
RGB 2
RGB 3
RGB 4
RGB 5
NIR 1
NIR 2
NIR 3
NIR 4
NIR 5


RGB 1
RGB 2
RGB 3
RGB 4
RGB 5
NIR 1
NIR 2
NIR 3
NIR 4
NIR 5
Nutrient Deficiency Label 1
Nutrient Deficiency Label 2
Nutrient Deficiency Label 3
Nutrient Deficiency Label 4
Nutrient Deficiency Label 5

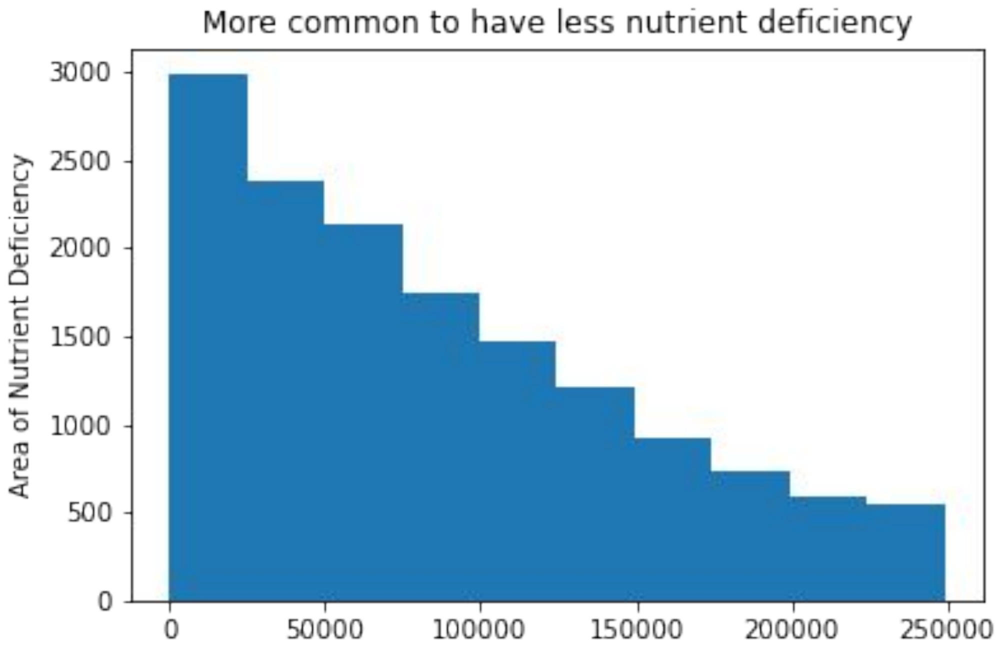
Limited EDA for a Binary Segmentation Task

- Each of the images in my dataset measure 512 pixels in width and height.
- Each image has an RGB image as well as a one-channel Near-Infrared image.

EDA: Insights from the Annotation metadata



EDA: Insights from the Annotation metadata



EDA: Clustering

Review of the images revealed there's significant variability in imaging characteristics.

- Some images are darker, some are brighter
- Some have lots of contrast, some less
- Other than getting experts to label the crop ailment areas, the publishers did not attempt to categorize the images by image characteristics like brightness or contrast

- Can **Clustering** and **Principal Component Analysis** assist us with understanding the data a little better?

EDA: Clustering

- Used a pre-trained VGG16 (a convolutional neural network) to extract each image's RGB data to 224x224 arrays

- Applied **PCA** demonstrates that the first 100 components explain about **84%** of the variance in the image features

- Using the first 100 components along with the image metadata (count of ND, ND area, etc.), employed **K-Means** clustering to attempt to group photographs into similar clusters

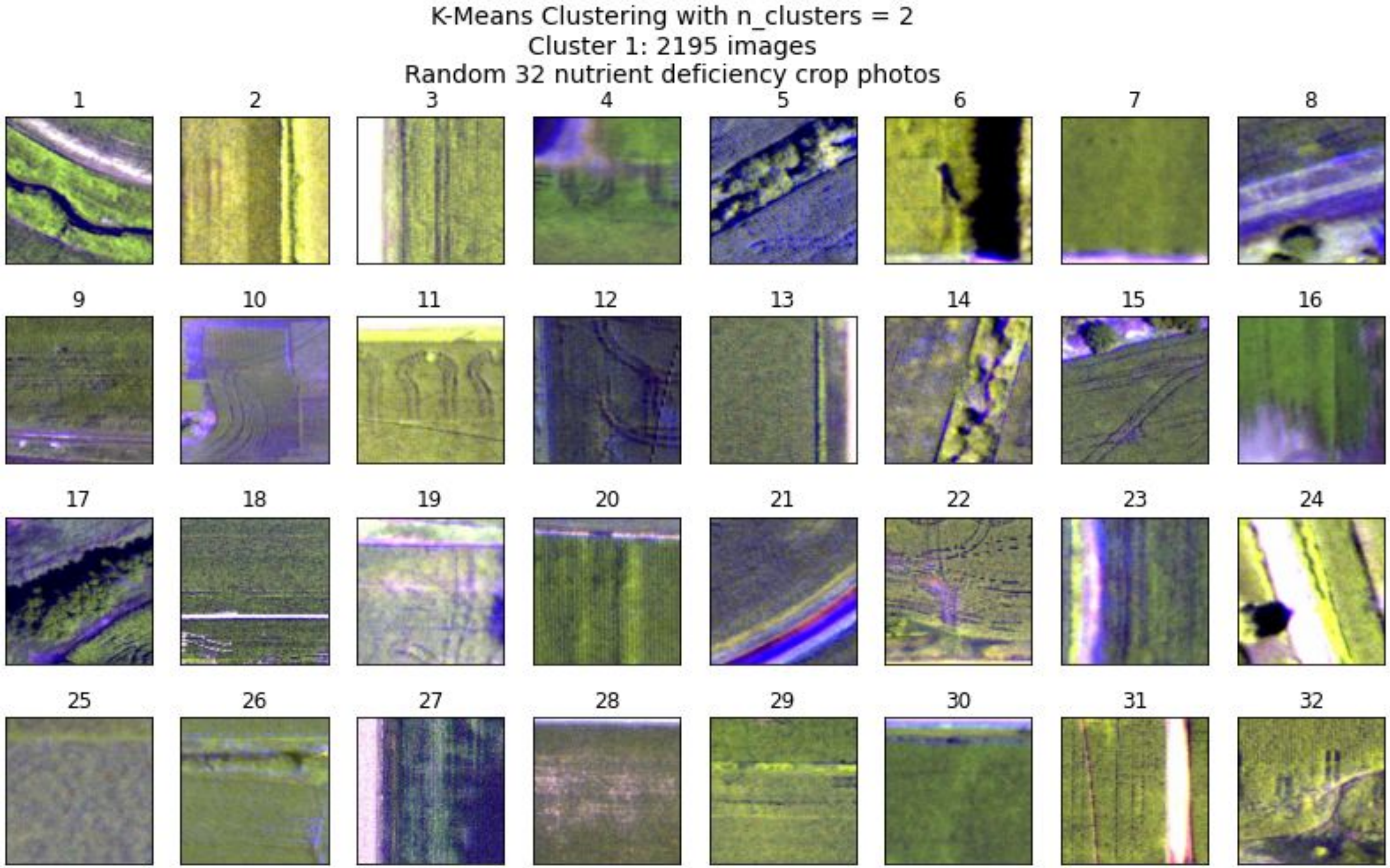
EDA: Clustering

- My initial intuition was that there might be a natural clustering of images by brightness or contrast, resulting in between 3 and 5 clusters.

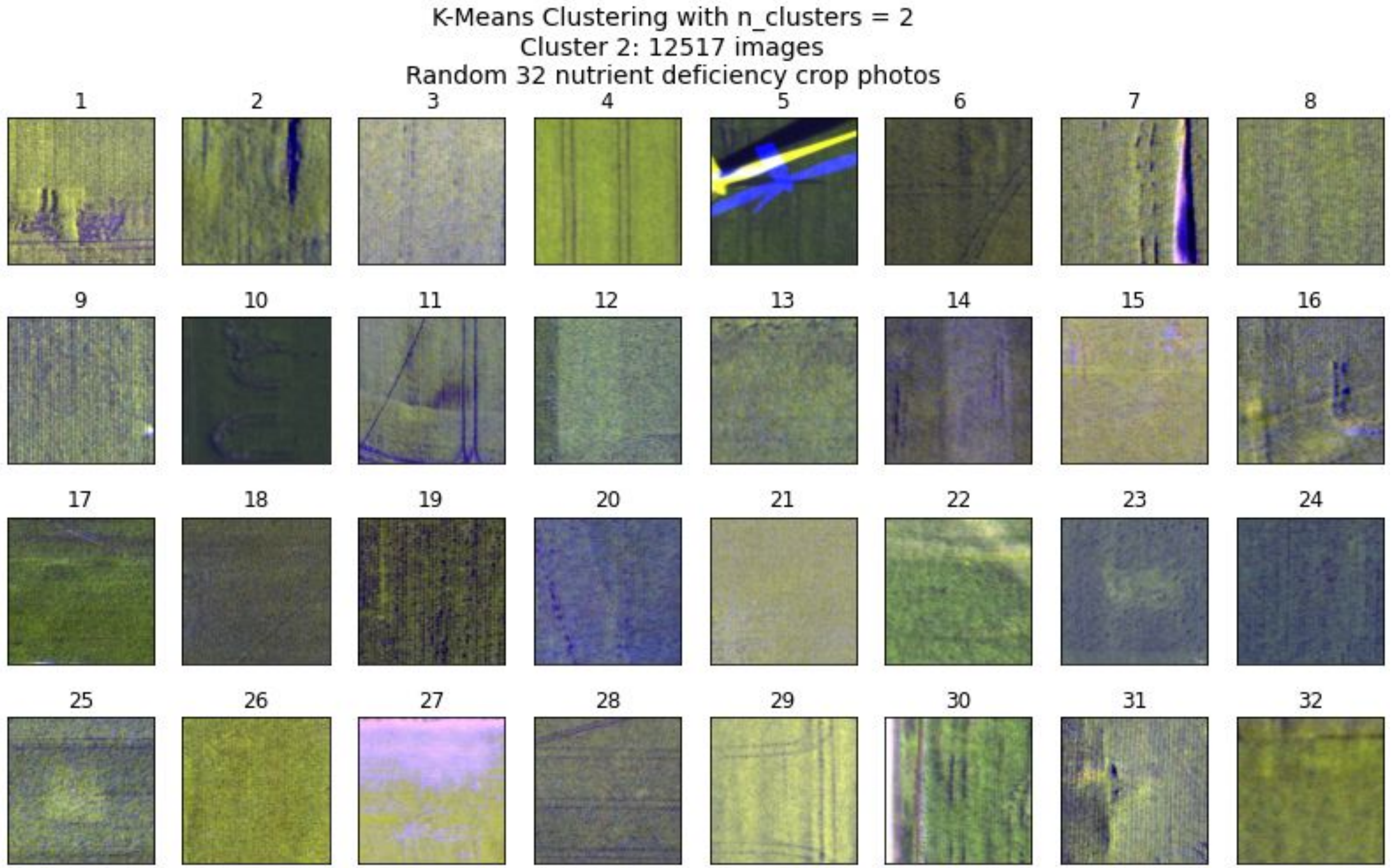
- Unfortunately, the **K-Means** clustering indicated that a clustering of **k=2** achieved the best **silhouette score**
- The Silhouette Score is a metric which describes how well a clustering schema groups observations (images in this case)
- The score is calculated using the average **intra**-cluster difference between points as well as the average **inter**-cluster distance between a point and points in all other clusters

k	s.score
2	0.174018
3	0.023372
4	-0.045856
5	-0.023108
6	-0.047234
7	-0.040236
8	-0.041729
9	-0.043397
10	-0.044324
11	-0.045785

EDA: Clustering with K=2



EDA: Clustering with K=2

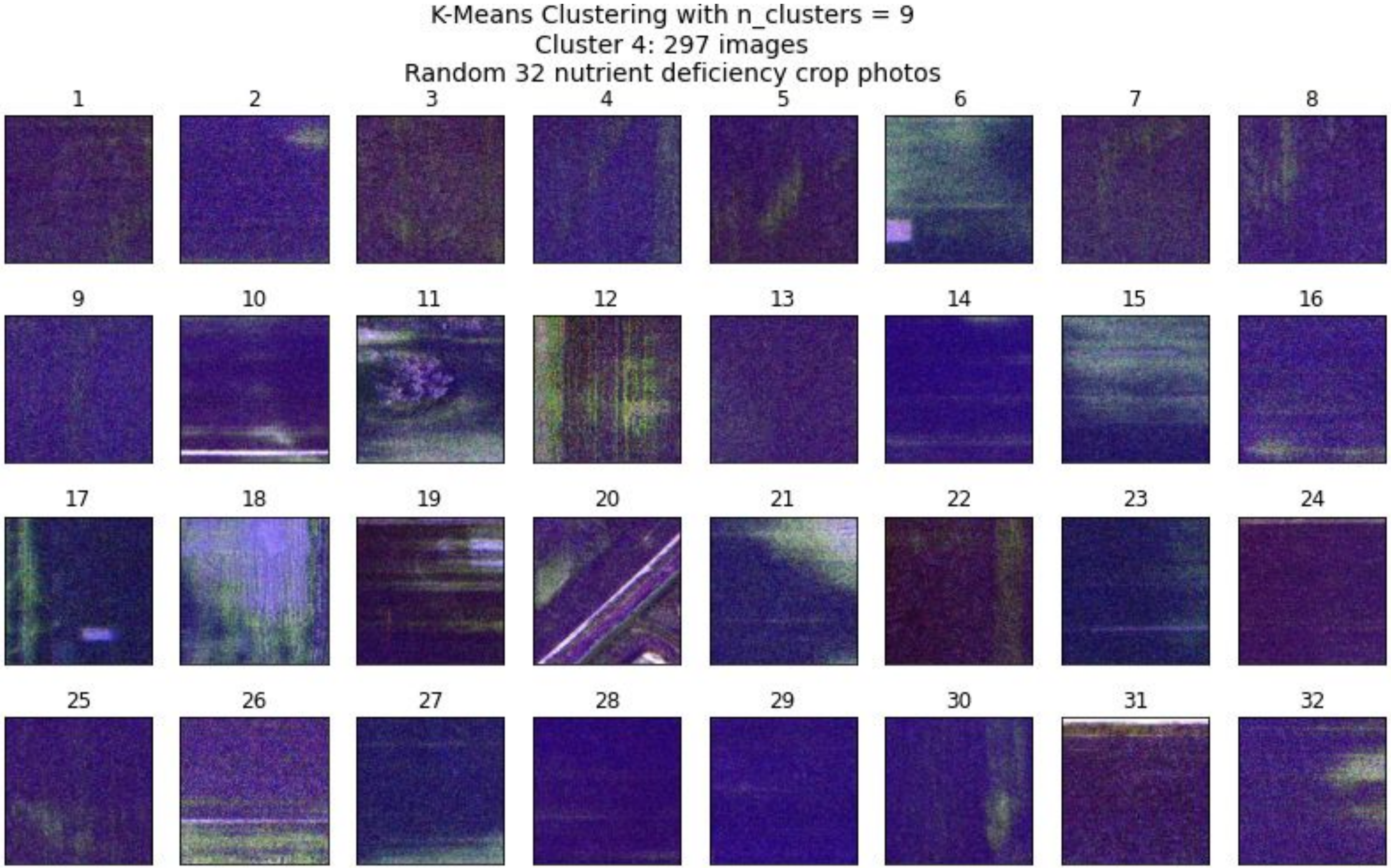


EDA: Conclusion on Clustering with K=2

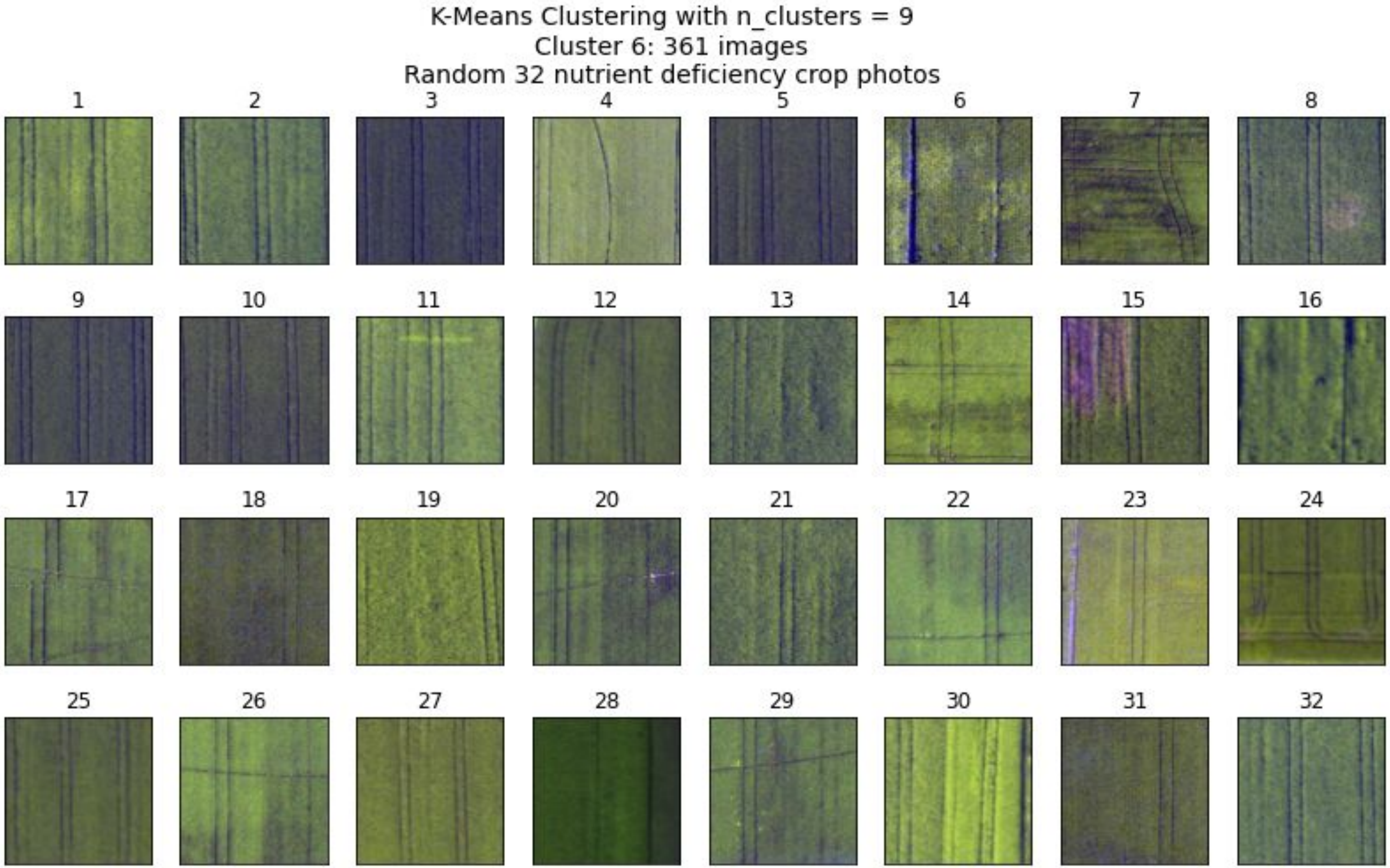
To my human and fallible eye, it looks like the images grouped into Cluster 1 had a greater range in color, accompanied by various land features like equipment tracks, waterways, patches of sand, etc.

By contrast, it seems that images in cluster 2 have reduced range in color.

Explore Another Day? Clustering with K=9



Explore Another Day? Clustering with K=9



With data in hand, a model must be chosen. . .

Prime candidates

In my research into available semantic segmentation models the two model architectures below seemed to be the best suited:

- **U-Net**
  - Great for semantic segmentation
  - Computationally inexpensive (by comparison)
  - Labelling requires region annotations only
- **Mask R-CNN**
  - Great for semantic and instance segmentation
  - Relatively simple to train (apparently)
  - Requires ground-truth bounding boxes

Prime candidates

In my research into available semantic segmentation models the two model architectures below seemed to be the best suited:

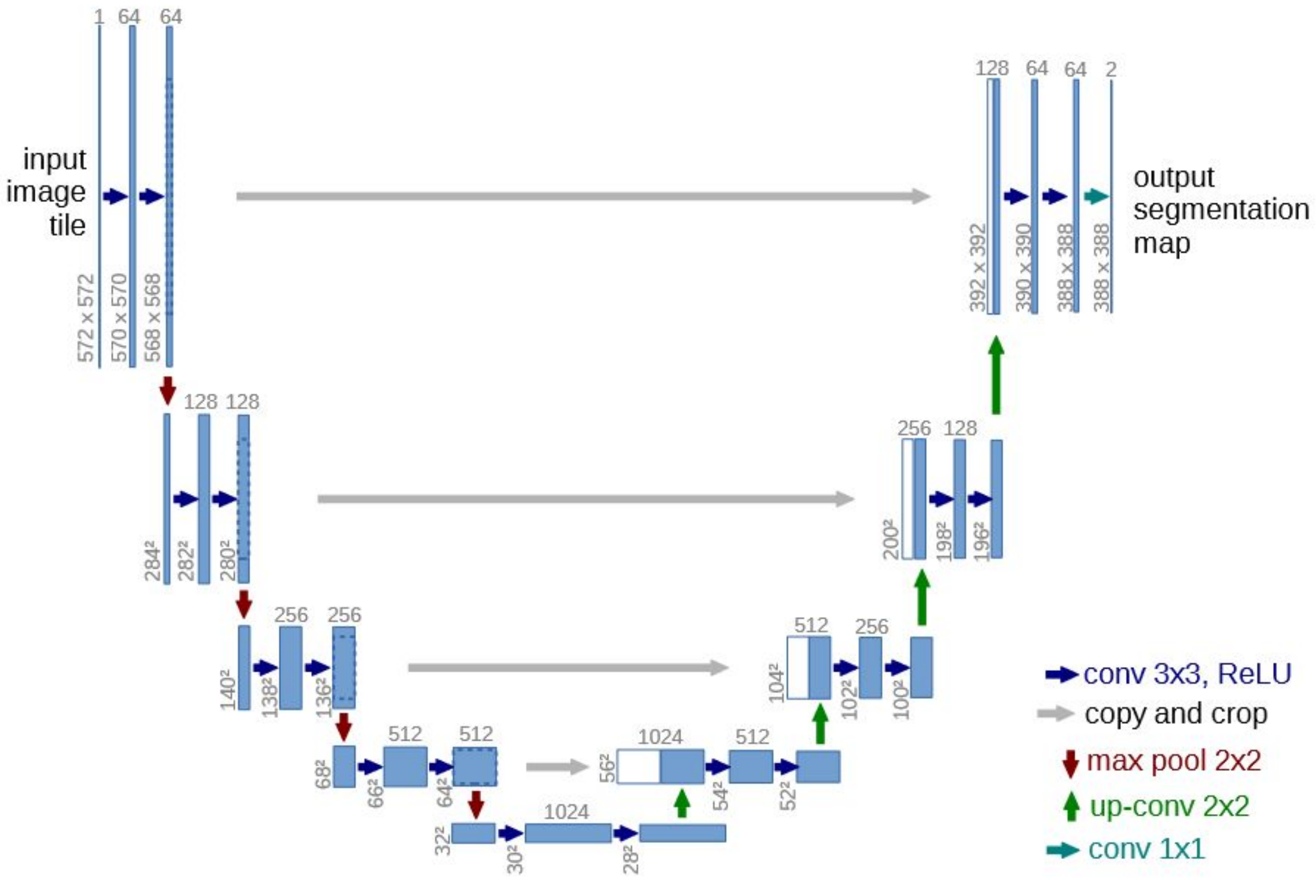
- U-Net - **For this project, I chose U-Net!**
  - Great for semantic segmentation
  - Computationally inexpensive (by comparison)
  - Labelling requires region annotations only
- Mask R-CNN
  - Great for semantic and instance segmentation
  - Relatively simple to train (apparently)
  - Requires ground-truth bounding boxes

About the **U-Net** architecture:

- Created in **2015** by Olaf Ronneberger, Philipp Fischer, and Thomas Brox from the **University of Freiburg** in Germany
- Originally designed for Biomedical Image Segmentation

Let’s visualize the model architecture.

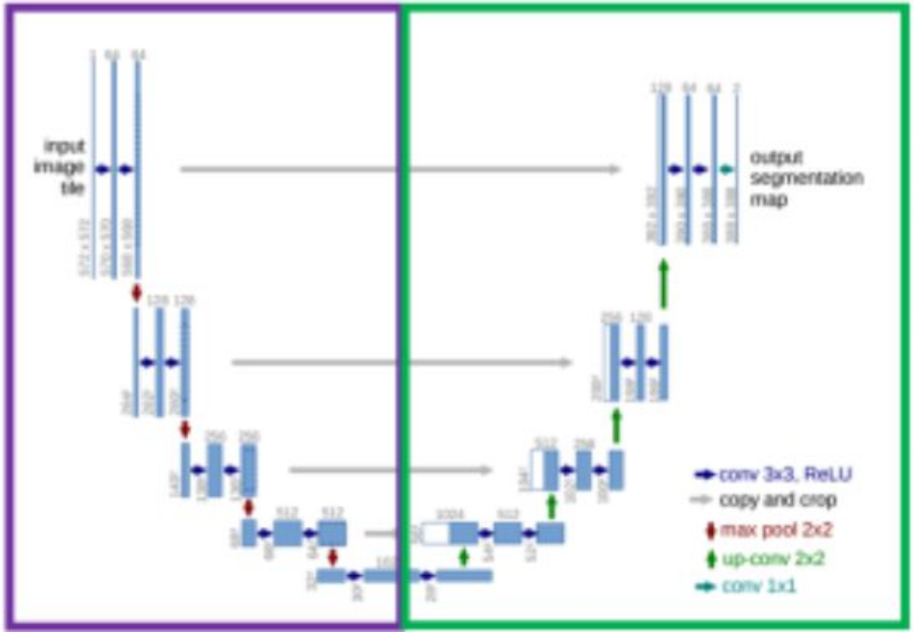
From Ronneberger et. al. (2015)



From Ronneberger et. al. (2015), annotations by Rachel Draelos

“contracting path”

- typical convolutional neural network
- 3x3 conv, ReLU
- 2x2 max pool for down-sampling
- At each down-sampling, double number of feature channels



“expansive path”

- Up-sampling of feature map
- 2x2 conv that halves # of feature channels
- concatenation with corresponding feature map from contracting path
- 3x3 conv, ReLU

Modeling goals

- Define a *baseline* ‘predictive’ model for comparison purposes
  - I’m going with a ‘model’ that predicts every pixel as belonging to the ND class
- Train a U-Net segmentation model which can on average, produce better predictions than the baseline model
- Explore whether using the fourth imaging channel, **Near-Infrared** can be helpful for the segmentation task

Modeling: Quick Note about Loss Functions

In lay terms, a **Loss Function** is how data scientists specify how models should optimize their learning.

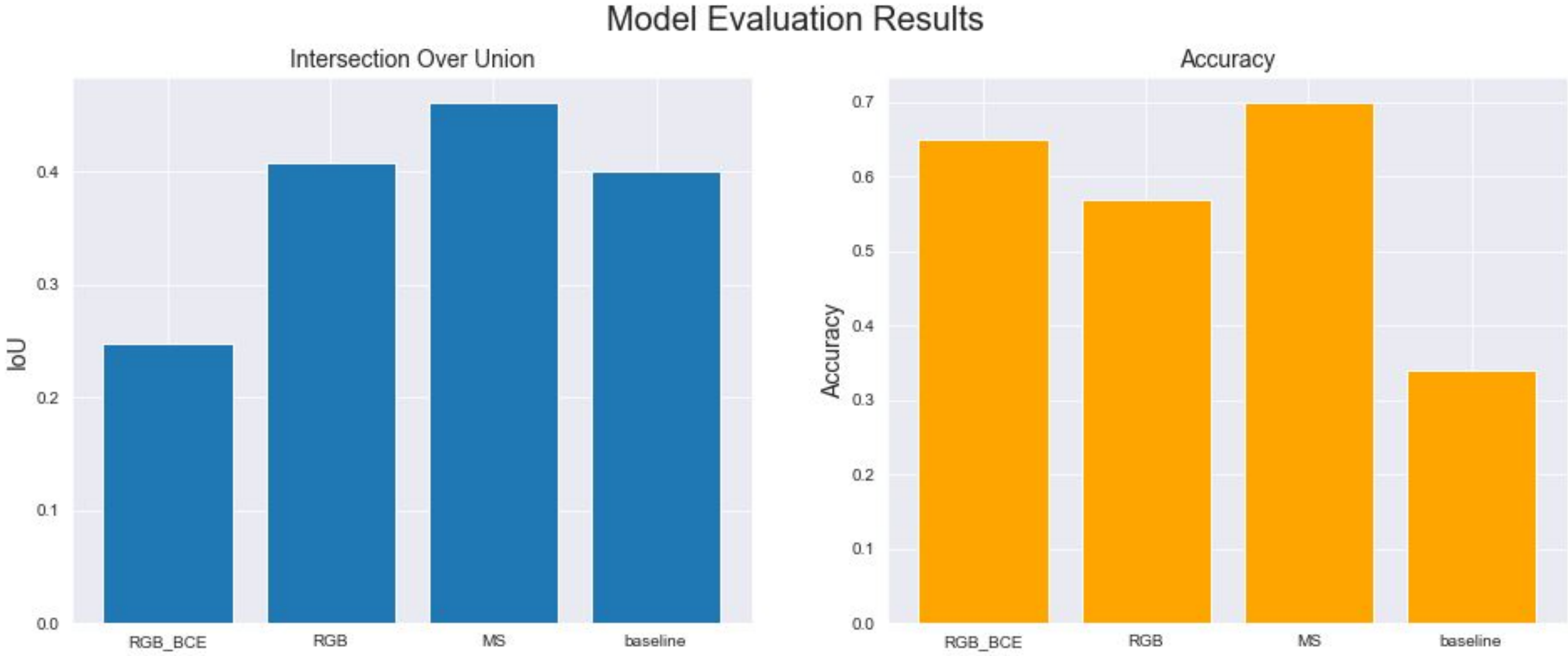
It's how the model penalizes bad predictions and rewards good predictions, acts which help it *learn* and make better predictions.

For this segmentation task, I'll be using **Intersection Over Union** or **Jaccard Similarity** (also referred to as Jaccard Distance) rather than the **Binary Crossentropy** typically used in binary classification problems.

Modeling process

- Recapping dataset statistics:
    - 14,712 images with Nutrient Deficiency annotations
    - 512x512 RGB, Near-Infrared channels
    - Training on 90% of the images, validation performed on remaining 10%
- Train the **U-Net** model
    - Once with RGB only, using BCE loss metric
    - Once with RGB only, using IoU loss metric
    - Once with RGB + Near-Infrared channels, using IoU loss metric
- Produce and evaluate ND predictions for the validation image set

Modeling Results



Modeling Results

The results demonstrate that

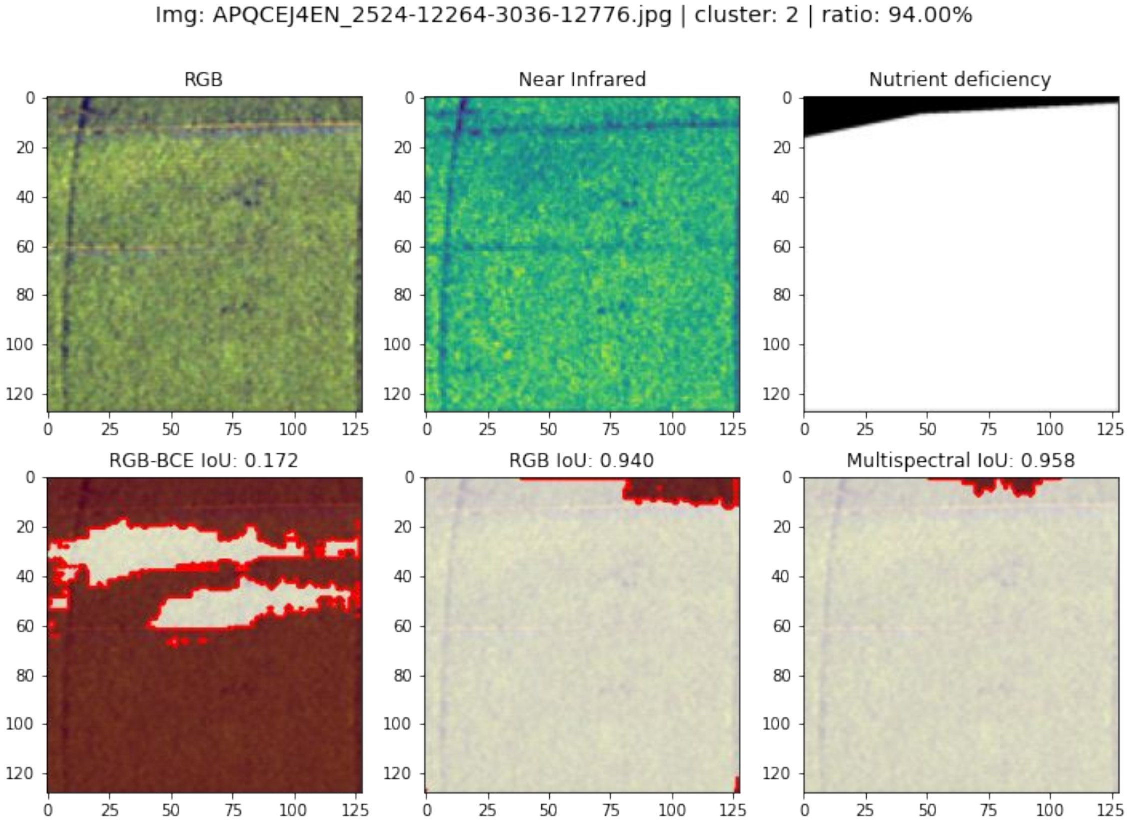
- Not only does the multispectral-input U-Net outperform the baseline in mean **Intersection over Union**,
- it outperforms in mean validation accuracy compared to the baseline and the RGB model which was trained on minimizing the binary crossentropy

- Compared to the RGB-only U-Net IoU model, the results demonstrate that the added information of having the near-infrared 'color' channel generally aides the model in identifying areas with Nutrient Deficiency

Review of Model Predictions

The Best

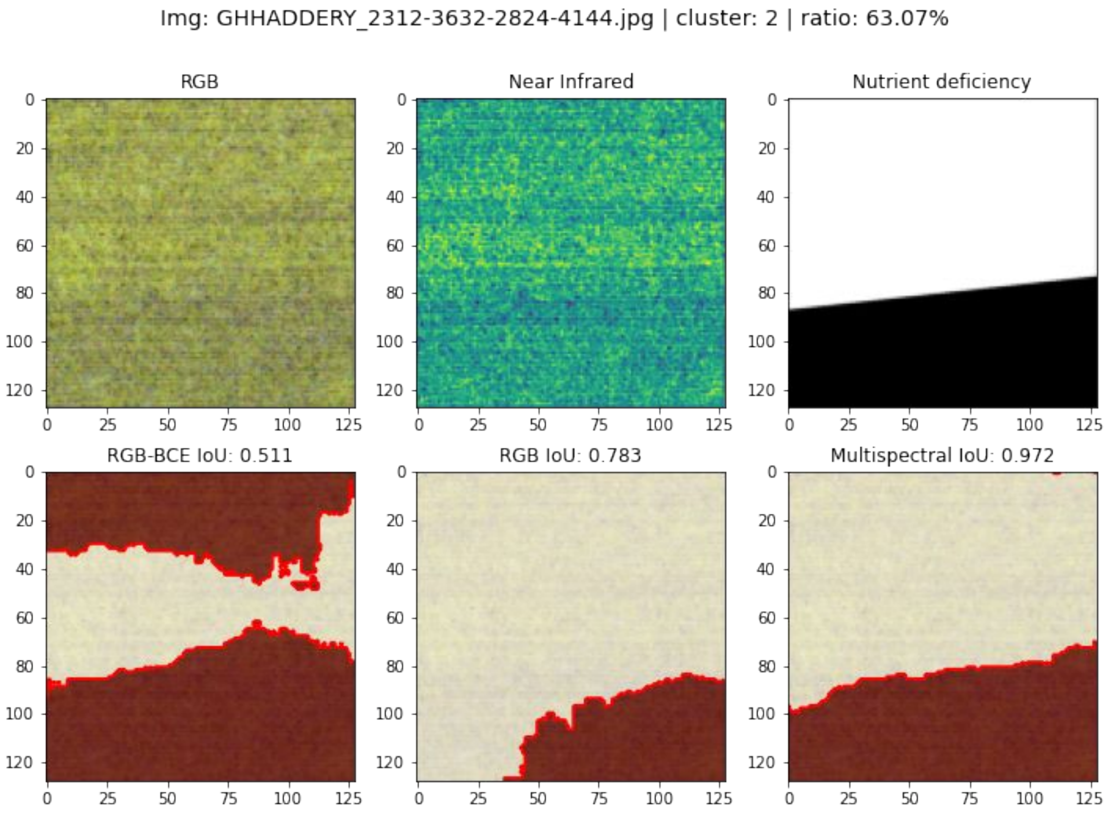
The Best (or maybe the luckiest?)



Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

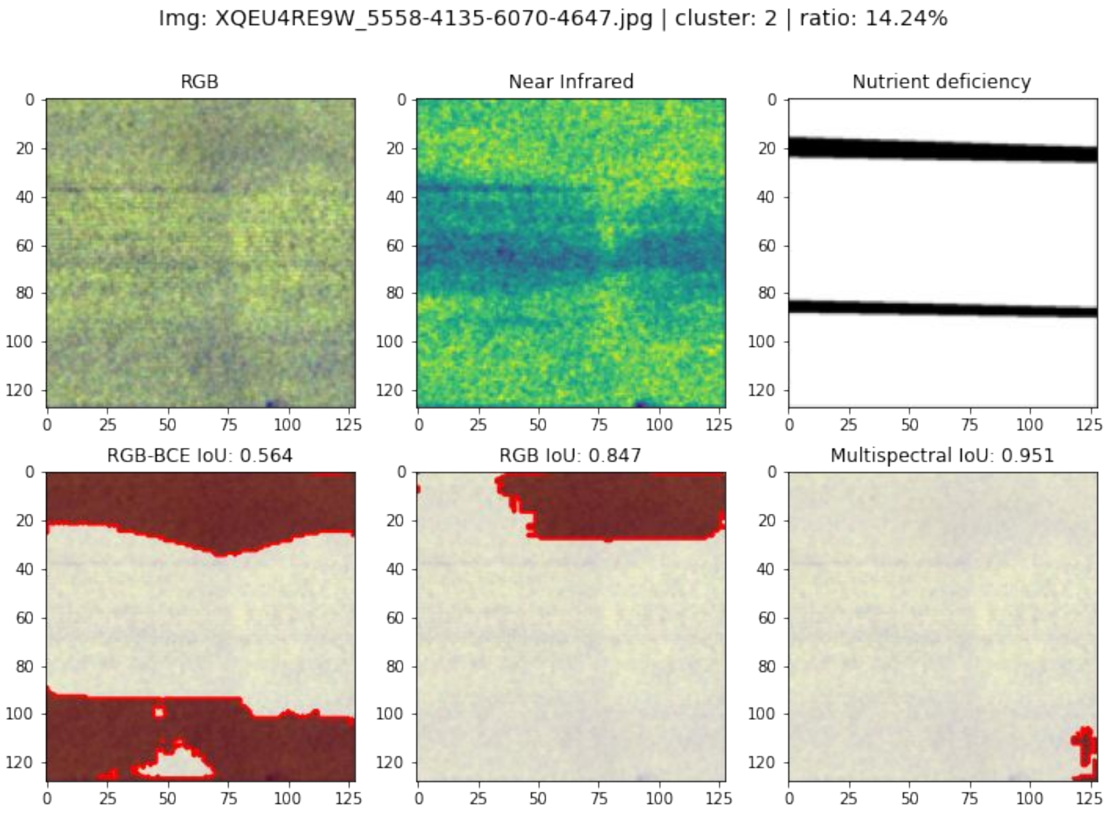
The Best



<u>Legend</u>
White: ND
Black: Not ND

<u>Legend</u>
Pale: ND
Dark: Not ND

The Best (or maybe the luckiest?)



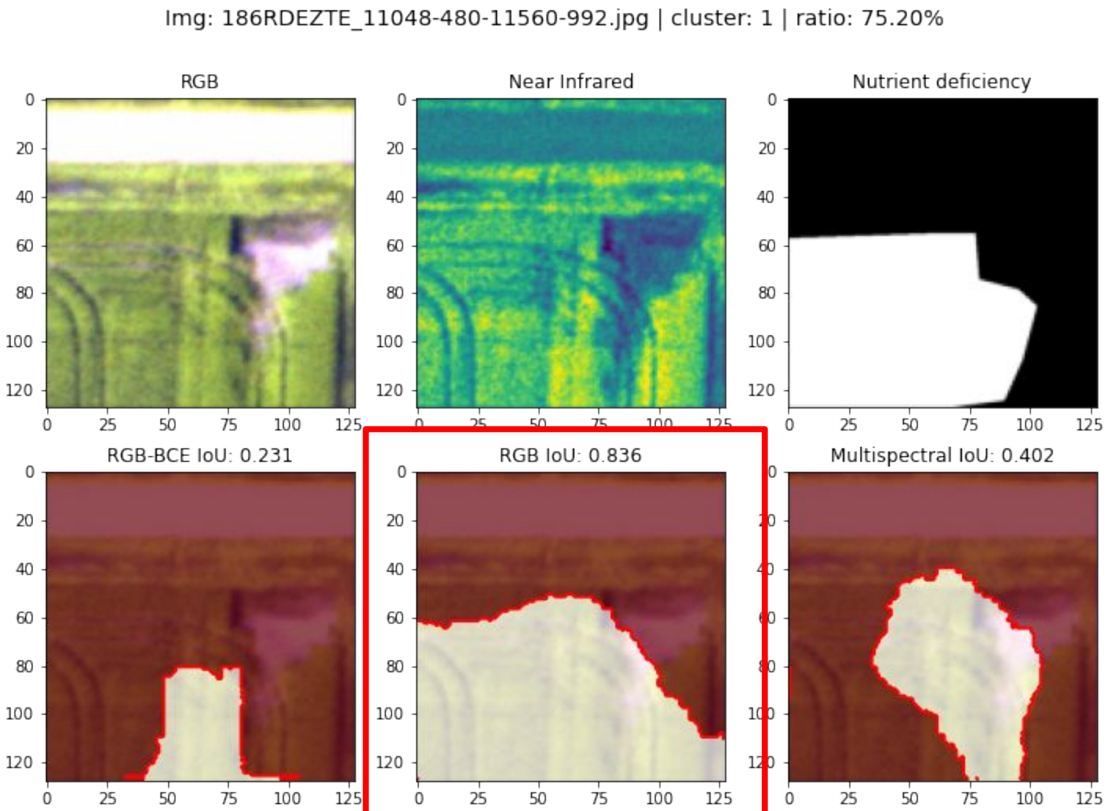
Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

Review of Model Predictions

The Okays

The Okays

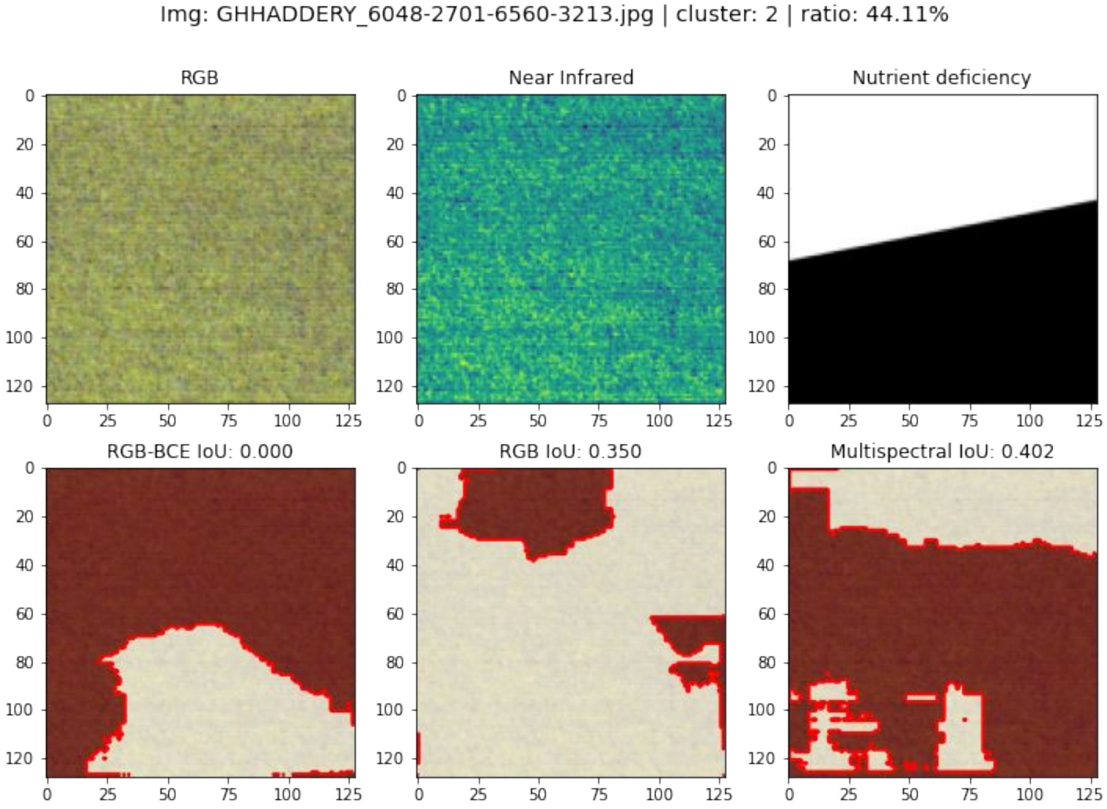


This is a case where not having NIR may have helped!

Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

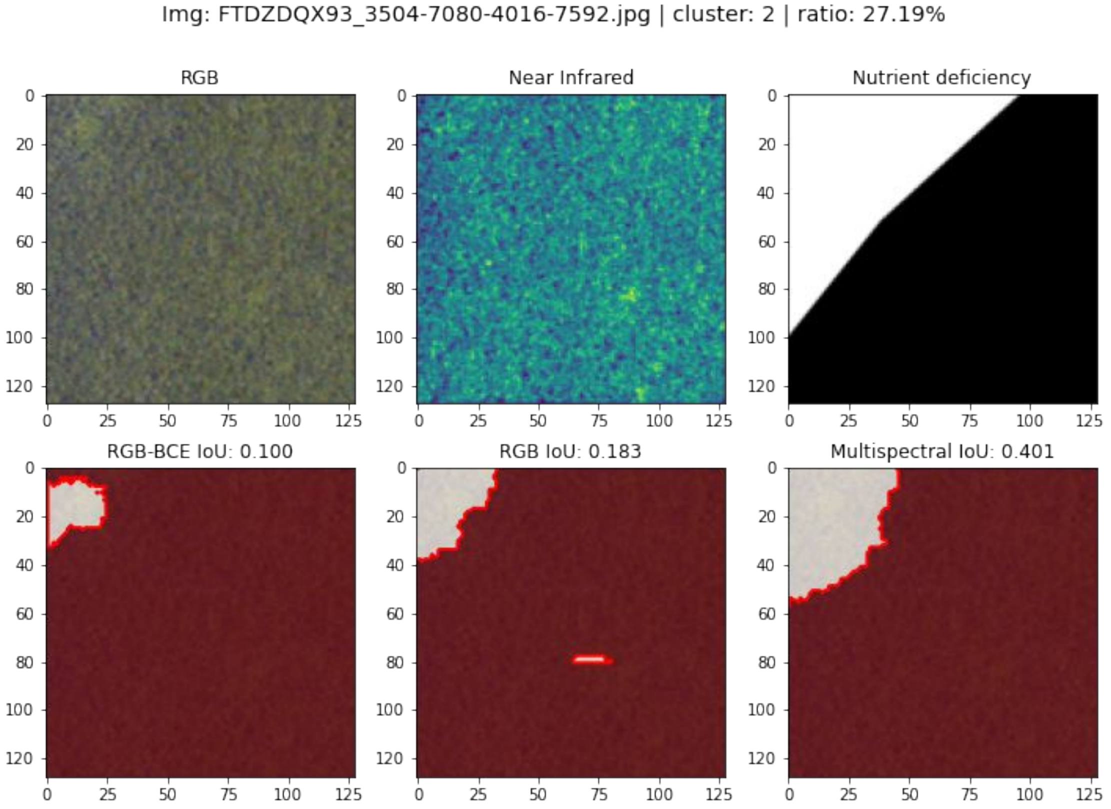
The Okays



<u>Legend</u>
White: ND
Black: Not ND

<u>Legend</u>
Pale: ND
Dark: Not ND

The Okays ⭐⭐⭐⭐⭐⭐⭐⭐⭐⭐



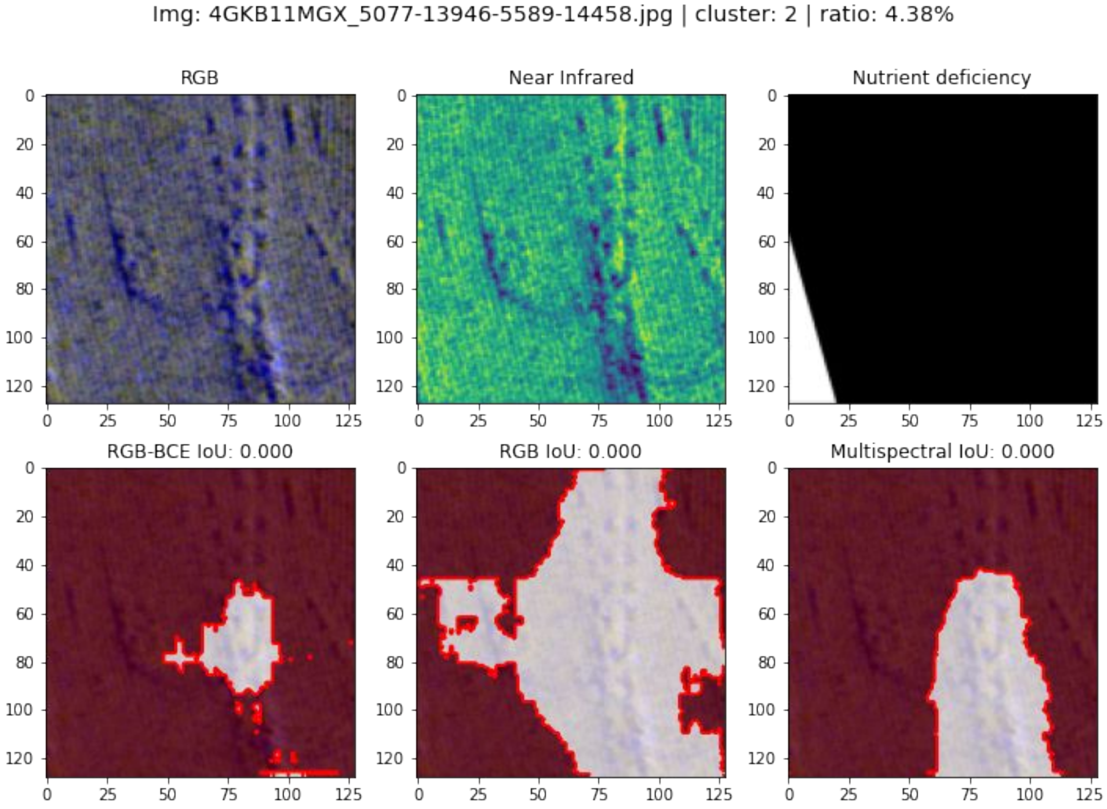
Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

Review of Model Predictions

The WORST

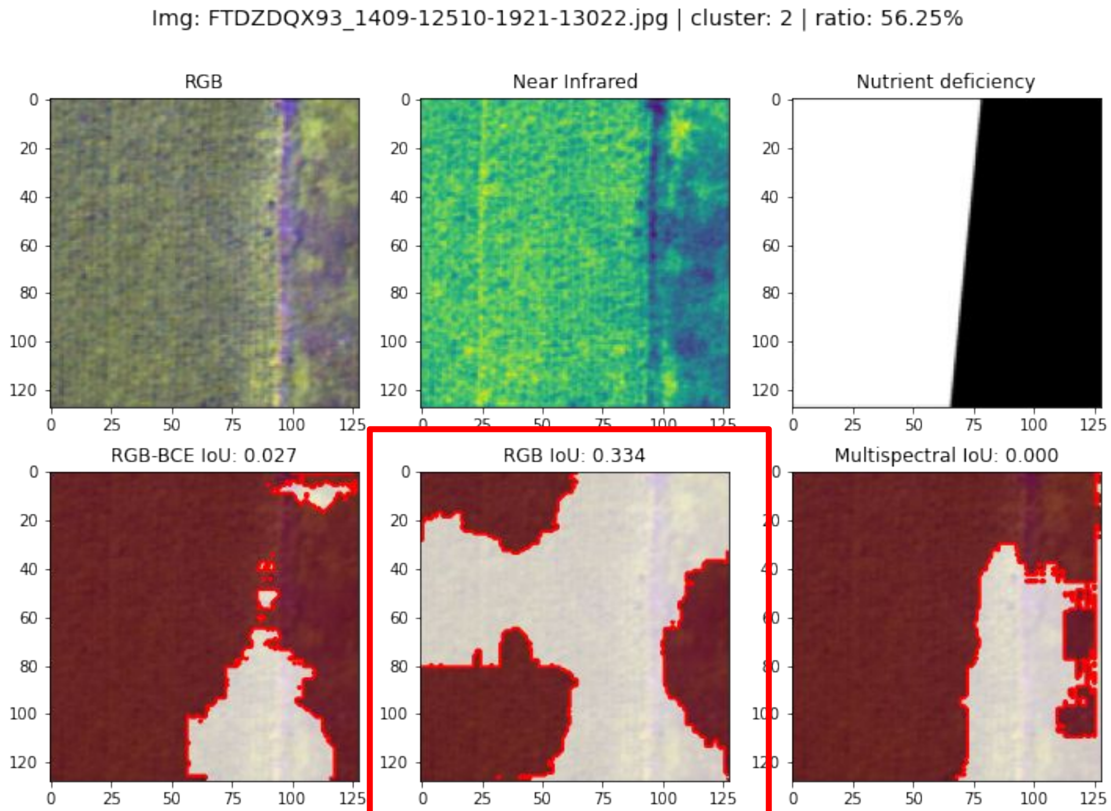
The WORST



Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

The WORST



Another example where having NIR may not have helped!

Legend
White: ND
Black: Not ND

Legend
Pale: ND
Dark: Not ND

Conclusions

- It’s critical to consider the appropriate loss function and evaluation metric
- Using the **Intersection over Union** metric, the multispectral U-Net outperforms a baseline 'predictive' model
- For agricultural applications in identifying nutrient deficiency, it's clear that sometimes the Near-Infrared spectral channel is helpful, but sometimes it's not
- In some of the best examples shown, appearance of nutrient deficiency may not be apparent to the human eye, but the U-Net models used in this analysis demonstrate that machine learning great has potential for aiding in protection and optimization of crop yields

Recommendations:

- The image data were originally provided in 512x512 size, but were resized to 128x128 for training (for computational expense minimization). With greater computing power and more time, the model could likely become even better at identifying nutrient deficiency.
- It's conceivable that an ensemble model might be beneficial for this type of task: in some cases the model using the near-infrared spectral channel performed worse at classification than the model that only used the RGB channels. **Ensemble Modeling** may be helpful.
- RGB and near-infrared may not be the only imaging techniques that could be useful in this type of task. It could be beneficial to explore usage of other imaging methods (like UV and Infrared).

Thank you!