Report

For this project, we are going to predict field anomalies for each pixel in an image. In this report, we did a thoroughly analysis on our training data set to see the data distribution pattern. Four features are used in the analysis, including 3 RGB values and 1 infrared value. For computational efficiency and data visualization purposes, we took one random pixel from each training image (512 x 512 pixels) and got a total of 12901 uncorrelated pixels. Table 1 shows the label distribution of the selected pixels.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | background | cloud\_shadow | double\_plant | planter\_skip | standing\_water | waterway | weed\_cluster |
| Number | 10173 | 396 | 96 | 23 | 204 | 149 | 1860 |

Table 1. label distribution of the selected pixels.

From Table 1, we can see that majority of the pixels in the training data are background label, which is reasonable from our inspection.

We plotted the data distribution of the selected pixels regarding to the RGB and infrared values.

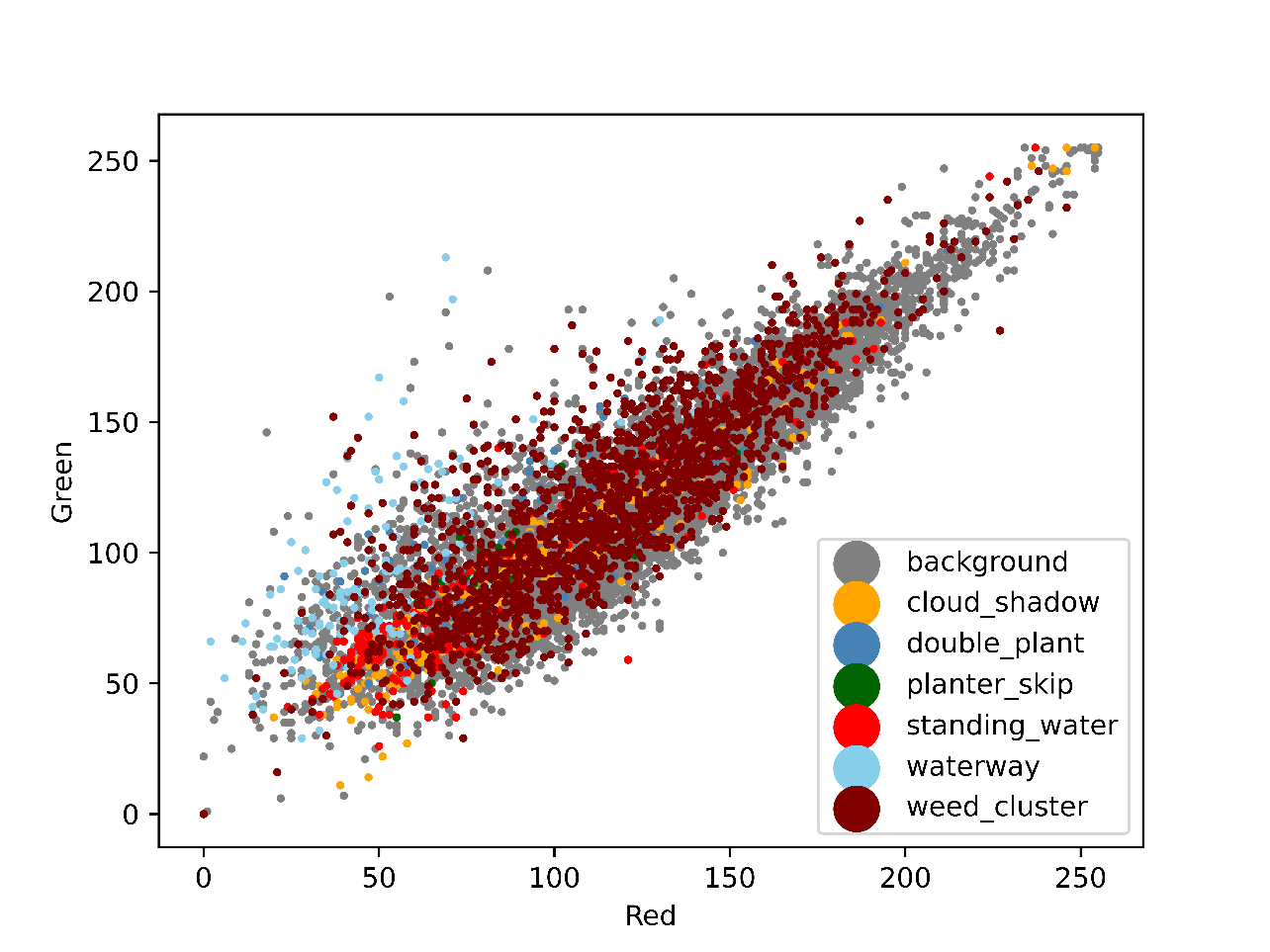


Figure 1. data distribution in red and green domain.

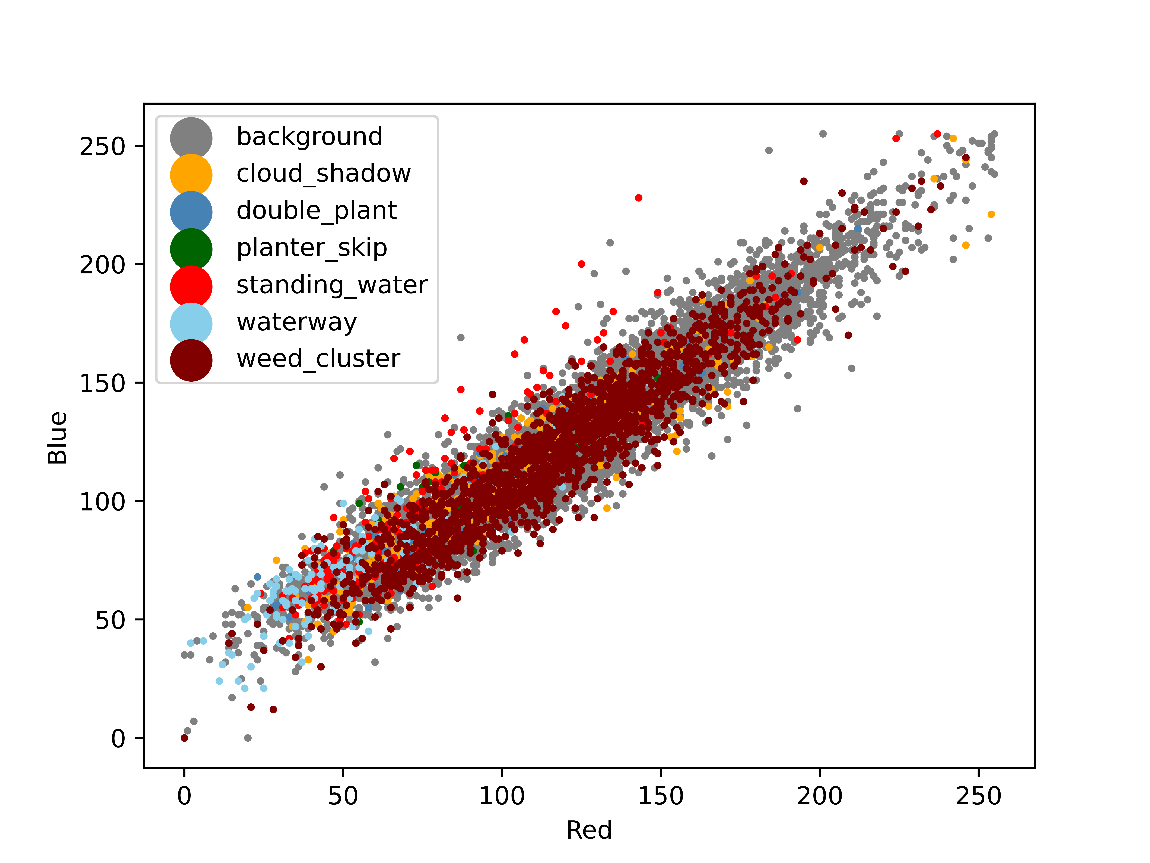


Figure 2. data distribution in red and blue domain.

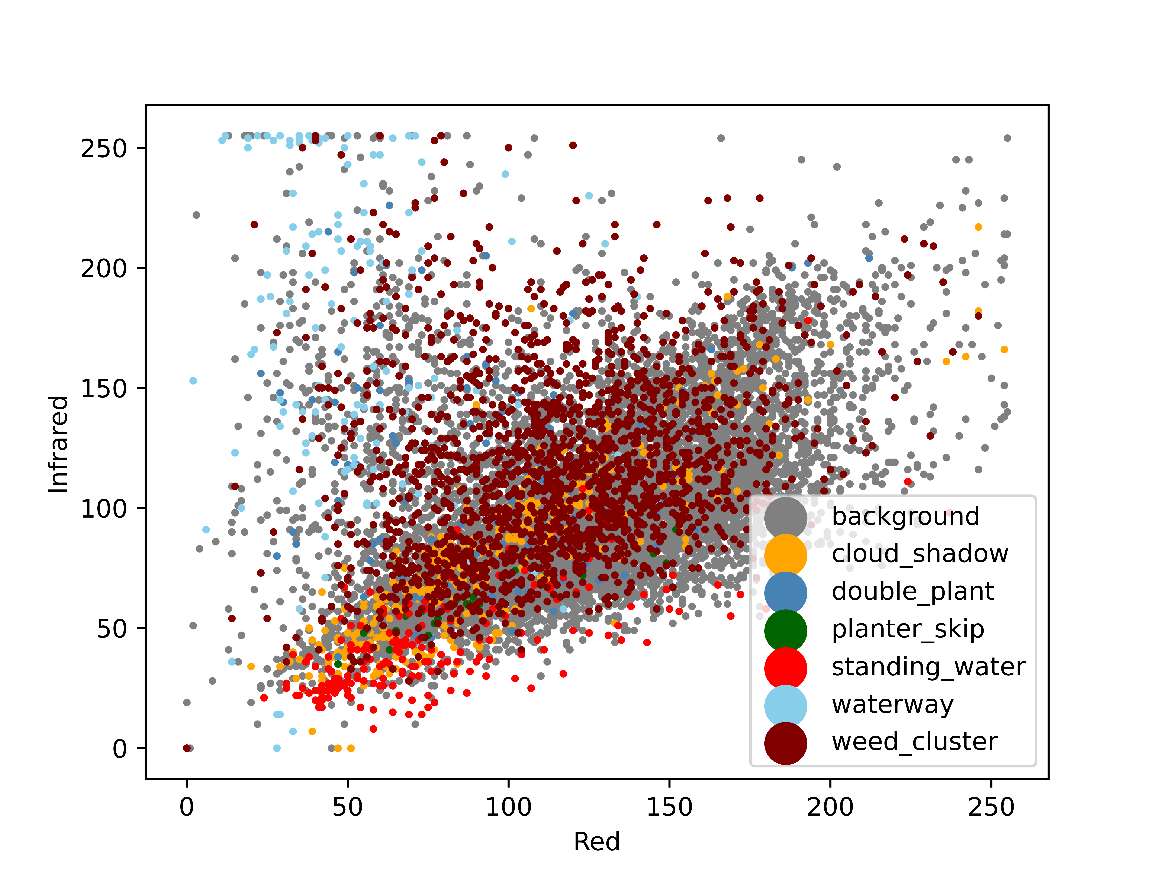


Figure 3. data distribution in red and infrared domain.

From Fig 1 and 2, we found that there is no clear clustering pattern and noticeable outliers for any label. Comparing to other labels, “standing\_water”(red) and “waterway”(lightblue) pixels generally have low RGB values. It is also noticeable that the background label spans across the figure, indicating the difficulty of separating it from other labels.

In the infrared domain, clearer clustering pattern can be seen for “standing\_water”(red) and “waterway”(lightblue) labels. We also saw outliers for “waterway” label with high infrared value.

From the data inspection, we can see that the RGB values of the pixels carry limited information. Data points show a linear trend with very few outliers and all labels show uniform distribution. On the contrary, the infrared value provides more information showing separation between labels. The mixed data could cause a lot of problems for machine to learn, because it would be hard for machine to separate each label, especially the “background” label shares similar characteristics as other labels.

After the general analysis, we further investigate pixels in a single specific image to see the differences between field anomaly and background labels.

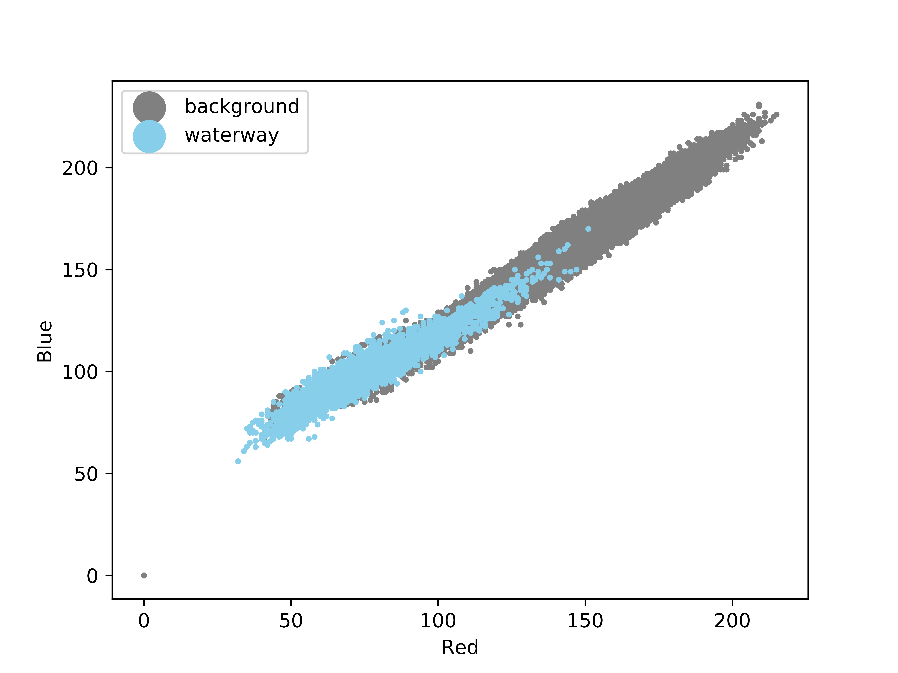


Figure 4. example image and waterway vs background RGB value distribution

From single image analysis, although pixels are still mixed in RGB values, better separation can be found for the two labels. However, for both labels we saw a wide range red and blue values (~50 – 200), indicating a weak clustering pattern. In conclusion, single image analysis results show a better label separation than the previous random pixel analysis, illustrating the importance of the neighboring effect of the pixel group instead of treating every pixel as independent.

At this stage, we only considered four features of the training data set (red, green, blue, and infrared values). However, red, green, and blue values share very similar characteristic, and could be reduce to the average of the three values. We believe that adding more features is necessary for better training result.

To start up, we used simple linear methods, softmax logistic regression and SVM, to train the dataset. In both methods, we randomly selected 600 images from the training dataset and then evenly choosing 600 pixels with each of the 6 labels.

**Training with logistic regression**

**1. Using 4 features: 3 RGB and 1 nir values**

We choose learning rates among 1e-7, 5e-7, 1e-6, 5e-6, and regularization values from 5e4, 1e5, 5e5, 1e8.

The loss of the model is 1.9459170318715031. The following table shows the training accuracy at each learning rate and regularization:

|  |  |  |
| --- | --- | --- |
| Learning rate | Regularization | Training Accuracy |
| 1e-7 | 5e4 | 0.113 |
| 1e-7 | 1e5 | 0.000278 |
| 1e-7 | 5e5 | 0.0 |
| 1e-7 | 1e8 | 0.33889 |
| 5e-7 | 5e4 | 0.20111 |
| 5e-7 | 1e5 | 0.180556 |
| 5e-7 | 5e5 | 0.16667 |
| 5e-7 | 1e8 | 0.336389 |
| 1e-6 | 5e4 | 0.16333 |
| 1e-6 | 1e5 | 0.348611 |
| 1e-6 | 5e5 | 0.328611 |
| 1e-6 | 1e8 | 0.336111 |
| 5e-6 | 5e4 | 0.336389 |
| 5e-6 | 1e5 | 0.252778 |
| 5e-6 | 5e5 | 0.336389 |
| 5e-6 | 1e8 | 0.252778 |

The highest accuracy trained by the multi-class logistic regression is 0.348611 using learning rate 1e-6 and reg 1e5. The confusion matrix for these parameters is:

[[190 0 458 0 0 2 0]

[ 43 0 457 0 0 0 0]

[ 0 0 599 0 0 1 0]

[ 0 0 100 0 0 0 0]

[ 1 0 41 0 457 1 0]

[ 63 0 456 0 0 131 0]

[ 0 0 600 0 0 0 0]]

From this matrix, we can clearly see the phenomenon that the trained model predicts most of the pixels to label 2, and almost no 1, 3, and 6. Beside, we observe that label 4 and 5 have a better prediction, where almost all points belonging to them are labeled correctly after training. This meets our expectation that ‘standing water’ and ‘waterway are more easily to be classified than other labels, as they have much lower RGB values.

**2. Using two features: the average of 3 RGB values and 1 nir values**

When plotting the cluster difference, we found that the plots for r&g and r&b have a high similarity. Thus, we supposed that the 3 RGB values are dependent with one another to some extent. Then, we reduced features to only the average value of the RGB and the nir value for the training model.

The loss of this model is 1.9459049171708622, which is very close to the previous model.

We again choose same learning rates and regularization values. The following table shows their training accuracy:

|  |  |  |
| --- | --- | --- |
| Learning rate | Regularization | Training Accuracy |
| 1e-7 | 5e4 | 0.183056 |
| 1e-7 | 1e5 | 0.135 |
| 1e-7 | 5e5 | 0.145 |
| 1e-7 | 1e8 | 0.18333 |
| 5e-7 | 5e4 | 0.04083 |
| 5e-7 | 1e5 | 0.18056 |
| 5e-7 | 5e5 | 0.21667 |
| 5e-7 | 1e8 | 0.18333 |
| 1e-6 | 5e4 | 0.18194 |
| 1e-6 | 1e5 | 0.02167 |
| 1e-6 | 5e5 | 0.18333 |
| 1e-6 | 1e8 | 0.18333 |
| 5e-6 | 5e4 | 0.19667 |
| 5e-6 | 1e5 | 0.18333 |
| 5e-6 | 5e5 | 0.18333 |
| 5e-6 | 1e8 | 0.18333 |

The highest accuracy trained by this multi-class logistic regression is 0.216667 using learning rate 1e-6 and reg 1e5. The confusion matrix for these parameters is:

[[ 67 0 387 0 0 196 0]

[ 37 0 99 0 0 64 0]

[ 0 0 431 0 0 169 0]

[ 0 0 250 0 0 150 0]

[ 0 0 447 0 50 3 0]

[ 57 0 361 0 0 232 0]

[ 0 0 355 0 0 245 0]]

From this matrix, we observe that label 4, ‘standing water’, still has a better prediction, with all data points being correctly classified. However, label 5 gets more wrong predictions. Thus, we can assume that even though RGB values for ‘waterway’ are lower than other labels, the average values might be similar, making it hard to separate from others.

The logistic regression fails capture the prediction.

**Training with SVM**

**Next Step**