

Final Project

UAV Imagery based Wetland Land Cover Classification

Qiao Hu, SNR, UNL

Paul Fletcher, CSE, UNL

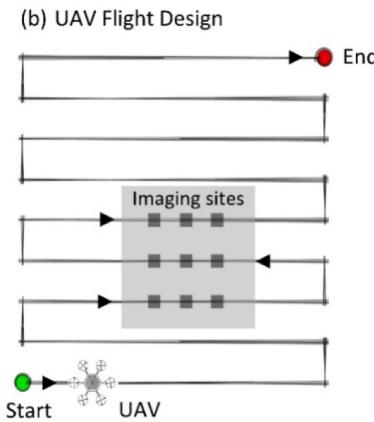
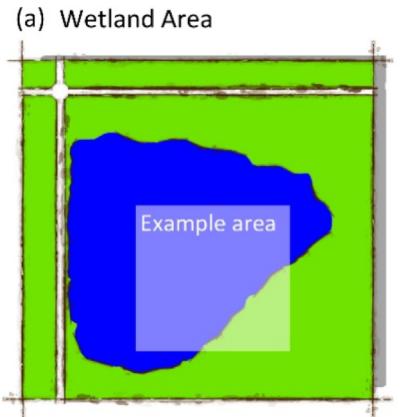
Jacob Gideon, CSE, UNL

Paul O'Dell, CSE, UNL



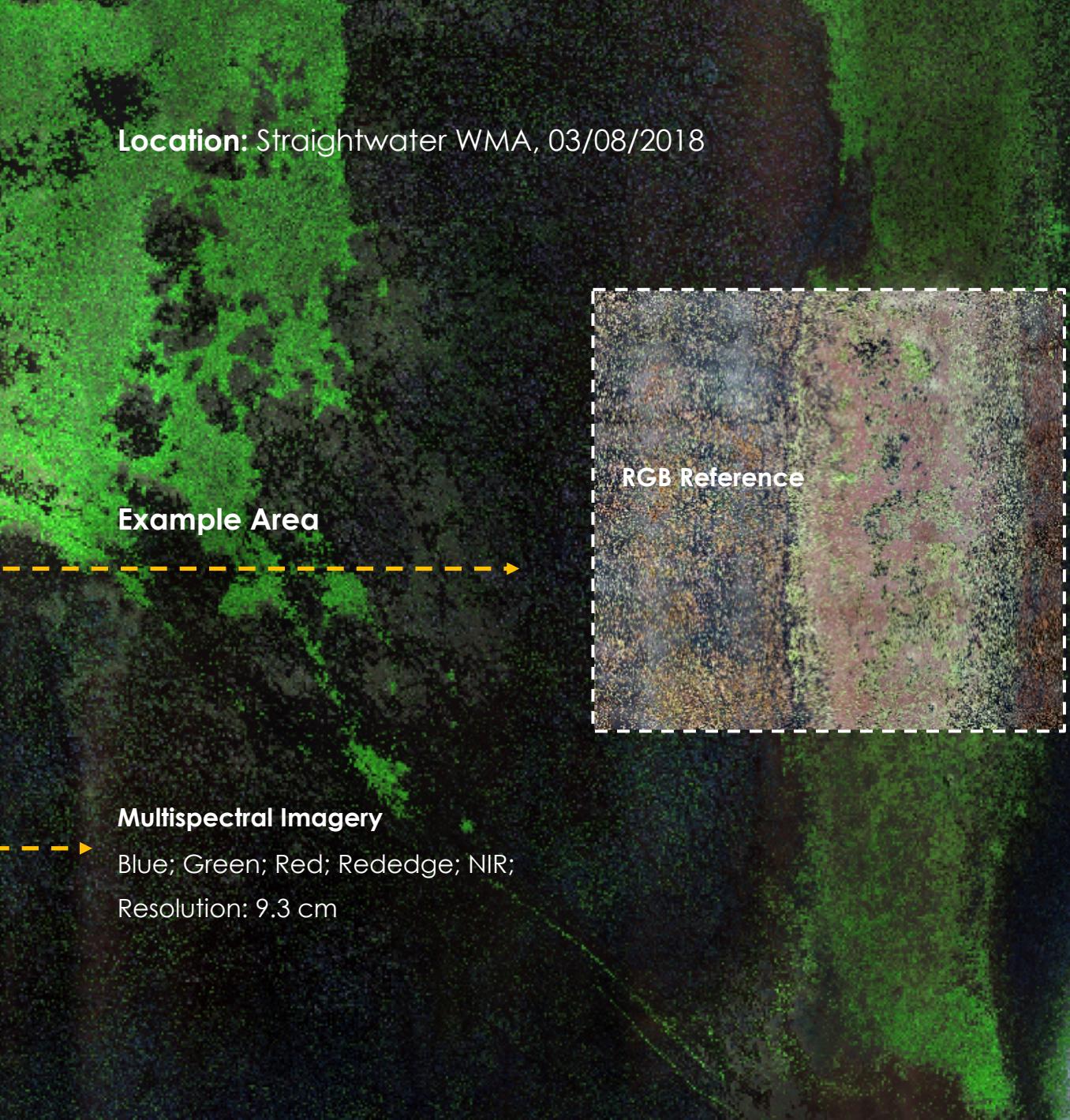
April 22, 2021

Data & Goals



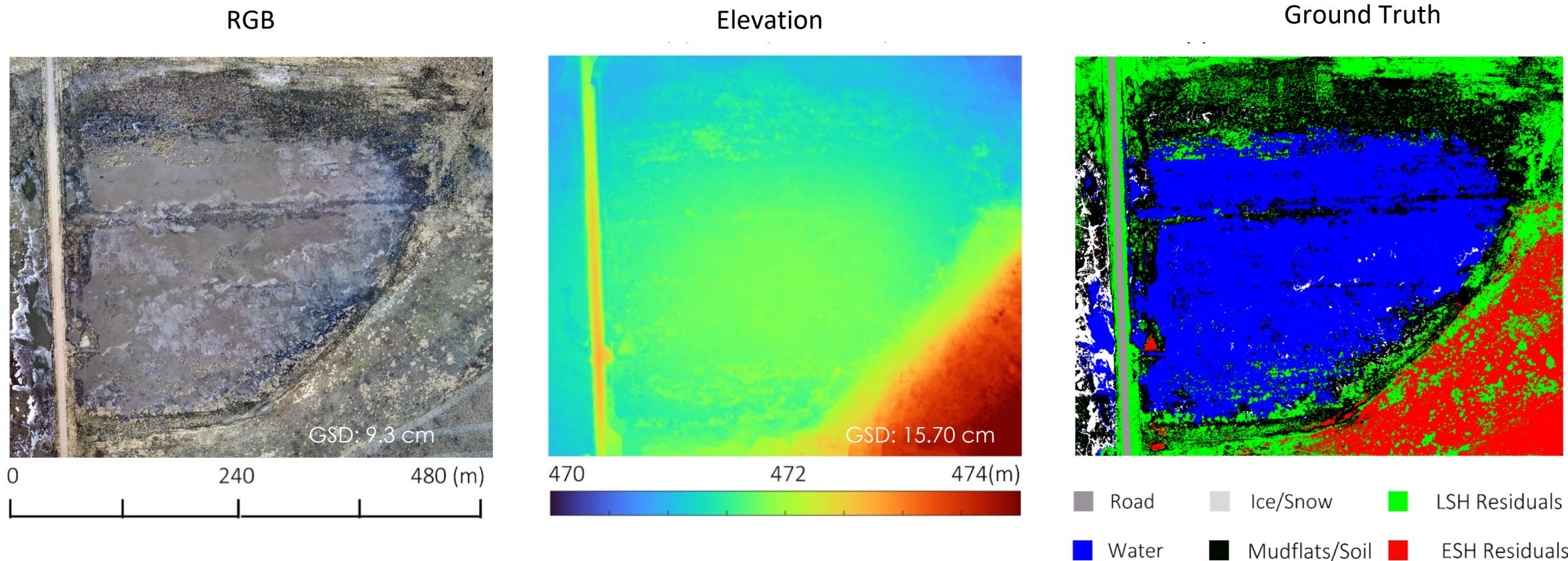
Goal: Develop machine learning method to automatically **delineate wetland area** based on UAV **multippectral data**.

Location: Straightwater WMA, 03/08/2018



Data

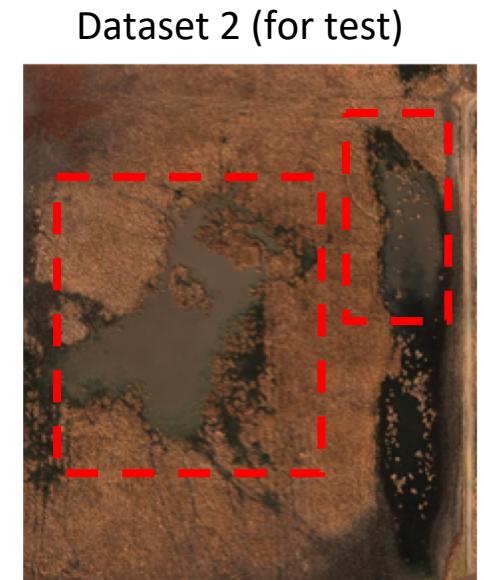
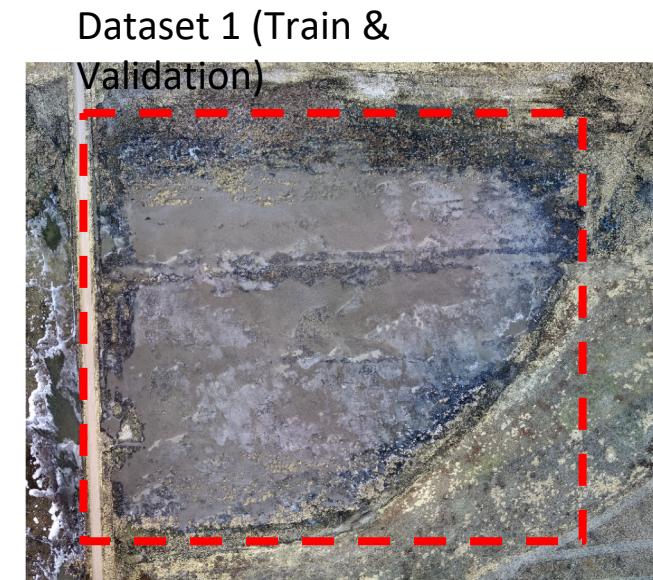
Multispectral Imagery: 6 bands = Blue, Green, Red, Rededge, NIR + DSM (5123x8045x6)



Data and Features

Multispectral Imagery: 6 bands = Blue, Green, Red, Rededge, NIR + DSM (5123x8045x6)

- Feature 1:** Multi-spectral Reflectance - Red (numeric 0-1)
- Feature 2:** Multi-spectral Reflectance - Blue (numeric 0-1)
- Feature 3:** Multi-spectral Reflectance - Green (numeric 0-1)
- Feature 4:** Multi-spectral Reflectance - NIR (numeric 0-1)
- Feature 5:** Multi-spectral Reflectance - Red Edge (numeric 0-1)
- Feature 6:** Elevation information --DSM--(numeric by meter)

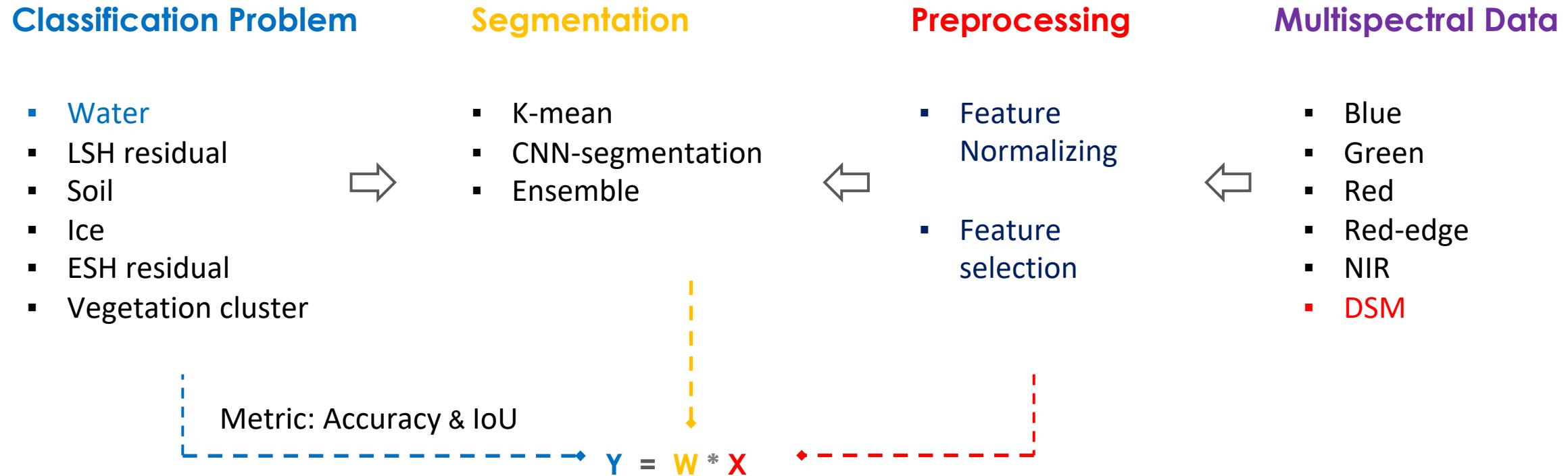


Water ponding area

Dataset 1: 3559 / 50 million (03/08/2018)

Dataset 2: 10117830 / 50 million (03/01/2018)

Workflow



Objective 1: Data preprocessing

Objective 2: K-means clustering

Objective 3: Using a single classifier and ensemble method to classify the image

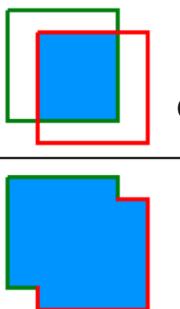
Objective 4: Using a CNN-based semantic segmentation to conduct semantic segmentation over an image

Metrics

Metrics

- **Accuracy:** Accuracy is the ratio of correct predictions to the total number of predictions.
- **Intersection over Union (IoU):** IoU is a metric that measures how similar our predicted boundary is to the ground truth boundary.

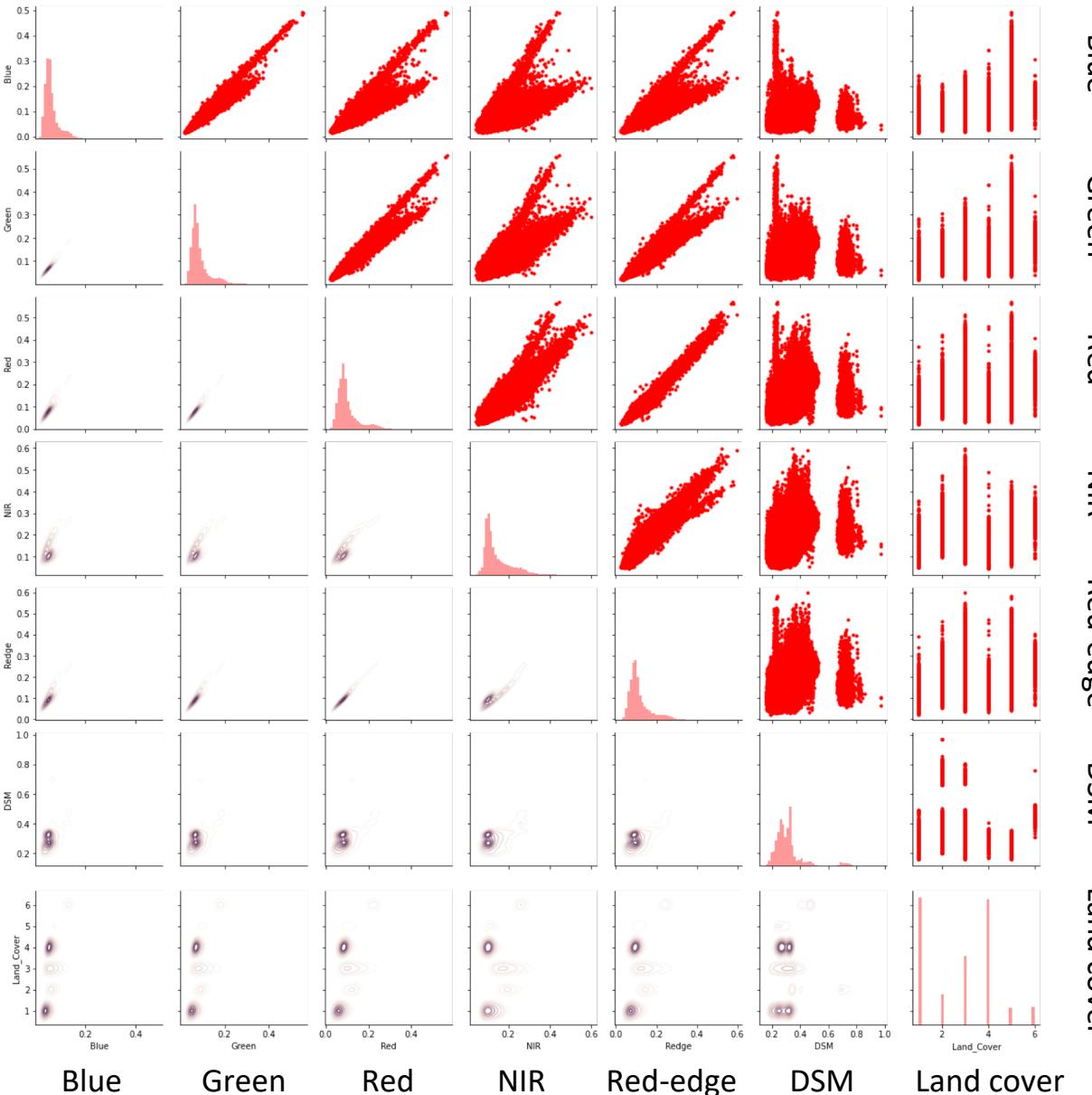
$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{Prediction} \cap \text{Ground truth}}{\text{Prediction} \cup \text{Ground truth}}$$


Preprocessing

Image Conversion

- Six Identical resolution images loaded into memory and converted to **single numerical value**
- Entries with six features and X vs Y coordinates
- Each feature **Z score normalized** with a normal distribution
- PCA Considered but not deemed necessary

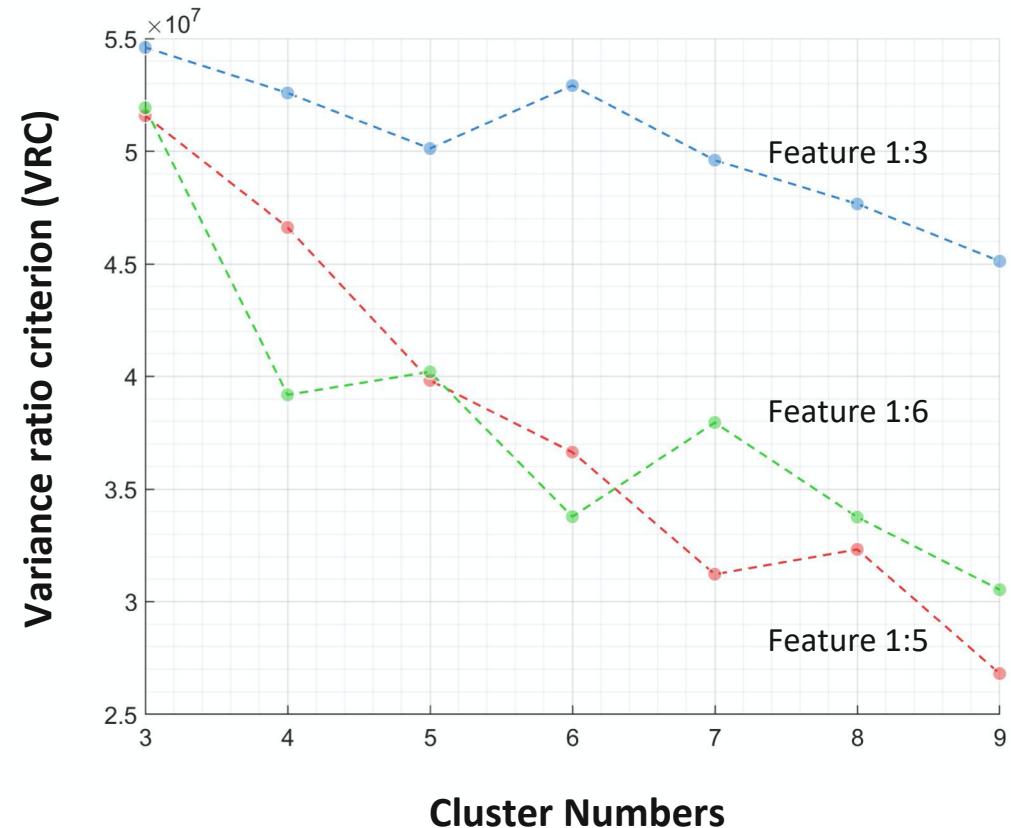


K-means

VRC ratio

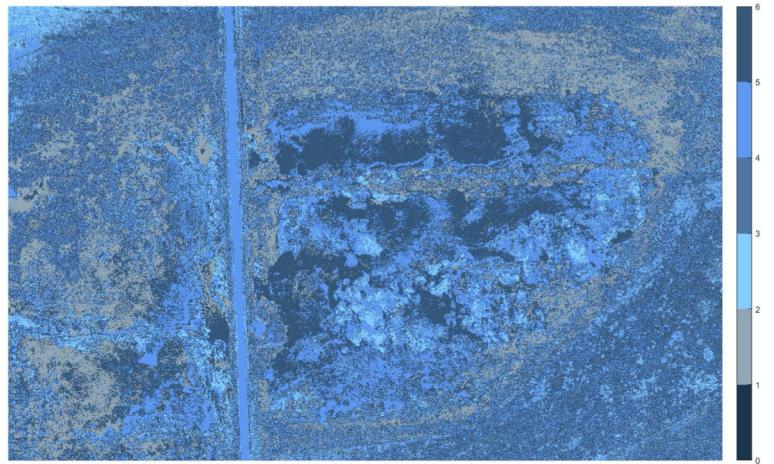
- Ratio of between-cluster variance to within-cluster variance
- Higher ratios indicate better clustering operation
- Elbow point indicate optimal clustering

Clustering with **5-7 clusters** show good VRC performance, so we choose 6 to conduct our clustering operation later.

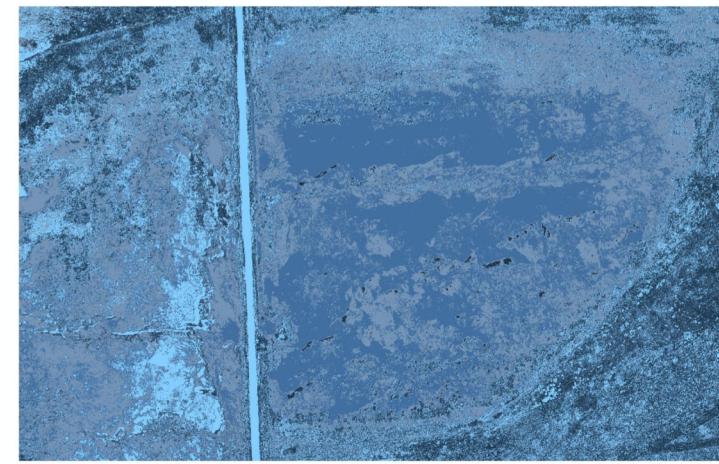


K-means

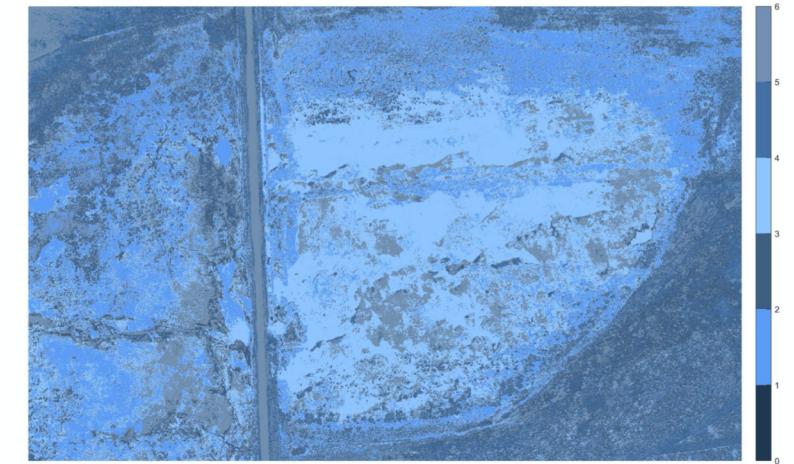
Feature 1:3



Feature 1:5



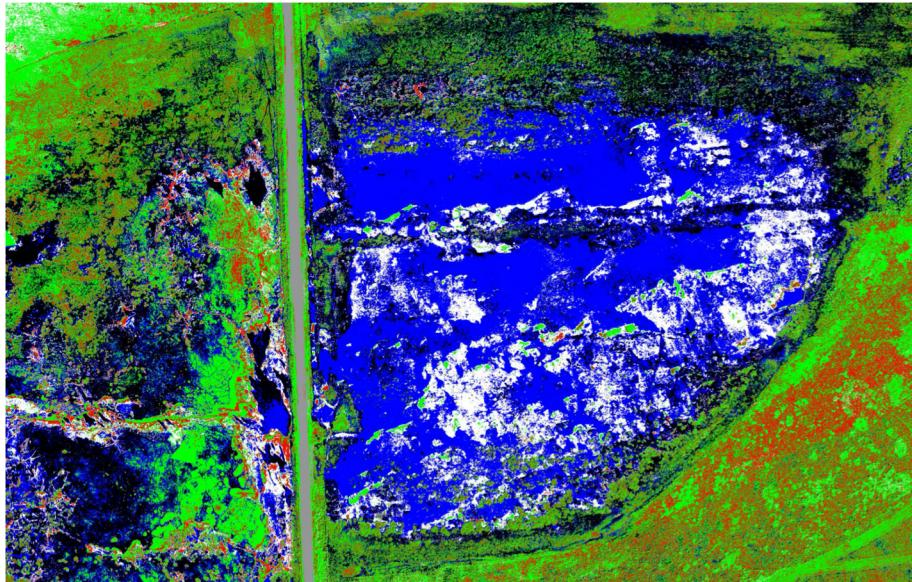
Feature 1:6



- Visually, adding NIR and red-edge band (Feature 1:5) and elevation (Feature 1:5) enhance the clustering cohesion.
- But it is hard to connect the unsupervised cluster to the really groundtruth directly.
- We select the result of feature 1:6 for later accuracy test as it return the best clustering performance visually.

K-means

Results from Feature 1:6 with 6 clusters



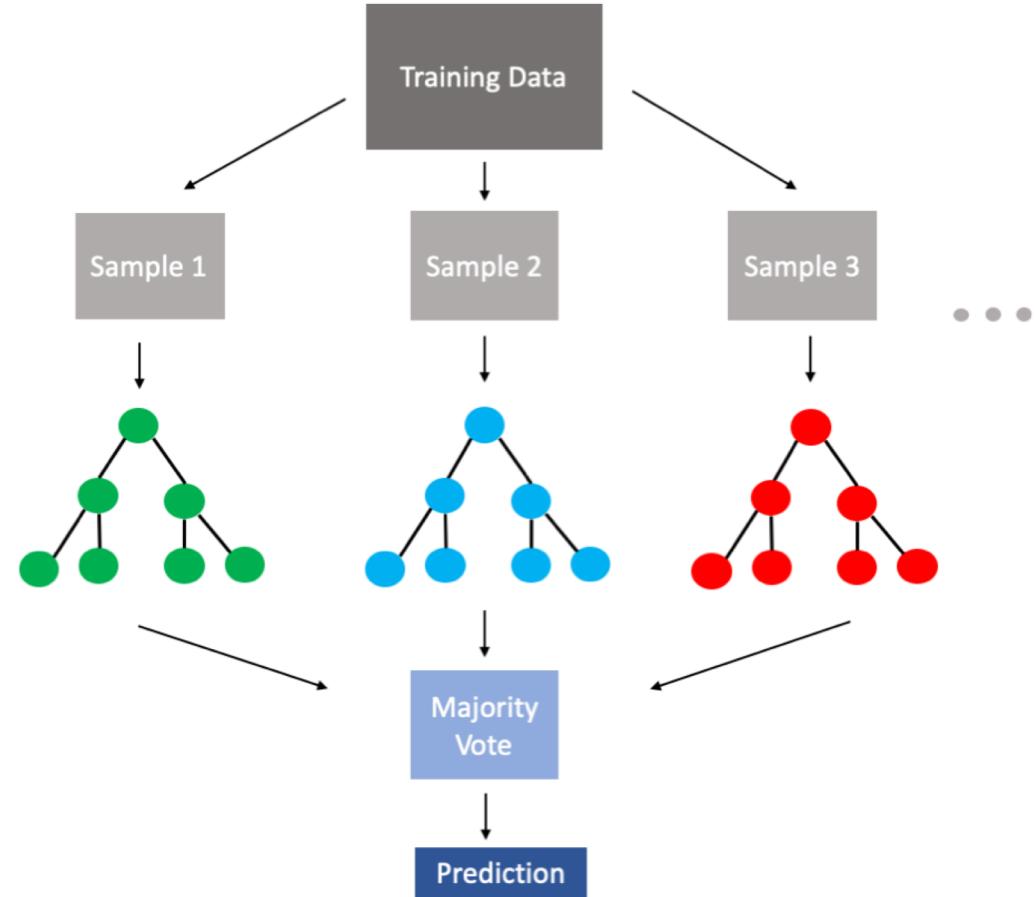
Category		Band 1:6	
Class	Label	Accuracy	IoU
Soil	1	61.2%	57.3%
ESH Residual	2	46.8%	35.7%
LSH Residual	3	77.8%	47.7%
Water	4	70.9%	60.8%
Ice	5	25.7%	0.04%
Road	6	91.1%	80.0%
Global	-	65.9%	60.0%

- We select the k-mean results from feature 1:6 and manually assign the closest label to each cluster.
- Accuracy of each class was measured.

Ensemble Method: Bagging Classifier

Decision Tree Bagging Classifier

- Comparison of single decision tree with bagging classifier.
- Bagging: sampling with replacement and training on multiple decision trees.
- Prediction made based on majority vote.
- Trained on 20% of the total data set (random pixels).
- Implemented in Scikit Learn [1].



Ensemble Method: Bagging Classifier for Site 1

Results

Category		Band 1:3		Band 1:5		Band 1:6	
Class	Label	DT Accuracy	Bag Accuracy	DT Accuracy	Bag Accuracy	DT Accuracy	Bag Accuracy
Soil	1	80.4%	84.0%	86.0%	89.2%	88.2%	90.1%
ESH Residual	2	48.6%	38.3%	55.9%	52.1%	88.2%	91.2%
LSH Residual	3	67.4%	66.1%	72.2%	71.3%	84.2%	83.0%
Water	4	82.1%	84.8%	93.2%	94.0%	94.5%	94.8%
Ice	5	77.0%	79.3%	79.8%	80.3%	82.8%	82.0%
Road	6	80.0%	79.8%	91.6%	92.1%	95.7%	95.2%
Global	-	74.7%	75.4%	81.7%	82.5%	89.0%	90.0%

	Band 1:3		Band 1:5		Band 1:6	
	DT	Bag	DT	Bag	DT	Bag
Training Runtime (seconds)	185.5	171.25	264.58	258.19	304.94	312.16

- Full band feature set resulted in best performance overall.
- Bagging classifier results in marginal increase in performance in each case ($\approx 1\%$).
- ESH Residual highly sensitive to inclusion of DSM (band 6).
- Water *not* sensitive to DSM; high accuracy from bands 1:5.

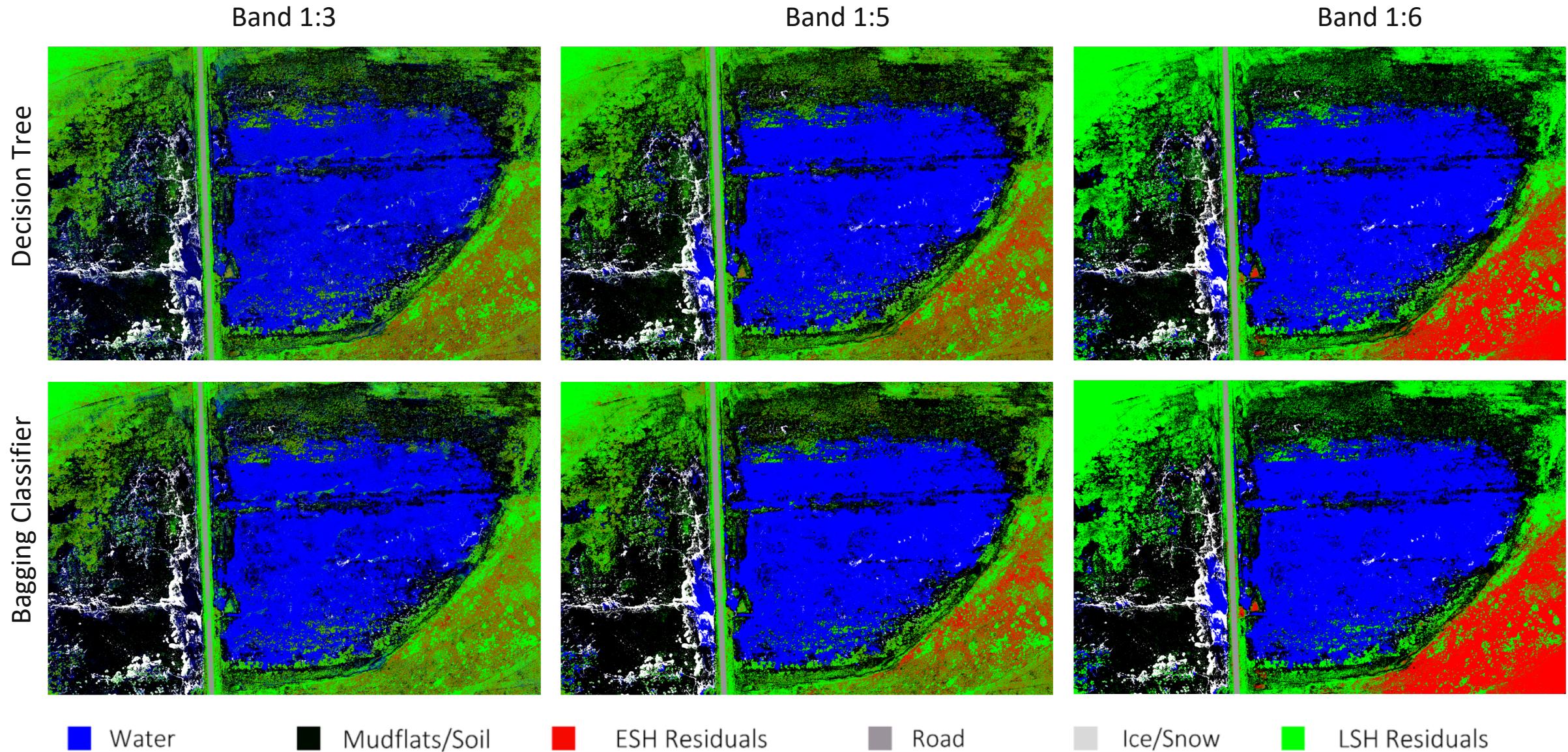
Ensemble Method: Bagging Classifier

Results

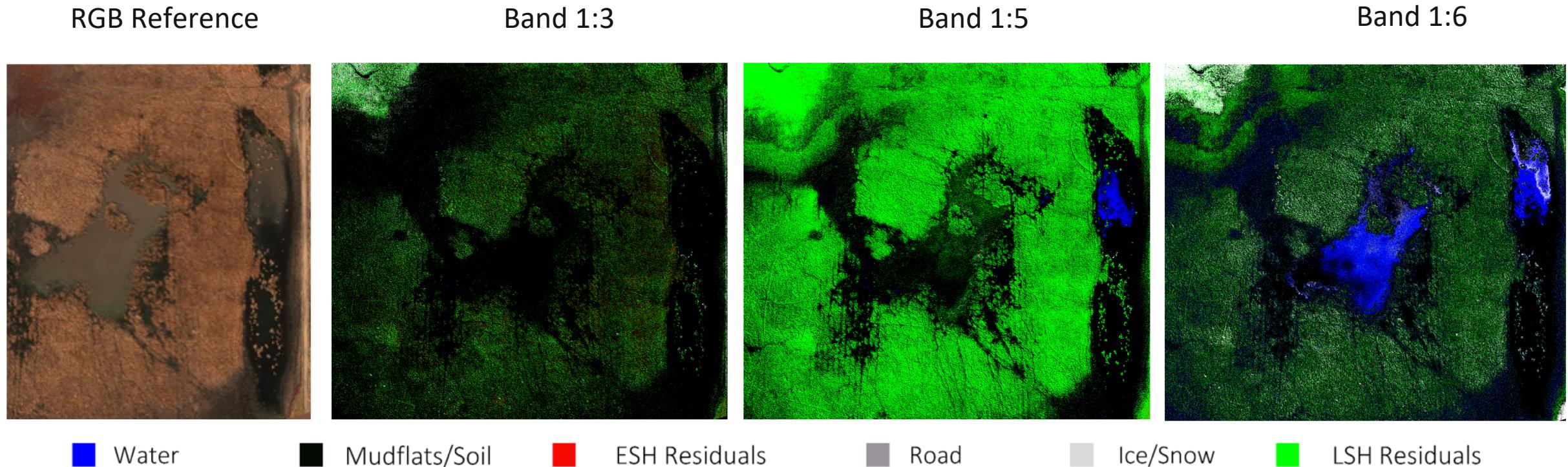
Confusion Matrix							
Class	Label	Soil	ESH Residual	LSH Residual	Water	Ice	Road
Soil	1	12,982,044	35,024	748,042	421,658	84,459	14
ESH Residual	2	49,048	3,599,014	297,269	102	1,187	1,353
LSH Residual	3	1,116,981	452,944	8,043,637	40,391	33,527	12,842
Water	4	493,806	74	43,283	10,917,645	54,993	24
Ice	5	118,370	335	37,923	76,052	1,066,777	33
Road	6	14	3,988	19,142	232	63	462,245

- LSH Residual, Water, and Ice were most often misclassified as Soil.
- Soil, ESH Residual, and Road were most often misclassified as LSH Residual.
- Road had lowest rate of misclassification and also lowest representation in the data set.

Ensemble Method: Bagging Classifier for Site 1



Ensemble Method: Bagging Classifier for Test



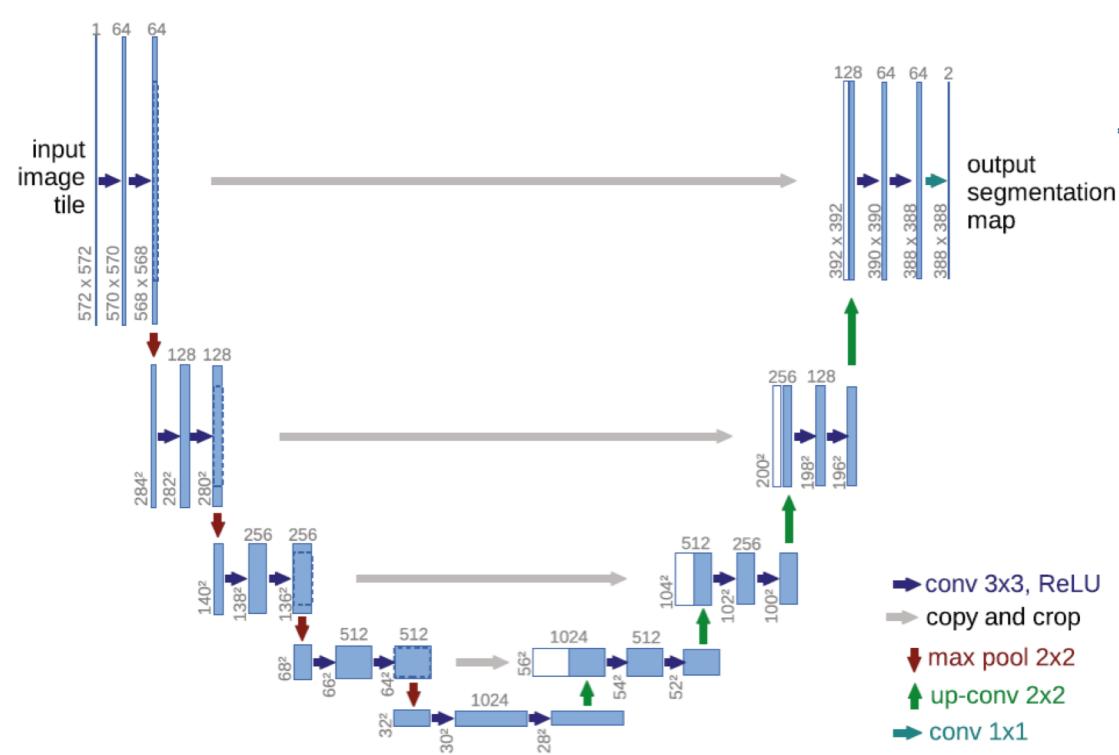
- RGB bands are insufficient alone to provide classification of varying land covers.
- Adding DSM improves the classification, especially with respect to water.

Semantic Segmentation: U-net

Architecture

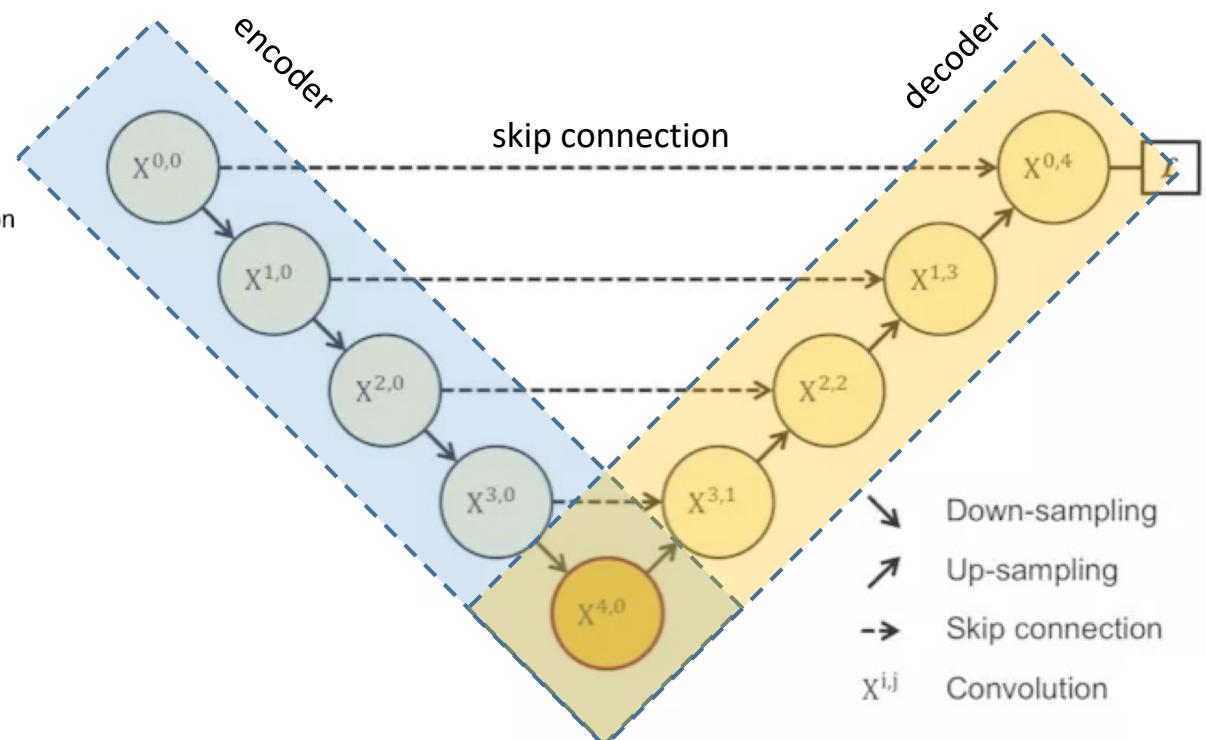
Key characteristics

- Encoder-decoder CNN architecture
- Skip connection
- Multi-scale spatial features



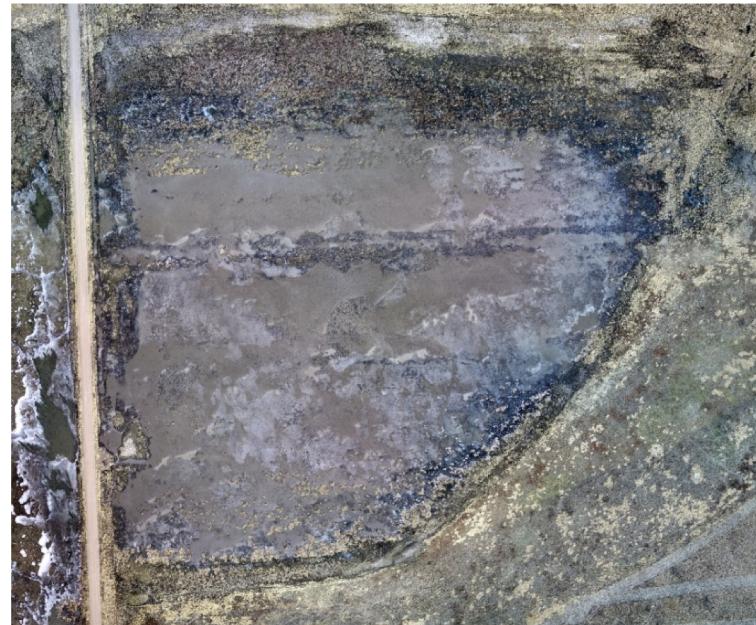
Input & output

- Given image with shape of (m,n,c)
- Produce label with shape of $(m,n,1)$
- Cross entropy loss

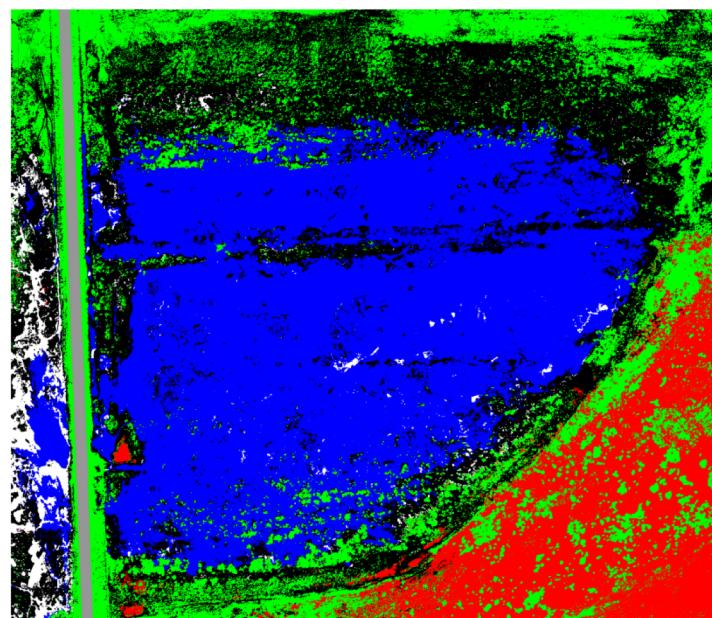


Semantic Segmentation: U-net Training

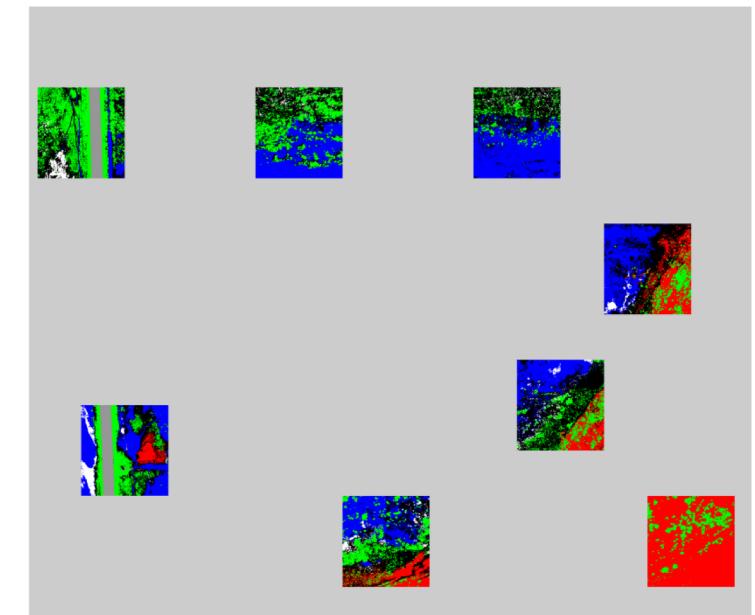
RGB reference



Ground truth



Sample region



■ Water

■ Mudflats/Soil

■ ESH Residuals

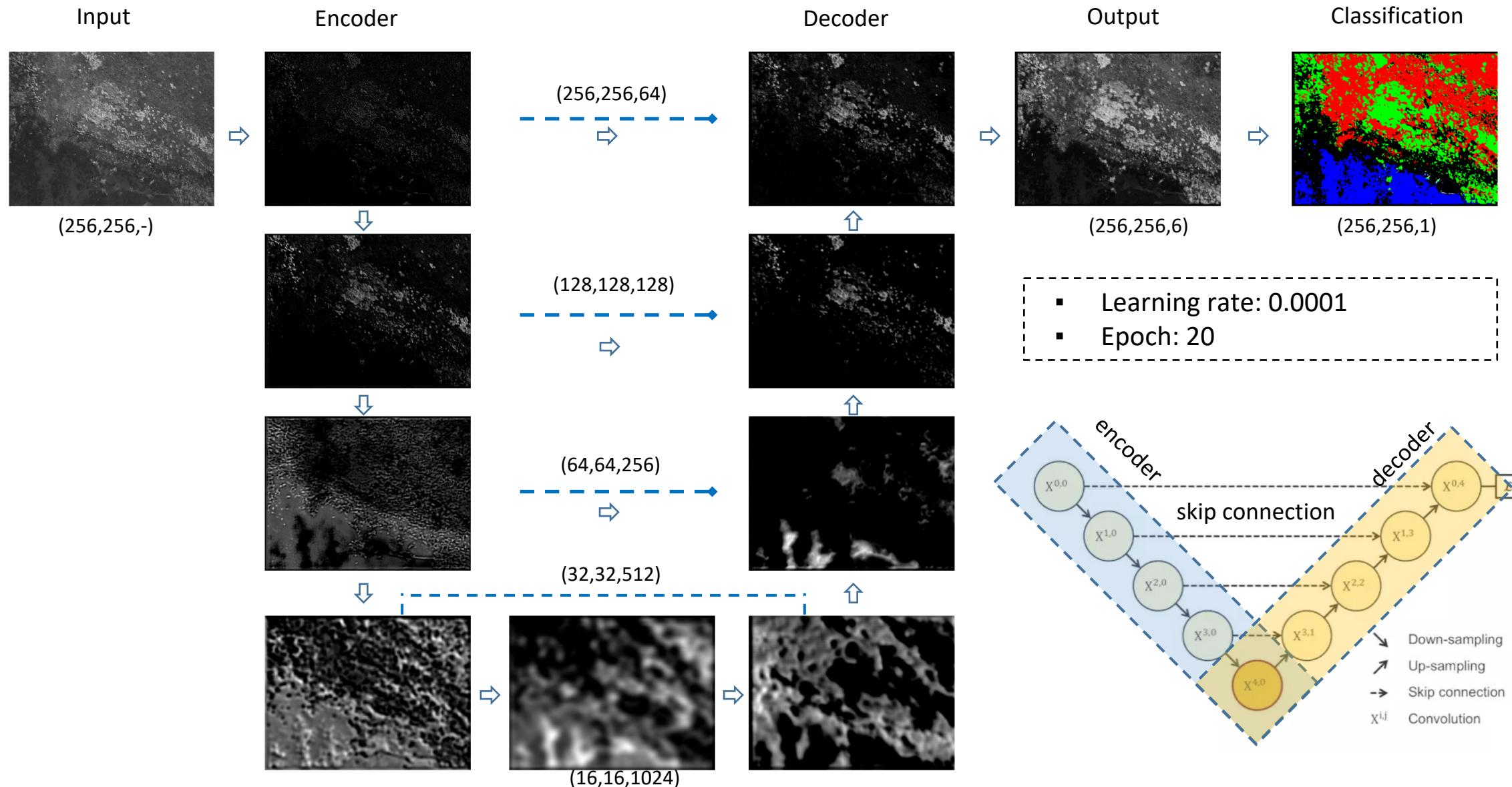
■ Road

■ Ice/Snow

■ LSH Residuals

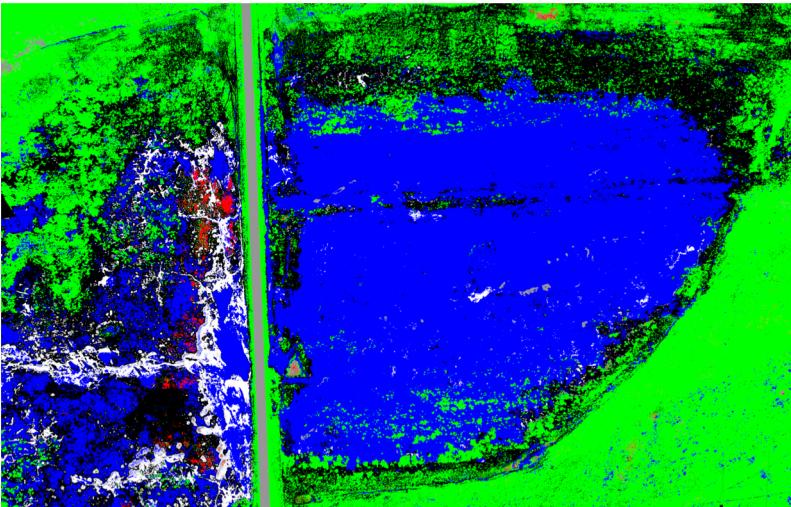
- Sample region: 8
- Training sample shape: 256x256x6
- Training sample number: 1600
- Data augmentation: reflection, translation and rotation

Semantic Segmentation: U-net Feature Flow



Semantic Segmentation Results for Site 1: U-Net

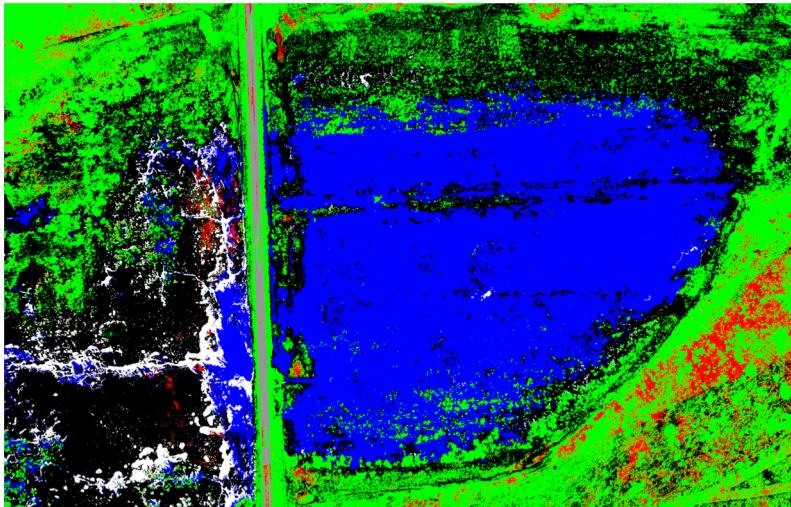
Feature 1:3



■ Water

■ Mudflats/Soil

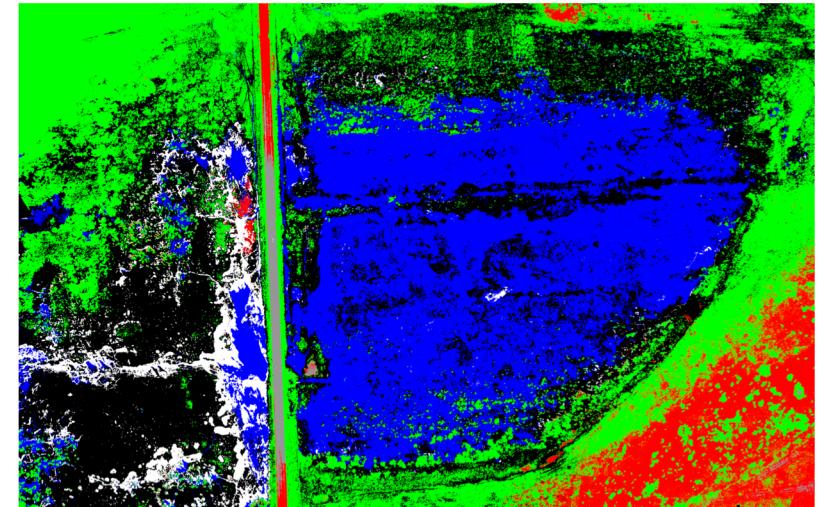
Feature 1:5



■ ESH Residuals

■ Road

Feature 1:6



■ Ice/Snow

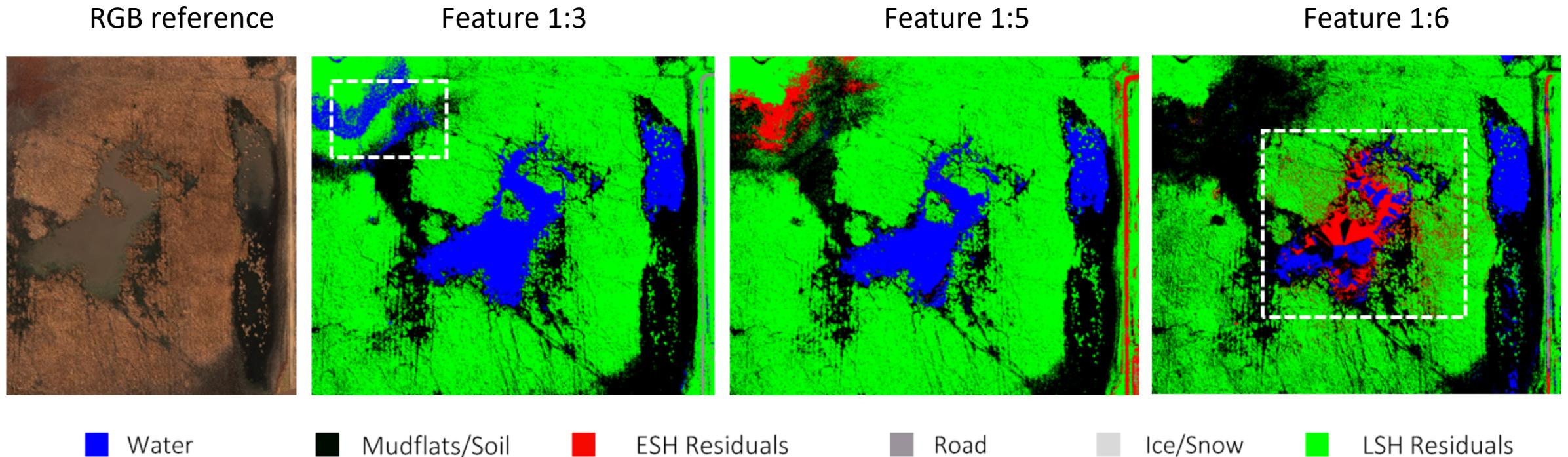
■ LSH Residuals

Semantic Segmentation Results for Site 1

Category		Band 1:3		Band 1:5		Band 1:6	
Class	Label	Accuracy	IoU	Accuracy	IoU	Accuracy	IoU
Soil	1	73.1%	67.8%	95.0%	75.0%	61.2%	59.5%
ESH Residual	2	0.4%	0.4%	55.4%	47.1%	71.2%	69.3%
LSH Residual	3	89.0%	51.8%	69.5%	56.6%	93.3%	70.5%
Water	4	99.1%	83.9%	95.4%	93.4%	99.5%	80.6%
Ice	5	74.2%	69.4%	89.0%	65.6%	60.0%	59.3%
Road	6	98.1%	86.9%	51.0%	50.7%	96.5%	95.0%
Global	-	72.3%	60.0%	86.2%	67.3%	83.7%	80.3%

- Generally, full band return significantly higher performance than three bands.
- Adding DSM band enhance the boundary matchness of the segmented land cover.

Semantic Segmentation Results for Site 2



- Adding NIR and rededge band enhance the classification performance.
- Adding DSM band decrease the performance of the land cover classification.
- CNN-semantic segmentation have better generalization than division tree.

Conclusion

Key points

- This study used multispectral UAV wetland image data to conduct a **pixel level semantic segmentation task**.
- **K-means clustering** returned weaker performance in terms of both accuracy and efficiency (labor investment required)
- **Ensemble Decision Tree** performed well with 90% overall accuracy and relatively fast training runtime (< 6 minutes in all cases) but have potential overfitting issue.
- Results showed **CNN semantic segmentation** effectively learn the spatial patterns and generalize well on different datasets.
- The study proved the potential of using **data mining/machine learning method** to automate the land cover classification task, enhancing the **efficiency and accuracy**, assisting future wetland delineation and management.

Pros & Cons

- *Kmean is efficient but additional human inspection have to be engaged*
- *CNN-based semantic segmentation training is time consuming, computing intensive during training (10h with 2 RTX 2060 GPU)*
- *Ensemble decision tree offer cost-effective approach for the delineation task*

Future perspective

- *Improving the model robustness by collecting data from a wider range of experiments*
- *Utilized PCA or other advanced feature selection approaches to preprocessing the data for potential efficiency and accuracy gain*

Cited Works

- [1] A. A. Bishop and M. V. Vrtiska, "Effects of the wetlands reserve program on waterfowl carrying capacity in the Rainwater Basin region of south-central Nebraska," *Nat. Resour. Conserv. Serv. Conserv. Eff. Assess.*, vol. 39, no. 2, p. 51, 2008.
- [2] C. L. Zweig, M. A. Burgess, H. F. Percival, and W. M. Kitchens, "Use of unmanned aircraft systems to delineate fine-scale wetland vegetation communities," *Wetlands*, vol. 35, no. 2, pp. 303–309, 2015
- [3] O. Mutanga, E. Adam, and M. A. Cho, "High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 18, no. 1, pp. 399–406, 2012.
- [4] A. L. Long, K. M. Kettenring, C. P. Hawkins, and C. M. U. Neale, "Distribution and Drivers of a Widespread, Invasive Wetland Grass, *Phragmites australis*, in Wetlands of the Great Salt Lake, Utah, USA," *Wetlands*, vol. 37, no. 1, pp. 45–57, 2017.
- [5] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and M. Motagh, "Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 13–31, 2017.
- [6] T. Liu and A. Abd-Elrahman, "Multi-view object-based classification of wetland land covers using unmanned aircraft system images," *Remote Sens. Environ.*, vol. 216, no. June, pp. 122–138, 2018.
- [7] T. Liu and A. Abd-Elrahman, "An object-based image analysis method for enhancing classification of land covers using fully convolutional networks and multi-view images of small Unmanned Aerial System," *Remote Sens.*, vol. 10, no. 3, 2018.
- [8] T. Liu, A. Abd-Elrahman, A. Zare, B. A. Dewitt, L. Flory, and S. E. Smith, "A fully learnable context-driven object-based model for mapping land cover using multi-view data from unmanned aircraft systems," *Remote Sens. Environ.*, vol. 216, no. July, pp. 328–344, 2018.
- [9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Van-derplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Van-derplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011

Thanks!