

*Assessing Classification Workflows for On-Pasture Forage Species ID
From High Resolution UAV Imagery*

Dylan Cicero
Yale School of Forestry and Environmental Studies
Master of Environmental Management, 2019
December 15, 2018

Background and Objectives

My goals for this research were (1) to assess the efficacy of classification workflows for on-pasture forage species identification based on high resolution UAV imagery, and (2) to contrast workflows and describe future steps for refinement.

Such a classification technique, if proven successful, could have substantial implications for both farmers and agricultural scholars. From a practical management standpoint, the savvy livestock grazer constantly assesses the quality and distribution of forages on her pasture. She then uses this information to both develop new management plans ex-ante, or evaluate prior management plans ex-post. A birds-eye view classification scheme might offer a more accurate inventory than a random sample or scattered observation. Likewise, for researchers, an accurate image-based species classification tool could be useful to investigate, for instance, how a given species responds to water stress, or how a given species responds following a grazing event.

Methodology

Data was acquired from an October 26, 2018 drone mission at Beavertides Farm in the northwest corner of Connecticut. The mission consisted of five flights at varying altitudes over a 1.2-acre parcel of pastureland. Data was collected using a DJI Phantom 4 Pro UAV (including an internal RGB camera), plus an attached MAPIR Survey 3 RGN camera. The DJI Ground Station Pro iPad application was used to program the flight route into the UAV (see Appendix 1 for flight settings). Five easily identifiable ground-control points were scattered in the flight path.

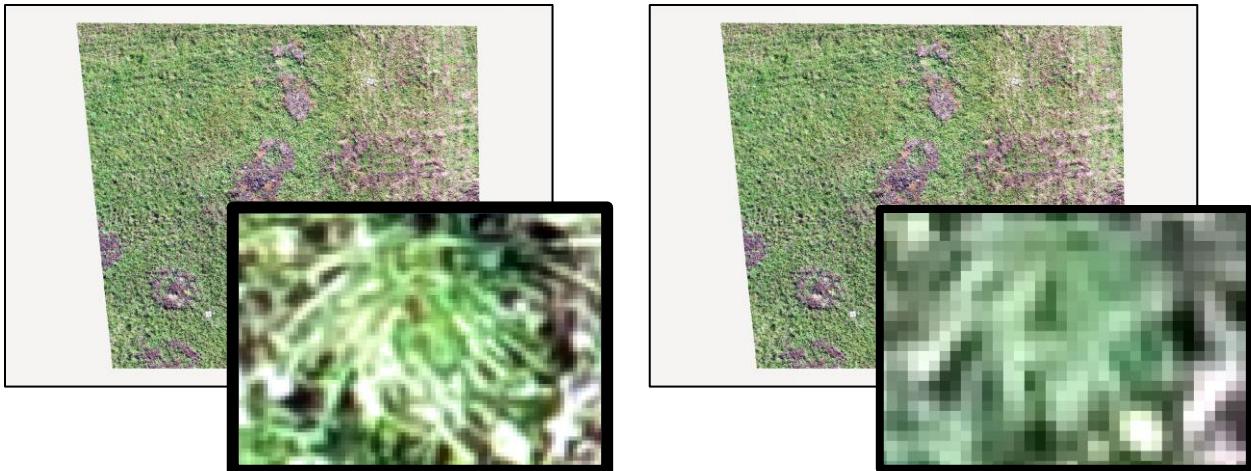
Pix-4D software was used to generate orthomosaics for the RGB and RGN images for each flight. Orthomosaics were brought into ArcGIS Pro, where red, green, blue, and near-infrared bands were composited for each flight based on ground control points. A $\frac{1}{4}$ -acre segment of the original flight path was chosen for the study site; here, a particularly heterogeneous species composition as well as patches of bare or non-vegetated soil, made it a prime candidate for classification workflow trials (Image 1). Ultimately, two sets of orthomosaics were selected for comparison: the higher resolution mosaic, derived from a flight altitude of 24.5 meters, consisted of 0.6cm pixels, while the lower resolution mosaic, derived from a flight altitude of 58.2 meters, consisted of 1.48cm pixels (Image 2).

Image 1



A $\frac{1}{4}$ -acre study site was selected from all pastureland across the 60-acre farm; here, a particularly heterogeneous species composition as well as patches of bare or non-vegetated soil made it a prime candidate for classification workflow trials

Image 2



Two sets of orthomosaics were selected for comparison. The higher resolution mosaic consisted of 0.6cm pixels, while the lower resolution mosaic consisted of 1.48cm pixels

Each orthomosaic was clipped to the study area. Then pixel-based and object-based supervised classification methods were applied to both. Pixel-based methods calculate spectral statistics for each pixel of a given image; classification is achieved for each pixel based on similarities to the spectral statistics of pixels in a pre-defined training set. Object-based methods involve the additional step of segmentation. First, neighboring pixels with similar spectral attributes are grouped together as discrete objects. Then spectral statistics, and also spatial and textural statistics are calculated for each object. Classification is achieved for each object based on similarities to the spectral, spatial, and textural statistics of objects in a pre-defined training set. The ability to consider spatial and textural attributes adds additional dimensionality to object-based classification.

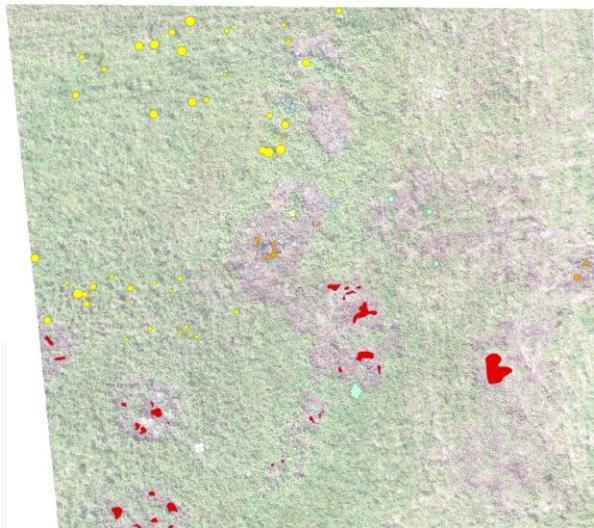
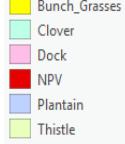
Furthermore, two different classification schemas were trialed for each workflow.

Schema 1 was more detailed, attempting to differentiate five different forage types-

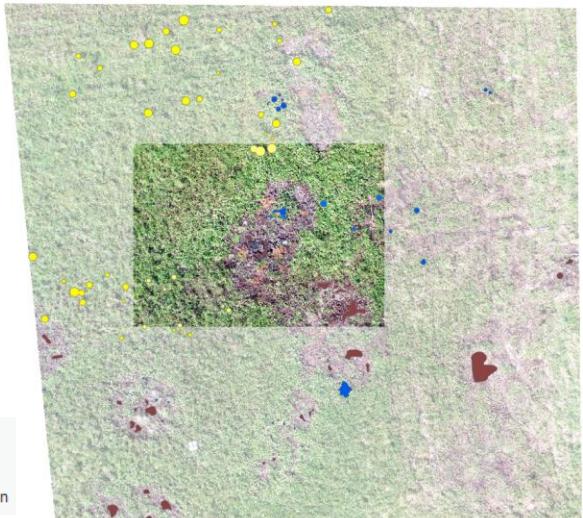
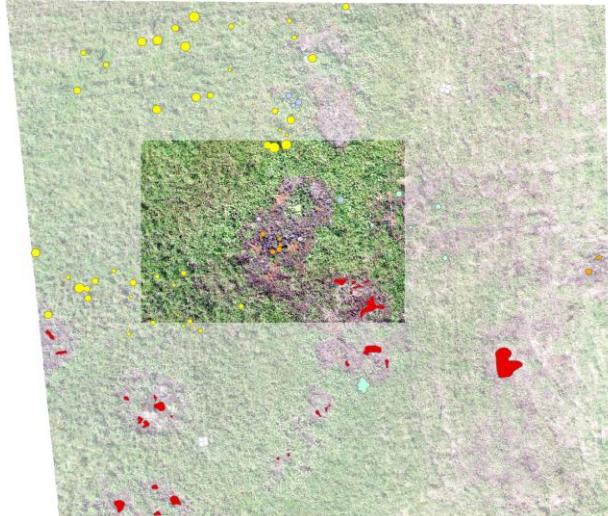
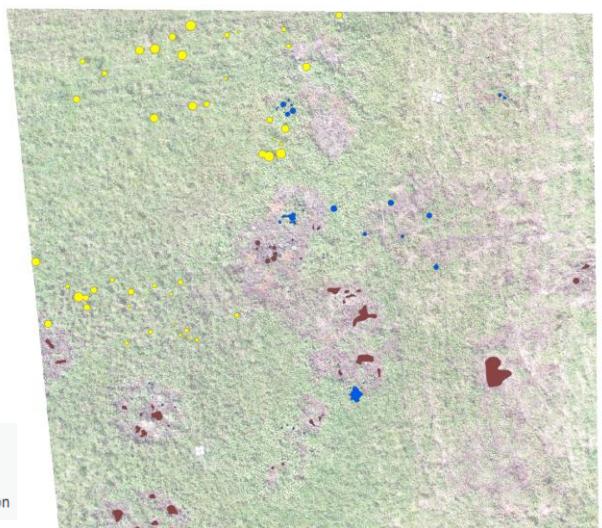
bunchgrasses, thistle, clover, plantain, and dock- as well as non-photosynthetic vegetation and bare soil. *Schema 2* was less detailed, attempting to differentiate bunchgrasses from other vegetation from non-vegetated surfaces. To develop training datasets and reference datasets (from which accuracy of each workflow would be assessed), I first scrutinized the higher resolution imagery, and attempted to manually label a large and representative set of pixels for each class. I then demarcated a smaller zone within the site that was representative of all labeled classes. Labeled pixels outside of the demarcation bounds were used for the training set. Labeled pixels inside of the demarcation bounds were used for the reference dataset. Image 3 depicts the above processes for both schemas. Tables 1 and 2 describe the number of pixels in the training and reference datasets for both schemas.

Image 3

Schema 1

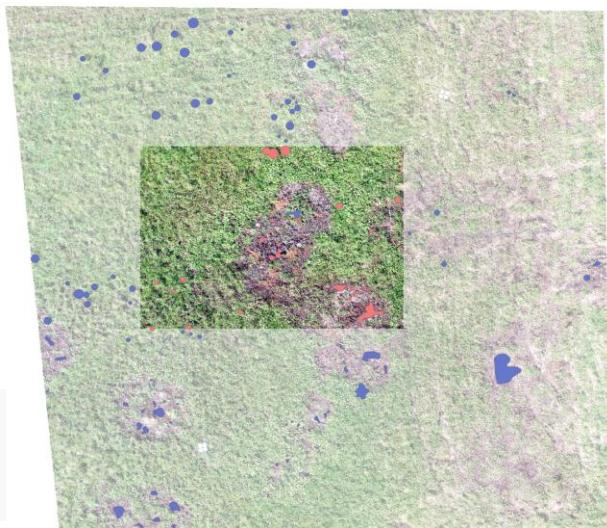


Schema 2



Training_Set

Reference_Set



Depiction of training and reference data for each schema. A large and representative sample of pixels were manually labeled for each class. Pixels outside of the demarcation zone were used for the training datasets. Pixels inside of the demarcation zone were used for the reference datasets.

Table 1

Class	Pixels Sampled (Training)	Pixels Sampled (Reference)
Bunch_Grasses	98885	25081
Thistle	4842	2493
Clover	17413	3479
Plantain	12580	4758
Dock	1685	1445
Bare_Soil	5002	10289
NPV	98053	23566
TOTAL	238,460	46,030

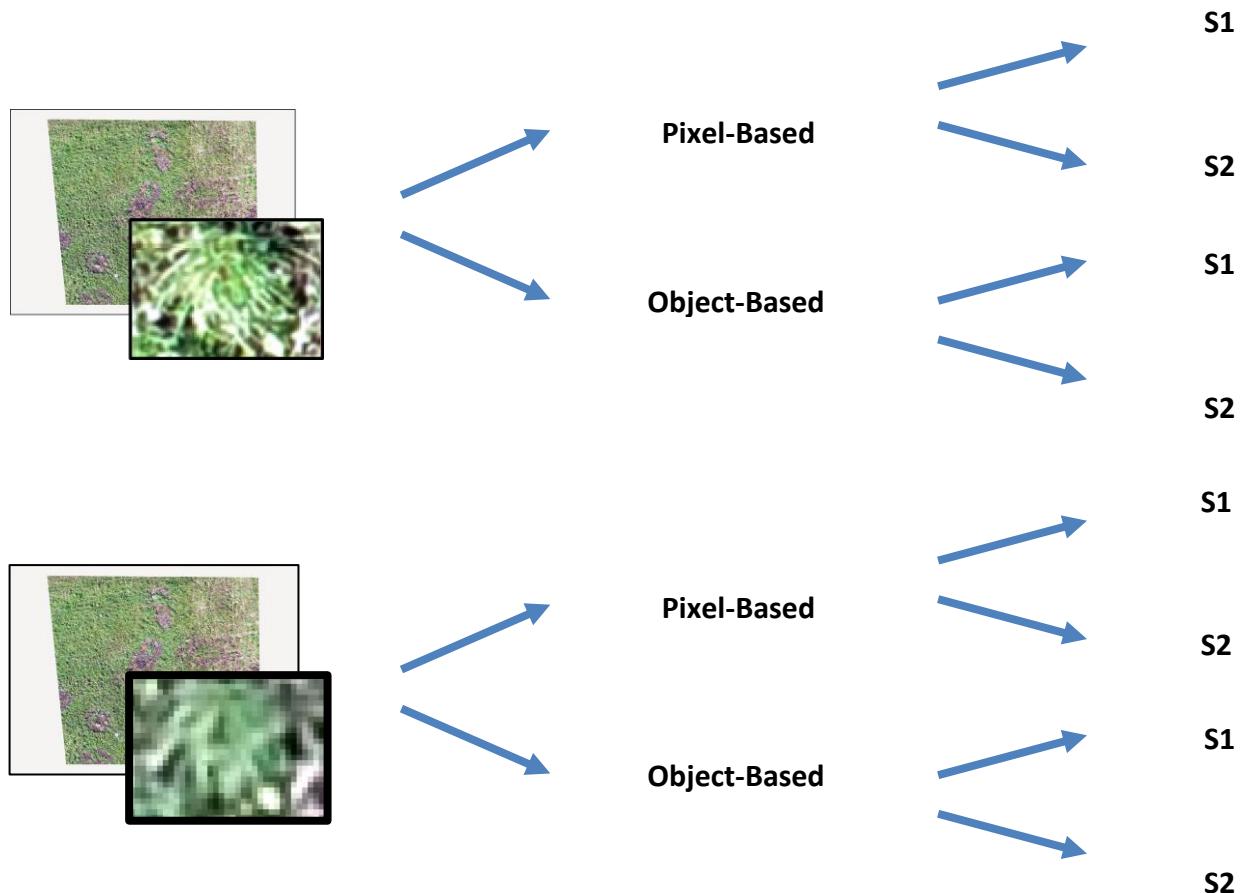
Table 2

Class	Pixels Sampled (Training)	Pixels Sampled (Reference)
Bunch_Grasses	98885	25081
Other Vegetation	36520	12175
Non-Vegetated	103055	33855
TOTAL	238,460	46,030

Description of the total number of pixels sampled in the training and reference datasets for each schema

In total, eight different classification workflows were trialed (Image 4) using ArcGIS Pro software. Training data along with the complete site orthomosaics were fed into a support-vector classifier (red, green, blue, and near-infrared bands were used). For the object-based workflows, default parameters were kept for segmentation. Likewise, for the object-based workflows, I selected “average chromaticity color”- a spectral measure of the proportion of each band to the total reflectance for all bands, “standard deviation”- a textural measure that assesses the smoothness or roughness of pixels in an object, and “compactness”- a spatial measure- as the attributes from which objects were classified.

Image 4



Eight different classification schemes were trialed

Results

Images 5 and 6 display the classification results for the *Schema 1* and *Schema 2* workflows. Tables 3 and 4 display the confusion matrices for the *Schema 1* and *Schema 2* workflows. Confusion matrices were computed by randomly sampling pixels from the reference datasets and seeing how those pixels compare to the classified dataset. The matrices detail the percentage of pixels accurately classified for each class (known as the “producer accuracy”) as well as the frequency that a given class was confused with another class in the schema. Most importantly, each matrix presents a kappa statistic, which indicates the overall likelihood of

Image 5

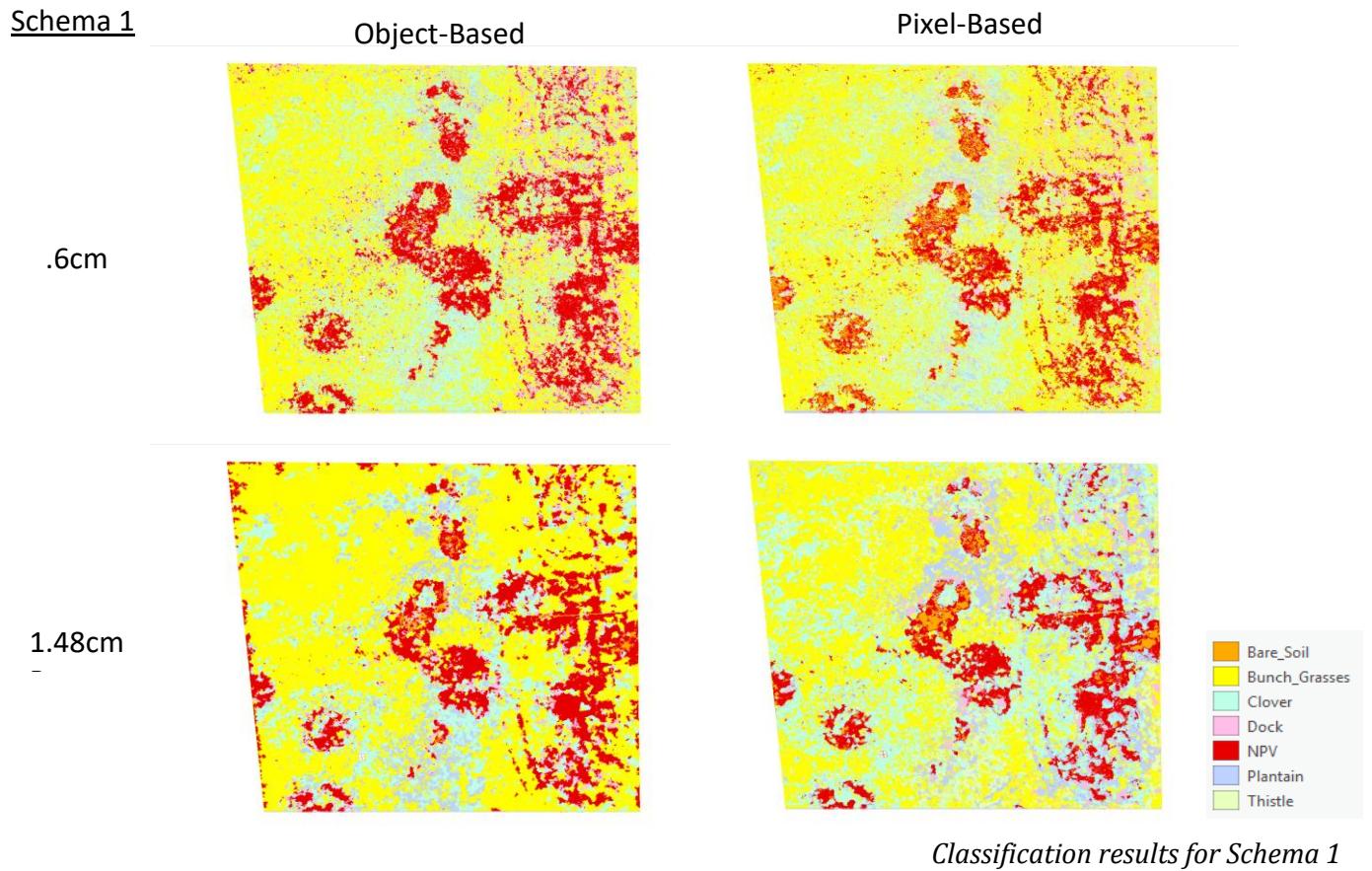


Image 6

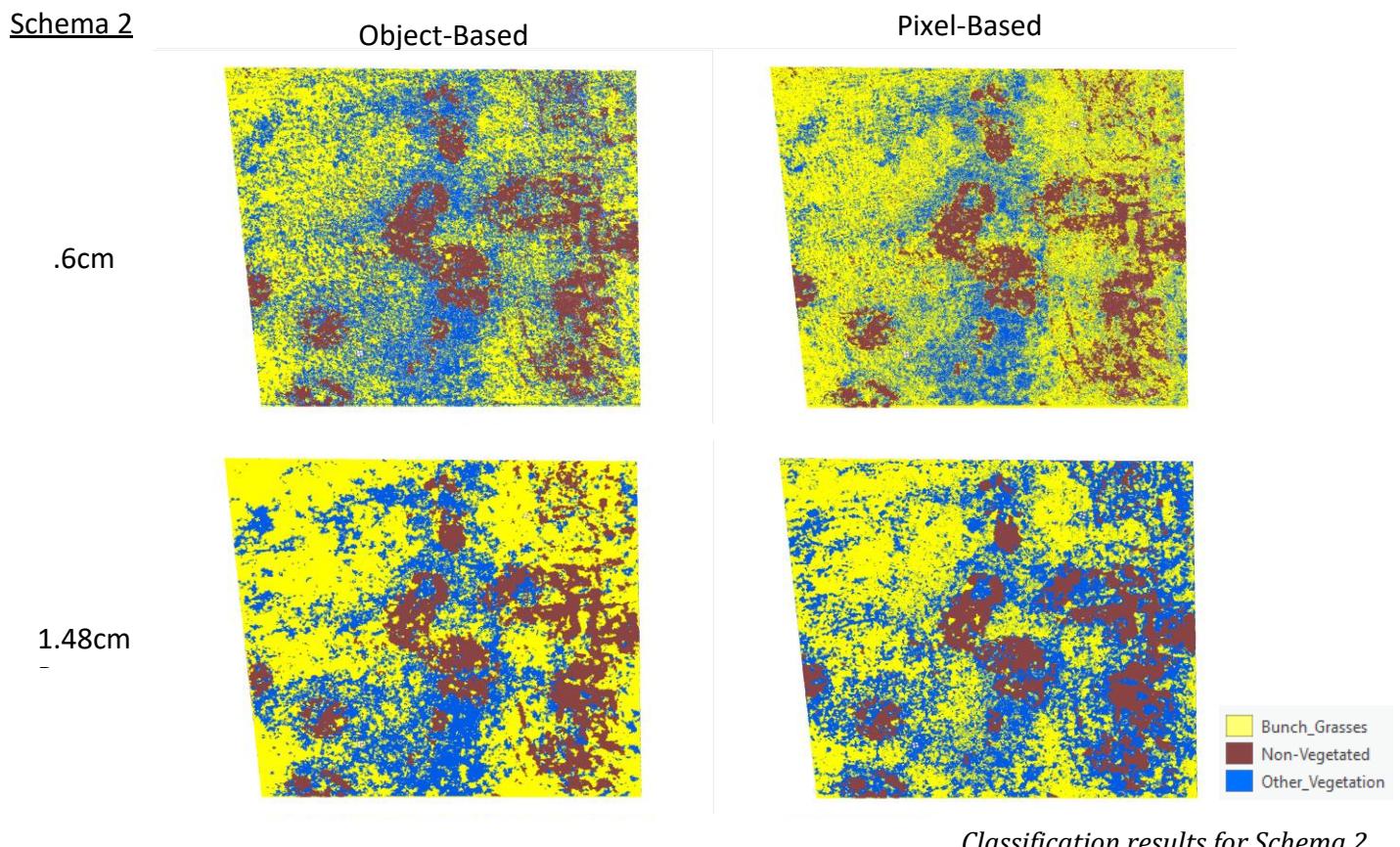


Table 3

Object-Based

Pixel-Based

	Bunch_Grasses	NPV	Bare_Soil	Thistle	Clover	Plantain	Dock	Total	U_Accuracy	Kappa	Reference
Bunch_Grasses	495	0	0	19	0	260	36	810	0.61	(a)	
NPV	0	496	177	0	1	26	349	1049	0.47	(a)	
Bare_Soil	0	0	259	0	0	0	41	300	0.86	(a)	
Thistle	0	0	0	126	18	0	0	144	0.88	(a)	
Clover	4	0	0	204	457	38	12	715	0.64	(a)	
Pantain	1	0	36	142	0	135	5	319	0.42	(a)	
Dock	0	0	28	9	24	5	57	123	0.46	(a)	
total	500	496	500	500	500	464	500	3460	0	(a)	
P_Accuracy	0.99	1.00	0.52	0.25	0.91	0.29	0.11	0.00	0.59	(a)	
Kappa	0	0	0	0	0	0	0	0	0.52	(a)	

Confusion matrices for Schema 1. Kappa statistic indicated in red

Table 4
Object-Based

Pixel-Based

	Reference				U_Accuracy	Kappa
	Bunch_Grasses	Other_Vegetation	Non_Vegetated	Total		
Bunch_Grasses	496	0	2	498	1.00	-
Other_Vegetation	0	413	62	475	0.87	-
Non-Vegetated	2	84	319	405	0.79	-
Total	498	497	383	1378	0.00	-
P_Accuracy	1.00	0.83	0.83	0.00	0.89	-
Kappa	0	0	0	0	0.84	-

	Bunch_Grasses	Other_Vegetation	Reference	Non_Vegetated	Total	U_Accuracy	Kappa
Bunch_Grasses	474	0	32	506	0.94	0.94	0.94
Other_Vegetation	16	464	194	674	0.69	0.69	0.69
Non-Vegetated	9	36	237	282	0.84	0.84	0.84
Total	499	500	463	1462	0.00	0.00	0.00
P_Accuracy	0.95	0.93	0.51	0	0.80	0.80	0.80
Kappa	0	0	0	0	0	0.70	0.70

Confusion matrices for Schema 2. Kappa statistics indicated in red

agreement of the classified image to the reference dataset normalized for chance occurrence. For the more detailed schema, the highest kappa statistic ($\kappa = 0.59$) was achieved for the lower resolution, pixel-based workflow. For the less detailed schema, the highest kappa statistic ($\kappa = 0.84$) was achieved for the higher resolution, object-based workflow. A close second was the lower resolution, pixel-based workflow statistic ($\kappa = 0.76$).

Discussion

A kappa statistic of 0.59 indicates moderate agreement between the classification workflow and training data. A kappa statistic of 0.84 indicates near-perfect agreement. Thus, we might conclude that as of now, none of the trialed classification workflows to differentiate specific types of forage are accurate enough to have any useful applications. However, classification workflows to more generally distinguish bunchgrasses, other types of vegetation, and non-vegetated surfaces might be worthy of application. A few points for discussion follow:

First, confusion matrices seem to indicate possibilities for further workflow improvement for both the more detailed and less detailed schemas. For *Schema 1*, though the pixel-based approaches achieved higher kappa scores, the object-based methods seem to hold more promise for future work if the schema is adjusted slightly. Here, the accuracy of the object-based methods was highly skewed by the poor ability of the model to discriminate dock and thistle. For example, in the lower resolution, object-based approach, only 11% of pixels that I labeled as dock in the reference data were actually classified as such, and only 25% of pixels that I labeled as thistle were actually classified as such. However, an outstanding 94% of clover pixels were correctly labeled in the lower resolution object-based workflow, and a respectable 63% of plantain pixels were correctly labeled in the high-resolution object-based workflow- numbers

that far exceed the classification accuracy of clover and plantain in the pixel-based workflows.

That is to imply that if dock and thistle were removed from the classification schema, differentiation of bunchgrasses from clover and plantain could probably be achieved with a high degree of accuracy using object-based methods.

For *Schema 2*, although a kappa of 0.84 is already very good, there still seems to be opportunities for further refinement. Consider that “other vegetation” was classified correctly 93% of the time in the lower resolution object-based workflow, but only 83% of the time in the higher resolution object-based workflow, and that “non-vegetated” surfaces were classified correctly only 51% of the time in the lower resolution workflow but 83% of the time in the higher resolution workflow. Perhaps, these statistics point towards a two-step classification approach. For example, if first, (1) object-based classification is run using the lower resolution imagery, then (2) the resultant classified image is used to mask out all “other vegetation” pixels from the high-resolution imagery, and finally (3) the masked high-resolution image is used to differentiate bunchgrasses from non-vegetated surfaces, I would expect a higher resultant kappa.

Second, it is worth emphasizing that my results are only valid to the extent that my accuracy assessment is representative of true on-the-ground conditions. This may or may not be the case. Though I aimed to select a large and representative set of pixels to train and evaluate the classifiers, my larger priority was to select “safe” clusters of pixels for each class. For instance, consider images 7, 8, 9, and 10, which showcase some non-UAV photographs of the study site captured from waist height. Image 7 is a cluster of bunchgrasses. Image 8 is (predominantly) a cluster of plantain. Image 9 is (predominantly) a cluster of clover. Image 10 is a cluster of mixed forages. Clusters as represented in images 7, 8, and 9 were the basis for both

Image 7



A cluster of bunchgrasses

Image 8



A cluster of (predominantly) plantain

Image 9



A cluster of (predominantly) clover

Image 10



A cluster of mixed forages

Clusters, like those seen in images 7, 8, and 9 were the basis for both training data and reference data. Exclusion of "messy" clusters, like that seen in image 10 might be a cause of bias in the assessment of accuracy for each classifier

training the classifiers and assessing their accuracy. Clusters as represented in image 10, however, were excluded. Ultimately, therefore, if the accuracy of the classifier (for whatever reason) varies according to the homogeneity of a given pixel in its neighborhood, the accuracy assessment would be biased. Future work should examine the degree to which this is true, by identifying reference data across a normal distribution of cluster extents, identifying reference data across a uniform distribution of cluster extents, and comparing accuracy assessment results for a given classifier based on differing reference datasets.

Third, while I was anticipating more accurate classification results from the more dynamic object-based workflows, there seems to be no such clear trend. In retrospect, the reasons for this are obvious. Consider that the segmentation step in the object-based classifier assumes that each discrete object in an image has a consistent (or near-consistent) spectral signature. This is indeed the case for perhaps black-shingled suburban rooftops, amid a landscape of green trees, blue water, and black-paved roads (here, an object-based classifier could easily distinguish the rooftops from the roads, because the rooftops tend to be more compact; a pixel-based classifier would miss this critical information). However, in attempting to classify species of forage, my “objects” of interest- clusters of clover, clusters of bunchgrasses, clusters of plantain- had variable spectral signatures. So, for instance, at its most effective, the segmentation step might segment a single plantain leaf or clover petal from surrounding plantain leaves or clover petals, but it would never segment clusters of plantain leaves or clover petals from surrounding clusters of bunchgrasses. (See Appendix 2 for segmented images used in object-based classification workflows). An innovative approach for future trials might seek to create a segmentation procedure in which objects are segmented based on the uniformity of their textural characteristics. For example, even though a cluster of clover

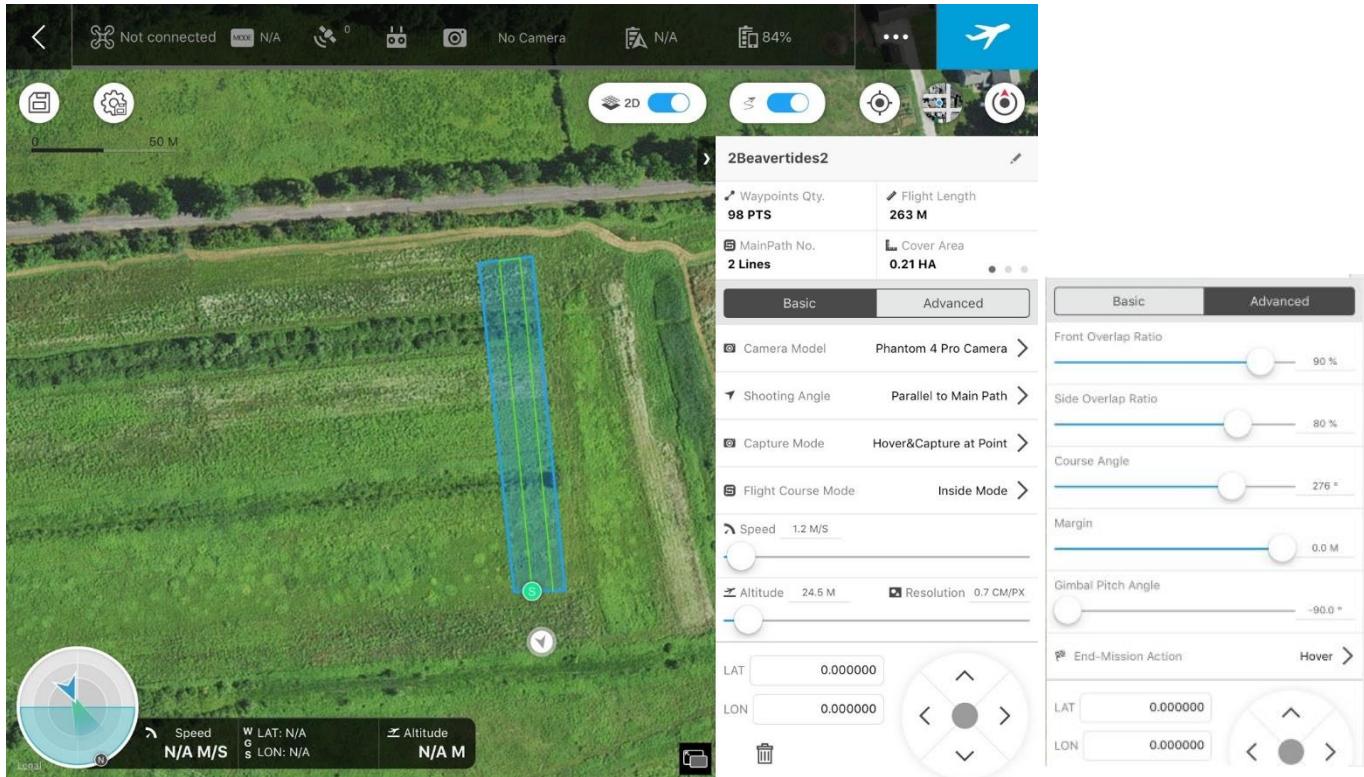
would never be segmented from surrounding bunchgrasses in pasture based on spectral consistency, that cluster of clover likely would have a uniform spectral standard-deviation across pixels (textural roughness) and on this basis might be differentiable from surrounding bunchgrasses.

Lastly, while I was anticipating more accurate classification results from the higher resolution orthomosaics, here too, there seems to be no such trend. This is particularly important due to the large differential in UAV data acquisition time for the higher resolution and lower resolution imagery. In fact, data acquisition was 5.5 times faster for the 1.48 cm/pixel orthomosaic (91 seconds/acre) than the 0.6cm/pixel orthomosaic (502 seconds/acre). Thus, inquiries with such ultra-high sub-centimeter UAV imagery for this purpose might be futile.

Conclusion

This study assessed on-pasture forage species classification workflows from high resolution UAV imagery. Object-based and pixel-based methods were contrasted for imagery at two different spatial resolutions and for two different classification schemas. Results indicate that such methods, especially for the less detailed classification schema are promising. Synthesis of results points to several recommendations for future workflow refinement.

Appendix 1



DJI Ground Station Pro flight settings

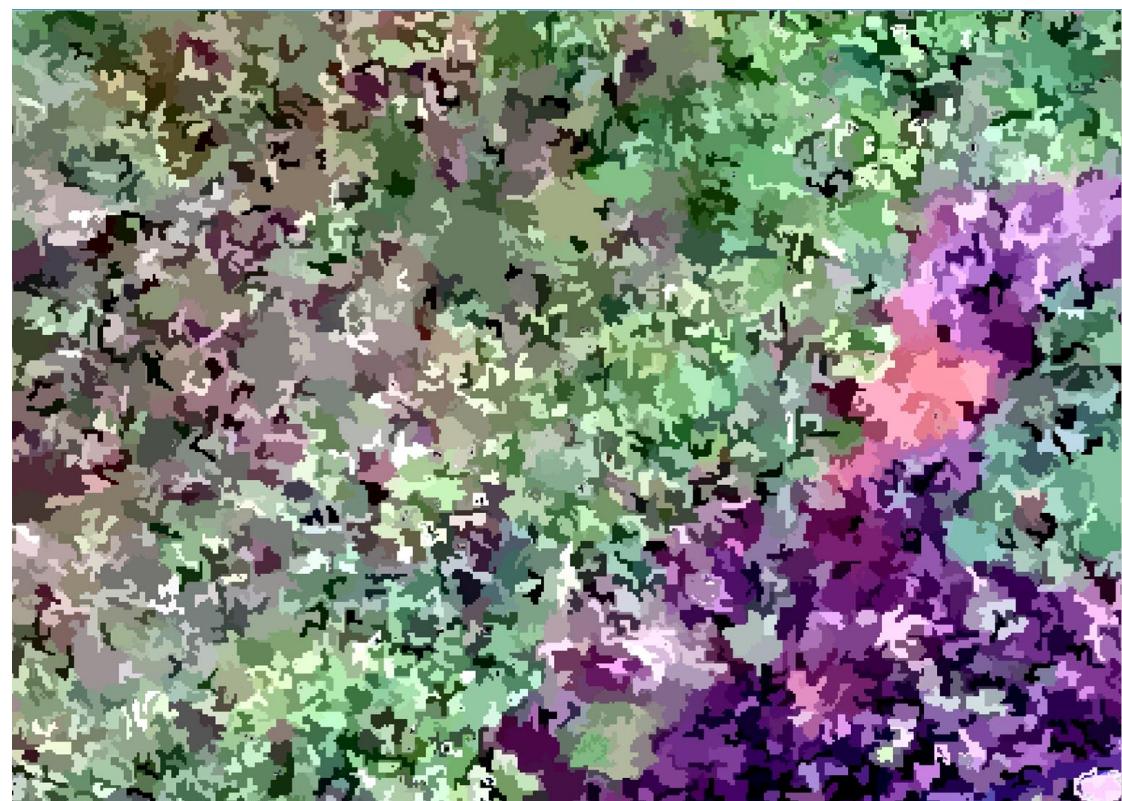
Appendix 2



Segmented
.6cm



Segmented
1.48cm



Example output from segmentation steps in object-based classification workflows