CLASSIFICATION OF CROP RESIDUE COVER IN HIGH-RESOLUTION RGB IMAGES USING MACHINE LEARNING

Parth C. Upadhyay[1], John A. Lory[2], Guilherme N. DeSouza[1],
Timotius A. P. Lagaunne[2], Christine M. Spinka[2]
[1]Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, Missouri, USA.
[2]Division of Plant Sciences, University of Missouri, Columbia, Missouri, USA.

Highlights
- A machine learning framework estimated residue cover in RGB images taken at three resolutions from 88 locations.
- Best results primarily used texture features, the RFE-SVM feature selection method, and the SVM classifier.
- Accounting for shadow and plants plus modifying and optimizing texture features may improve performance.
- An automated system developed using machine learning is a viable strategy to estimate residue cover from RGB images obtained by hand-held or UAV platforms.

ABSTRACT. *Maintaining plant residue on the soil surface contributes to sustainable cultivation of arable land. Applying machine learning methods to RGB images of residue could overcome the subjectivity of manual methods. The objectives of this work were to use supervised machine learning while identifying the best feature selection method, the best classifier, and the most effective image feature types for the task of classifying residue levels from RGB imagery. Imagery was collected from 88 locations in 40 row crop fields in five Missouri counties between mid-April and early July in 2018 and 2019 using a tripod mounted camera (0.014 Ground Sampling Distance (GSD; cm pixel^{-1})) and unmanned aerial vehicle (0.05 and 0.14 GSD). At each field location, 50 contiguous 0.3 × 0.2 m region of interest (ROI) images were extracted from imagery resulting in a dataset of 4,400 ROI images at each GSD. Residue percentage for ground truth was estimated using a bullseye grid method (n=100 points) based on the 0.014 GSD images. Representatives of color, texture and shape features were extracted and evaluated using four feature selection methods and two classifiers. Recursive Feature Elimination using Support Vector Machine (RFE-SVM) turned out to be the best feature selection method, and the SVM classifier performed the best for classifying residue for a three-class problem. All best features for this application were associated with texture, with Local Binary Pattern (LBP) features being the most prevalent for all three GSD*

images. Shape features were irrelevant. Three residue level classes were correctly identified with 88%, 84% and 81% 10-fold cross validation scores on the 2018 training data and 81%, 69% and 65% accuracy on the 2019 testing dataset in decreasing resolution order. Converting image-wise data (0.014 GSD) into location residue estimates using a Bayesian model documented good agreement with the location-based ground truth ($r^2 = 0.90$). This initial assessment documented the potential of RGB images to match other methods of estimating residue with potential to replace or be used as quality control for transect methods.

A key indicator of sustainability in row crop systems is maintenance of residue on the soil surface (Bronick et al., 2005; Singh et al., 2007; Ranaivoson et al., 2017; Searle et al., 2017; Cherubin et al., 2018). Plant-derived residues on the surface of the soil have many benefits including reduced impact of rain drops on the soil surface and increased soil carbon sequestration. Ultimately, these benefits decrease the breakdown of soils aggregates leading to reduced soil erosion and increased infiltration with associated improvements in soil conservation, productivity, and surface water quality (Lal, 2008; 2009; Ranaivoson et al., 2017; Cherubin et al., 2018).

In the US, The Food Security Act of 1985 (U.S.C, 1985) established a requirement to maintain "sustainable erosion rates" on cropland, hayland and pasture defined as "highly erodible land" (HEL). First, the Natural Resource Conservation Service (NRCS) makes the determination if an agricultural tract is HEL. Next, a key element of this assessment is determining residue cover on the field, excluding green plant cover, after planting but before the grain crop obscures the soil surface (USDA-NRCS, 2011). Trained NRCS employees visit thousands of randomly selected tracts per year to assess HEL compliance. There are known issues with the line-transect method

including a bias to overestimate residue, reader bias, and the inability to provide continuous estimates of residue across a field (Laflen et al., 1981; Richards et al., 1984; Laamrani et al., 2017; Lory et al., 2021). There is an extensive history of evaluating alternatives to the line-transect method. Other techniques using human judgement to estimate crop residue include visual estimate, point intercept, meter stick, spiked wheel, photograph comparison, and photographic-grid methods (Laflen et al, 1981; Dickey et al., 1989; Morrison Jr. et al., 1993, 1995; Laamrani et al., 2017; Lory et al., 2021). All these approaches are also time consuming assessments of residue that are subjective, and heavily dependent on an individual's experience and focus. A reliable, consistent and automated approach that eliminates subjectivity and provides documentation of the assessment would be of great benefit to government technical staff and others assessing residue cover.

The availability of economical options for unmanned aerial vehicles (UAV's) and cell phones has renewed interest in applying modern machine learning techniques to high-resolution (<0.2 cm pixel^{-1} ground sampling distance (GSD; cm pixel^{-1})) RGB images to estimate residue cover in agricultural systems. Kavoosi et al. (2020) obtained a correlation of r^2=0.84 with crop residue using a segmention strategy on RGB images collected from 10 farms at mid-day using a UAV flying at 5- to 10-m above the ground. Riegler-Nurscher et al. (2018) obtained an r^2=0.84 for residue classification in a 99-image testing dataset for a pixel-by-pixel classifier trained on around 200 training images. Images were obtained using multiple hand-held cameras under cloudy light conditions. Bauer and Strauss (2014) obtained a correlation of r^2=0.75 using an object-based image analysis methodology (OBIA) to quantify residue from the 61 RGB images collected by a hand-held camera in diffuse sunlight conditions. Laamrani et al. (2018) developed a mobile phone application for crop residue cover mapping using RGB images based on the algorithm that used image processing techniques and an automated color threshold for image classification, and they

achieved a correlation of r^2=0.86 on a dataset of 54 images collected from 18 fields. In summary, research to date assessing residue in high-resolution RGB images from hand held cameras and UAV's have shown potential for success but the work has been on smaller datasets and has not fully explored opportunities with supervised machine learning. More broadly, multiple researchers have had success using color, shape and texture features from RGB images to answer agriculture related questions (Schmittmann et al., 2017; Yang et al., 2015; Yuan et al., 2019; Ojala et al., 2002; Le et al., 2019; Chaugule et al., 2014). A more aggressive application of supervised machine learning techniques has potential to improve residue estimates from high-resolution RGB images.

Our goal was to use a classification-based supervised machine learning system to estimate percent residue cover in row crop fields from high-resolution ($\leq$0.15 GSD) RGB images. This required: (i) determining the best feature selection tool and model classifier for a residue application; (ii) investigate the efficacy of a wide range of extracted image feature types for their utility to classify residue; (iii) determine if GSD/sensor type affected the types of selected features; and (iv) identify and test a model to convert multiple classified images into a location-wise estimate of percent residue cover.

MATERIALS AND METHODS

ACQUISITION OF RESIDUE IMAGERY

Imagery used in this project was collected as part of an associated project. An abbreviated description of the key details is included here; additional information on acquisition, processing and estimating percent residue cover for the ground truth images can be found in Lory et al. (2021). Imagery used in this project was obtained from 40 Missouri row-crop fields between early May and early July of 2018 and 2019 (Table 1). Imagery was collected from multiple locations in each field, actively seeking locations in each field respresenting different residue conditions. Ultimately,

imagery from one to four locations in each field was retained for use in this project to provide a diverse range of residue types and residue coverage (Table 1). The final dataset included 88 field locations, 60 field locations in 2018 (32 planted to corn and 28 planted to soybean) and 28 field locations in 2019 (6 planted to corn and 22 planted to soybean).

At each of the 88 field locations, a 15.24-m tape was placed at 45 degrees to the planted row direction. Three sets of images were then obtained of each tape. The first set of images were obtained at 1.0-m above the ground surface on a tripod-mounted Canon (Canon USA, Melville, NY) EOS Rebel T6i Digital single lens reflex (DSLR) camera equiped with a 24 mm stepper motor technology lens. The sensor had 24.2 MP resolution which generated image size of 6000 x 4000 pixels. Estimated GSD for these images was 0.014 cm pixel^{-1}. Typically, 50 images were obtained by moving the tripod 30 cm between images along the tape, starting with the first image centered over the 0 to 30 cm section of the tape.

The second and third sets of images were obtained at 2- and 5-m using DJI Phantom 4 PRO UAV (DJI, Shenzhen, China) including a camera with a 24-mm lens and a sensor with 20 MP resolution which generated an image size of 5472 x 3648 pixels. The UAV was flown manually over the tapes on days when wind speeds were below 7 m s^{-1}, starting at the 0 point on the tape, with at least 65% overlap between consecutive images along the tape. In the 2-m images, typically 17 images were taken along each tape and in the 5-m image typically 12 images were taken along each tape. Estimated GSD of the 2-m images was 0.05 cm pixel^{-1} and 5-m images was 0.14 cm pixel^{-1}. To minimize parallax effects, all images were taken horizontal to the soil surface (90-degree angle to the ground). Images were collected under a diversity of light conditions, typically between 8:00 AM and 5:00 PM central daylight time. Figure 1 includes an example of each of the three image types.

Table 1. Approximate location, sampling date, crop, and residue details for 2018 and 2019 study sites. Exact locations were not provided to protect farmer privacy.

ID	County	Approximate Site Location	Sampling Date	Locations at Site	Crop (Stage)[1]	Dominant Residue[1]	Other Issues	Location Residue Ground Truth (%)
2018-01	Audrain	39.260, -92.272	8-May-18	3	C (V2)	SB	-	14, 9, 5
2018-02	Audrain	39.257, -92.267	20-May-18	3	NE	C, TL	-	48, 57, 54
2018-03	Audrain	39.264, -92.270	8-May-18	4	C (V2)	SB	-	10, 9, 21, 1
2018-04	Audrain	39.269, -92.265	8-May-18	3	NE	C, SB	-	41, 44, 49
2018-06	Audrain	39.267, -92.266	10-May-18	4	NE	C	OE (4)	57, 58, 54, 27
2018-07	Audrain	39.270, -92.212	14-May-18	3	C (V2)	SB	S (1,2), UE (3)	13, 13, 5
2018-09	Audrain	39.257, -92.150	29-May-18	3	SB (V2.5)	C	S (1,2,3)	23, 23, 23
2018-11	Callaway	39.029, -92.078	30-May-18	3	SB (V1)	WSG, C	WL (1,2)	97, 91, 89
2018-15	Boone	39.215, -92.204	23-May-18	3	SB (V1)	C	S (1)	57, 66, 51
2018-16	Boone	39.212, -92.203	30-May-18	3	SB (V1)	C	S (1,2,3)	57, 60, 57
2018-20	Cooper	38.778, -92.657	17-May-18	3	C (V3)	SB	-	21, 27, 26
2018-21	Cooper	38.782, -92.654	17-May-18	3	C (V3)	WSG	-	25, 27, 24
2018-22	Cooper	38.817, -92.623	18-May-18	3	SB (VE)	C, W, SB	WL (1)	85, 88, 94
2018-25	Cooper	38.815, -92.623	18-May-18	3	SB (VE)	C, W	S (1,2), WL (1)	92, 85, 89
2018-26	Callaway	39.032, -92.079	4-Jun-18	3	NE	WSG	-	94, 91, 96
2018-27	Cooper	38.822, -92.620	18-May-18	3	C (V3.5)	SB, W	-	18, 20, 15
2018-40	Boone	38.901, -92.214	14-Jun-18	3	NE	SB, W	S (2,3)	3, 4, 2
2018-41	Boone	38.901, -92.210	14-Jun-18	2	SB (VE)	C, W	S (2)	44, 41
2018-42	Boone	38.900, -92.206	14-Jun-18	4	NE	SB, W	S (4), WL (3,4)	19, 12, 7, 15
2018-43	Boone	38.901, -92.210	14-Jun-18	1	SB (VE)	C, W	-	25
2019-01	Howard	39.037, -92.694	14-May-19	2	C (V1.5)	SB, C	-	9, 24
2019-02	Howard	39.039, -92.694	14-May-19	1	C (V1.5)	SB, C	S	15
2019-03	Cooper	38.914, -92.736	15-May-19	1	C (V3)	SB	-	7
2019-05	Howard	38.864, -92.843	17-May-19	1	C (V2)	SB	S	9
2019-06	Cooper	38.864, -92.843	20-May-19	1	C (V2)	SB	S, WL	19
2019-07	Cooper	38.864, -92.843	20-May-19	1	C (V2)	SB, C	-	11
2019-08	Howard	39.020, -92.573	29-May-19	2	SB (VC)	SB	S (1), WL (1,2)	68, 60
2019-09	Cooper	38.774, -92.940	30-May-19	1	SB (VE)	C, W	S, WL	7
2019-10	Cooper	38.791, -92.931	30-May-19	1	SB (VE)	C, W	S	6
2019-11	Audrain	39.283, -91.880	31-May-19	1	C (V2)	SB	WL	6
2019-12	Callaway	38.969, -92.103	3-Jun-19	2	C (V2)	SB, W, C	S (1), WL (1)	73, 71
2019-13	Callaway	39.040, -91.838	3-Jun-19	1	NE	SB	-	3
2019-14	Audrain	39.299, -91.970	4-Jun-19	2	NE	SB, C	S (1,2), WL	13, 42
2019-15	Audrain	39.241, -91.752	5-Jun-19	1	NE	SB, W	S	4
2019-16	Audrain	39.223, -91.830	5-Jun-19	2	NE	WSG	S (1,2)	27, 40
2019-17	Audrain	39.334, -91.740	11-Jun-19	1	SB (VE)	SB, W	S	41
2019-18	Audrain	39.332, -91.756	11-Jun-19	2	SB (VE)	SB, W	S (1,2)	51, 77
2019-19	Cooper	38.915, -92.739	18-Jun-19	3	SB (V1)	WSG, C	WL (2)	66, 80, 82
2019-20	Audrain	39.178, -91.742	18-Jun-19	1	NE	C, W	WL	56
2019-21	Audrain	32.299, -91.972	28-Jun-19	1	SB (V1)	C, W	S	51

[1] Crops and residue types: C=corn; SB=soybean; TL=tree leaves; WSG=winter small grain; W=weeds. Note for residue, the first entry is residue from the previous year's crop. Growth Stages: NE=not emerged; VE=vegetative stage emerged; V#= vegetative stage. Other issues: WL=weeds live; S=significant shadow; OE/UE=over/under exposure in 0.014 GSD images; (numbers in parenthesis are location numbers affected).
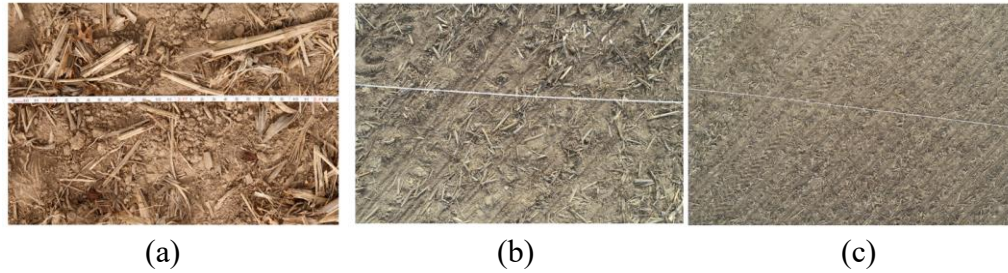
(a)   (b)   (c)

Figure 1. Photographs taken from: (a) 1m; (b) 2m and (c) 5m from the ground with 0.014, 0.05 and 0.14 cm pixel^{-1} ground sampling distance.

The region of interest (ROI) images were cropped from the larger images using the software package Adobe Photoshop (Adobe Inc., San Jose, CA). The cropped ROI images were obtained contiguous to the tape from the side of the tape determined in the field to have an acute sun angle to minimize shadows from the tape in the ROI image, starting at the 0 point on the tape. The ROI image was cropped from the picture where the ROI image was most central to the larger image to minimze the parallax effect. Each cropped ROI image contained approximately 2,400 × 1,600, 610 × 410 or 220 × 150 pixel area (approximately 0.305-m × 0.020-m surface area of the soil) for 0.014, 0.05 and 0.14 GSD images, respectively. There was no effort to adjust or normalize any images based on a gray card. This resulted in a dataset of 4400 ROI images (88 tape locations × 50 image per location) for each GSD.
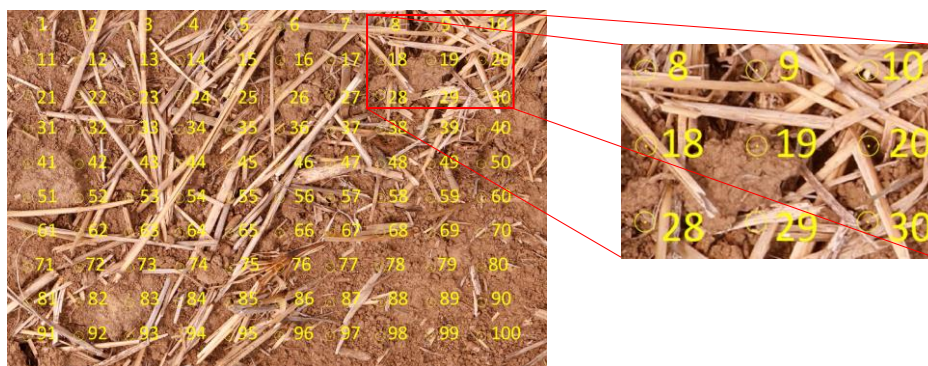

Figure 2. An examples of a 2,400 × 1,600 pixel region of interest image with 0.014 cm pixel^{-1} ground sampling distance showing the randomly assigned 100-point grid (left) and an enalarged region showing the bullseye points used for assessing residue. Radius of the circle around the assessment point is scaled to be equal to the minimum dimension for residue cover (2.4 mm) to facilitate assessment of residue. As per Bullseye method (Lory et al., 2021), to be classified as residue, residue must touch the center point or touch residue that touches the center point and cover > 50% of the area of the circle. An average of all points classified as residue was assigned as the residue cover ground truth for the image.

To determine the residue cover for each of the 4400 ROI images, a bullseye grid method with n=100 grid points was used on 0.014 GSD ROI images (Lory et al., 2021; Figure 2). The ground-truth determined on the 0.014 GSD was then assigned to associated 0.05 and 0.14 GSD cropped ROI images that shared the same surface area of the soil. Average residue cover for the field location was calculated as the mean of the 50 ROI images along the tape.

Based on the residue cover ground truth of the ROI images, they were then divided into three classes based on residue range (Table 2). In the analysis, ROI images taken in year 2018 were used as the training dataset for the model and ROI images from year 2019 were used as the testing dataset. Total number of training and testing ROI images, and associated locations is summarized in Table 2.

Table 2. Number of training and testing ROI images and locations per class.

Class No. (Residue range)	Training Dataset (2018)		Testing Dataset (2019)	
	No. of Training ROI Images	No. of Training Locations	No. of Testing ROI Images	No. of Test Locations
Low (<33.3%)	1549	31	772	14
Medium (33.3% to 66.6%)	626	17	302	8
High (>66.6%)	825	12	326	6

MACHINE LEARNING OPERATIONS ON ROI IMAGES

Overview

The project workflow to develop and evaluate a classification model based on the ROI images is summarized in Figure 3. In step one, we identified 70 typical image features of interest from the literature and extracted them from 3000 training and 1400 testing ROI images. Extracted image features were normalized using a min.-max. scaling method. In step two, the 70 normalized image features from the training dataset were ranked using four feature selection methods. In step three, the feature importance rank orders of each of the four feature selection methods from Step 2 were evaluated using a 10-fold cross validation through accuracy assessment using both the Support Vector Machine (SVM) and Random Forest (RF) classifiers. The best combination of feature selection method and classifier was determined by comparing 10-fold cross-validation scores of

the eight options (4 feature section methods × 2 classifiers). In step 4, the optimal subset of features determined in Step 3 were used to generate the final classification model using the full training dataset and the superior classifier. The resulting model was then tested using the testing dataset.
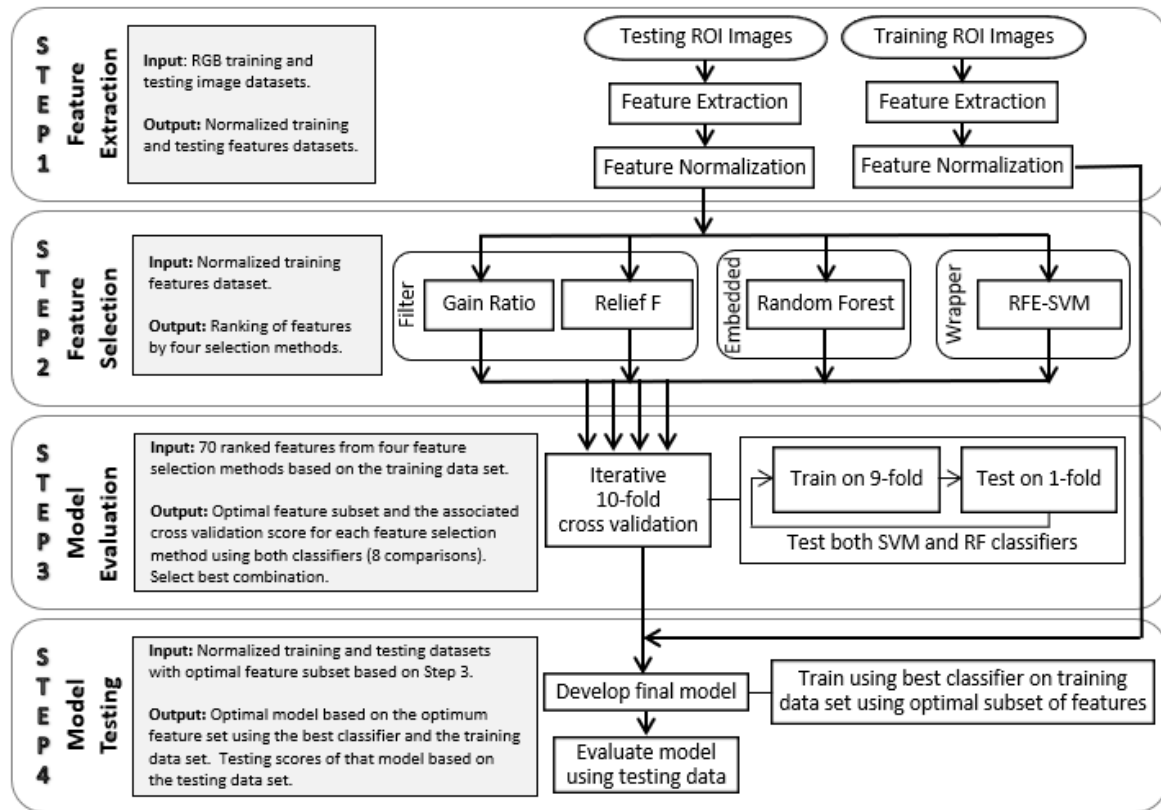

Figure 3. Flowchart for machine learning operations on ROI images.

This process was repeated on each of the ROI images obtained at each of the three resolution levels (0.014, 0.05 and 0.14 GSD). All machine learning operations were performed using the software development environment Jupyter Notebook (ver. 6.2.0, Kluyver et al., 2016) and statistical analysis was done using the software SAS (ver. 9.4, SAS Institute Inc., Cary NC). Specific libraries used for each step of the protocol are available in Supplemental table S1 https://doi.org/10.13031/16920373.v1. Additionally, access to the code and data used in this project can be requested at http://vigir.missouri.edu/Research/crop_residue.htm.

Feature Extraction (Step 1)

We used image features which have been successfully used in agricultural applications including color, texture, and shape features (Schmittmann et al., 2017; Yang et al., 2015; Yuan et al., 2019; Ojala et al., 2002; Le et al., 2019; Chaugule et al., 2014; Herrera et al., 2014).

Color features (n=24) were the image-wise mean, median, standard deviation and skewness of each of the color bands from the HSV and CIE-LAB color-spaces (24 = 6 color bands × 4 statistical representations). These color spaces were selected because the resulting features are uncorrelated with brightness. The CIE-LAB color-space contains information about coloration and luminance (Schmittmann et al., 2017), and single color features in the HSV color space are invariant to brightness (Yang et al., 2015).

Global texture features (n=13) were based on the gray-level co-occurrence matrix (GLCM) as defined by Haralick et al. (1973). The GLCM describes the joint probability of pixel pairs at any gray level that represents the texture of an image statistically. The 13 Haralick statistical features calculated from GLCM were angular second moment, contrast, correlation, sum of squares (variance), inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measure of correlation, and information measure of correlation squared (Haralick et al., 1973). These global features of the greyscale image are popular because of their simplicity and adaptability (Yuan et al., 2019).

Local texture features (n = 26) were also based on the image converted to grayscale and were represented using a local binary pattern (LBP; Ojala et al., 2002). In this study, the same LBP pattern was used for all three GSD's. The feature was calculated based on 24 equally-spaced pixels in a circle around the central pixel with a radius of three pixels; the resulting pattern will be one of 26 uniform patterns (Ojala et al., 2002). The LBP features for an image was based on the proportion of the by-pixel assessments in the image that fell in each of the 26 bins of the histogram. For

example, LBP-1 is the first bin of the histogram. The LBP was selected for local texture because it is a simple yet efficient operator and has been widely used in plant discrimination applications (Ojala et al., 2002; Le et al., 2019).

For shape features, Hu moments were used, which are a set of seven invariant moments with respect to translation, scale and rotation (Hu, 1961). From a mathematical point of view, invariant I is a function defined on the space of all admissible image functions which does not change its value under degradation operator D, i.e. which satisfies the condition $I(f) = I(D(f))$ for any image function f. To extract Hu moments features, gray-scale images were further transformed into binary images using an Otsu filter-based optimal threshold (Hu, 1961).

After extraction of the 70 features from the 4,400 images, each feature was then scaled to a range of [0,1] based on the min-max scaled values from the 2018 training dataset.

Feature Selection (Step 2)

Feature selection is used to find the most relevant features and avoid redundant and irrelevant features by selecting a subset that is most useful to the model accuracy and gives the optimal classification result. There are three major categories of feature selection techniques – filter, wrapper and embedded (Chandrashekar et al.,2014; Li, J. et al., 2016). With considerations of the distinct characteristics and advantages as shown in Table 3, we implemented four feature selection methods, including one embedded method (Random Forest), two filter methods (Gain ratio and Relief-F) and one wrapper method (RFE-SVM). In this step, we used the 2018 training dataset for ranking the features using the above four feature selection methods.

Table 3. List of feature selection methods used and their description

Feature Selection Method	Description
Random Forest	An embedded method. Employs a random method to establish a forest comprised of many mutually independent decision trees (Ho,1995).
	It is very straightforward to derive the importance of each variable on the tree decision where each tree is grown using a bagging or bootstrap sample from the training set (Hapfelmeier et al., 2013).
Gain Ratio	A filter method. Is the extension of the Information Gain (IG) and overcome the drawbacks of the IG by selecting features with large number of values (Harris, 2001).
	Gain ratio measures how much "information" a feature gives about the class. Features that perfectly partition should

	give maximal information meaning those features provide the greatest information gain; unrelated features should give no information (Novakovic et al., 2011).
Relief-F	A filter method. Uses the concept of nearest neighbors to derive feature statistics that indirectly account for interactions (Robnik et al., 2003). Relief-F does not remove feature redundancies, i.e. it seeks to select all features relevant to the endpoint regardless of whether some features are strongly correlated with others (Urbanowicz et al., 2018).
Recursive Feature Elimination-SVM (RFE-SVM)	A wrapper method. A backwards selection learning scheme is implemented to evaluate feature sets, and the accuracy of the learning scheme is estimated using cross-validation to detect the best subset (Guyon et al., 2002). Here, SVM with linear kernel as an external estimator and 10-fold cross validation was used.

Model Evaluation (Step 3)

Classification Techniques

Previous research suggested that SVM and RF are highly suitable for many agricultural applications (Guerrero et al.,2012; Le et al., 2019; Ma et al.,2017; Li, M. et al., 2016). Support vector machine is a non-parametric supervised learning classifier and one of the first classification algorithms to exploit the idea of kernel functions to build a high-dimension, non-linear space for drawing decision boundaries between classes in a pair-wise fashion (Vapnik, 1995; Burges, 1998). In this study, the 'SVC' package from the "sklearn" library (Pedregosa et al., 2011) was implemented to carry out the SVM algorithm using the radial bases function (RBF) as the kernel. For the RBF kernel we used default hyperparameters; two key parameters are the penalty parameter where C was selected as '1'and the kernel parameter where γ was selected as 'auto'.

Random forest is a popular machine learning algorithm with generally good predictive performance, low overfitting, and easy interpretability (Hapfelmeier et al., 2013; Ma et al.,2017). It uses a bagging method to generate a training dataset to grow each tree and unlabeled data are classified by assigning them to the most frequently voted class. In this study, "RandomForestClassifier" package from the "sklearn" library (Pedregosa et al., 2011) was implemented using the default hyperparameters.

Selecting optimum feature selection and classifier method

For each of the four ranked feature sets (based on 2018 data), each derived from one of the four selection methods (Step 2), model building was completed using both of the classifier methods. Ten-fold cross validation was used to calculate the mean cross validation score for the addition of

each ranked feature into the model, starting from the first rank feature, for each of the eight combinations of feature selection rank and classifier methods. Cross validation is a statistical method used to estimate the skill of machine learning models (Kohavi, R., 1995). It is a popular method because it generally results in a less biased or optimistic estimate of the model skill than other methods. For 10-fold cross validation, the 3000 images from the 2018 training dataset were randomly split into $k =10$ groups. Each unique group was held out as the test dataset and the remaining nine groups were used as the training dataset. The resulting model was evaluated using the held out test set and the test data accuracy score was retained. For 10-fold cross validation, this resulted in 10 accuracy scores that were averaged and reported as the mean cross-validation score. The optimum number of features was determined by selecting the number of features that maximized the 10-fold cross validation score. Comparisons of feature selection methods and classifiers were based on the maximum mean cross-validation score associated with the optimum number of features.

Model Testing (Step 4)

The best feature set from Step 3 was then used to develop the final model using the 3000 training ROI images and the best classifier identified also in Step 3. To assess model accuracy, this model was applied back to the 3000 ROI images of the 2018 training dataset to determine the proportion of the images correctly classified, defined here as the image-wise classification training score. Then model accuracy was evaluated on the 1400 ROI images from 2019 testing dataset to obtain a similar image-wise classification testing score.

TRANSLATING ROI IMAGE CLASSIFICATION SCORES INTO LOCATION ESTIMATES OF RESIDUE

While image-wise classification scores can be useful in the machine learning realm, a key element of this project is determining the level of residue on a location-in-the-field basis. A Bayesian multinomial Gaussian response model was used to estimate location percentage residue

from the three-class estimates of residue of the 50 ROI images at each location. The approach was based on the Bummer model (Vasko et al., 2000), controlled by parameters α_k (scaling factor), β_k (optimum) and γ_k (tolerance):

$$E\left(\tilde{y}_{ik}\big|x_i, \theta\right) = \frac{\lambda_{ik}}{\lambda_{i+}} \tag{1}$$

where,
$$\lambda_{ik} = E(\tilde{y}_{ik}|x_i, \theta_k) = \alpha_k \exp\left[\left(-\frac{\beta_k - x_i}{\gamma_k}\right)^2\right] \tag{2}$$

$$\lambda_{i+} = \sum_{h=1}^{m} \lambda_{ih} \tag{3}$$

Here, $\theta_k = (\alpha_k, \beta_k, \gamma_k)$; $\tilde{y}_{ik}$ is the abundance of class k at location i; x_i is the ith location percentage; m is the number of classes, which is three in this case.

A framework was created using the 2018 training dataset based on equation (1) using the probability distributions of classifier model-based ROI image class categories and the location-wise ground truth percentage to create the model. We set the scaling factor (α_k) to one, assumed the optimum factors (β_k) were normally distributed, the tolerance factors (γ_k) had a gamma distribution, and used non-informative priors consistent with Vasko et al., 2000. The performance of the model developed using 2018 data was reported as the correlation between the location-wise estimate of the Bayesian model compared to location-wise ground truth using the 2019 testing dataset.

RESULTS AND DISCUSSION

CLASSIFICATION OF ROI IMAGES THROUGH MACHINE LEARNING

Selecting the Feature Selection and Classifier Methods

Our first objective was to determine the best combination of a feature selection method and model classifier for a residue problem. The SVM classifier with the default hyper-parameters, consistently obtained higher 10-fold cross validation scores compared to the RF classifier for all feature selection methods (data not shown; paired t-test, $P<0.01$). Optimizing the hyper-parameters

with all features included did not affect this outcome. Consequently, we reported the outcome of feature selection methods using only the SVM classifier.

Among the four feature selection methods, RFE-SVM always maximized the 10-fold cross-validation score compared to the other three methods for all three GSD's (Table 4; $P \leq 0.01$). For both 0.05 and 0.14 GSD, the RFE-SVM method obtained the highest cross validation score using the fewest number of features (Table 4); for 0.014 GSD, the features selected by ReliefF and Gain Ratio were also included by RFE-SVM, but RFE-SVM identified additional features capable of improving model fit. Because the RFE-SVM feature selection always obtained the highest 10-fold cross-validation score, typically using the fewest features, we concluded it was the superior feature selection method for RGB residue images.

Table 4. The maximum 10-fold cross-validation score and associated number of selected features for four feature selection methods evaluated using the support vector machine classifier tested on RGB images obtained at three mean ground sampling distances using the 2018 training dataset.

Mean Ground Sampling Distance (cm pixel[-1])	Feature Selection Method[1]	Support Vector Machine Classifier	
		Maximum Cross-Validation Score	Number of Selected Features
0.014	RFE-SVM	87.8	13
	RF	86.9	26
	Relief F	85.2	5
	GR	85.3	4
0.05	RFE-SVM	84.1	7
	RF	82.9	15
	Relief F	80.7	68
	GR	81.6	20
0.14	RFE-SVM	80.9	6
	RF	78.3	18
	Relief F	77.6	68
	GR	77.5	70

[1]RFE-SVM=recursive feature elimination using support vector machine; RF=random forest; GR=gain ratio.

Similar to our results, Ma et al. (2017) concluded RFE-SVM was an appropriate and reliable feature selection method for both the RF and the SVM classifiers when using RGB imagery from an UAV for object-based classification of agricultural patterns such as bare ground, cropland, woodlands, roads, and buildings. Moghimi et al. (2018) concluded RFE-SVM feature selection

method alone was not the most effective method to select features from hyperspectral imaging in a plant phenotyping problem; they recommended an ensemble approach. While our results and those of Ma et al. (2017) imply RFE-SVM may be the feature selection method of choice for RGB imagery in agricultural applications we recommend testing feature selection approaches remain a part of a machine learning protocol.

Selected Features

Figure 4 shows the relationship between the mean 10-fold cross-validation score and the number of features included in the model based on the RFE-SVM feature ranking method for three resolutions of RGB images of residue. Table 5 reports the selected features in rank order each of the three GSD's and the associated classification accuracy scores.
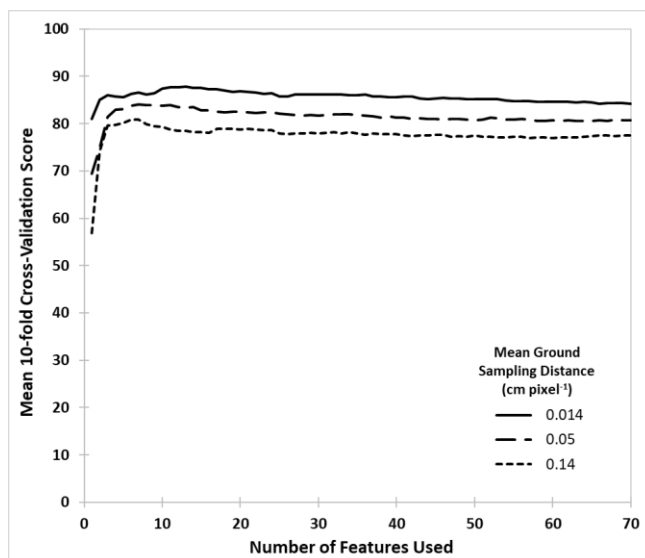
Figure 4. The relationship between the mean 10-fold cross-validation score and the number of features used in the three-class classification model estimating residue cover in RGB imagery obtained at three resolutions. At each resolution, features were ranked using the Recursive Feature Elimination using Support Vector Machine method and model accuracy was tested with 10-fold cross validation using the support vector machine classifier.

Local texture features, represented by LBP features were important at all three GSD's, and multiple LBP features (bins) were selected at each GSD (Table 5). Additionally, the number of LBP features selected increased with higher resolution images. All resolutions selected low-numbered bins (LBP 1 and 3 for 0.014 GSD, LBP 3 for 0.05 GSD and LBP 2 and 3 for 0.14 GSD)

associating residue variability with these indicators of uniform pixel patterns. For example, for LBP bin 1, all neighbor pixels are lighter than the reference pixel, a condition associated with flatter areas of an image (Ojala et al., 2002). Middle bins are likely to be associated with detection of edges and were well represented at all resolutions.

At all GSD's, color features were also important. The standard deviation of the S band of the HSV color space was selected at all three GSD's. Saturation in the HSV color-space describes the intensity of a color; white at S=0 and the pure hue at S=1 (Loesdau, M. et al., 2014). The importance of the standard deviation of S in RGB residue ROI images may be analogous to a texture feature, capturing the variability in saturation with mixed soil-residue images. In the highest resolution ROI images (0.014 GSD) only, two additional color features were selected from the LAB color space (Table 5). Again, the selected features captured the variability of the color feature in the image through the standard deviation.

Table 5. Selected features and the associated model performance statistics for predicting residue cover class from RGB region of interest (ROI) images at three ground sampling distances. Selected features included texture features (local binary patterns (LBP) bins, Haralick-5 feature (inverse difference moment); and color features, including the standard deviation (STD) of saturation (S) from HSV color space, and the STD of A (Red/Green value) and B (Blue/Yellow value) from the LAB color space. Classification accuracy scores include the 10-fold cross validation score and the training score using the 2018 ROI images and the testing score using the 2019 ROI images.

Feature Rank	Mean Ground Sampling Distance (cm pixel^{-1})		
	0.014	0.05	0.14
1	LBP-20	S-STD	Haralick-5
2	LBP-14	LBP-13	LBP-12
3	S-STD	LBP-7	LBP-2
4	LBP-1	LBP-25	LBP-13
5	LBP-19	LBP-5	S-STD
6	B-STD	LBP-12	LBP-3
7	A-STD	LBP-3	-
8	LBP-8	-	-
9	LBP-15	-	-
10	LBP-4	-	-
11	LBP-21	-	-
12	LBP-12	-	-
13	LBP-3	-	-
ROI image-wise classification scores			

Cross Validation	87.8	84.1	80.9
Training	91.6	85.5	83.7
Testing	80.9	68.4	65.4

At the lowest resolution (0.14 GSD), a global texture feature, the inverse difference moment Haralick texture feature (labeled Haralick-5) was also selected (Table 5). Cooper (2004) stated inverse difference moment measures the homogeneity of an image, which attains a maximum value when all image pixels that are compared to the reference pixel have the same value. Its inclusion in the 0.14 GSD may reflect variation across the image and may become more important than capturing local complexity in patterns compared to the higher resolution images.

Shape features, represented by Hu moments, were not selected at any GSD. Hu moments are effective measures for tracing redundant image patterns with continuous functions (without abrupt changes in value of the image pixels) and are noise-free (Huang et al., 2010; Papakostas, 2014). Our results imply that it is more effective to consider variability of residue as a texture; apparently differences among residue types and some of the added complexity of residue images interfered with effectively using shape features to characterize residue.

Analysis on Feature Types

In a second analysis, we used boxplots to document the capacity of feature types to predict residue cover to further contrast the role of the features (Figure 5). In this analysis, 70 single-feature models (one for each feature) were tested using 10-fold cross validation and reported by feature category using boxplot (Figure 5). The boxplot documents the range of success these single-feature models had classifying residue by feature type. The individual LBP features typically provided substantially higher cross-validation accuracy for the highest resolution images compared to lower resolution images. In contrast, Haralick features typically performed better in low resolution images. A potential explanation is lower resolution images contain more information at the global than the local level. For the highest resolution imagery, a pixel is 6% of

the smallest dimension defined as residue; for the lowest resolution images, a pixel is 62% of the smallest defined dimension of residue. The LBP features apparently captured greater local complexity of the highest resolution images.

The boxplot analysis confirmed only a few color features were closely associated with residue prediction, implying that selected standard deviation features were uniquely applicable (Figure 5). The lack of importance of color, compared to the variability of color, may be due to our choice to not calibrate color in our images plus the expectation that soil and residue color can vary dramatically with moisture conditions in the field. The poor performance of all Hu moment features is also clear in Figure 5.
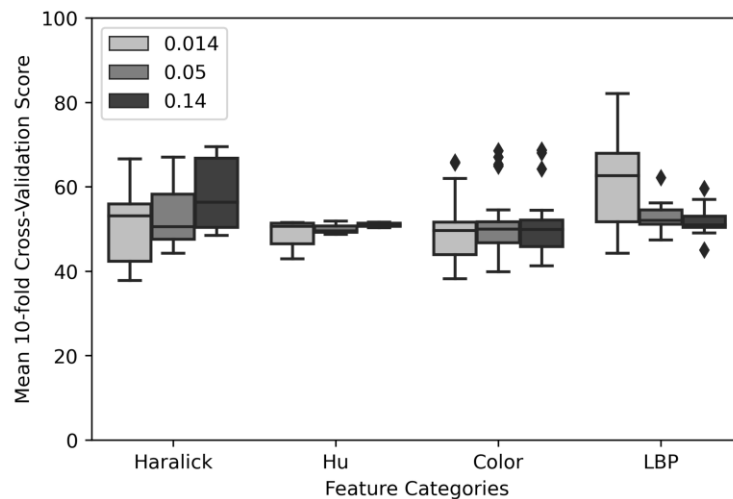
Figure 5. A boxplot for the relationship between the mean 10-fold classification score and the individual distinct feature used in the three-class classification model estimating residue cover in RGB imagery obtained at three resolutions (0.014, 0.05 and 0.14 cm pixel^{-1}). There was a total of 70 features tested in four categories: Haralick (n=13), Hu (n=7), color (n=24), and Local Binary Pattern (LBP; n=26).

In summary, texture features, as represented by LBP, Haralick and the standard deviation of color explained residue variability at three resolutions of imagery of residue. These results emphasize the dominating influence of local invariant patterns captured by LBP and color standard deviation and the potential importance of global features as resolution decreases. This research also suggests the best opportunity to improve estimates of residue is to evaluate more variants of LBP features with optimized parameters like radius and consider their interaction with image GSD.

ROI Image-wise Model Accuracy

Table 5 reports image-wise classification accuracy based on the 10-fold cross validation score and the final model training score using the 2018 dataset of 3000 images, and the final model testing score using the 2019 dataset of 1400 images; results are reported for each of the three GSD's. Both cross validation and testing scores decreased as resolution decreased. The decrease in model testing score performance with increasing GSD was over twice as much with the lower resolution images compared to the 0.014 GSD images (Table 5). The biggest difference was between the testing score of the highest resolution GSD and the two UAV-obtained GSD's. This suggested that lower resolution was not solely responsible for the decrease in model performance.

The 0.014 GSD images were captured under more controlled conditions including using a tripod to enforce a 1.0-m distance from the ground, and using a higher quality DSLR camera and lens. The lower two resolution images were captured using a lower quality camera deployed on a UAV, typically flown by hand, targeting either 2- or 5-m height over uneven terrain, and dealing with variable wind conditions. This resulted in more variation in GSD among images at a given elevation and our analysis of the distribution of colors in the ROI images for all three GSD suggested that drone images had a narrower distribution of color histogram compared to the higher-quality DSLR camera images (data not shown). The cross validation scores for all three GSD's documented successful feature selection with the 2018 dataset. We cannot say definitively why testing scores were substantially worse with the 2019 UAV testing datasets, but it likely reflects year-to-year differences associated with our limited effort to optimize image acquisition using the UAV. Our results suggest that images captured with better control of altitude (e.g. using Spidercam, or higher-quality UAV) and/or a higher quality cameras might improve estimation results associated with our UAV datasets.

ASSESSING LOCATION-WISE RESIDUE COVER BASED ON BAYESIAN MODEL

To address our last objective of estimating the location percentage residue cover based on the 50 individual ROI image classes of a location, a Bayesian-based method was used based on the Bummer model from Vasko et al. (2000). First, an ideal Bayesian model was built using the ground-truth based classification of the 3000 ROI images from 2018 training dataset and the location-wise ground-truth residue values. The resulting ideal Bayesian model predicted accurately location-wise residue for both the 2018 training and the 2019 testing datasets (training $r^2 = 0.99$, testing $r^2 = 0.99$). Given the success of the ideal model, we then developed a Bayesian model using the predicted classes from the trained classifier model and the ground truth location-wise percentage residue; this was done for all three GSD's. Location-wise agreement with ground truth (r^2) for the training dataset was 0.97, 0.95, and 0.87 and for the testing dataset was 0.90, 0.52, and 0.43 for the 0.014, 0.05 and 0.14 GSD images, respectively. The testing dataset model fit for GSD=0.014 is presented in Fig. 6a. Poor performance on the two lower GSD testing datasets was consistent with the poor image-wise testing classification accuracy (Table 5).
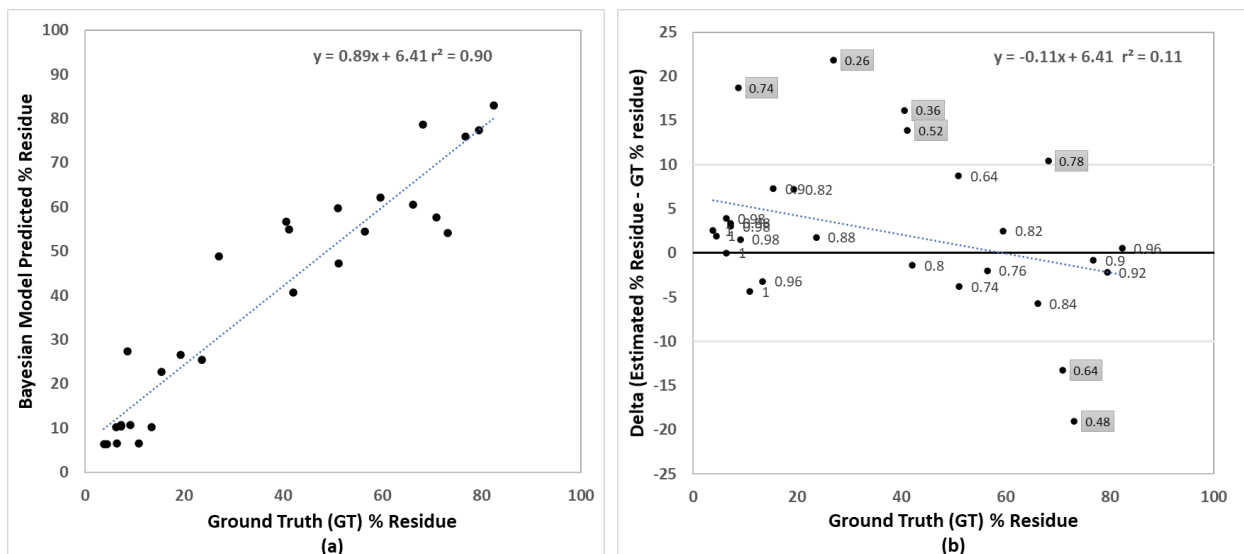

Figure 6- (a) The relationship between predicted % residue for location reading using a Bayesian model and the ground truth (GT) location reading. (b) Relationship between delta (estimated % residue minus GT % residue) and the GT location % residue. Outlier points (highlighted in gray) have a delta $>\pm10$ percentage units from GT. Data is for the 28 tape locations associated with the 2019 testing dataset.

To better understand deviations of location-wise estimates (0.014 GSD), we plotted the differences from the ground truth values (Figure 6b). Location-wise estimates systematically over-estimated low residue locations and under-estimated high residue locations (Figure 6b). This is likely a limitation of moving from a three-class image-wise estimate to a continuous estimate of residue cover. When ±10 % delta was set as a threshold for the optimum accuracy, seven outliers locations were observed (highlighted points in Figure 6b). If we don't consider the seven outliers for 0.14 GSD model, r^2 for the testing dataset is 0.98. In Figure 6(b), the numbers associated with the data points are the location-wise classification accuracy of the classifier for the 50 ROI images at the location. The ROI images contributing to the outlier locations had mean classification accuracy of 54% (range 26% to 78%) compared to 90% (range 64% to 100%) for the 21 test locations predicted within a delta of 10%. For comparison, the 60 locations from the training dataset had the mean classification score of 92% (range 74% to 100%).

One possible contributer to outlier locations was shadows. Five out of seven outlier locations had significant shadow from tripod and/or plants. There is some indication that the locations with shadow showed a significant rise in residue estimation, whereas locations with plant but no shadow show significant drop in residue estimation. This trend was mainly seen for the locations whose % residue is near to threshold (33%, 66%).

Evaluating the 0.05 GSD images, five of the seven outlier locations of 0.014 GSD were also outlier locations at this resolution plus one additional outlier location. If these outlier locations were elimited from the 0.05 GSD testing dataset then the model, r^2 for the testing dataset is 0.90. The additional outlier had a lot of small-sized weed residue which was associated with an underestimate of residue level by the model. The same phenomenon was observed in 2 outliers of the 0.14 GSD model, which had a total of five outlier locations, including three outlier locations

from 0.014 GSD model. If we don't consider the outliers for 0.14 GSD model, r^2 for the testing dataset is 0.81. Images contributing to the outlier locations had a mean classification accuracy of 24% and 25% for 0.05 and 0.14 GSD respectively.

We conclude that there were locations in 2019 where the model derived from the 2018 dataset was not able to predict residue correctly. As discussed earlier, the exact reasons will require further investigation beyond the scope of this exploratory project. Given the prominence of shadow and plants in the outlier locations, additional work is needed to evaluate the impact of these common components of residue images. Green plants represented ground cover that was not a component of the ground truth residue estimate. Dark shadow is the area in an image where residue cover cannot be known. Both of these were prevelant in outlier images suggesting steps need to be taken to address these different potential sources of error in the current approach. Possible strategies include removing plant and shadow from the images or considering a separate variant which would represent the amount of shadow and plant in the image during a classifier model building process. Additionally, some outlier location images from all GSD datasets were comparatively different with respect to illumination and color from the ones used for the model training. This suggests the importance of a diverse range of residue images from a wide range of soil types to ensure a resilient model under a wide range of conditions.

Results from our location estimates based our highest-resolution imagery (r^2=0.90; Figure 6a) compared favorably with residue estimates from previous research. Correlation coeficients from previous research estimating residue cover from high-resolution RGB imagery ranged from 0.75 to 0.86 (Bauer and Strauss, 2014; Laamrani et al., 2018; Riegler-Nurscher et al., 2018; Kavoosi et al., 2020). These studies typically used limited datasets compared to 88 locations and 4400 ROI images representing a diverse range of fields and conditions in our dataset collected over two years

in mid-Missouri. Residue estimates from satellite imagery often performed as well or better than previous reports from high-resolution RGB imagery. For example, Daughtry et al. (2006) obtained average r^2=0.81 for estimates of crop residue cover as a function of cellulose absorption index using EO-1 Hyperion imaging spectrometer data. Beeson et al. (2016) used multiple sources of satellite images and obtained classification accuracies ranging from 61% to 92% for two-class models compared to visual and line transect estimates of residue in Iowa fields over three years. Hively et al. (2019) were able to map crop residue cover with 92% (+/- 10%) accuracy using WorldView-3 imagery.

This study confirmed the potential of high-resolution RGB images to provide accurate estimates of residue cover that meets or potentially exceed other strategies. There were some obvious limitations of our current data collection protocols, features and classification models with potential for improvement. In particular, poor performance of UAV-collected data with the 2019 testing dataset implied that there are critical quality control criteria missing that were present in the tripod-mounted higher resolution system.

CONCLUSION

This study provided new insights into the use of machine learning methods for the estimation of crop residue level from the RGB images with three different GSDs. The SVM classifier performed superior to the RF classifier and the RFE-SVM feature-selection method outperformed other feature selection methods for crop residue images. For all three GSD images, texture related features such as local texture features and standard deviation in color were most important. Shape features were irrelevant for residue images. While the classifier achieved poorer testing scores for the lower resolution images obtained using a UAV, the consistantly good results throughout cross-validation and testing on the 2018 dataset show that the SVM learned well and did not overtrain.

Poorer performance in 2019 testing scores with the UAV images may be due to some external uncontrolled factors from 2018 to 2019, and not the different image resolutions.

Location-wise estimates of residue cover from the classified images using a Bayesian model was highly dependent on the accuracy of the classification model which was poor for lower resolution datasets associated with the UAV platform. The Bayesian model was able to estimate testing location residue percentage from 0.014 GSD images in good agreement with the ground truth ($r^2 = 0.90$). Seven outlier locations had the common phenomenon of lower classification accuracy which was in the range of 26% to 78%, compared to 74% to 100% for other locations. Accounting for shadow and plants may offer a way to reduce the issues associated with these outlier locations. Besides validating this claim, future work should also focus on improving texture features. Potential strategies include assessing different variants of LBP features and optimizing LBP parameters such as radius for different resolution images.

In-field systems based on RGB imagery obtained by handheld or UAV-based platforms have potential as method to document residue cover for applied applications such as NRCS compliance review (USDA-NRCS, 2011). Additionally, a RGB based successful model could help provide better ground truth for satellite imagery.

REFERENCES

Bauer, T., & Strauss, P. (2014). A rule-based image analysis approach for calculating residues and vegetation cover under field conditions. *Catena, 113*, 363-369. http://dx.doi.org/10.1016/j.catena.2013.08.022

Beeson, P.C., Daughtry, C.S.T., Hunt, E.R., Akhmedov, B., Sadeghi, A.M., Karlen, D.L., & Tomer, M.D. (2016). Multispectral satellite mapping of crop residue cover and tillage intensity in Iowa. *J. Soil Water Conserv*. 71 (5) 385-395. https://doi.org/10.2489/jswc.71.5.385

Bronick, C.J., & Lal, R. (2005). Soil structure and management: a review. *Geoderma. Vol. 124, Issues 1–2,* 3-22. https://doi.org/10.1016/j.geoderma.2004.03.005

Burges, C.J. (1998). A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov. 2,* 121–167. https://doi.org/10.1023/A:1009715923555

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Comput. Electr. Eng. 40*, 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chaugule, A., & Mali, S.N. (2014). Evaluation of the texture and shape features for classification of four paddy varieties. *J. Eng., Article ID 617263*, 1-8. http://dx.doi.org/10.1155/2014/617263

Cherubin, M.R., Oliveira, D.M.S., Feigl, B.J., Pimentel, L.G., Lisboa, I.P., Gmach, M.R., … Cerri, C.C. (2018). Crop residue harvest for bioenergy production and its implications on soil functioning and plant growth: A review. *Sci. Agric., 75*, 255-272. https://doi.org/10.1590/1678-992x-2016-0459

Cooper, G.R.J. (2004). The textural analysis of gravity data using co-occurrence matrices. *Comput. Geosci., 30*. 107-115. https://doi.org/10.1016/j.cageo.2003.08.001

Daughtry, C. S. T., Doraiswamy, P. C., Hunt Jr., E. R., Stern, A. J., McMurtrey III, J. E., & Prueger, J. H. (2006). Remote sensing of crop residue cover and soil tillage intensity. *Soil Tillage Res., 91,* 101-108. https://doi.org/10.1016/j.still.2005.11.013

Dickey, E.C., Shelton, D.P., Meyer, G.E., & Fairbanks, K.T. (1989). Determining crop residue cover with electronic image analysis. *Biol. Syst. Eng.: Pap. Publ., 238,* 3056–3064. https://digitalcommons.unl.edu/biosysengfacpub/238/

Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., & Guijarro, M. (2012). Support vector machines for crop/weeds identification in maize fields. *Expert Syst. Appl., 39(12),* 11149-11155. https://doi.org/10.1016/j.eswa.2012.03.040

Guyon, I., Weston, J., Barnhill, S., & Vapnik V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn., 46,* 389–422. https://doi.org/10.1023/A:1012487302797

Hapfelmeier, A., & Ulm, K. (2013). A new variable selection approach using random forests. *Comput. Stat. Data Anal., 60*, 50–69. https://doi.org/10.1016/j.csda.2012.09.020

Haralick, R. M., Shanmugam K., & Dinstein I. (1973). Textural features for image classification. *IEEE Trans. Syst. Man Cybern. Vol. SMC-3, No. 6,* 610-621. https://doi.org/10.1109/TSMC.1973.4309314

Harris Jr., E. (2001). Information gain versus gain ratio: A study of split method biases. *The MITRE Corp.,* 1-20. Retrieved from https://www.mitre.org/sites/default/files/pdf/harris_biases.pdf

Herrera, P.J., Dorado, J., & Ribeiro, A. (2014). A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method. *Sensors, 14,* 15304-15324. https://doi.org/10.3390/s140815304

Hively, W.D., Shermeyer, J., Lamb, B.T., Daughtry, C.T., Quemada, M., & Keppler, J. (2019). Mapping crop residue by combining landsat and worldview-3 satellite imagery. *Remote Sens., 11, 1857.* https://doi.org/10.3390/rs11161857

Ho, T. K. (1995). Random decision forests. *Proc. 3rd Int. Conf. Doc. Anal. Recognit. 1*, pp. 278–282. https://doi.ieeecomputersociety.org/10.1109/ICDAR.1995.598994

Hu, M.K. (1961). Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory*, 179-187. https://doi.org/10.1109/TIT.1962.1057692

Huang Z., & Leng, J. (2010). Analysis of Hu's moment invariants on image scaling and rotation. *2010 2nd Int. Conf. Comput. Eng. Technol. Chengdu 7,*pp. 476- 480. https://doi.org/10.1109/ICCET.2010.5485542

Kavoosi, Z., Raoufat, M.H., Dehghani, M., Abdolabbas, J., Kazemeini, S.A., & Nazemossadat, M.J. (2020). Feasibility of satellite and drone images for monitoring soil residue cover. *J. Saudi Soc. Agric. Sci. 19*(1), 56-64. https://doi.org/10.1016/j.jssas.2018.06.001

Kluyver, T., Ragan-Kelley, B., Perez, F., Granger, B., Bussonnier, M., Frederic, J., … Willing, C. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Loizides, F. & Schmidt, B. (Eds.), *Position. Power Acad. Publ.: Players, Agents and Agendas*, 87–90. https://doi.org/doi:10.3233/978-1-61499-649-1-87

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc. 14th Int. Jt. Conf. Artif. Intell 2.,* pp. 1137–1143. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. https://dl.acm.org/doi/10.5555/1643031.1643047

Laamrani, A., Joosse, P., & Feisthauer, N. (2017). Determining the number of measurements required to estimate crop residue cover by different methods. *J. Soil Water Conserv. 72*(5), 471-479. https://doi.org/10.2489/jswc.72.5.471

Laamrani, A., Lara, R.P., Berg, A.A., Branson, D., & Joosse, P. (2018). Using a mobile device "App" and proximal remote sensing technologies to assess soil cover fractions on agricultural fields. *Sensors. 18*(3). https://doi.org/10.3390/s18030708

Laflen, J.M., Amemiya, M., & Hintz, E.A. (1981). Measuring crop residue cover. *J. Soil Water Conserv. 36*(6), 341-343.

Lal, R. (2008). Soils and sustainable agriculture. A review. *Agron. Sustain. Dev. 28,* 57–64. https://doi.org/10.1051/agro:2007025.

Lal, R. (2009). Soil quality impacts of residue removal for bioethanol production. *Soil Tillage Res. 102*(2), 233-241. https://doi.org/10.1016/j.still.2008.07.003

Le, V.N.T., Apopei, B., & Alameh, K. (2019). Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods. *Inf. Process. Agric. 6*, 116-131. https://doi.org/10.1016/j.inpa.2018.08.002

Li, J., Wang, S., Cheng, K., & Morstatter, F. (2016). Feature selection: A data perspective. *ACM Comput. Surv. 50*. https://doi.org/10.1145/3136625

Li, M., Ma, L., Blaschke, T., Cheng, L., & Tiede, D. (2016). A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *Int. J. Appl. Earth Obs. Geoinf. 49,* 87–98. https://doi.org/10.1016/j.jag.2016.01.011

Loesdau, M., Chabrier, S., & Gabillon, A. (2014). Hue and saturation in the RGB color space. In: Elmoataz, A., Lezoray, O., Nouboud, F., & Mammass, D. (Eds.), *Image Signal Process. ICISP. Lect. Notes Comput. Sci. 8509.* Springer, Cham. https://doi.org/10.1007/978-3-319-07998-1_23

Lory, J.A., Upadhyay, P.C., Lagaunne, T.A.P, Spinka, C., Miller, R., Davis, G., & DeSouza, G.N. (2021). Capability of high-resolution RGB imagery to accurately document residue in row-crop fields. *J. Soil Water Conserv*. 00193. https://doi.org/10.2489/jswc.2021.00193

Ma, L., Fu, T., Blaschke, T., Li, M., Tiede, D., Zhou, Z., Ma, X., & Chen, D. (2017). Evaluation of feature selection methods for object-based land cover mapping of unmanned aerial vehicle imagery using random forest and support vector machine classifiers. *ISPRS Int. J. Geo-Inf. 6*(51). https://doi.org/10.3390/ijgi6020051

Moghimi, A., Yang, C., & Marchetto, P. M. (2018). Ensemble feature selection for plant phenotyping: A journey from hyperspectral to multispectral imaging. *IEEE Access, 6,* 56870-56884. https://doi.org/10.1109/ACCESS.2018.2872801

Morrison Jr., J.E., Huang, C., Lightle, D.T., & Daughtry, S.T. (1993). Residue measurement techniques. *Soil Water Conserv. Soc. 48(6)*, 478-483.

Morrison Jr., J.E., Lemunyon, J., & Bogusch Jr., H.C. (1995). Sources of variation and performance of nine devices when measuring percent residue cover. *Trans. ASAE., 38*(2), 521-529. https://doi.org/10.13031/2013.27861

Novakovic, J., Strbac, P., & Bulatovic, D. (2011). Toward optimal feature selection using ranking methods and classification algorithms. *Yugoslav J. Oper. Res., 21*(1)*,* 119-135. https://doi.org/10.2298/YJOR1101119N

Ojala, T., Peitikainen M., & Maenpaa T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell., 24*(7), 971-987. https://doi.org/10.1109/TPAMI.2002.1017623

Papakostas G.A. (2014). Chapter1: Moments and moment invariants. In *Theory and Applications, GCSR,* Vol. 1, 3-32. Science Gate Publishing. https://doi.org/10.15579/gcsr.vol1.ch1

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B.,… Duchesnay E. (2011). Scikit-learn: Machine learning in python," *J. Mach. Learn. Res. 12*, 2825-2830.

Ranaivoson, L., K. Naudin, A. Ripoche, F. Affholder, L. Rabeharisoa, & M. Corbeels. 2017. Agro-ecological functions of crop residues under conservation agriculture. A review. *Agron. Sustain. Dev.,37*(26). doi.10.1007/s13593-017-0432-z.

Richards, B.K., Walter, M.F., & Muck, R.E. (1984). Variation in line transect measurements of crop residue cover. *J. Soil Water Conserv.* 60-61.

Riegler-Nurscher, P., Prankl, J., Bauer, T., Strauss, P., & Prankl, H., (2018). A machine learning approach for pixel wise classification of residue and vegetation cover under field conditions. *Biosyst. Eng., 169,* 188-198. https://doi.org/10.1016/j.biosystemseng.2018.02.011

Robnik-Sikonja, M., & Kononenko, I. (2003). Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn. J., 53,* 23-69. https://doi.org/10.1023/A:1025667309714

Schmittmann, O., & Lammers, P. (2017). A true-color sensor and suitable evaluation algorithm for plant recognition. *Sensors 17(8), 1823*. https://doi.org/10.3390/s17081823

Searle, S., & Bitnere, K. (2017). Review of the impact of crop residue management on soil organic carbon in Europe. The International Council on Clean Transportation. Retrieved from https://theicct.org/publications/impact-of-crop-residue-mgmt-EU

Singh B., & Rengel Z. (2007). The role of crop residues in improving soil fertility. In Marschner, P., & Rengel, Z. (Eds.), *Nutrient Cycling in Terrestrial Ecosystems. Soil Biol., 10. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/978-3-540-68027-7_7

USDA. (2011). Subpart 503E crop residue. Natural Resources Conservation Service. *Natl. Agron. Man. 503*, 126–128.

United States Code (U.S.C.). (1985). The food security act of 1985. *U.S.C., 16*, 3801-3862.

Urbanowicz, R. J.,Meeker, M., LaCava, W., Olson, R. S., & Moore J. H. (2018). Relief-based feature selection: Introduction and review. *J. Biomed. Inform., 85,* 189-203. https://doi.org/10.1016/j.jbi.2018.07.014

Vapnik, V. (1995). The nature of statistical learning theory. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4757-2440-0

Vasko, K., Toivonen, H.T.T., & Korhola, A. (2000). A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. *J.Paleolimnol., 24*, 243-250. https://doi.org/10.1023/A:1008180500301

Yang, W., Wang, S., Zhao, X., Zhang, J., & Feng, J. (2015). Greenness identification based on HSV decision tree. *Inf. Process. Agric., 2,* 149-160. https://doi.org/10.1016/j.inpa.2015.07.003

Yuan, W., Wijewardane, N.K., Jenkins, S., Bai, G., Ge, Y., & Graef, G.L. (2019). Early prediction of soybean traits through color and texture features of canopy RGB imagery. *Nature Sci. Rep., 9*(14089). https://doi.org/10.1038/s41598-019-50480-x