



Identification of Weeds Based on Hyperspectral Imaging and Machine Learning

Yanjie Li^{††}, Mahmoud Al-Sarayreh¹, Kenji Irie², Deborah Hackell³, Graeme Bourdot⁴, Marlon M. Reis¹ and Kioumars Ghamkhar^{1*}

¹ AgResearch Ltd., Grasslands Research Centre, Palmerston North, New Zealand, ² Red Fern Solutions Ltd., Christchurch, New Zealand, ³ AgResearch Ltd., Ruakura Research Centre, Hamilton, New Zealand, ⁴ AgResearch Ltd., Christchurch, New Zealand

OPEN ACCESS

Edited by:

Wanneng Yang,
Huazhong Agricultural University,
China

Reviewed by:

Lingfeng Duan,
Huazhong Agricultural University,
China

Haiyan Cen,
Zhejiang University, China

*Correspondence:

Kioumars Ghamkhar
Kioumars.Ghamkhar@
agresearch.co.nz

† Present address:

Yanjie Li,
Research Institute of Subtropical
Forestry, Chinese Academy
of Forestry, Hangzhou, China

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 29 September 2020

Accepted: 30 December 2020

Published: 25 January 2021

Citation:

Li Y, Al-Sarayreh M, Irie K,
Hackell D, Bourdot G, Reis MM and
Ghamkhar K (2021) Identification
of Weeds Based on Hyperspectral
Imaging and Machine Learning.
Front. Plant Sci. 11:611622.
doi: 10.3389/fpls.2020.611622

Weeds can be major environmental and economic burdens in New Zealand. Traditional methods of weed control including manual and chemical approaches can be time consuming and costly. Some chemical herbicides may have negative environmental and human health impacts. One of the proposed important steps for providing alternatives to these traditional approaches is the automated identification and mapping of weeds. We used hyperspectral imaging data and machine learning to explore the possibility of fast, accurate and automated discrimination of weeds in pastures where ryegrass and clovers are the sown species. Hyperspectral images from two grasses (*Setaria pumila* [yellow bristle grass] and *Stipa arundinacea* [wind grass]) and two broad leaf weed species (*Ranunculus acris* [giant buttercup] and *Cirsium arvense* [Californian thistle]) were acquired and pre-processed using the standard normal variate method. We trained three classification models, namely partial least squares-discriminant analysis, support vector machine, and Multilayer Perceptron (MLP) using whole plant averaged (Av) spectra and superpixels (Sp) averaged spectra from each weed sample. All three classification models showed repeatable identification of four weeds using both Av and Sp spectra with a range of overall accuracy of 70–100%. However, MLP based on the Sp method produced the most reliable and robust prediction result (89.1% accuracy). Four significant spectral regions were found as highly informative for characterizing the four weed species and could form the basis for a rapid and efficient methodology for identifying weeds in ryegrass/clover pastures.

Keywords: hyperspectral imaging, weeds classification, superpixel, PLS-DA, multilayer perceptron

INTRODUCTION

Pastures based on perennial ryegrass (*Lolium perenne*) and white clover (*Trifolium repens*) are the main source of forage for animal production in New Zealand (McClearn et al., 2020). Weeds are a major economic constraint. Within the primary sector alone, weeds cost farmers NZ\$50M in actual expenditure on chemical herbicides and labor (Bourdôt et al., 2007). Technologies that reduce these costs, and help minimize the use of synthetic herbicides, would improve the value of forage production (Bacco et al., 2018).

Recently, technologies such as hyperspectral imaging (HSI) systems are providing opportunities for rapid classification of plant species both in the laboratory and the field (Griffel et al., 2018; Liu and Zhang, 2018; Xu et al., 2018; Ferreira et al., 2019). The advantage of HSI is the provision of a combined spectroscopy and relationships between various chemical components and the absorption of spectra (Curran, 1989; Ebberts et al., 2002). The principle of HSI spectroscopy is based on molecular vibrations in the IR region (Youngentob et al., 2012). Therefore, absorbance at specific wavelengths, which might be related to specific chemical bands, can be used for different materials classification and quality determination (Vaiphasa et al., 2007; Schwanninger et al., 2011).

Many attempts have been made to use the visible light imaging or Red-Green-Blue (RGB) to identify weeds (Ahmad et al., 2018; Raja et al., 2020). However, shape, color and size, are limiting constraints of RGB imaging for the identification of species with similar phenotype (Wang et al., 2019). HSI can overcome these limitations by capturing spectral and spatial information simultaneously. It has a proven history of widespread use in materials discrimination and quality estimates including in meat science (Al-Sarayreh et al., 2018; Reis et al., 2018), forestry (Te et al., 2019), and land cover mapping (Jiang et al., 2017; Xu et al., 2018). Shorten et al. (2019) showed the potential of HSI for the measurement of a few components of forage quality. Other studies (Ahmad et al., 2019; Farooq et al., 2019b) indicate promising potentials for plant identification based on chemical signatures of different species.

Numerous analytical methods for HSI data classification have been reported (Lunga et al., 2014; Li et al., 2017; Audebert et al., 2019). Among these techniques, support vector machine (SVM; Yu et al., 2012; Peng et al., 2015) and partial least squares discriminant analysis (PLS-DA; Yang et al., 2015; Carreiro Soares et al., 2016; Walter et al., 2019) are considered as the most reliable techniques. This is specifically the case when limited training data are available (Melgani and Bruzzone, 2004; Chevallier et al., 2006).

Machine learning has been widely used for image classification (Chen et al., 2014). The Multilayer Perceptrons (MLP) methods have the advantage of handling a large number of training data (Golhani et al., 2018; Taneja, 2020). These methods could automatically learn features, while yielding comparable results on HSI classification process to other methods (Sutskever and Hinton, 2008).

Generally, the original HSI image contains the target (i.e., weeds) as well as the background and other components that could affect the labeling accuracy of the target species. To remove the background and obtain the region of interest (ROI), a segmentation strategy is required (Ren and Malik, 2003). Single or multistep thresholding algorithms are commonly used for obtaining the ROI and extracting the average spectrum of a sample (Sharma and Bhavya, 2020). Building prediction models based on the extracted spectra is commonly used for indoor applications (Mishra et al., 2017; Yuan et al., 2019).

Spatial variation in spectra requires more attention in pixel-wise prediction and outdoor applications for achieving high prediction accuracy (Vaughn et al., 2016).

Image segmentation is a crucial step in analyzing and understanding the contents of an image. It can be used to extract a wide range of image features including spatial features and superpixel (Sp) segmentation is one of these segmentation methods (Ren and Malik, 2003). In this method the pixels are grouped into many small segments adhering to the target boundaries where each segment shares the same spectral and spatial features of a common target (Li and Chen, 2015). It provides a compact and uniform segmentation for the target and extracts the spatial spectra from the image (Fan et al., 2017).

Therefore, we hypothesize that there are unique spectral signatures in each weed species, which are detectable by HSI and modeling. To test our hypothesis, we used three common weeds and a proxy weed species in ryegrass paddocks of New Zealand (NZ). The three weed species were the annual winged thistle (*Carduus tenuiflorus* Curtis), the annual yellow bristle grass [*Setaria pumila* (Poir.) Roem. & Schult.] and the perennial giant buttercup (*Ranunculus acris* L.). Winged thistle is a problematic weed of drought-prone low-fertility sheep and beef cattle pastures while yellow bristle grass and giant buttercup are weeds of high-fertility dairy pastures (Bourdôt et al., 2003; Lamoureaux, 2014). The fourth species, wind grass [*Anemanthele lessoniana* (Steud.) Veldkamp], is a NZ native species. It was used as a proxy for Chilean Needle Grass [*Nassella nessiana* (Trin. & Rupr.)], which cannot legally be cultivated in NZ. This grass is currently limited in its geographical distribution in New Zealand but threatens vast tracts of low-fertility drought prone hill-country pasture land (Bourdôt et al., 2010, 2012). Proven true, this hypothesis will allow future development of database for spectral signatures of weeds, which is valuable to detect weeds independently of the type of fields they are found in (i.e., independently of the type of plant species that are surrounding the weeds).

MATERIALS AND METHODS

Weed Sample Preparation

Four criteria were considered to choose the weed species, i.e., cover grass, broadleaf, perennial, and annual weeds. Selected weed species included three weeds of ryegrass pastures [thistle (TT), yellow bristle grass (YBG), and buttercup (BC)] and one proxy endemic species [wind grass (WG)] (**Figure 1**). Weed seeds were sourced from Margot Forde Germplasm Centre (MFGC), Palmerston North, NZ. For each weed species, 30 single seeds were planted in pots (one plant per pot) on 2 October 2018 at AgResearch Ruakura campus (Hamilton, New Zealand). The pots were standard 7 cm (7 × 7 × 8 cm) plastic pots placed on tables in the open-air greenhouse with a temperature between 18 and 25 degrees and watered as required (2–3 times a week). The standard Daltons potting mix soiled was used (40% bark fiber, 20% C.A.N Fines A grade, 20% Coco fiber classic, 20% pumice 7 mm plus fertilizer containing lime, permawet, osmocote, microplus (Te, Mg, and Fe), Gypsum, dolomite, and coated ExteNd). After 3 weeks, all plants were transferred to the lab for HSI scanning.



FIGURE 1 | The four weed species (a) Thistle (TT), (b) Yellow bristle grass (YBG), (c) Buttercup (BC), and (d) wind grass (WG).

Hyperspectral Imaging

A line scan HSI spectrograph system (Extended VNIR, Headwall Photonics, Fitchburg, MA, United States), with a 320×240 pixels camera was used for the HSI data collection. This system covered the range of 550–1,700 nm spectra with 5 nm spectral resolution and 235 wavelengths from the visible and NIR range of the electromagnetic spectrum. A halogen lamp light source (JCR 21V 150W/AL Japan 2DB) was set up on one side of the camera's lens, at 30° from the vertical plane as the illumination system. The light power was adjusted using a white reference tile (Labsphere Inc., North Sutton, NH, United States) where the highest intensity detected in the white reference tile (Labsphere Inc., North Sutton, NH, United States) was set as 85% of the saturation of the detector to prevent areas where the sample may saturate the detector. The distance between the plant sample and the camera was adjusted to 25 cm and the plants were placed directly below the HSI system with the camera exposure time set on 25 ms. The translation speed of the linear stage was set to 11.1 mm/s. The white reference image was captured by placing a white tile under the hyperspectral camera. Dark reference images were acquired with the lens cap on the hyperspectral camera.

Single weed-pots were placed on the linear stage to capture the hyperspectral images when they pass under the camera. Considering that each pot was scanned individually, the presence of shadows was not a major issue. For cases where the scanning is performed in the field, it is possible to use a different illumination system to reduce the presence of shadows (Bateman et al., 2020).

Three steps of the workflow for identification of weeds are shown in **Figure 2**. These three steps are:

Step 1: Acquiring hyperspectral images

Hyperspectral images were captured for 30 samples of each weed. Four types of weeds samples were present so in total, 120 hyperspectral images of weeds have been captured.

Step 2: Image Processing

Captured hyperspectral images were go through a series of steps to process them for extracting the spectral data for modeling.

Step 3: Modeling

Spectral data extracted from the image processing step pre-processed and then it was used as an input for model development.

Image Processing

This was the second step of the workflow where ROI extracted by employing segmentation on calibrated hyperspectral images.

Calibration

Each captured hyperspectral image was calibrated, using dark reference (D), and white reference image (W). Hyperspectral images raw intensity values were used to calculate reflectance by using Eq. 1.

$$R_c = \frac{X_{raw} - X_D}{X_w - X_D} \quad (1)$$

where R_c is the absolute reflectance, X_{raw} stands for the intensity value of sample weeds scanned, X_w symbolizes the intensity value

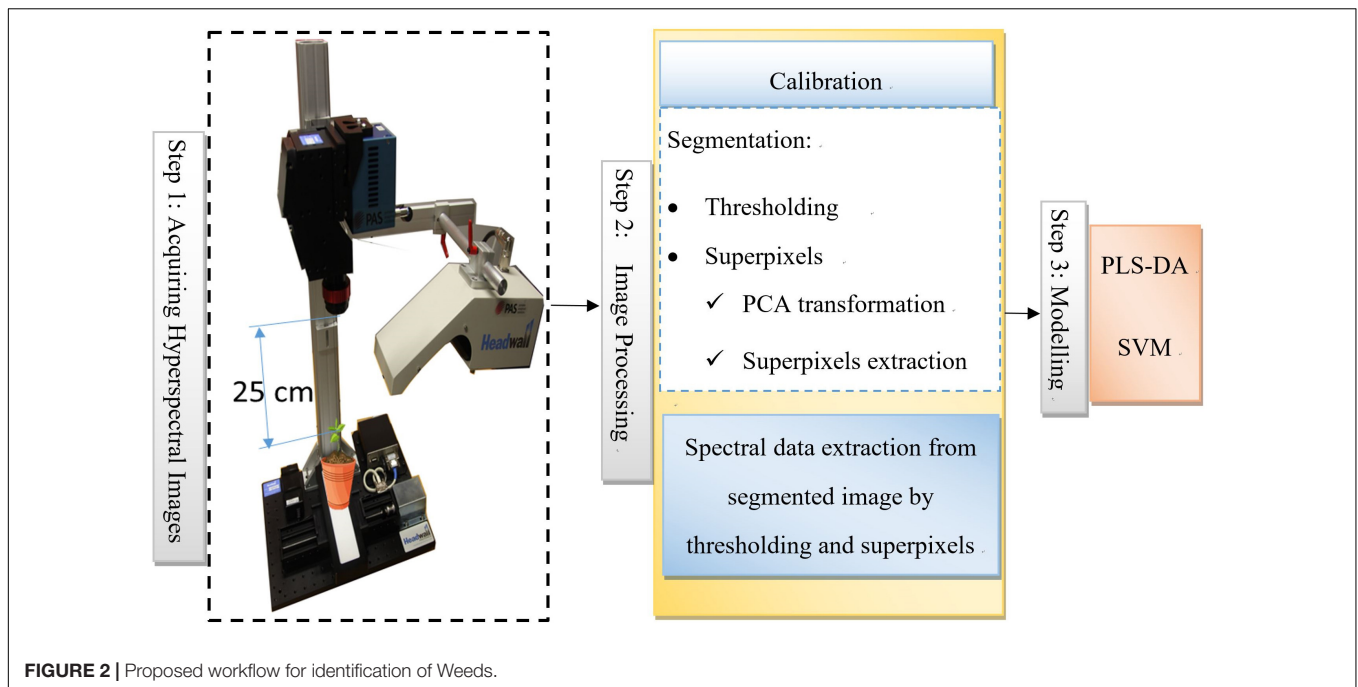


FIGURE 2 | Proposed workflow for identification of Weeds.

of captured white reference and X_D represents the intensity of the dark reference.

Segmentation

The aim of segmentation was to extract plant ROI by segregating the background (i.e., soil, stones, etc.) from the vegetation (i.e., leaves of different weeds). Custom code was generated in-house using R for thresholding segmentation and superpixel segmentation. Details of this are provided below.

- Thresholding segmentation:** A thresholding algorithm was developed by applying threshold value of 0.19 at 950 nm wavelength. This generated a mask which was then multiplied with the original HSI image to create an image of vegetation material only (Figure 3). The spectral data extracted after the thresholding segmentation was averaged and we named these averaged spectra as “Av” spectra which is the spectra for each plant with 235 components corresponding to 235 wavelengths. To obtain the Av, the mean spectrum of 120 segmented HSI images (one HSI image for each potted plant) was calculated on weed leaves resulting in the collection of 120 samples, which were used for training (96 samples), and validating (24 samples) the PLS-DA, SVM, and MLP models.
- Superpixel segmentation (Figure 4):** The simple linear iterative clustering (SLIC) algorithm (Achanta et al., 2012) was used to divide the plant image into non-overlapping patches S_p (superpixels). This was achieved by taking the similarity in spectral and spatial domains into account when grouping pixels into clusters. Principal components analysis (PCA) was used to transform the original HSI image (where each pixel contained 235 wavelengths) into three channels (each pixel contained 3 principal

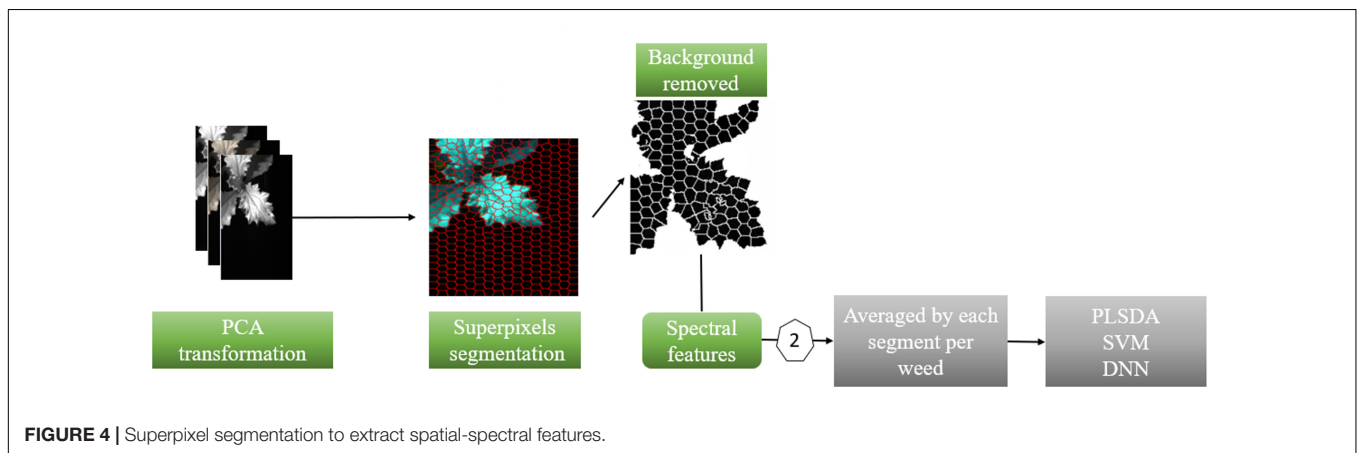
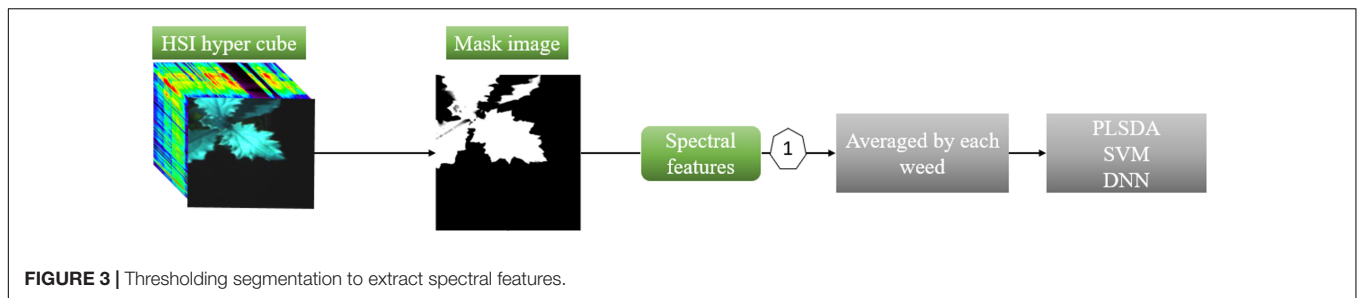
components). In total, 120 PCA images were created. This was followed by the segmentation of each PCA image into 400 patches using SLIC algorithm. The patches that contained leaves were extracted and the mean spectrum of each valid patch was extracted from the HSI image and used as the spectral and spatial (Spectro-spatial) features of the weed leaves. All patches in each plant were marked with the same label. We denoted it as “ S_p ” averaged spectra which mean each plant has “ n ” “ S_p ” spectra, each containing 235 elements corresponding to 235 wavelengths. “ n ” is the number of super-pixels for an individual plant. This number of clusters used in SLIC (i.e., 400 patches) was enough to avoid patches mixing regions from weed leaves and the background. While for this study the valid patches (which contained leaves) were separated manually, in a practical application a model may be used to separate between patches representing the plant leaves and the background.

Modeling

Three commonly used discrimination models for object classification were used in this study. These were partial least squares-discriminant analysis (PLS-DA), SVM, and MLP.

Partial Least Squares-Discriminant Analysis

Partial least squares-discriminant analysis is a very popular linear classification method in chemometrics and is based on the PLS regression algorithm (Lee et al., 2018). In PLS-DA, the output (y variable) of PLS regression is transferred into a categorical structure as reference value and descriptor matrix x is used for discrimination analysis. It typically produces the lowest within-class variability and therefore maximum separation. The scores of latent variables (LVs) from the resulting PLS-DA



model were used to identify groups of samples representing the four types/classes of weeds. The regression coefficients corresponding to those LVs with discrimination power between classes were then evaluated to identify the spectral regions potentially associated with discrimination of weeds (Wold et al., 2001; Barker and Rayens, 2003). This method has been widely used for identifying chemical traits and species classification in food and agriculture sciences (Bassbasi et al., 2014; Botelho et al., 2015; Lenhardt et al., 2015).

Support Vector Machine

Support vector machine (Boser et al., 1992), is a powerful technique and ideal for data classification, especially for the high-dimensional data with a limited number of training samples (Tarabalka et al., 2010). This method was originally defined for binary classification and has been also extended to form a multi-class classification (Pérez-Cruz and Artés-Rodríguez, 2002). This extension allows for a broad application in hyperspectral image analysis (Fauvel et al., 2008; Pal and Foody, 2010) and remote sensing (Melgani and Bruzzone, 2004; Mountrakis et al., 2011).

Multilayer Perceptron

Multilayer perceptron is a powerful machine learning technique that can characterize the features of the samples and learn the appropriate classification features from the samples (Goodfellow et al., 2016). The MLP model is dependent on multiple sets of parameters, such as the number of hidden layers, regularization parameter, and activation epoch (Ramchoun et al., 2017). Activation function allows the introduction of non-linear function to the neural network. Activation epoch also

prevents the MLP model from becoming a simple linear function with limited learning power. There are three main activation functions: hyperbolic tangent (Tanh; Kalman and Kwasny, 1992), rectifier (Xavier et al., 2011), and maxout (Goodfellow et al., 2013). The two types of regularization (L1 and L2) are useful functions in the MLP model to reduce the effect of overfitting.

Multivariate Data Analysis

Data analysis was conducted in R software version 3.1.2 (R Core Team, 2017). The “mdatools” package (Kucheryavskiy, 2019) was applied for the PLS-DA model, and the “e1071” package (Meyer et al., 2019) for SVM model construction. The “h2o” package (Erin et al., 2019) was used for MLP modeling and variable selection.

Two types of data, i.e., (a) spectral data, i.e., “Av” and (b) combined spectral and spatial data “Sp” were used for training the model by using PLS-DA, SVM, and MLP methods. These included average of a leaf spectra for each weed and the average of each selected leaf patch of a weed species (Figures 3, 4). We chose 80% of each data set for model calibration and the remaining 20% for validation and elementary testing. Two assessments were used: (1) all the pixels from segmented plant; and (2) the averaged spectra of segmented plants. This generated two datasets: “all pixels dataset” and “average dataset.” Then each dataset was split into two sets. The “all pixels” generated “all pixels calibration dataset” and “all pixels validation dataset.” Similarly, the “averaged dataset” generated “averaged calibration dataset” and “averaged validation dataset.” Each calibration dataset was used to fit a model independently. Each model was then applied to the corresponding validation dataset. The assessment of

predictions resulting from these two independent validation datasets resulted in two set of accuracies.

The pre-processing method SNV (Standard Normal Variate) was applied to the spectra before model calibration which has been shown to be a reliable pre-processing method on weed classification (Shirzadifar et al., 2018). The number of significant LVs for the PLS-DA, and the parameters of epsilon and cost for SVM models were determined using the leave-one-out cross-validation method (Sudheer et al., 2014; Vehtari et al., 2017).

Model Performance Metrics and Optimization

The parameters recall (R), precision (P), average accuracy (AA), and overall accuracy (OA) were used for PLS-DA, SVM, and MLP model performance. Four quantities from the performance of a classification process in the population of all instances were used to calculate R, P, AA, and OA: True positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) using below the equations:

$$R = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

$$AA = \frac{TP + TN}{TP + TN + FP + FN}$$

$$OA = \frac{AA_1 + AA_2 + \dots + AA_n}{n}$$

where n is the number of the classes [thistle (TT), yellow bristle grass (YBG), buttercup (BC), and wind grass (WG)]. To qualitatively evaluate the predictability power of models for weed classification, the t-SNE algorithm (Maaten and Hinton, 2008) was applied. The accuracy of a classification process was defined as the portion of true positives and true negatives in all instances.

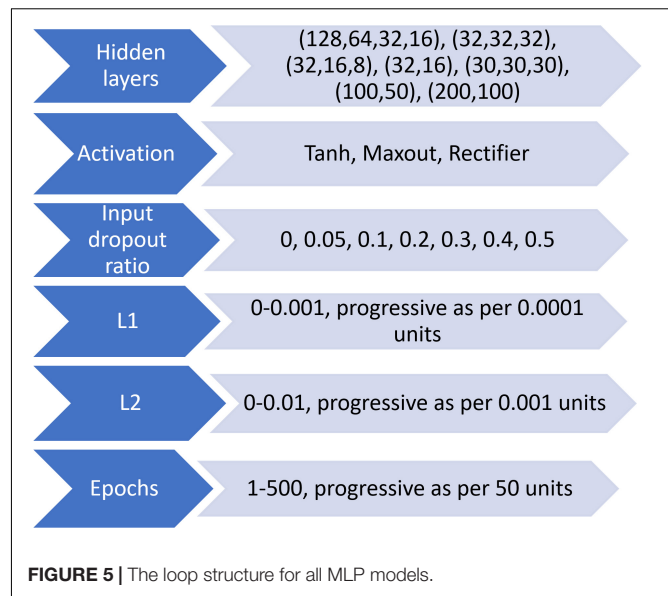
Model optimization

To find the best parameters of the MLP model for Av and Sp data, a grid loop with different hidden layers, activation function, epochs and different l1 or l2 regularization parameters was set up (Figure 5). A five-fold cross-validation was used for fine-tuning these parameters. Models with the highest OA and the lowest loss values were chosen as the final model, and the feature weight from the final best performance model was used for feature evaluation.

RESULTS

Mean Raw and Standard Normal Variate Spectra

The raw and SNV mean of 30 weeds for each species are plotted in Figure 6. While all weeds showed similar general patterns in both types of spectra, large variability between the four weeds was also observed. The SNV spectra highlights regions that can be used for discrimination between the four weed species. Three important regions were identified based on these two types of spectra: 550 to 700 nm, 1,000 to 1,200 nm, and 1,300 to 1,500 nm.



These regions are the only regions in the spectra used in this study with detectable difference in reflectance value among the four weed species.

Models Evaluation

The notations used for models developed with Av and Sp data using PLS-DA, SVM, and MLP are given below:

- Av_PLS-DA: PLS-DA model developed with Av data
- Sp_PLS-DA: PLS-DA model developed with Sp data
- Av_SVM: SVM model created with Av data
- Sp_SVM: SVM model created with Sp data
- Av_MLP: MLP model generated with Av data
- Sp_MLP: MLP model generated with Sp data

The optimal number of LVs was chosen as 10 for both Av and Sp_PLS-DA models based on cross validation. The optimal value of epsilon and cost for SVM model using Av and Sp data were 0 and 32, and 0 and 4, respectively. For the MLP model, Tanh activation with two hidden layers (32, 16) were selected for final application. The full-length spectra with 5% dropout was set as the input layer, and the four classification classes were set as the output layer. The validation set of both Av and Sp data were used to test the ability of our model for weed classification. The results of the modeling are presented in Table 1. Overall, the PLS-DA, SVM and MLP models yielded relatively high classification results based on both the Av and Sp data with an overall accuracy (OA) of 70–100%. MLP model yielded the highest recall (R), precision (P). Furthermore, average accuracy (AA) and OA with Av and Sp data were 1, 1, and 0.89, 0.90, respectively.

Multilayer Perceptron models were best performing models with Av and Sp data. The t-SNE algorithm was applied to the features that were extracted from Av_MLP model hidden layer and the raw Av spectral data for comparison. The results showed that the raw Av data did not discriminate the four species

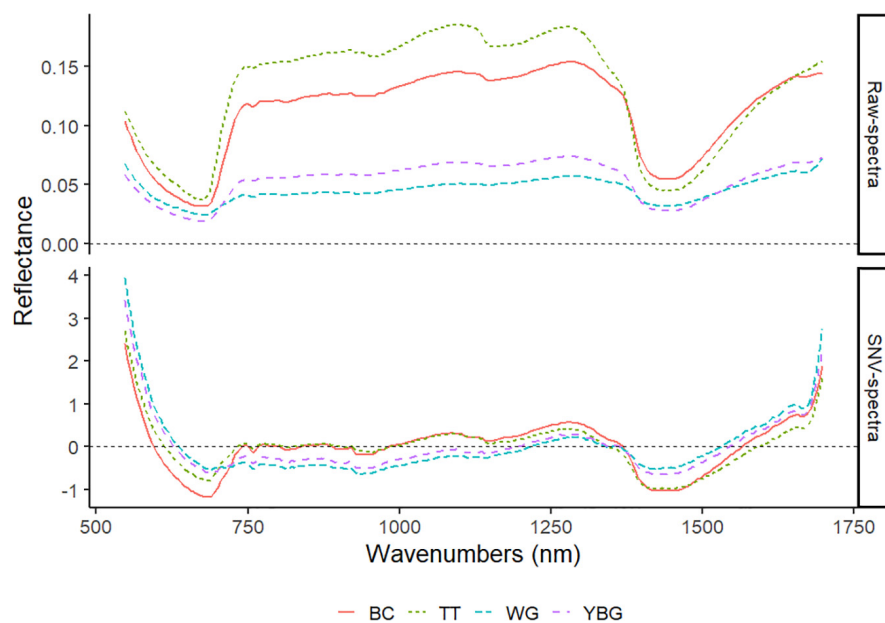


FIGURE 6 | The raw and SNV VIS-NIR mean spectra of four weed species. TT: thistle, YBG: yellow bristle grass, BC: buttercup, and WG: wind grass.

TABLE 1 | Evaluation of the performance of the proposed PLS-DA SVM, MLP models for weed recognition in the validation set, using four parameters: average accuracy (AA), overall accuracy (OA), recall (R), and precision (P), based on Av and Sp data.

		Av data				Sp data			
		R (%)	P (%)	AA (%)	OA (%)	R (%)	P (%)	AA (%)	OA (%)
PLS-DA	YBG	100	83	87	91	70	96	70	70
	BC	80	100			63	64		
	TT	100	100			72	88		
	WG	67	100			76	25		
SVM	YBG	100	100	92	92	92	90	84	86
	BC	80	80			73	81		
	TT	100	100			86	92		
	WG	89	89			86	61		
MLP	YBG	100	100	100	100	94	91	89	90
	BC	100	100			82	81		
	TT	100	100			90	95		
	WG	100	100			89	79		

TT: thistle, YBG: yellow bristle grass, BC: buttercup, and WG: wind grass.

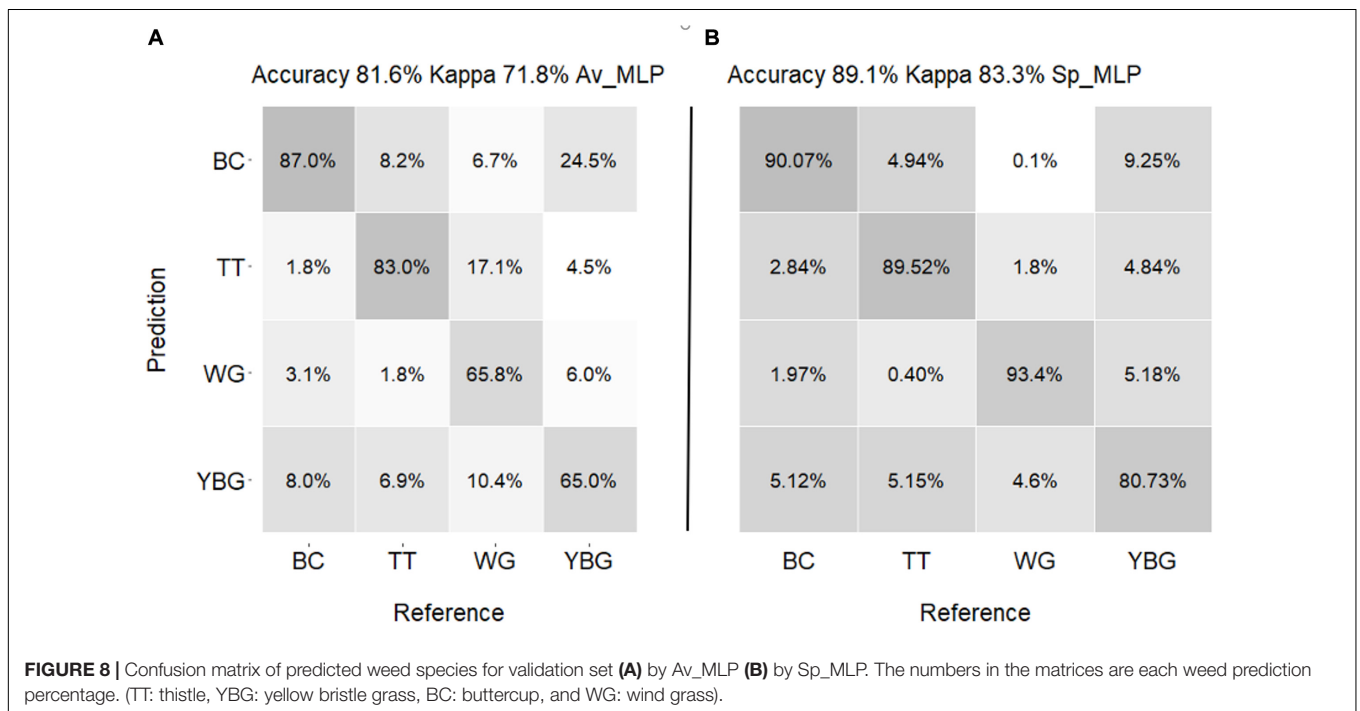
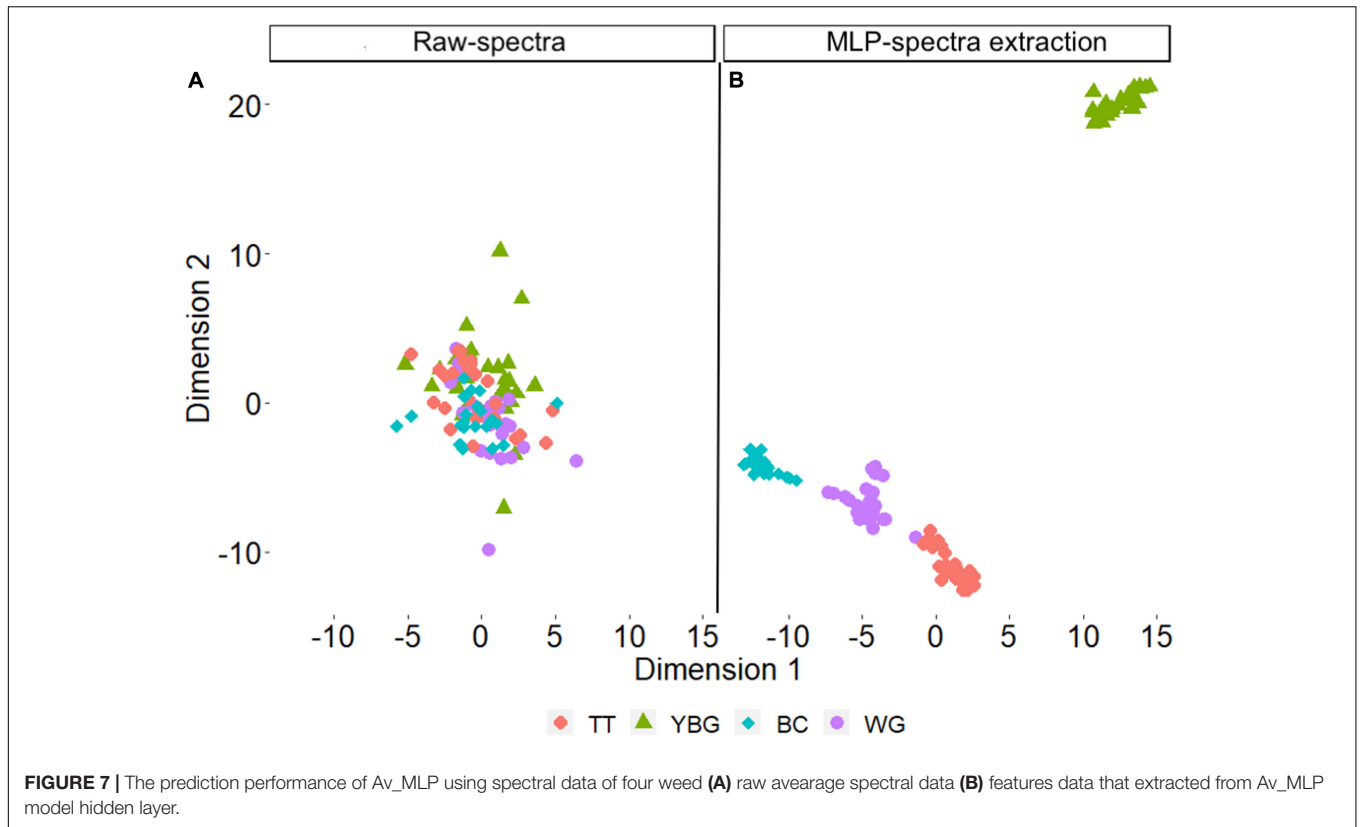
(Figure 7A). However, they were distinctly classified after the application of the MLP extraction model (Figure 7B).

Performance of the Models on All Hyperspectral Pixels

The confusion matrix of the actual weeds and predicted weeds by Av_MLP and Sp_MLP models are shown in Figure 8. The Sp_MLP model produced higher prediction accuracy than the Av_MLP model. The prediction accuracy of Av_MLP model (81.6%) using all pixels validation dataset was lower than the accuracy of using averaged validation dataset (100%; Table 1). The Sp_MLP and Av_MLP models yielded similar accuracy for the prediction of all pixels with 89.1 and 81.6% accuracy,

respectively. Av_MLP performed lower accuracies (65 and 65.8%) than the Sp_MLP model (80.7 and 93.4%) for the identification of YBG and windgrass. These results suggest that averaging all spectra across the weed species enhances the amount of information captured about the weeds. However, when spectra come from smaller regions (e.g., single pixel, from Sp) this ability is reduced.

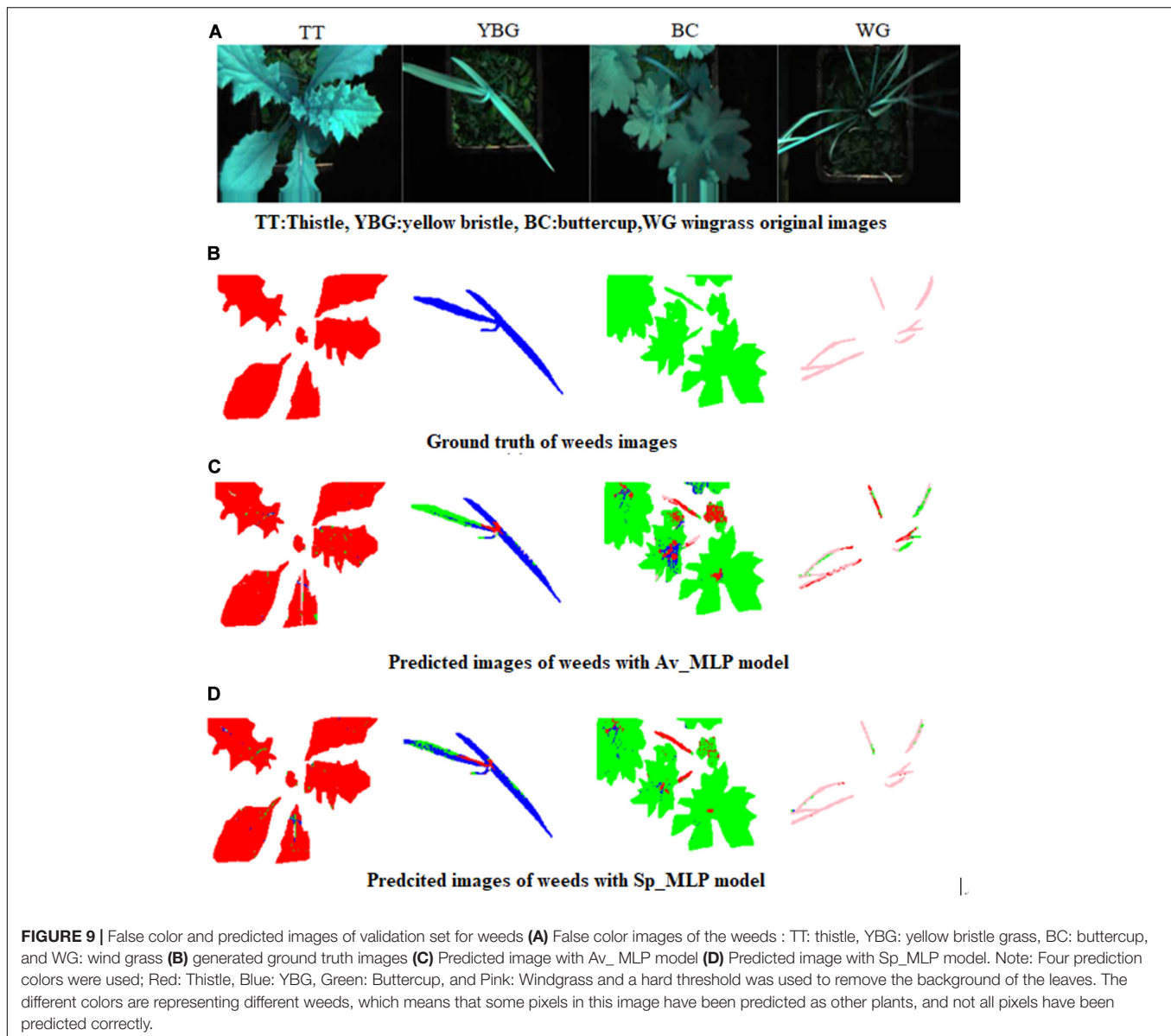
The predicted, ground truth and false color images of the weeds are presented in Figure 9. From Figures 9C,D we can conclude that Sp_MLP model showed better precision in overall weed recognition (i.e., for YBG, BC, and WG weeds) than the Av_MLP model. Furthermore, in the case of thistle, similar prediction was observed by Sp_MLP and Av_MLP models.



DISCUSSION

Targeted weed control could increase the speed and accuracy and reduce costs to farmers (Komi et al., 2007). The prerequisite of

targeted weed control is a reliable weed identification system. Many techniques have been used in recent years to increase the accuracy and speed of weed identification mainly applying RGB imaging (Solahudin et al., 2018). HSI, however, has the



advantage of identifying weeds based on their reflectance data, which is an indicator of the plant's chemical composition (Farooq et al., 2019a). Using HSI will add value to the weed identification techniques based on RGB imaging, which is hinged on shape, size, and color discrimination. Wei et al. (2014) applied canonical discriminant analysis and the PLS-DA model to data from five wavelengths (672, 757, 897, 1,117, and 1,722 nm) to discriminate soil and five weed species from winter rape. They achieved this with a high accuracy of 90.91%, which is slightly higher than our result. The reason for this might be that they only discriminated the broad leaf species, whereas we have very narrow-leaf species of grass as well. Broad leaf species are slightly simpler to identify than the narrow leaf species because of the fact of more uniform spectral data collection.

The MLP method has been widely used for classification in agricultural research (Golhani et al., 2018; Kiani et al., 2018).

In our study, four important spectral regions (550–750, 995–1,005, 1,110–1,220, and 1,380–1,470 nm) have been identified by Sp_MLP model as “the best model” with high weed identification performance. In general, the 500–750 nm spectral region has been reported as important in vegetation discrimination (Cochrane, 2000; Smith and Blackshaw, 2003; Frels et al., 2018). Further, the region around 700 nm is known to be highly informative for vegetation discrimination due to its association with chlorophyll content (Gitelson and Merzlyak, 1997). The spectra in the ranges of 880–1,000, 1,050–1,200, and 1,250–1,550 nm has been mostly associated with the third and second overtone of C-H stretching and second overtone of O-H stretching (Burns and Ciurczak, 2007; Schwanninger et al., 2011). Danson et al. (1992) described that the bands at 970, 1,200, and 1,450 nm wavelengths are water absorption bands.

All three identification models used in this study yielded high prediction accuracy both based on the Av and Sp selected spectra. However, the MLP model produced the highest accuracy, sensitivity and speciality compared to the other identification models both in the calibration and validation sets using the Av and Sp data. The t-SNE method has been recognized as a very powerful approach for data exploration and visualizing high-dimensional data (Maaten and Hinton, 2008). In this study, the first two dimensions of t-SNE extracted from original spectra and Av_MLP showed that the original spectra do not efficiently discriminate the four weed species. However, the discrimination power was improved by the Av_MLP model, with four weeds identified and discriminated.

The Av spectral models also distinctly identified the four weed species. This is likely due to chemical composition of each species (Vengris et al., 1953). However, the data could not be simply averaged for model calibration when the model was used for classification across all pixels obtained from hyperspectral images. According to Fang et al. (2015) the shape of the hyperspectral images should be assessed based on the different structures of HSI specifically if the heterogeneous spatial area is large. Averaging each selected Sp area as the input data for model calibration has the advantage of including overall and distributed leaf spectral information. It will also reduce the size of the input data for training the model. This method has been widely used for the RGB and HSI imaging (Achanta et al., 2012; Fang et al., 2015).

Overall, the Sp_MLP model showed the best predictive results, followed by Sp_SVM and Sp_PLS-DA. Silva et al. (2013) also suggested that PLS-DA has a lower classification power and is not suitable for weed identification.

The novel approach introduced in this study based on superpixels (Sp) allows the detection of weeds even where only few parts of the plant are visible, for instance in pastures where these weeds are mixed with other plant species. Thus, the introduced approach helps to overcome the challenging situation where the incomplete visibility of plant's morphology is a limiting factor for RGB imaging. Also, it is worth noting that spectral signatures could be obtained with non-imaging approaches, but this would have practical challenges in large grazing fields. The HSI used for detection of weeds in our study is either based on unique spectral signature and/or morphological features extracted from the hyperspectral images.

CONCLUSION

This study demonstrated the ability of HSI to detect unique spectral signatures of a diverse group of weed species including

grass and broadleaf as well as annual and perennial weeds. Models developed with Sp spectral data can provide better results in comparison to averaged spectral data for weed classification. Compared to the traditional classification methods, MLP is a more robust and reliable method when developed with Sp data. This novel approach based on Sp will significantly advance the applicability of HSI in plant identification. This is especially useful when it is applied in the grazing field including in mixed swards of a few plant species. Future work should focus on the development of a system that provides classification using spectral signature and/or morphological features aligned with decision tree strategies to deal with complex systems such as mixed swards.

DATA AVAILABILITY STATEMENT

Data for this manuscript is available upon request from the corresponding author.

AUTHOR CONTRIBUTIONS

KG and MR conceived and planned the experiments. MR and DH carried out the experiments and contributed to sample preparation. YL, MA-S, KI, MR, and KG contributed to the interpretation of the results. YL processed the experimental data and wrote the manuscript with input from all authors. MA-S, KI, GB, MR, and KG contributed to the final version of the manuscript. KG supervised the project. All authors provided critical feedback and helped shape the research.

FUNDING

Funding was provided by the Ministry of Business, Innovation, and Employment through the "Smart Ideas" scheme.

ACKNOWLEDGMENTS

We would like to thank Dr. Trevor James for the generous dedication of his time and commitment to scientific advice, Jack Somerton for acquisition of the hyperspectral images, and Claire Dowsett and Elizabeth North for preparation of the weeds and Trevor James and Abhimanyu Garhwal for their input as internal reviewers.

REFERENCES

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *ITPAM* 34, 2274–2282.
- Ahmad, A., Guyonneau, R., Mercier, F., and Belin, É (2018). "An Image processing method based on features selection for crop plants and weeds discrimination using RGB images," in *Image and Signal Processing*, eds A. Mansouri, A. El Moataz, F. Nouboud, and D. Mammas, (Berlin: Springer International Publishing), 3–10. doi: 10.1007/978-3-319-94211-7_1
- Ahmad, M., Khan, A. M., Mazzara, M., and Distefano, S. (2019). "Multi-layer extreme learning machine-based autoencoder for hyperspectral image classification," in *Proceedings of the 14th International Conference on Computer Vision Theory and Applications (VISAPP'19)*, Prague.
- Al-Sarayreh, M., Reis, M. M., Yan, W. Q., and Klette, R. (2018). Detection of red-meat adulteration by deep spectral-spatial features in hyperspectral images. *J. Imaging* 4:63. doi: 10.3390/jimaging4050063
- Audebert, N., Le Saux, B., and Lefèvre, S. (2019). Deep learning for classification of hyperspectral data: a comparative review. *IEEE Trans. Geosci. Remote Sens.* 7, 159–173. doi: 10.1109/MGRS.2019.2912563

- Bacco, M., Berton, A., Ferro, E., Gennaro, C., Gotta, A., Matteoli, S., et al. (2018). "Smart farming: opportunities, challenges and technology enablers," in *2018 IoT Vertical and Topical Summit on Agriculture - Tuscany (IOT Tuscany)*, Siena, doi: 10.1109/IOT-TUSCANY.2018.8373043
- Barker, M., and Rayens, W. (2003). Partial least squares for discrimination. *J. Chemometrics* 17, 166–173. doi: 10.1002/cem.785
- Bassbasi, M., De Luca, M., Ioele, G., Oussama, A., and Ragno, G. (2014). Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data. *J. Food Compos. Anal.* 33, 210–215. doi: 10.1016/j.jfca.2013.11.010
- Bateman, C. J., Fourie, J., Hsiao, J., Irie, K., Heslop, A., Hilditch, A., et al. (2020). Assessment of mixed sward using context sensitive convolutional neural networks. *Front. Plant Sci.* 11:159. doi: 10.3389/fpls.2020.00159
- Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, PA, 144–152. doi: 10.1145/130385.130401
- Botelho, B. G., Reis, N., Oliveira, L. S., and Sena, M. M. (2015). Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. *Food Chem.* 181, 31–37. doi: 10.1016/j.foodchem.2015.02.077
- Bourdôt, G., Saville, D., and Crone, D. (2003). Dairy production revenue losses in New Zealand due to giant buttercup (*Ranunculus acris*). *N. Z. J. Agric. Res.* 46, 295–303.
- Bourdôt, G. W., Fowler, S. V., Edwards, G. R., Kriticos, D. J., Kean, J. M., Rahman, A., et al. (2007). Pastoral weeds in New Zealand: status and potential solutions. *N. Z. J. Agric. Res.* 50, 139–161. doi: 10.1080/00288230709510288
- Bourdôt, G. W., Lamoureaux, S. L., Kriticos, D. J., Watt, M. S., and Brown, M. (2010). "Current and potential distributions of *Nassella neesiana* (Chilean needle grass) in Australia and New Zealand," in *Proceedings of the 17th Australasian Weeds Conference*, Christchurch.
- Bourdôt, G. W., Lamoureaux, S. L., Watt, M. S., Manning, L. K., and Kriticos, D. J. (2012). The potential global distribution of the invasive weed *Nassella neesiana* under current and future climates. *Biol. Invasions* 14, 1545–1556.
- Burns, D. A., and Ciurczak, E. W. (2007). *Handbook of Near-Infrared Analysis*. Boca Raton, FL: CRC press.
- Carreiro Soares, S. F., Medeiros, E. P., Pasquini, C., De Leis Morello, C., Harrop Galvão, R. K., and Ugulino Araújo, M. C. (2016). Classification of individual cotton seeds with respect to variety using near-infrared hyperspectral imaging. *Anal. Methods* 8, 8498–8505. doi: 10.1039/C6AY02896A
- Chen, Y., Lin, Z., Zhao, X., Wang, G., and Gu, Y. (2014). Deep learning-based classification of hyperspectral data. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 7, 2094–2107. doi: 10.1109/JSTARS.2014.2329330
- Chevallier, S., Bertrand, D., Kohler, A., and Courcoux, P. (2006). Application of PLS-DA in multivariate image analysis. *J. Chem.* 20, 221–229. doi: 10.1002/cem.994
- Cochrane, M. A. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. *Int. J. Remote Sens.* 21, 2075–2087. doi: 10.1080/01431160050021303
- Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30, 271–278. doi: 10.1016/0034-4257(89)90069-2
- Danson, F. M., Steven, M. D., Malthus, T. J., and Clark, J. A. (1992). High-spectral resolution data for determining leaf water content. *Int. J. Remote Sens.* 13, 461–470. doi: 10.1080/01431169208904049
- Ebbers, M. J. H., Wallis, I. R., Dury, S., Floyd, R., and Foley, W. J. (2002). Spectrometric prediction of secondary metabolites and nitrogen in fresh eucalyptus foliage: towards remote sensing of the nutritional quality of foliage for leaf-eating marsupials. *Aust. J. Bot.* 50, 761–768. doi: 10.1071/BT02025
- Erin, L., Navdeep, G., Spencer, A., Anqi, F., Arno, C., Cliff, C., et al. (2019). *h2o: R Interface for H2O, R Package Version 3.24.0.5*. Available at: <https://github.com/h2oai/h2o-3> (accessed May 25, 2019).
- Fan, F., Ma, Y., Li, C., Mei, X., Huang, J., and Ma, J. (2017). Hyperspectral image denoising with superpixel segmentation and low-rank representation. *Inf. Sci.* 397–398, 48–68. doi: 10.1016/j.ins.2017.02.044
- Fang, L., Li, S., Duan, W., Ren, J., and Benediktsson, J. A. (2015). Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels. *IEEE. Trans. Geosci. Remote. Sens.* 53, 6663–6674. doi: 10.1109/TGRS.2015.2445767
- Farooq, A., Hu, J., and Jia, X. (2019a). Analysis of spectral bands and spatial resolutions for weed classification via deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* 16, 183–187. doi: 10.1109/LGRS.2018.2869879
- Farooq, A., Jia, X., Hu, J., and Zhou, J. (2019b). Multi-resolution weed classification via convolutional neural network and superpixel based local binary pattern using remote sensing images. *Remote Sens.* 11:1692. doi: 10.3390/rs11141692
- Fauvel, M., Benediktsson, J. A., Chanussot, J., and Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE. Trans. Geosci. Remote. Sens.* 46, 3804–3814. doi: 10.1109/TGRS.2008.922034
- Ferreira, M. P., Wagner, F. H., Aragão, L. E. O. C., Shimabukuro, Y. E., and De Souza Filho, C. R. (1992). Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. *ISPRS J. Photogr. Remote Sens.* 149, 119–131. doi: 10.1016/j.isprsjprs.2019.01.019
- Frels, K., Guttieri, M., Joyce, B., Leavitt, B., and Baenziger, P. S. (2018). Evaluating canopy spectral reflectance vegetation indices to estimate nitrogen use traits in hard winter wheat. *Field Crops Res.* 217, 82–92. doi: 10.1016/j.fcr.2017.12.004
- Gitelson, A. A., and Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697. doi: 10.1080/014311697217558
- Glorot, X., Bordes, A., and Bengio, Y. (2011). "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings*, 15, 315–323. Available online at: <http://proceedings.mlr.press/v15/glorot11a.html>
- Golhani, K., Balasundram, S. K., Vadmalai, G., and Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Inf. Process. Agric.* 5, 354–371. doi: 10.1016/j.inpa.2018.05.002
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. Cambridge, MA: MIT press.
- Goodfellow, I., Warde-Farley, D., Mirza, M., Courville, A., and Bengio, Y. (2013). "Maxout networks," in *Proceedings of the 30th International Conference on Machine Learning (PMLR)*, 28, 1319–1327. Available online at: <http://proceedings.mlr.press/v28/goodfellow13.html>
- Griffel, L. M., Delparte, D., and Edwards, J. (2018). Using support vector machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. *Comput. Electron. Agric.* 153, 318–324. doi: 10.1016/j.compag.2018.08.027
- Jiang, J., Chen, C., Yu, Y., Jiang, X., and Ma, J. (2017). Spatial-aware collaborative representation for hyperspectral remote sensing image classification. *IEEE Geosci. Remote. Sens. Lett.* 14, 404–408. doi: 10.1109/LGRS.2016.2645708
- Kalman, B. L., and Kwasny, S. C. (1992). "Why tanh: choosing a sigmoidal function," in *Proceedings 1992 of International Joint Conference on Neural Networks (IJCNN)*, Baltimore, MD. doi: 10.1109/IJCNN.1992.227257
- Kiani, S., Van Ruth, S. M., Minaei, S., and Ghasemi-Varnamkhashti, M. (2018). Hyperspectral imaging, a non-destructive technique in medicinal and aromatic plant products industry: current status and potential future applications. *Comput. Electron. Agric.* 152, 9–18. doi: 10.1016/j.compag.2018.06.025
- Komi, P. J., Jackson, M. R., and Parkin, R. M. (2007). "Plant classification combining colour and spectral cameras for weed control purposes," in *IEEE International Symposium on Industrial Electronics*, Vigo, 2039–2042. doi: 10.1109/ISIE.2007.4374921
- Kucheryavskiy, S. (2019). *mdatools: Multivariate Data Analysis for Chemometrics, R Package Version 0.9.4*. Available at: <https://CRAN.R-project.org/package=mdatools> (accessed October 22, 2020).
- Lamoureaux, S. (2014). The potential distribution of yellow bristle grass (*Setaria pumila*) in New Zealand. *N. Z. Plant Protect.* 67, 226–230.
- Lee, L. C., Liang, C.-Y., and Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* 143, 3526–3539. doi: 10.1039/C8AN00599K
- Lenhardt, L., Bro, R., Zeković, I., Dramićanin, T., and Dramićanin, M. D. (2015). Fluorescence spectroscopy coupled with PARAFAC and PLS DA for characterization and classification of honey. *Food Chem.* 175, 284–291. doi: 10.1016/j.foodchem.2014.11.162

- Li, Y., Xie, W., and Li, H. (2017). Hyperspectral image reconstruction by deep convolutional neural network for classification. *Pattern Recognition* 63, 371–383. doi: 10.1016/j.patcog.2016.10.019
- Li, Z., and Chen, J. (2015). “Superpixel segmentation using linear spectral clustering” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, 1356–1363.
- Liu, Y., and Zhang, Y. (2018). “Hyperspectral image refined plant classification by graph-based composite kernel,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Valencia, doi: 10.1109/IGARSS.2018.8518585
- Lunga, D., Prasad, S., Crawford, M. M., and Ersoy, O. (2014). Manifold-learning-based feature extraction for classification of hyperspectral data: a review of advances in manifold learning. *IEEE. Signal. Process. Mag.* 31, 55–66. doi: 10.1109/MSP.2013.2279894
- Maaten, L. V. D., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
- McClearn, B., Gilliland, T., Guy, C., Dineen, M., Coughlan, F., and McCarthy, B. (2020). The effect of perennial ryegrass ploidy and white clover inclusion on milk production of dairy cows. *Anim. Prod. Sci.* 60, 143–147. doi: 10.1071/AN18539
- Melgani, F., and Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE. Trans. Geosci. Remote. Sens.* 42, 1778–1790. doi: 10.1109/TGRS.2004.831865
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. (2019). “Package ‘e1071’”, in: *The R Journal*. Available at: <https://CRAN.R-project.org/package=e1071> (accessed October 14, 2020).
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., and Scheunders, P. (2017). Close range hyperspectral imaging of plants: a review. *Biosys. Eng.* 164, 49–67. doi: 10.1016/j.biosystemseng.2017.09.009
- Mountrakis, G., Im, J., and Ogole, C. (2011). Support vector machines in remote sensing: a review. *ISPRS J. Photogramm. Remote. Sens.* 66, 247–259. doi: 10.1016/j.isprsjprs.2010.11.001
- Pal, M., and Foody, G. M. (2010). Feature selection for classification of hyperspectral data by SVM. *IEEE. Trans. Geosci. Remote. Sens.* 48, 2297–2307. doi: 10.1109/TGRS.2009.2039484
- Peng, J., Zhou, Y., and Chen, C. L. P. (2015). Region-kernel-based support vector machines for hyperspectral image classification. *IEEE. Trans. Geosci. Remote. Sens.* 53, 4810–4824. doi: 10.1109/TGRS.2015.2410991
- Pérez-Cruz, F., and Artés-Rodríguez, A. (2002). “Puncturing multi-class support vector machines,” in *Artificial Neural Networks — ICANN 2002*, ed. J. R. Dorronsoro, Berlin: Springer Berlin Heidelberg. doi: 10.1007/3-540-46084-5_122
- R Core Team, (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Raja, R., Nguyen, T. T., Slaughter, D. C., and Fennimore, S. A. (2020). Real-time weed-crop classification and localisation technique for robotic weed control in lettuce. *Biosys. Eng.* 192, 257–274. doi: 10.1016/j.biosystemseng.2020.02.002
- Ramchoun, H., Idrissi, M., Ghanou, Y., and Ettaouil, M. (2017). New modeling of multilayer perceptron architecture optimization with regularization: an application to pattern classification. *IAENG.Int. J. Comput. Sci.* 44, 261–269.
- Reis, M. M., Van Beers, R., Al-Sarayreh, M., Shorten, P., Yan, W. Q., Saeys, W., et al. (2018). Chemometrics and hyperspectral imaging applied to assessment of chemical, textural and structural characteristics of meat. *Meat. Sci.* 144, 100–109. doi: 10.1016/j.meatsci.2018.05.020
- Ren, and Malik. (2003). “Learning a classification model for segmentation,” in *Proceedings Ninth IEEE International Conference on Computer Vision*, Nice, doi: 10.1109/ICCV.2003.1238308
- Schwanninger, M., Rodrigues, J. C., and Fackler, K. (2011). A review of band assignments in near infrared spectra of wood and wood components. *J. Near Infrared Spectrosci.* 19, 287–308. doi: 10.1255/jnirs.955
- Sharma, S., and Bhavya, S. (2020). Thresholding algorithms for image segmentation - entropy based comparison. *Glob. J. Enterp. Inf. Syst.* 11, 19–23.
- Shirzadifar, A., Bajwa, S., Mireei, S. A., Howatt, K., and Nowatzki, J. (2018). Weed species discrimination based on SIMCA analysis of plant canopy spectral data. *Biosyst. Eng.* 171, 143–154. doi: 10.1016/j.biosystemseng.2018.04.019
- Shorten, P. R., Leath, S. R., Schmidt, J., and Ghamkhar, K. (2019). Predicting the quality of ryegrass using hyperspectral imaging. *Plant Methods* 15:63.
- Silva, L. O. L. A., Koga, M. L., Cugnasca, C. E., and Costa, A. H. R. (2013). Comparative assessment of feature selection and classification techniques for visual inspection of pot plant seedlings. *Comput. Electron. Agric.* 97, 47–55. doi: 10.1016/j.compag.2013.07.001
- Smith, A. M., and Blackshaw, R. E. (2003). Weed–Crop discrimination using remote sensing: a detached leaf experiment. *Weed Technol.* 17, 811–820. doi: 10.1614/WT02-179
- Solahudin, M., Widodo, S., and Wahyu, W. (2018). Development of weeds density evaluation system based on RGB sensor. *IOP Conf. Ser. Earth Environ. Sci.* 147:012047. doi: 10.1088/1755-1315/147/1/012047
- Sudheer, C., Maheswaran, R., Panigrahi, B. K., and Mathur, S. (2014). A hybrid SVM-PSO model for forecasting monthly streamflow. *Neural Comput. Appl.* 24, 1381–1389. doi: 10.1007/s00521-013-1341-y
- Sutskever, I., and Hinton, G. E. (2008). Deep, narrow sigmoid belief networks are universal approximators. *Neural Comput.* 20, 2629–2636. doi: 10.1162/neco.2008.12-07-661
- Taneja, K. (2020). *Machine Learning Based Data Driven Modelling of Time Series of Power Plant Data*. Waterloo, ONT: UWSpace.
- Tarabalka, Y., Fauvel, M., Chanussot, J., and Benediktsson, J. A. (2010). SVM- and MRF-based method for accurate classification of hyperspectral images. *IEEE Geosci. Remote. Sens. Lett.* 7, 736–740. doi: 10.1109/LGRS.2010.2047711
- Te, M., Tetsuya, I., Mayuka, B., and Satoru, T. (2019). Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI). *Holzforchung* 73, 323–330. doi: 10.1515/hf-2018-0128
- Vaiphasa, C., Skidmore, A. K., De Boer, W. F., and Vaiphasa, T. (2007). A hyperspectral band selector for plant species discrimination. *ISPRS J. Photogramm. Remote. Sens.* 62, 225–235. doi: 10.1016/j.isprsjprs.2007.05.006
- Vaughn, A. B., Neilsen, T. B., Gee, K. L., Okamoto, K., and Akamine, M. (2016). Near-field spatial variation in similarity spectra decomposition of a Mach 1.8 laboratory-scale jet. *Proc. Meet. Acoust.* 29:045004. doi: 10.1121/2.000456
- Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27, 1413–1432. doi: 10.1007/s11222-016-9696-4
- Vengris, J., Drake, M., Colby, W. G., and Bart, J. (1953). Chemical composition of weeds and accompanying crop Plants1. *Agron. J.* 45, 213–218. doi: 10.2134/agronj1953.00021962004500050009x
- Walter, C., Onisimo, M., and Chandrashekhar, B. (2019). Phenology-based discrimination of maize (*Zea mays* L.) varieties using multitemporal hyperspectral data. *J. Appl. Remote Sens.* 13, 1–19. doi: 10.1117/1.JRS.13.017504
- Wang, A., Zhang, W., and Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* 158, 226–240. doi: 10.1016/j.compag.2019.02.005
- Wei, X., Bai, J., Zhang, J., and Wu, S. (2014). Weed identification from winter rape at seedling stage based on spectral variable selection. *Sensor Lett.* 12, 530–536. doi: 10.1166/sl.2014.3107
- Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics Intellig. Lab. Syst.* 58, 109–130. doi: 10.1016/S0169-7439(01)00155-1
- Xu, X., Li, W., Ran, Q., Du, Q., Gao, L., and Zhang, B. (2018). Multisource remote sensing data classification based on convolutional neural network. *IEEE. Trans. Geosci. Remote. Sens.* 56, 937–949. doi: 10.1109/TGRS.2017.2756851
- Yang, X., Hong, H., You, Z., and Cheng, F. (2015). Spectral and image integrated analysis of hyperspectral data for waxy corn seed

- variety classification. *Sensors* 15, 15578–15594. doi: 10.3390/s150715578
- Youngentob, K. N., Renzullo, L. J., Held, A. A., Jia, X., Lindenmayer, D. B., and Foley, W. J. (2012). Using imaging spectroscopy to estimate integrated measures of foliage nutritional quality. *Methods Ecol. Evol.* 3, 416–426. doi: 10.1111/j.2041-210X.2011.00149.x
- Yu, X., Liu, K., Wu, D., and He, Y. (2012). Raisin quality classification using least squares support vector machine (LSSVM) based on combined color and texture features. *Food. Bioproc. Tech.* 5, 1552–1563. doi: 10.1007/s11947-011-0531-9
- Yuan, L., Yan, P., Han, W., Huang, Y., Wang, B., Zhang, J., et al. (2019). Detection of anthracnose in tea plants based on hyperspectral imaging. *Comput. Electron. Agric.* 167:105039. doi: 10.1016/j.compag.2019.105039
- Conflict of Interest:** MA-S, DH, GB, MR, and KG were employed by the company AgResearch Ltd. KI was employed by the company Red Fern Solutions Ltd.
- The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2021 Li, Al-Sarayreh, Irie, Hackell, Bourdot, Reis and Ghamkhar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*