



## OPEN ACCESS

EDITED BY  
Cataldo Pulvento,  
University of Bari Aldo Moro, Italy

REVIEWED BY  
Leiqing Pan,  
Nanjing Agricultural University,  
China  
Roxana Vidican,  
University of Agricultural Sciences and  
Veterinary Medicine of Cluj-Napoca,  
Romania

\*CORRESPONDENCE  
Xiaowen Hu

✉ huxw@lzu.edu.cn

## SPECIALTY SECTION

This article was submitted to  
Crop and Product Physiology,  
a section of the journal  
Frontiers in Plant Science

RECEIVED 01 December 2022

ACCEPTED 16 January 2023

PUBLISHED 07 February 2023

## CITATION

Fu X, Bai M, Xu Y, Wang T, Hui Z and Hu X  
(2023) Cultivars identification of oat  
(*Avena sativa* L.) seed via  
multispectral imaging analysis.  
*Front. Plant Sci.* 14:1113535.  
doi: 10.3389/fpls.2023.1113535

## COPYRIGHT

© 2023 Fu, Bai, Xu, Wang, Hui and Hu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Cultivars identification of oat (*Avena sativa* L.) seed via multispectral imaging analysis

Xiuzhen Fu<sup>1,2,3,4</sup>, Mengjie Bai<sup>1,2,3,4</sup>, Yawen Xu<sup>1,2,3,4</sup>, Tao Wang<sup>1,2,3,4</sup>, Zhenning Hui<sup>1,2,3,4</sup> and Xiaowen Hu<sup>1,2,3,4\*</sup>

<sup>1</sup>State Key Laboratory of Herbage Improvement and Grassland Agro-ecosystems, Lanzhou University, Lanzhou, China, <sup>2</sup>Key Laboratory of Grassland Livestock Industry Innovation, Ministry of Agriculture and Rural Affairs, Lanzhou University, Lanzhou, China, <sup>3</sup>Engineering Research Center of Grassland Industry, Ministry of Education, Lanzhou University, Lanzhou, China, <sup>4</sup>College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou, China

Cultivar identification plays an important role in ensuring the quality of oat production and the interests of producers. However, the traditional methods for discrimination of oat cultivars are generally destructive, time-consuming and complex. In this study, the feasibility of a rapid and nondestructive determination of cultivars of oat seeds was examined by using multispectral imaging combined with multivariate analysis. The principal component analysis (PCA), linear discrimination analysis (LDA) and support vector machines (SVM) were applied to classify seeds of 16 oat cultivars according to their morphological features, spectral traits or a combination thereof. The results demonstrate that clear differences among cultivars of oat seeds could be easily visualized using the multispectral imaging technique and an excellent discrimination could be achieved by combining data of the morphological and spectral features. The average classification accuracy of the testing sets was 89.69% for LDA, and 92.71% for SVM model. Therefore, the potential of a new method for rapid and nondestructive identification of oat cultivars was provided by multispectral imaging combined with multivariate analysis.

## KEYWORDS

**oat, seed feature, nondestructive identification, linear discriminant analysis, support vector machine**

## 1 Introduction

Oat (*Avena sativa* L.), an annual grass species, is widely cultivated worldwide as one of the most important traditional human food as well as a healthy source of proteins, unsaturated fatty acids, vitamins, phenolics and significant quantities of dietary fibre, contributing to the beneficial health effects (Marshall et al., 2013; Singh et al., 2013; Menon et al., 2016). Meanwhile, oat is an important forage in cool semiarid regions (Carr et al., 2004). It also can be used as silage, pasture or hay and plays a vital role in sustainable agriculture related to animal production by providing forage in winter when warm-season grass production is limited by freezing temperatures (Maloney et al., 1999; Kim et al., 2017).

It is well-known that with the continuous application of new breeding techniques, more and more cultivars oat available for cultivation are produced worldwide. Different oat cultivars show diverse adaptability and growth potential for their different genetic basis (Wang et al., 2019). For example, significant differences were found in seed yield in 196 native Turkish oats cultivars in 2-year field experiments (Dumlupinar et al., 2011). Likewise, the genotype (cultivar/strain) effect on the oat grain yield was the greatest, accounting for 61% of yield variation (Tobiasz-Salach et al., 2019). In addition, different cultivars showed different tolerance under salt stress (Zhang et al., 2022), heat stress (Rauf et al., 2020) and drought stress (Benlioglu and Ozkan, 2021). Studies have also shown that oat cultivars have a significant effect on the chemical component of oat (Silva et al., 2012; Tong et al., 2014). According to the purpose of cultivation and the different growth period of cultivars, selecting proper cultivars to cultivate in different places is an important means to ensure the quality of oat production. Therefore, effective cultivar identification to sort and ensure the purity of oat seeds is not only conducive to the healthy development of the seed industry to protect the breeding of new cultivars, but also conducive to the healthy development of agriculture to ensure the interests of producers.

Generally, cultivar identification is carried out on the basis of specific traits specified in official protocols during the registration process, such as seed shape, size, color and field seedling growth characteristics (Grillo et al., 2011). However, this identification process requires experienced and professional technicians, and the whole process is expensive and time-consuming. The high number of new cultivars and high similarity do not allow classified only based on phenotype (Dvořáček et al., 2003). Due to the continuous development and application of molecular marker technology in cultivar identification, molecular detection has been considered as an effective identification method with high precision (Saccomanno et al., 2020). RFLP (restriction fragment length polymorphism), RAPD (randomly amplified polymorphic DNA), SSR (simple sequence repeat) and other molecular markers have been successfully applied in oat cultivars identification (O'Donoghue et al., 1994; Paczos-Grzeda, 2004; Wight et al., 2010; Pipan et al., 2021). However, the lack of validated markers is one of the constraints of using molecular method (Sakiyama et al., 2014), and it is generally conducted destructive. Therefore, the potential of using non-destructive methods such as imaging techniques, NIR spectroscopy or precise remote sensors to overcome the limitations of traditional methods is gaining increasing attention.

The use of image analysis technology and measurement of the geometric shape of single seed can provide an alternative to manual inspection of seed for characteristic properties, thus the information could be visually obtained much faster (Geetha et al., 2011; Grillo et al., 2011; Sumathi and Balamurugan, 2013). Near-infrared spectroscopy (NIRS), as a sensing technology, is for the determination of chemical composition, both on the seed surface and inside the seeds (Boelt et al., 2018), providing potential for cultivar identification. Multispectral imaging, as an emerging technology, combines imaging and spectroscopy techniques, which can not only provide spatial information of seeds such as area, perimeter, length, shape, color, etc., but also provide detailed information about chemical composition, structure and other internal characteristics, determining various traits

of seeds simultaneously. The merits of non-destructive, easy and no sample pre-treatment make it widely applied to seed quality testing, such as identifying rice seeds (Liu et al., 2016b), classifying different tomato seed cultivars (Shrestha et al., 2016), discriminating alfalfa cultivars seed (Yang et al., 2020; Jia et al., 2022), classifying *Jatropha curcas* seed health (Barboza da Silva et al., 2021). Sumathi and Balamurugan applied machine vision technology to identify the morphology of 11 oat cultivars, and the observation was more accurate and objective than manual method (Sumathi and Balamurugan, 2013). However, there is no report on the application of multispectral imaging technique for rapid recognition for oat cultivars.

From the above, our study aimed to develop a non-destructive, rapid and high-throughput oat cultivars identification method based on multispectral imaging technology combined with multivariate analysis.

## 2 Material and methods

### 2.1 Seed samples

Sixteen oat cultivars were Blade, Deon, Jerry, Kona, Longyan1, Longyan2, Longyan3, Longyan4, Brave1, Morgan, Monica, Tanke, Youmu1, Baiyan7, Dingyan2 and Quebec. The seeds of Dingyan2, Baiyan7 and Quebec were provided by the Academy of Agricultural Science, Dingxi, Gansu Province; and the other cultivars were provided by Beijing Best Grass Industry Co., Ltd. Photo of the seeds is displayed in Figure 1.

For each cultivar, 200 seeds were used for classification experiments in January 2022. For each species, 140 seeds for each sample were randomly selected as training set and remaining 60 seeds were used for independent testing set.

### 2.2 Multispectral imaging system and seed imaging

Multispectral images were captured with a VideometerLab4 (Videometer, Hørsholm, Denmark) multispectral imaging system which is the same as the study by Hu et al. (2020). Samples were placed beneath a hollow integrating sphere, with a camera located in the top of the sphere. During image capture, the sphere closes over the sample stage to create optically closed conditions, allowing even lighting with minimal shadows and specular reflection. Samples were illuminated by 19 high power light emitting diodes (LEDs) at specific wavelengths: 365, 405, 430, 450, 470, 490, 515, 540, 570, 590, 630, 645, 660, 690, 780, 850, 880, 890, and 970 nm. The detailed information about VideometerLab4 correction have been described in their study.

Seeds were kept in water-proof bags in a storage room with average temperature of  $-18^{\circ}\text{C}$  before imaging. During seed imaging, disease-free oat seeds were selected. For 200 seeds of each cultivar, 25 seeds (for 8 times) laid in a petri dish were used for image acquiring with the VideometerLab4 by controlling the ball falling and closing the top camera. The seeds were evenly arranged without contact to ensure that the extraction of the characteristic value of a single seed was not affected in the later stage.

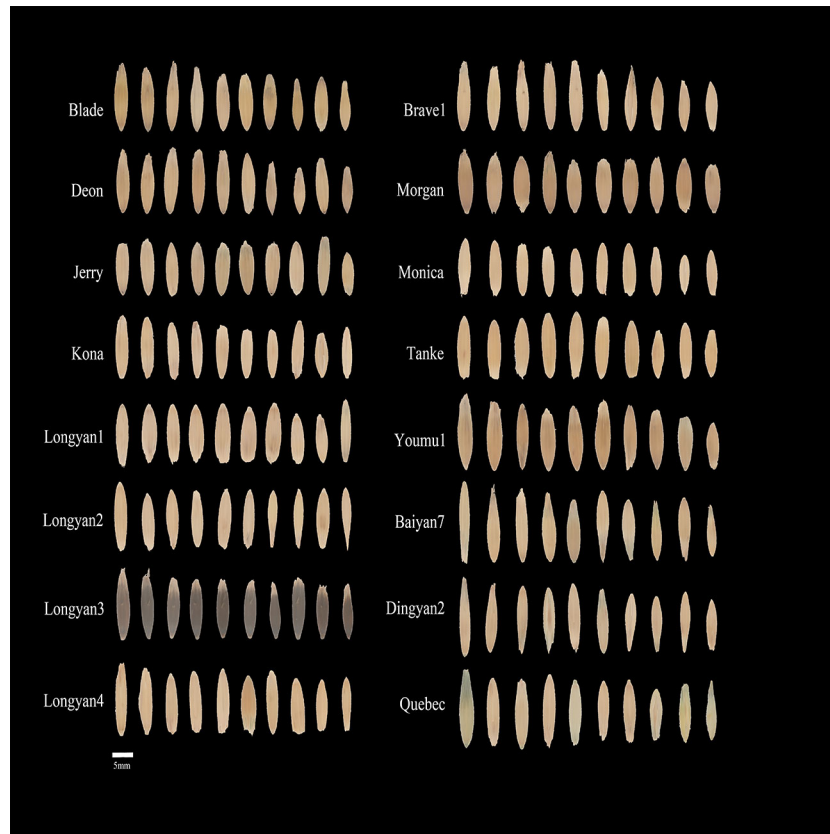


FIGURE 1  
Seed image of 16 cultivars for *Avena sativa* L.

## 2.3 Multispectral image analysis

The main objects appearing in the multispectral image are the 25 oat seeds in addition to some other objects, such as the petri dish and its surrounding background that should be removed from the image before extracting morphological spectral information of the individual seed. Thus, use a normalized canonical discriminant analysis (nCDA) in the VideometerLab software version 3.12 to segment the acquired image preliminarily and produce an image mask where only the oat seeds were isolated from non-seed pixels. The nCDA is a dimension-reduction method that maximized variation between the classes and reduced variation within a class (Olesen et al., 2015).

Then, attributes of the seeds such as morphological traits and main spectral features of all individual seeds were extracted from the image analysis and processed. The morphological traits included area, length, width, perimeter, width/length ratio, compactness circle, compactness ellipse, BetaShape\_a, BetaShape\_b, vertical skewness, CIELab L\*, CIELab a\*, CIELab b\*, saturation and hue. The area, length, width and perimeter are binary features of the seeds, which are used to show seed size. Width/length ratio is the ratio of width to length of the sample. Compactness circle and compactness ellipse represent respectively the ratio of the measured object area to the area of a circle of the same length and the ratio to the area of an ellipse of the same length and width. BetaShape\_a and BetaShape\_b can be understood as endpoint parameters at both ends of the image.

Vertical skewness refers to the skewness around the horizontal central axis. As shape features, all of them are described the shape and outlines characteristics of measured object. The saturation is defined as the brilliant of the sample color, and hue, as the biggest feature of color, is the appearance of the color. The  $L^*a^*b^*$  is an international standard for color measurement proposed by Commission International d'Eclairage (CIE) in 1976.  $L^*$  is the luminance component, and  $a^*$  represents from green to red color and  $b^*$  from blue to yellow are the two chromatic components (Papadakis et al., 2000). The extracted seed spectral signatures represent the mean intensity of the reflected light for each single wavelength calculated from all seed pixels in the image.

## 2.4 Multivariate data analysis

Multivariate analysis, including principal component analysis (PCA), linear discrimination analysis (LDA) and support vector machines (SVM), was conducted using *FactoMineR*, *MASS*, and *e1071* packages in R 2.4.1 respectively, to classify and distinguish the oat cultivar seed (Hu et al., 2020).

### 2.4.1 PCA

Principal component analysis (PCA), the most commonly used unsupervised exploratory multivariate data analysis technique, was carried out to identify the patterns hidden in the morphological

features and spectral data of all extracted seeds (ElMasry et al., 2019). PCA is often used to reduce the dimension of the dataset and resolve most of the information in the original variables with fewer variables, thus transforming multiple variables into few comprehensive indicators, namely principal components, without losing important information. The complex factors are grouped into several PCs by PCA, and unknown samples to be classified. In this work, PCA was employed to investigate the classification of the oat seeds.

## 2.4.2 LDA

Linear discriminant analysis (LDA), as a classical and supervised machine learning algorithm which minimizes the within-class distance and maximizes the between-class distance simultaneously, thus achieving maximum discrimination, mainly used for prediction and sample recognition and classification (ElMasry et al., 2019; Yu et al., 2021). In this study, seeds of each cultivar were randomly divided into 140 seeds (70% of total samples) as training and 60 seeds (remaining 30%) as testing sets. The models for LDA classification were built using the training packages.

For each discriminant model, the recognition levels of samples in the testing set were the proportion of the number of seeds correctly identified to the total number of seeds in the testing set, which are calculated using the following equation.

$$\text{Accuracy}(\%) = \frac{\text{Correctly classified seeds}}{\text{Total number of seeds}} \times 100 \quad (\text{eq.1})$$

## 2.4.3 SVM

Support vector machine (SVM), as a supervised machine learning model, has been widely used in classification and nonlinear function

estimation (Wang and Hu, 2005). As for multivariate function estimation or non-linear classification tasks, SVM performed effectively. Different from other analysis methods, SVM can use fewer training variables or samples in high-dimensional characteristic space. It has proven to be powerful in NIR spectra classification (Devos et al., 2009).

The “Accuracy” of discrimination was determined according to Equation (1).

## 3 Results

### 3.1 Seed morphologic features

For binary features, there were significant differences among 16 cultivars. The seeds of Longyan4 in area were the largest, and of Monica were the smallest. The length and perimeter of Monica were significantly lower than those of the other cultivars (Figure 2).

For color features, 16 cultivars were different. The mean values of CIELab L\* of Monica, CIELab a\* of Morgan, CIELab b\* and saturation of Tanke were the highest, the hue values of Baiyan7, Brave1 and Blade were significantly higher than the other cultivars. Longyan3, whose mean values of CIELab L\*, CIELab a\*, CIELab b\* and saturation were the lowest, and the hue was only higher than that of Morgan (Figure 3).

For shape features, significant differences were found among 16 cultivars. The width/length ratio, BetaShape\_b and compactness circle of Morgan had the largest values. Baiyan7 had the highest value of BetaShape\_a, but the lowest values of compactness circle, compactness ellipse and vertical skewness (Figure 4).

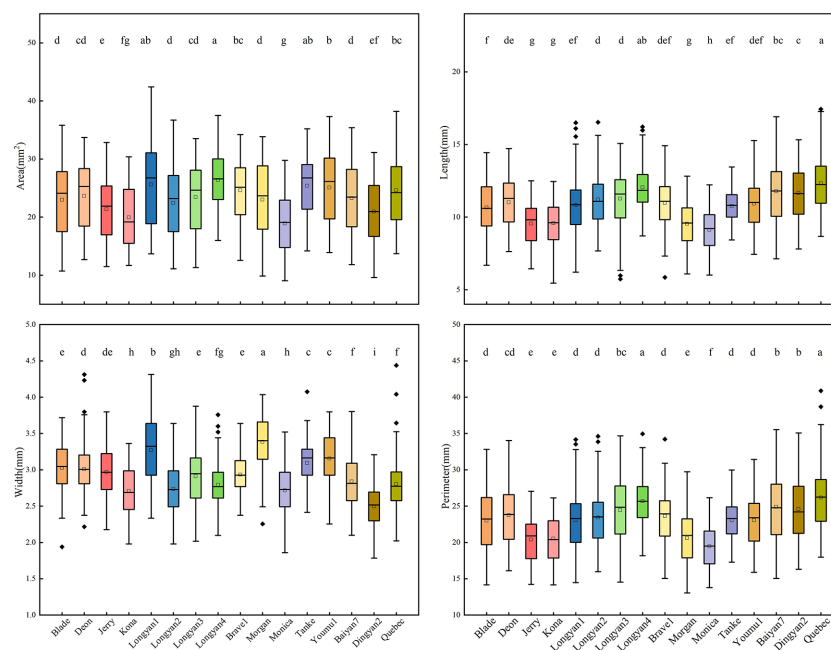
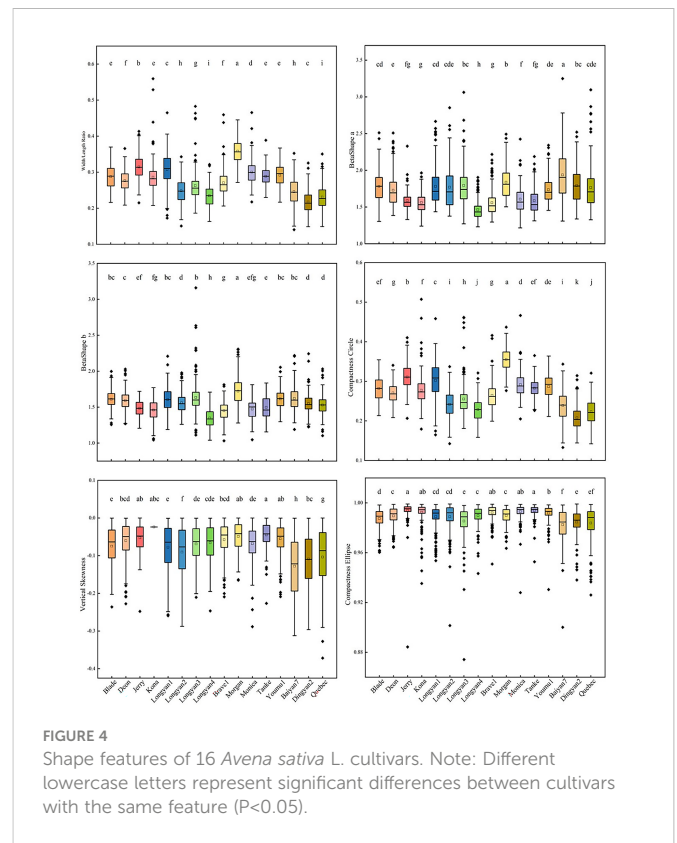
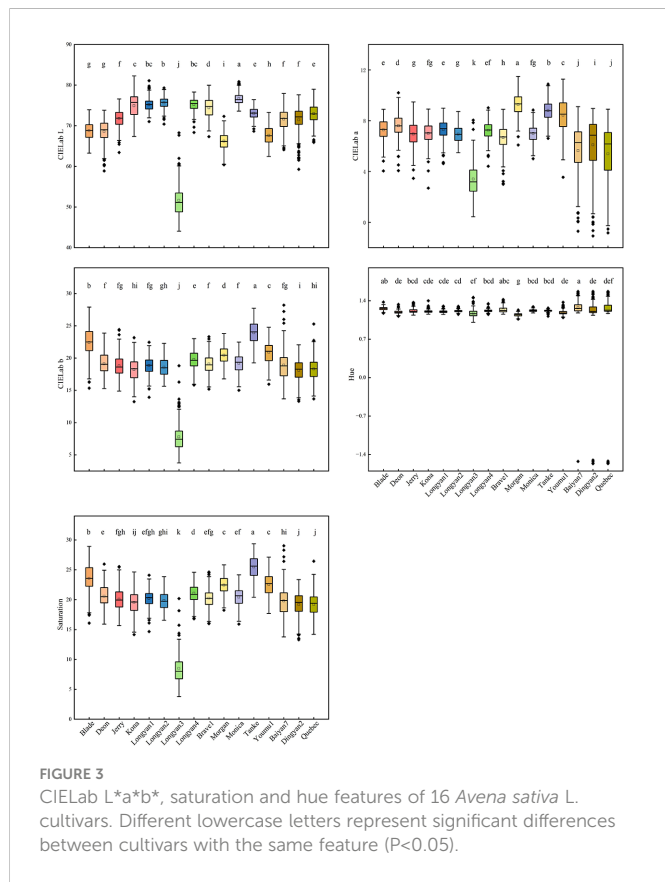


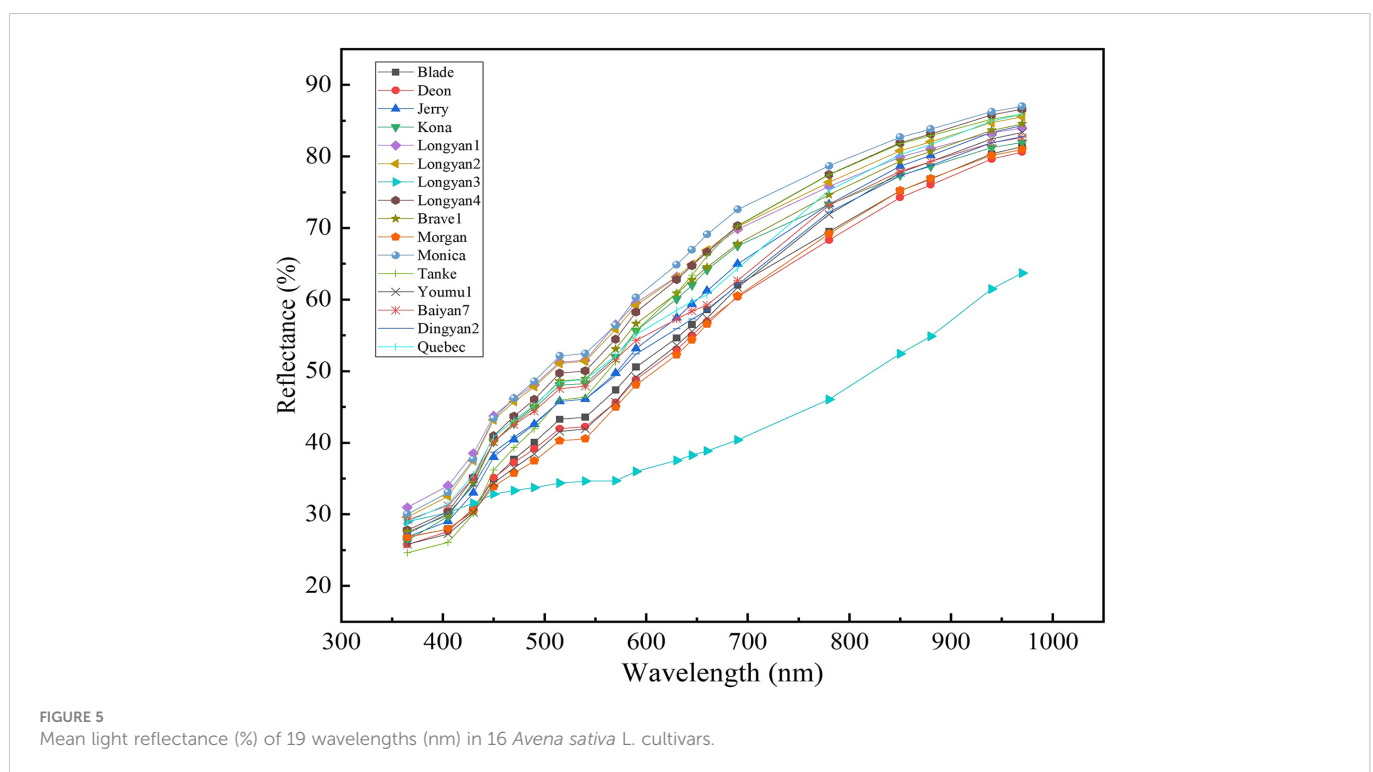
FIGURE 2  
Binary features of 16 *Avena sativa* L. cultivars. Different lowercase letters represent significant differences between cultivars with the same feature (P < 0.05).



### 3.2 Spectroscopic analysis

The spectral reflectance curves of oat seeds of different cultivars were smooth and had the similar trend across the whole wavelength

region, the longer the wavelength, the higher the average reflectance (Figure 5). The reflectivity of different cultivars differed significantly at the same wavelength (Supplementary Table 1). For example, the average reflectivity of Monica was significantly higher than that of the other cultivars in the range of 515-970nm; while that of Longyan3 is





significantly lower in the spectral range from 450 to 970 nm than the other cultivars.

### 3.3 PCA analysis

For morphological traits, the first three principal components could explain 61.12% of the variation among cultivars with 24.95, 21.93 and 14.24%, respectively (Figure 6A). For spectral features, the explained variance rates for the first three principal components were 80.11, 16.32 and 2.49%, respectively (Figure 6B). In addition, PCA results based on morphological characteristics and seed spectra showed that the first three principal components could explain 72.44% of the original variance among seeds with 48.08, 14.93 and 9.43% for PC1, PC2 and PC3, respectively (Figure 6C). However, either the score plot of PCA based on morphological, spectral data or the combination of the two failed to distinguish seeds of different cultivars into 16 groups.

### 3.4 Identification of cultivars by LDA model

The LDA model based on morphological features data had an average classification accuracy of 61.16% and 60.63% for training and independent testing datasets, respectively (Supplementary Table 2; Figure 7A). In contrast, the LDA model based on spectral data had a great improvement in the accuracy of cultivars discrimination, the overall discrimination accuracy was as high as 83.88% and 80.92% for the training and testing datasets, respectively (Supplementary Table 3; Figure 7B). The LDA model combining morphological features and spectral data had the highest discrimination accuracy, with the value of 89.69% for testing datasets (Table 1; Figure 7C).

However, the discriminant accuracy of different cultivars varied greatly. The discrimination accuracy of Longyan3 in the testing datasets reached 100%, while that of Brave1 was low as to 31.67%, when the model based on the morphology features data (Supplementary Table 2). Similarly, the discriminant accuracy of Longyan3 and Morgan was 100%, while the discrimination accuracy in both Kona and Quebec was less than 80% (Table 1).

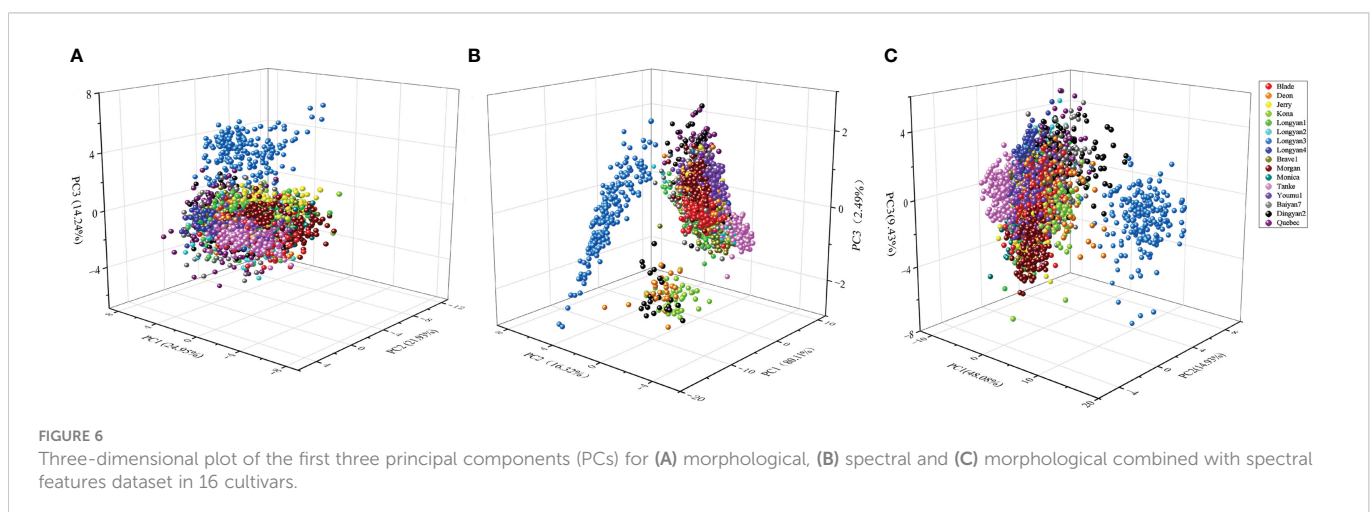
## 3.5 Identification of cultivars by SVM model

The classification accuracy of the training and the testing datasets of SVM model based on the morphological feature data of the 16 cultivars was 66.83% and 66.25%, respectively (Supplementary Table 4). When the discrimination is based on the spectral data, the accuracy reached 88.79% and 88.65%, respectively (Supplementary Table 5). The overall discrimination accuracy of training and testing set based on the combination of morphological and spectral data was 95.49% and 92.71%, respectively (Table 2).

Likewise, SVM model based on different types of data has different classification accuracy. For the model with morphology data, the accuracy of Longyan3 was 100%, while that of Brave1 was only 30% (Supplementary Table 4). When based on the spectral traits, there were 13 cultivars with more than 80% accuracy, while that of Jerry, Brave1 and Quebec was 76.67%, 78.33% and 66.67%, respectively (Supplementary Table 5). Using the combination of morphology features and spectral data, the accuracy of all cultivars reached more than 80%, the classification accuracy of Longyan3 and Tanke were 100%. Either using LDA or SVM model, Dingyan2 and Quebec have cross-error classification (Table 2).

## 4 Discussion

Although seeds of different oat cultivars differ a lot in morphological and spectral traits, any single seed trait failed to distinguish the seeds of different cultivars well due to a large variation within cultivars. PCA is an effective tool to classify oat cultivars into several groups by integrating several seed traits, while it failed to distinguish similar one, especially when trait variation is mainly from within cultivars rather than between cultivars. In contrast, the supervised learning algorithms, both LDA and SVM, using morphological, spectral features or their combination data all have a high accuracy in cultivar classification, in particularly, the overall accuracy of SVM model based on the combined data is higher than 90%.



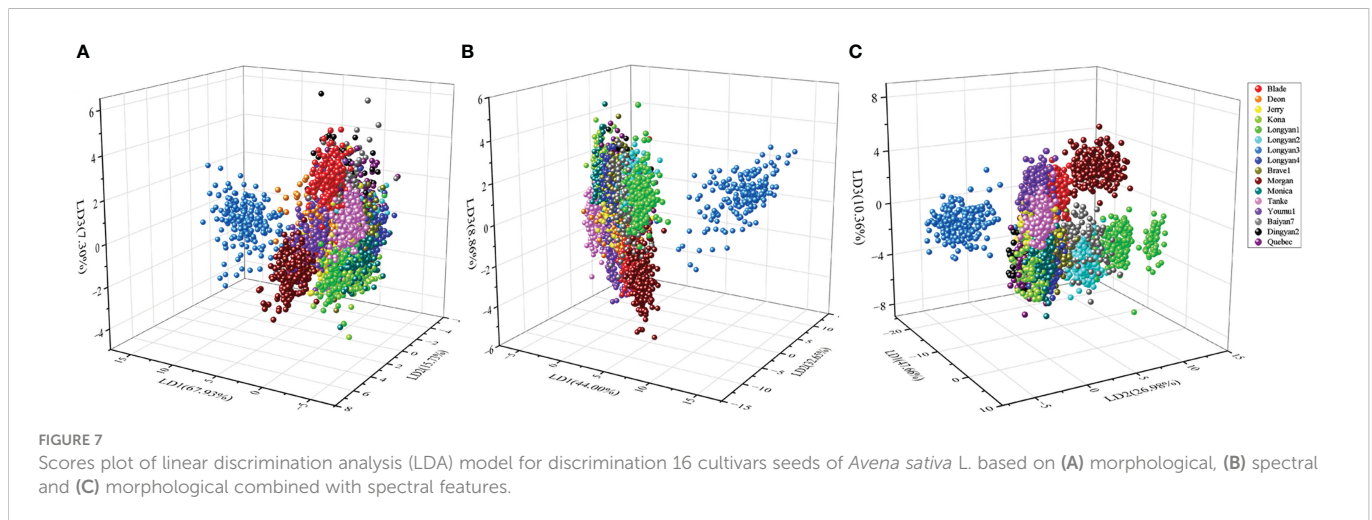


FIGURE 7

Scores plot of linear discrimination analysis (LDA) model for discrimination 16 cultivars seeds of *Avena sativa* L. based on (A) morphological, (B) spectral and (C) morphological combined with spectral features.

## 4.1 Seed features variation

Whilst the morphological features clearly showed significant differences in binary, color and shape in terms of mean value among 16 cultivars, a large variation generally occurs within each cultivar, and thus neither of these traits are cultivar-specific, except for the dark seed coat color of Longyan3 which shows a distinct difference from the other 15 cultivars. This means cultivar classification based on single morphological traits is generally hard since variation within cultivars overrides the difference among cultivars. The result is similar with the study on seed identification of alfalfa cultivars (Yang et al., 2020). Moreover, as we expected, for physical eyes, morphological features similarity between cultivars was too high to be identified by any single feature.

Similar to morphological traits, the average spectral reflectance of different cultivars is different, which may relate to different physical and chemical components of seeds (Liu et al., 2016b). For example, the difference in visible light region may be concerned with the color of the seeds, and the difference in near-infrared region (NIR) may be related to variation in seed chemical composition among cultivars. Therefore, seed spectral trait provides abundant information for seed physical and chemical attributes, thus offers the potential in cultivar classification.

## 4.2 Unsupervised classification analysis

PCA, as an unsupervised exploratory pattern recognition technique, is commonly used to get an overview of the systematic variation in data, exploring the possibility of grouping the seeds (ElMasry et al., 2019; Yang et al., 2020). Previous studies have shown that conventional soybean, glyphosate-resistant soybean seeds and hybrid progeny were significantly separated from each other based on the distribution scatter plot of three-dimensional principal components (Liu et al., 2016a). In contrast, although PCA could distinguish the cultivars into several categories in our study, most of the cultivars cannot be successfully identified (Figure 6). It indicates that traits of the cultivars are differentiation, which were better than the use of a single trait to distinguish, and demonstrate that PCA has the potential for preliminary classification of multiple cultivars. However, the method failed to separate similar cultivars. A possible reason is that PCA method aims to maximize the variance of

cultivars rather than to maximize the ability to distinguish oat cultivars. Under the circumstances, if the mean of the between-group variables is very similar and the variance is large, the total variance will be mainly composed of within-group variance, not those between groups. Thus, PCA would not detect inter-group differences. In reality, both morphological and spectral features have very close mean values, and there is a large overlap and cross distribution among the seeds of oat cultivars. Besides, some important information was lost in the process of PCA dimensionality reduction, which is also one of the reasons why seeds failed to be discriminated well.

## 4.3 Supervised classification analysis

Supervised methods aim to minimize the distance within class and to maximize the distance between groups, thus models can show good discriminability among groups. As supervised discriminant models, LDA and SVM performed well in the process of identifying oat cultivars, especially when based on the combination of morphological and spectral features data of seeds for discrimination, the average accuracy of LDA and SVM models for oat cultivars identification was 89.69% and 92.71%, respectively. Yang et al. used LDA and SVM models to classify 12 alfalfa cultivars, and the classification accuracy of the testing set was 91.53% and 93.47%, respectively (Yang et al., 2020). However, it is also worth noting that there is a relatively low discrimination accuracy due to cross-error classification for some cultivars, such as Quebec and Dingyan2. In this case, a part of seeds of Quebec are misclassified as Dingyan2, and vice versa. The possible reason is that seed characteristics of different oat cultivars are affected by environmental factors as well as genetic factors, and the stability of the traits of different oat cultivars is diverse. For example, the protein content of wheat grain varied with the production region (Zhou et al., 2021). Oat seeds from the same cultivated area have similar growth environment and seed processing, which results in similar seed morphology and chemical composition. For instance, both Quebec and Dingyan2 seeds in our study are produced in Dingxi, which may lead to a similar maternal environmental effect for seed traits.

Although our study provided a rapid way to screen seeds of different oat cultivars using multispectral imaging technique and multivariate analysis, some limitations for present study is worthy noticed for its further application. Firstly, compared to the large number of oat cultivars

TABLE 1 Discrimination performance based on LDA with morphology and spectral features of 16 *Avena sativa* L. cultivars.

	Predict	Actual																Total (%)
		Blade	Deon	Jerry	Kona	Longyan1	Longyan2	Longyan3	Longyan4	Brave1	Morgan	Monica	Tanke	Youmu1	Baiyan7	Dingyan2	Quebec	
Testing (60)	Blade	53	0	2	0	0	0	0	0	0	0	0	0	0	2	0	0	
	Deon	4	57	3	0	0	0	0	0	2	0	0	0	0	0	1	3	
	Jerry	0	3	48	0	0	0	0	0	0	0	0	0	2	0	0	2	
	Kona	0	0	1	46	0	0	0	2	0	0	2	0	0	0	1	0	
	Longyan1	0	0	0	0	58	1	0	0	0	0	0	0	0	0	0	0	
	Longyan2	2	0	0	0	2	55	0	0	0	0	0	0	0	2	0	0	
	Longyan3	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	
	Longyan4	0	0	0	1	0	0	0	49	4	0	0	1	0	0	1	1	
	Brave1	1	0	1	0	0	2	0	6	51	0	0	0	0	0	1	4	
	Morgan	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	
	Monica	0	0	4	13	0	0	0	0	3	0	58	0	0	0	0	0	
	Tanke	0	0	0	0	0	0	0	0	0	0	0	59	0	0	0	0	
	Youmu1	0	0	1	0	0	0	0	0	0	0	0	0	56	0	0	0	
	Baiyan7	0	0	0	0	0	2	0	0	0	0	0	0	0	56	0	0	
	Dingyan2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	49	4	
	Quebec	0	0	0	0	0	0	0	2	0	0	0	0	2	0	7	46	
	<b>Accuracy (%)</b>	88.33	95.00	80.00	76.67	96.67	91.67	100.00	81.67	85.00	100.00	96.67	98.33	93.33	93.33	81.67	76.66	89.69



TABLE 2 Discrimination performance based on SVM with morphology and spectral features of 16 *Avena sativa* L. cultivars.

	Predict	Actual																Total (%)
		Blade	Deon	Jerry	Kona	Longyan1	Longyan2	Longyan3	Longyan4	Brave1	Morgan	Monica	Tanke	Youmu1	Baiyan7	Dingyan2	Quebec	
Testing (60)	Blade	56	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	
	Deon	4	52	1	0	1	0	0	0	1	1	0	0	0	2	1	2	
	Jerry	0	3	53	0	0	0	0	0	1	0	0	0	1	0	0	0	
	Kona	0	0	0	54	0	0	0	0	1	0	3	0	0	0	0	0	
	Longyan1	0	0	0	0	55	0	0	0	0	0	0	0	0	0	0	0	
	Longyan2	0	0	0	0	4	58	0	0	0	0	0	0	0	1	0	0	
	Longyan3	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	
	Longyan4	0	0	0	0	0	0	0	57	2	0	0	0	0	1	2	2	
	Brave1	0	1	2	0	0	2	0	2	53	0	0	0	0	0	0	1	
	Morgan	0	0	0	0	0	0	0	0	0	59	0	0	0	0	0	0	
	Monica	0	0	0	6	0	0	0	0	1	0	57	0	0	0	0	0	
	Tanke	0	0	1	0	0	0	0	0	0	0	0	60	0	0	0	0	
	Youmu1	0	1	1	0	0	0	0	0	0	0	0	0	59	0	0	0	
	Baiyan7	0	0	1	0	0	0	0	0	0	0	0	0	0	53	0	0	
	Dingyan2	0	2	1	0	0	0	0	0	0	0	0	0	0	1	54	5	
	Quebec	0	0	0	0	0	0	0	1	1	0	0	0	0	0	3	50	
	Accuracy (%)	93.33	86.67	88.33	90.00	91.67	96.67	100.00	95.00	88.33	98.33	95.00	100.00	98.33	88.33	90.00	83.33	92.71

across the world, 16 is a quite small number and thus limits the discrimination model application in a broader range. Thus, incorporating a huge sample size in future study will help to produce a universal, sharp and steady oat cultivar identification model *via* multispectral imaging techniques. Secondly, as we discussed above, seed traits could be affected by growth conditions as well as harvest and processing methods even for the same cultivar, thus, including various seed lots of each cultivar is necessary to account for genetic effect in the discrimination model. Thirdly, the VideometerLab4 multispectral imaging system used in our study including 19 specific wavelengths ranged from 365 to 970 nm only, the limited band range will inevitably lead to the loss of some information, thus limiting its application. In particular for seed chemical composition, the spectral information is generally correlated with the reflectance at longer wavelengths. For example, the peak reflectance at 1,100nm may relate to the second overtone of C-H stretching, and so is the valley near 1,200nm (Zhang et al., 2016), the peak at 1,304.6 nm may be assigned to combination between the first with N-H in-plane bending vibrations (Zhang et al., 2017a), the valley near 1,450 nm is related to the first overtone of the tension and vibration inherent in O-H and N-H (Zhang et al., 2017b). These are related to the composition and content of chemical substances, such as fatty, protein and starch, resulting in different reflectivity. Thus, a wide range wavelength such as near infrared spectroscopy may help to improve the data quality and favor discrimination model building.

## 5 Conclusion

In brief, our study clearly shows that differences in seed features among oat cultivars did exist and could be visualized *via* the multispectral imaging technique. Both LDA and SVM, as supervised classification models, have high accuracy in classifying oat cultivars, especially when based on the combination of seed morphological and spectral features. It suggests that multispectral imaging together with multivariate analysis could be an efficient and non-destructive way for oat cultivar identification.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

## References

- Barboza da Silva, C., Bianchini, V., de, J. M., Medeiros, A. D., de, Moraes, M. H. D., de, Marassi, A. G., et al. (2021). A novel approach for jatropha curcas seed health analysis based on multispectral and resonance imaging techniques. *Ind. Crop Prod.* 161, 113186. doi: 10.1016/j.indcrop.2020.113186
- Benlioglu, B., and Ozkan, U. (2021). The influence of salinity and drought stress on some oat cultivars (*Avena sativa* L.) by determining some stress indexes and growth performances at the germination stage. *Fresen Environ. Bull.* 30, 771–778.
- Boelt, B., Shrestha, S., Salimi, Z., Jørgensen, J., Nicolaisen, M., and Carstensen, J. (2018). Multispectral imaging – a new tool in seed quality assessment? *Seed Sci. Res.* 28 (3), 222–228. doi: 10.1017/S0960258518000235
- Carr, P. M., Horsley, R. D., and Poland, W. W. (2004). Barley, oat, and cereal-pea mixtures as dryland forages in the northern great plains. *Agron. J.* 96, 677–684. doi: 10.2134/agronj2004.0677
- Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., and Huvenne, J.-P. (2009). Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemometr Intell. Lab. 96* (1), 27–33. doi: 10.1016/j.chemolab.2008.11.005
- Dumlupinar, Z., Maral, H., Kara, R., Dokuyucu, T., and Akkaya, A. (2011). Evaluation of Turkish oat landraces based on grain yield, yield components and some quality traits. *Turkish J. Field Crops.* 16, 190196.

## Author contributions

XH conceived and designed the experiment. XF performed the experiments and analyzed the data. MB, YX, TW and ZH contributed in the experiment. XF wrote and XH revised the paper. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Fundamental Research Funds for the Central Universities under the Grant (lzujbky-2021-ct09).

## Acknowledgments

We thank to Zuxin Zhang for providing the method of data processing, and Riyao Liu for language polish.

## Conflict of interest

The 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1113535/full#supplementary-material>

- Dvořáček, V., Čurn, V., and Moudrý, J. (2003). Suitability of oat-seed storage-protein markers for identification of cultivars in grain and mixed flour samples. *Plant Soil Environ.* 49 (11), 486–491. doi: 10.17221/4182-psc
- ElMasry, G., Mandour, N., Wagner, M.-H., Demilly, D., Verdier, J., Belin, E., et al. (2019). Utilization of computer vision and multispectral imaging techniques for classification of cowpea (*Vigna unguiculata*) seeds. *Plant Methods* 15 (1), 24. doi: 10.1186/s13007-019-0411-2
- Geetha, V., Balamuruga, P., and Bhaskaran, M. (2011). Characterization of mustard genotypes through image analysis. *Res. J. Seed Science.* 4, 192–198. doi: 10.3923/rjss.2011.192.198
- Grillo, O., Miceli, C., and Venora, G. (2011). Computerised image analysis applied to inspection of vetch seeds for varietal identification. *Seed Sci. Technol.* 39 (2), 490–500. doi: 10.15258/sst.2011.39.2.20
- Hu, X., Yang, L., and Zhang, Z. (2020). Non-destructive identification of single hard seed via multispectral imaging analysis in six legume species. *Plant Methods* 16, 116. doi: 10.1186/s13007-020-00659-5
- Jia, Z., Sun, M., Ou, C., Sun, S., Mao, C., Hong, L., et al. (2022). ). single seed identification in three *Medicago* species via multispectral imaging combined with stacking ensemble learning. *Sensors (Basel)* 22 (19), 7521. doi: 10.3390/s22197521
- Kim, K., Webb, S., Newell, M., Anderson, J., and Butler, T. (2017). Variation of winter forage production in four small grain species: oat, rye, triticale and wheat. *Pak J. Bot.* 49 (2), 553–559.
- Liu, C., Liu, W., Lu, X., Chen, W., Chen, F., Yang, J., et al. (2016a). Non-destructive discrimination of conventional and glyphosate-resistant soybean seeds and their hybrid descendants using multispectral imaging and chemometric methods. *J. Agr. Sci.* 154 (1), 1–12. doi: 10.1017/S0021859614001142
- Liu, W., Liu, Ch., Ma, F., Lu, X., Yang, J., and Zheng, L. (2016b). Online variety discrimination of rice seeds using multispectral imaging and chemometric methods. *J. Appl. Spectrosc.* 82 (6), 993–999. doi: 10.1007/s10812-016-0217-1
- Maloney, T. S., Oplinger, E. S., and Albrecht, K. A. (1999). Small grains for fall and spring forage. *J. Production Agriculture.* 12 (3), 488–494. doi: 10.2134/jpa1999.0488
- Marshall, A., Cowan, S., Edwards, S., Griffiths, I., Howarth, C., Langdon, T., et al. (2013). Crops that feed the world 9. oats- a cereal crop for human and livestock feed with industrial applications. *Food Secur.* 5 (1), 13–33. doi: 10.1007/s12571-012-0232-x
- Menon, R., Gonzalez, T., Ferruzzi, M., Jackson, E., Winderl, D., and Watson, J. (2016). Oats—from farm to fork. *Adv. Food Nutr. Res.* 77, 1–55. doi: 10.1016/bs.afnr.2015.12.001
- O'Donoghue, L. S., Souza, E., Tanksley, S. D., and Sorrells, M. E. (1994). Relationships among north American oat cultivars based on restriction fragment length polymorphisms. *Crop Sci.* 34 (5), 1251–1258. doi: 10.2135/cropsci1994.0011183X003400050021x
- Olesen, M., Nikneshan, P., Shrestha, S., Tadayyon, A., Deleuran, L., Boelt, B., et al. (2015). Viability prediction of *Ricinus communis* L. seeds using multispectral imaging. *Sensors.* 15 (2), 4592–4604. doi: 10.3390/s150204592
- Paczos-Grzeda, E. (2004). Pedigree, RAPD and simplified AFLP-based assessment of genetic relationships among *Avena sativa* L. cultivars. *Euphytica.* 138 (1), 13–22. doi: 10.1023/B:EUPH.0000047055.99322.7a
- Papadakis, S. E., Abdul-Malek, S., Kamdem, R. E., and Yam, K. L. (2000). A versatile and inexpensive technique for measuring color of foods. *Food Technol.* 54 (12), 48–51. doi: 10.1177/108201320000600608
- Pipan, B., Sinkovič, L., Rutar, R., and Meglič, V. (2021). The use of DNA markers for genetic differentiation of common (*Avena sativa* L.) and naked oat (*Avena nuda* L.). *J. Cent. Eur. Agric.* 22, 329–340. doi: 10.5513/JCEA01/22.2.3114
- Rauf, S., Niazi, I., Anwar, S., and Shahzad, M. (2020). Oat (*Avena sativa* L.) advanced lines outperform existing cultivars for forage yield and its components under terminal heat stress. *Pak J. Agr. Sci.* 57 (2), 327–331. doi: 10.21162/PAKJAS/2020.8557
- Saccomanno, B., Wallace, M., O'Sullivan, D. M., and Cockram, J. (2020). Use of genetic markers for the detection of off-types for DUS phenotypic traits in the inbreeding crop, barley. *Mol. Breeding.* 40 (1), 13. doi: 10.1007/s11032-019-1088-y
- Sakiyama, N., Ramos, H. C., Caixeta, E., and Pereira, M. (2014). Plant breeding with marker-assisted selection in Brazil. *Crop Breed Appl. Biot.* 14, 54–60. doi: 10.1590/S1984-70332014000100009
- Shrestha, S., Deleuran, L., and Gislum, R. (2016). Classification of different tomato seed cultivars by multispectral visible-near infrared spectroscopy and chemometrics. *J. Spectral Imaging.* 5, 1–9. doi: 10.1255/jsi.2016.a1
- Silva, C., Menezes, L., Ziech, M., Kuss, F., Ronsani, R., Biesek, R., et al. (2012). Overseeded oat cultivars in grazing African star with different wastes managed forage. *Semin: Cienc. Agrar.* 33 (6), 2441–2450. doi: 10.5433/1679-0359.2012v33n6p2441
- Singh, R., De, S., and Belkheir, A. (2013). *Avena sativa* (Oat), a potential nutraceutical and therapeutic agent: an overview. *Crit. Rev. Food Sci.* 53 (2), 126–144. doi: 10.1080/10408398.2010.526725
- Sumathi, S., and Balamurugan, P. (2013). Characterization of oats (*Avena sativa* L.) cultivars using machine vision. *Pakistan J. Biol. Sci.* 16 (20), 1179–1183. doi: 10.3923/pjbs.2013.1179.1183
- Tobiasz-Salach, R., Kalaji, H. M., Mastalerzczuk, G., Baba, W., Bobrecka-Jameo, D., and Noras, K. (2019). Can photosynthetic performance of oat (*Avena sativa* L.) plants be used as bioindicator for their proper growth conditions? *Chiang Mai J. Sci.* 46 (5), 880–895.
- Tong, L., Liu, L., Zhong, K., Wang, Y., Guo, L., and Zhou, S. (2014). Effects of cultivar on phenolic content and antioxidant activity of naked oat in China. *J. Integr. Agr.* 13 (8), 1809–1816. doi: 10.1016/S2095-3119(13)60626-7
- Wang, H., and Hu, D. (2005). Comparison of SVM and LS-SVM for regression. *Int. Conf. Neural Networks Brain.* doi: 10.1109/CNNB.2005.1614615
- Wang, Q., Li, Z., Li, J., and Zhou, B. (2019). Evaluation of agronomic and forage quality traits of a range of oat cultivars. (In Chinese, with English abstract.). *Acta Prataculturae Sinica.* 28 (12), 149. doi: 10.11686/cyxb2019040
- Wight, C. P., Yan, W., Fetch, J. M., and Deyl, J. (2010). And tinker, n A set of new simple sequence repeat and avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena* spp.). *A. Crop Sci.* 50 (4), 1207–1218. doi: 10.2135/cropsci2009.09.0474
- Yang, L., Zhang, Z., and Hu, X. (2020). Cultivar discrimination of single alfalfa (*Medicago sativa* L.) seed via multispectral imaging combined with multivariate analysis. *Sensors* 20 (22), 6575. doi: 10.3390/s20226575
- Yu, Z., Fang, H., Zhangjin, Q., Mi, C., Feng, X., and He, Y. (2021) Hyperspectral imaging technology combined with deep learning for hybrid okra seed identification. *Biosyst. Eng.* 212, 46–61. doi: 10.1016/j.biosystemseng.2021.09.010
- Zhang, M., Bai, R., Nan, M., Ren, W., Wang, C., Shabala, S., et al. (2022). Evaluation of salt tolerance of oat cultivars and the mechanism of adaptation to salinity. *J. Plant Physio.* 273, 153708. doi: 10.1016/j.jplph.2022.153708
- Zhang, C., Guo, C., Liu, F., Kong, W., He, Y., and Lou, B. (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *J. Food Eng.* 179, 11–18. doi: 10.1016/j.jfoodeng.2016.01.002
- Zhang, C., Jiang, H., Liu, F., and He, Y. (2017a). Application of near-infrared hyperspectral imaging with variable selection methods to determine and visualize caffeine content of coffee beans. *Food Bioprocess Tech.* 10 (1), 213–221. doi: 10.1007/s11947-016-1809-8
- Zhang, C., Wang, Q., Liu, F., He, Y., and Xiao, Y. (2017b). Rapid and non-destructive measurement of spinach pigments content during storage using hyperspectral imaging with chemometrics. *Measurement.* 97, 149–155. doi: 10.1016/j.measurement.2016.10.058
- Zhou, W., Liu, Y., Ata-Ul-Karim, S. T., and Ge, Q. (2021). Spatial difference of climate change effects on wheat protein concentration in China. *Environ. Res. Lett.* 16 (12), 124011. doi: 10.1088/1748-9326/ac3401