Computers and Electronics in Agriculture 156 (2019) 585–605

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Review

Current and future applications of statistical machine learning algorithms for agricultural machine vision systems

Tanzeel U. Rehman[a,1], Md. Sultan Mahmud[b,1], Young K. Chang[b,*], Jian Jin[a], Jaemyung Shin[b]

[a] Department of Agricultural and Biological Engineering, Purdue University, West Lafayette 47907, USA
[b] Department of Engineering, Faculty of Agriculture, Dalhousie University, Truro, NS B2N5E3, Canada

ARTICLE INFO

Keywords:
Machine vision
Statistical machine learning
Naïve Bayes
Discriminant analysis
k-Nearest Neighbour
Support vector machines
K-means clustering
Fuzzy clustering
Gaussian mixture model

ABSTRACT

With being rapid increasing population in worldwide, the need for satisfactory level of crop production with decreased amount of agricultural lands. Machine vision would ensure the increase of crop production by using an automated, non-destructive and cost-effective technique. In last few years, remarkable results have been achieved in different sectors of agriculture. These achievements are integrated with machine learning techniques on machine vision approach that cope with colour, shape, texture and spectral analysis from the image of objects. Despite having many applications of different machine learning techniques, this review only described the statistical machine learning technologies with machine vision systems in agriculture due to broad area of machine learning applications. Two types of statistical machine learning techniques such as supervised and unsupervised learning have been utilized for agriculture. This paper comprehensively surveyed current application of statistical machine learning techniques in machine vision systems, analyses each technique potential for specific application and represents an overview of instructive examples in different agricultural areas. Suggestions of specific statistical machine learning technique for specific purpose and limitations of each technique are also given. Future trends of statistical machine learning technology applications are discussed.

1. Introduction

The world's population is assumed to be nearly 10 billion by 2050, boosting agricultural order-in a situation of humble financial development by somewhere in the range of 50% contrasted with 2013 (FAO, 2017). At present, about 37.7% of total land surface is used for crop production (World Bank, 2015). People would require raising overall food production by around 70 percent between 2005/07 and 2050 (FAO, 2009). However, farmers in worldwide will need to ensure the projected amount of crop production, either by increasing amount of agricultural land or by enhancing productivity on existing agricultural cultivable land by the adoption of precision or smart farming (Elferink and Schierhorn, 2016).

Present agriculture technologies are mostly diverted to machine learning (ML) algorithms because it has maximized crop yield with minimizing the input costs. ML algorithm enables the farmer to enhance the crop selection and crop yield prediction, crop diseases prediction, weather forecasting, minimum support price and smart irrigation system (Kaur, 2016). The ML techniques are not and will never be the solution to all the problems risen by agricultural cropping systems. However, these techniques provide a powerful set of tools that applied for different field applications (Nieuwenhuizen et al., 2007; Tagarakis et al., 2013), weed detection (Burks et al., 2000; Pereira et al., 2012; Mursalin and Mesbah-Ul-Awal, 2014; De Rainville et al., 2014), yield prediction (Dey et al., 2012; Liu et al., 2017), crop management (Gili et al., 2017), plant phenotyping (stress prediction) (Kirk et al., 2009; Rousseau et al., 2013; Park et al., 2017), disease detection in fruits (Dubey and Jalal, 2014), disease detection in plants (Sankaran and Ehsani, 2012; Bandi et al., 2013), fruit grading (Blasco et al., 2007; Gómez-Sanchis et al., 2008), soil analysis (Haghverdi et al., 2015), management zone clustering (Boydell and McBratney, 2002), irrigation/ ET/ water productivity (Casanova et al., 2014) in agriculture. Advantages of ML technology for agricultural crop are that it can be fairly accurate (Puerto et al., 2015), non-destructive (Munera et al., 2017) and yields consistent (Qureshi et al., 2017) results. Dingle Robertson and King (2011) used k-Nearest Neighbour (kNN) based ML algorithm for classifying broad agricultural land cover types with Landsat-5 TM imagery and found that the difference in overall accuracy between these classification approaches was not statistically significant. The naïve Bayes, k-mean clustering, support vector machines (SVMs) and kNN based ML algorithms were utilized for weed detection along with machine vision system and yielded as a successful solution having

* Corresponding author.
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.compag.2018.12.006
Received 23 March 2018; Received in revised form 2 October 2018; Accepted 2 December 2018
0168-1699/

higher success rate for all cases (Åstrand and Baerveldt, 2002; Nieuwenhuizen et al., 2007; Tellaeche et al., 2011; De Rainville et al., 2014). Dey et al. (2012) applied Gaussian Mixture Model (GMM) based ML algorithm to estimate yield of grapevine and observed accuracies of 98% prior to ripening and 96% during ripening. ML techniques also showed higher success rate above 95% in plant disease detection with different algorithms (Mishra et al., 2011; Bandi et al., 2013; Mondal and Kole, 2016). Gili et al. (2017) applied fuzzy k-means cluster-based ML technique in site specific crop management having coincidence in the classification of 98% of the sampled grid points. Yao et al. (2009) used SVMs based ML classifier to detect rice diseases with 97.2% of accuracy. Granitto et al. (2005) suggested the naïve Bayes based ML approach which would reduce the operational complexity and hardware cost of a commercial system.

Previous review of ML approaches exposed in agricultural data mining (Garner et al., 1995; Frank et al., 2009), plant stress phenotyping (Singh et al., 2016), food quality evaluation (Du and Sun, 2006), detection of biotic stress in crop (Behmann et al., 2015) etc. However, limited reviews have been conducted on the application of ML in agriculture machine vision system that cover the wide overview of cropping system to find out the effective ML algorithm for each criterion. This review intends to present a comprehensive view of ML applications in agricultural machine vision system. However, we restrict our study to statistical ML techniques as including other non-statistical (artificial intelligence/ deep learning) can diverge the domain of paper considerably. We have selected the statistical ML because of its historic and vast applications in the field of agriculture. This study suggested specific ML algorithm for specific application in agricultural machine vision use. The objectives of this review are (i) to provide an overview of work done in the field of agricultural machine vision system using statistical ML algorithms (ii) to highlight the limitations of different statistical ML algorithms for specific agricultural applications and (iii) to suggest the effective statistical ML algorithms in each specific area under agricultural domain.

2. Statistical machine learning techniques

2.1. Supervised machine learning algorithms

2.1.1. Naïve bayes

Naïve Bayes (NB) algorithm is a generative probabilistic model based on an assumption of conditional independence among the predictor variables/ features in a way that the presence of a particular feature in a class is not related with any other features (Frank et al., 2000). The naïve Bayes' assumption of conditional independence helps to calculate the class-conditional probabilities of the sample data that can be directly estimated from the training data rather than evaluating all the possibilities of a feature (Bishop, 2007). The posterior probability for every observation belonging to a class is calculated by using prior class probabilities and class conditional probabilities of features given the class level (Bishop, 2007) followed by assigning the observation to class having highest posterior probability. Additionally, Bayes classifier assumes that observations in every class are independent and identically distributed (drawn from a similar probability distribution) (Raschka, 2014). The probability distribution of each class can be considered as unimodal Gaussian for numerical data, while can be represented with multinomial or Bernoulli distribution for the categorical data (Raschka, 2014) (Table 1).

Fresh agricultural product (fruits, vegetables and grains) grading and sorting are among common applications of NB algorithm (Asadollahi et al., 2009; Siedliska et al., 2014; Kurtulmus et al., 2014; Ravikanth et al., 2015; Cen et al., 2016; Veernagouda Ganganagowder and Kamath, 2017). Asadollahi et al. (2009) compared the performance of NB with seven other classification techniques Multiple Layer Perception (MLP), Neural Binary Tree, Random Forest (RF), Random Tree (RT), Radial Basis Function Network, kNN and K-star) for tomato grading using colour, shape and textural features. The results indicated that MLP, Neural Binary tree, RF and RT performed better by achieving accuracies of 82.2%, 83.3%, 85.3% and 87.3%, respectively, compared to NB with an accuracy of 80.1%. Siedliska et al. (2014) used NB and mean spectral reflectance extracted from visible and near infrared (VINIR) imaging (400–1000 nm) and short wavelength infrared (SWIR) imaging (1000–2500 nm) for grading of the apples based on bruises and detection of five different cultivars. The raw spectral data was pre-processed with 2nd derivative function for avoiding the offset and linear 'tilt' present in data followed by feature selection (wavelength/ band selection) by using correlation-based feature selection. The results indicated accuracies of 83.3% and 86.7% for bruise and cultivar detection, respectively, on the test dataset. Infected cucumbers from chilling injury were detected for sorting, using mean spectral reflectance (500–675 nm) and transmittance (675–1000 nm) with NB (Cen et al., 2016). Optimal bands were selected by using three different feature selection (mutual information selection, max-relevance min-redolence and sequential forward selection) protocols followed by the extraction of textural features from selected bands. The results reported that majority of selected bands from all three methods were from transmittance in short near-infrared region. The overall accuracies of reflectance features were in a range of 91.6% to 97.6% and 81% to 85.7% for two class and three class problems, respectively. Ravikanth et al. (2015) used a similar approach for extracting the mean spectral reflectance to differentiate the contaminants in wheat grains using near-infrared (NIR) imaging (1000–1600 nm). The raw spectral response was pre-processed with 1st derivative, 2nd derivative, Savitzky-Golay (SG) smoothing and differentiation, multiplicative scatter correction and standard normal variate for reducing the signal noise. The accuracy of the developed techniques was in a range of 77.9% to 100% depending upon the types of contaminants and pre-processing approach (Table 1).

Pereira et al. (2012) classified three different species of the aquatic weeds using NB for aquatic weed control. The shape analysis of each type was performed by extracting the 180, 300, 126, 14, 100 and 180 features from beam angle statistics (BAS)-60, BAS-100, Fourier descriptors, 14 Moment Invariants descriptor (Hu, 1962), multiscale fractal dimensions and tensor scale descriptors, respectively, for all collected images. The results indicated that algorithm achieved a mean classification accuracy of 69.43%, 72.01%, 93.27%, 80.01%, 85.54% and 77.95% for beam angle, Fourier, Moment Invariants moments, multiscale fractal dimension, and tensor scale descriptors, respectively. Mursalin and Mesbah-Ul-Awal (2014) identified four different weeds in capsicum fields using nine different shape features in conjunction with NB. The average accuracy rate was found to be 98.9%. Laursen et al. (2014) used different colour indices (R, G, B, NIR, normalized r, g and b, ExG, NDVI, ExR and ExGN) to segment the green plant pixels from the soil background. A NB classifier was trained by extracting the colour indices from manually annotating the plant and soil pixels, thus the resulting in maximum classification accuracies of 84.71%, 99.61% and 1.26% for vegetation, soil and missed pixels, respectively. The fusion of shape and colour features were used for differentiating the 32 different types of leaves for plant recognition (Caglayan et al., 2013). The shape, colour features and their combination were able to correctly classify 79.92%, 88.77% and 89.25% of the leaves, respectively (Table 1).

Sankaran and Ehsani (2012) detected the citrus leaves infected by Huanglongbing disease by using the spectral features extracted from the fluorescence imaging under both laboratory and field conditions. The results showed that fluorescence spectral features with NB have an ability to correctly classify the 90.1% of observations (infected and healthy both) in laboratory dataset, while the accuracy was dropped to 68.3% for field dataset. Bandi et al. (2013) used textural features (Haralick et al., 1973) extracted from HSI colour co-occurrence matrices (CCMs) with NB to identify three common diseases on citrus leaves with an overall accuracy of 95%. Stegmayer et al. (2013) combined the colour, shape and textural features to train a NB for

Table 1
Overview of example applications of naïve Bayes algorithm in agriculture.

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
57 different seeds of weed species	Mean, variance and skewness of red, green and intensity, Short run emphasis etc. from grey level run length matrix (GLRM), 17 textural features from grey level co-occurrence matrix (GLCM) (Conners et al., 1984), moments of planar mass distribution, ratio of seed and enclosing box areas, ratio of semi-axis lengths of the main principal axis, square root of seed area.	Colour, morphology and texture features	Weed seed classification	99.2% for test data set	Granitto et al. (2002)
A dataset of 32 different leave types (Wu et al., 2007)	Smoothness, aspect ratio, form factor, rectangularity, narrow factor, ratio of perimeter to longest distance, ratio of perimeter to sum of main vein length and widest distance, five structural features, mean and standard deviation of red, green, blue channel, and average of three channels, and colour histograms. SVMs, kNN, and Random Forest were also used.	Shape and colour features	Identification of different leaf types	79.92% for shape, 88.77% for shape and 1st set of colour features and 89.25% by combining all the features.	Caglayan et al. (2013)
A FoodCast Research Image Database	Mean, standard deviation, range, luminance, chrominance, colour distance metric, major and minor axis length, area, perimeter, equivalent diameter, convex area, solidity, extent, eccentricity, roundness, compactness, contrast, correlation, angular second moment, energy, homogeneity, dissimilarity, entropy, cluster shade, cluster performance, smoothness, third moment and maximum probability.	Colour, shape and textural features	Fruit, vegetables, leafy vegetables, biscuits grains and nuts grading	72%, 73%, 73%, 70%, 70% and 70% for fruits, grains, nuts, biscuits, vegetables, and leafy vegetables, respectively.	Veernagouda Ganganagowder and Kamath (2017)
Apples	Mean reflectance spectra extracted using VINIR (400–1000 nm) and SWIR (1000–2500 nm) hyperspectral camera.	Spectral features	Bruise and cultivar detection	83.3% for bruise and 86.7% for cultivar detection, respectively.	Siedliska et al. (2014)
Capsicum	Area, perimeter, thickness, convex area, convex perimeter, form factor, elongation, convexity, solidity, mean and standard deviation of R, G, B channels, four Hu moments.	Shape features	Weeds identification	Average 98.9% accuracy with 10-fold-cross validation.	Mursalin and Mesbah-Ul-Awal (2014)
Cereal	Raw R, G, B, NIR, normalized r, g, b, NIR, ExG, NDVI, ExR and ExGNIR.	Colour features	Vegetation Segmentation	84.71%, 99.61%, 1.26% for vegetation, soil and missed pixels with different prior probability.	Laursen et al. (2014)
Citrus	Yellow, red and far-red fluorescence from R, G, B and UV wavelengths, simple fluorescence ratio with green and red, fluorescence excitation ratio with red-UV and red-green excitation, flavonols, anthocyanins, nitrogen balance index with UV-green and UV-red excitation, fluorescence excitation ratio and anthocyanin relative index.	Spectral features from fluorescence imagery	Disease detection in leaves	Over all classification accuracy of 90.1% for laboratory and 68.3% for field dataset.	Sankaran and Ehsani (2012)
Citrus	Uniformity, contrast, mean intensity, variance correlation, product moment, inverse difference moment, entropy, sum entropy, difference entropy, information measures of correlation I and II and modus.	Textural features	Disease detection on leaves	95%	Bandi et al. (2013)
Citrus (Mandarins)	Shape, topography, deepness, transition zone, colour of transition zone, central colour, ruggedness of central surface, pattern of central zone, central texture and presence of fruiting bodies from infected area.	Colour, shape and textural features	Disease classification on fruits	Overall accuracy of 78.72%	Stegmayer et al. (2013)
Corn	Volumetric fractal dimension, Gabor wavelet and volumetric fractal dimension with canonical analysis	Textural features	Nitrogen nutrition status	82.5% for nitrogen deficiency at V4 and 87.5% at V7 stage.	Romualdo et al. (2014)
Cucumber	Mean reflectance and transmittance spectra extracted using VINIR hyperspectral camera (400–1000 nm), 1st order statistics (Energy, entropy, standard deviation, skewness, 2nd moment) and energy, contrast, homogeneity and correlation from GLCM.	Spectral and textural features	Vegetable grading on the basis chilling injury detection	Over all accuracy of 91.6–97.6% for two-class, 81–85.7% for three-class problem with spectral features.	Cen et al. (2016)
E. crassipes, *P. stratiotes*, and *S. auriculata*	Beam angle statistics, Fourier descriptors, Hu moments, multiscale fractal dimensions and tensor scale descriptor.	Shape features	Aquatic weed classification	Mean accuracy of 69.43–93.27%	Pereira et al. (2012)
Okra and bitter gourd	Mean, standard deviation, entropy, number of histogram peaks, inertia, homogeneity, correlation, energy from HSI and RGB colour planes.	Colour and textural features	Disease detection in plant leaves caused by virus	95% for okra and 82.67% for bitter gourd	Mondal et al. (2017)
Peaches	Eigenfruit (Kurtulmus et al., 2011) features from hue and saturation along with the circular Gabor texture.	Colour and textural features	Immature fruit detection in peach orchards	Maximum accuracies of 69% and 56% for training and validation datasets	Kurtulmus et al. (2014)

(*continued on next page*)

Table 1 (*continued*)

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
Rice	Mean and standard deviation of infected spot, background and colour change of infected region with respect to background from R, G and B colour planes, radius, major axis length, minor axis length, length, width and orientation of region.	Colour and shape features	Leaf disease classification	Maximum accuracy of 91.39%	Phadikar et al. (2013)
Tomatoes	Greenness, redness, yellowness grade, average of greenness, redness and yellowness grade, entropy, energy, contrast, sympathetic, circularity and area.	Colour, shape, and textural features	Fruit grading	Maximum classification accuracy of 80.1%	Asadollahi et al. (2009)
Wheat	Mean reflectance spectra extracted using NIR hyperspectral camera (950–1700 nm), 1stderivative, 2nd derivative, SG smoothing and derivative, MSC and SNV.	Spectral features	Grain cleaning on the basis contamination removal	77.9–100%	Ravikanth et al. (2015)

classifying the citrus fruits infected by three different type of diseases with an overall accuracy of 78.72%. Phadikar et al. (2013) identified different types of diseases on rice leaves using Fermi energy (Gao and Sammes, 1999) based segmentation method followed by the extraction of colour and shape features. The shape of infected region was used to determine the shape of spot based on two steps genetic algorithm consisting of locating the centre of spot followed by positioning the set of primitive shapes at the centre of spot. The NB classifier trained on the extracted colour and shape features indicated a maximum classification accuracy of 91.39%. Mondal et al. (2017) identified the okra and bitter gourd leaves infected by yellow vein mosaic virus using a combination of colour and textural features along with NB classifier. The success rates for the correct identification of diseases were found to be 95% and 82.67% for okra and bitter gourd, respectively. The major drawback of NB is that it cannot learn the interaction between two predictor variables/features because of its conditional independence assumption (Bishop, 2007). Therefore, adding more predictor variable might not increase but decrease the overall classification accuracy because of their correlation (Rennie et al., 2003) (Table 1).

2.1.2. Discriminant analysis

Discriminant analysis (DA) is a generative model that classifies an observation by estimating the posterior probability of which observation belonging to every class (Lachenbruch, 1975). The posterior probabilities are calculated from the prior probabilities of class membership and estimating the multivariate normal distribution (conditional) for every class using squared Mahalanobis distance (Klecka, 1980). The observation is assigned to the class that has the highest posterior probability or has the smallest square Mahalanobis distance (Lachenbruch, 1975). During the training phase, the parameters of conditional distribution functions, i.e., population mean vectors and covariance matrices (from sample statistics) were estimated (Bishop, 2007). The classification criterion in DA is based on assumption that classes have unimodal Gaussian conditional densities (Klecka, 1980; Bishop, 2007). In addition to the Gaussian density assumption, the linear discriminant analysis (LDA) also assume equal covariance thus resulting to the use of pooled covariance matrix. While, the quadratic discriminant analysis (QDA) does not require assumption of equal variance resulting in use of within group covariance matrices (Morrison, 1976) (Table 2).

A large number of DA applications in agriculture were focused on weed detection for different cropping system (Shearer and Holmes, 1990; Zhang and Chaisattapagon, 1995; Meyer et al., 1998; Lee et al., 1999; Perez et al., 2000; Burks et al., 2000; Gebhardt et al., 2006; Okamoto et al., 2007; Piron et al., 2008 and 2009; Chang et al., 2012; Kazmi et al., 2015; Rehman, 2017; Rehman et al., 2018). Shearer and Holmes (1990) used textural features extracted from HSI based CCMs followed by QDA to differentiate between seven different cultivars of nursery stock with an overall classification accuracy of 91%. Meyer et al. (1998) extracted the similar features from grey scale images of plant and soil followed by a feature reduction procedure. They concluded that canonical discriminant analysis with four features was able to achieve a maximum classification accuracy of 96.7% for soil class. Burks et al. (2000) trained a classification criterion based on LDA and textural features extracted from CCMs to achieve an overall accuracy of 93% (using hue and saturation statistics) for differentiating the five-weed species. In addition to the HSI colour space, the concept of CCM was also implemented on the luminance colour space for the identification of different weeds in wild blueberry cropping system (Chang et al. 2012). The highest classification accuracy of reduced feature set (94.9%) was achieved by HSI colour space. The authors concluded that addition of the luminance did not show any significant improvement in results. Zhang and Chaisattapagon (1995) used shape and geometrical features along with multivariate discriminant analysis for the identification of different weed species in wheat. Gebhardt et al. (2006) combined the colour, gradient and shape features for the detection of

Table 2
Overview of example applications of discriminant analysis algorithm in agriculture.

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
A database of broad and narrow leave weed	Angular second moment, inertia, entropy and local homogeneity.	GLCM matrices based textural features	Weed detection	96.7% for soil	Meyer et al. (1998)
Apple	Mean, Standard deviation, median, minimum, maximum of grey values at 450, 500, 750 and 800 nm, Perimeter, circularity, defect ratio, angular second moment, contrast, variance and inverse difference moment.	Colour, shape and textural features	Fruit grading based on surface defects	93.5% for two class grading system	Unay et al. (2011)
Bell pepper	Hue mapping of Orange, yellow, cyan, green, blue, violet, magenta and red colour.	Colour features	Vegetable grading based on colour change and defects	96% for colour and 63% for defects	Shearer and Payne (1990)
Carrots	Spectral response of the vegetation for 22 interference filters in range of 450–950 nm.	Spectral features	Weed detection	Overall accuracy of 72%	Piron et al. (2008)
Citrus	Colour values from RGB, HSI, Lab, LUV, XYZ, NIR, FL and UV images.	Colour features	Fruit grading and sorting on the basis of surface defects including disease and mould	87.2%, 83.3%, 83.7%, 82.1%, 71.1%, 63.4%, 79.5% and 92.9% for HSI, Lab, RGB, LUV, XYZ, FL, UV and NIR, respectively	Blasco et al. (2007)
Corn	Perimeter, area, length, width, major axis length, minor axis length, 1st, 2nd, 3rd, 4th, 5th, 6th, 7th equivalent width, maximum, minimum, mean radius, standard deviation of all radii, circularity, aspect ratio, radius ratio, Haralick ratio, four Hu moments (φ_1, φ_2, φ_3, φ_4) and Four Fourier descriptor (R_1, R_2, R_3, R_4).	Shape, geometrical and colour features	Grain variety identification	94%-100% for different varieties	Chen et al. (2010)
Grasslands	Mean values of intensity, green, red, mean value of gradient, shape factor, Circularity, Eccentricity, Area and Perimeter.	Colour, gradient and shape features	Weed detection	71% to 95%	Gebhardt et al. (2006)
Mandarins (Citrus)	Mean reflectance spectra extracted using VINIR hyperspectral camera (400–1000 nm).	Spectral features	Fruit grading based on stem detection and fungus infection	98.2% and 80.1% for sound and rotten mandarin	Gómez-Sanchis et al. (2008)
Mushrooms	Mean red spectral reflectance for 10 regions of interest and L channel of Hunter Lab colour space.	Colour features	Vegetable grading based on defects caused by freezing	100% of undamaged and 97.9% of freeze-damaged samples	Gowen et al. (2009)
Olives	Gradient images for wrinkle identification and difference of colour images.	Colour features	Fruit grading based on wrinkle detection	100%	Puerto et al. (2015)
Peaches	Mean reflectance spectra extracted using VINIR hyperspectral camera (400–1000 nm).	Spectral features	Chilling injury detection	92.96% to 97.28% for different level of injuries	Sun et al. (2017)
Pomegranate arils	Segmentation of arils in Red channel followed by the size and centroid estimation to remove extra small and large objects.	Colour and shape features	Fruit grading and sorting for rotten arils and membrane pieces	An average success rate was 90%	Blasco et al. (2009b)
Satsuma (Mandarin) segments	Area, axis of inertia, roundness, elongation, compactness, symmetry and 10 harmonics of FFT of shape signature.	Shape and Fourier transform based textural features	Fruit grading/ sorting	93.2% of complete fruit segments	Blasco et al. (2009a)
Seven different weeds found in nursery stock	Angular second moment, mean intensity, variance correlation, product moment, inverse difference moment, entropy, sum entropy, difference entropy and information measures of correlation I and II.	Co-occurrence matrices based textural features	Weed detection	Overall accuracy was 91%	Shearer and Holmes (1990)
Sugar beet	240 approximation coefficients extracted from Daubechies wavelet at level 4.	Wavelet transform based textural features	Weed detection	81.30%	Okamoto et al. (2007)
Sugar beet	Mean values of R, G, B, grey, ExG, ExR, CIVE, ExGR, NDI, GB, RBI, ERI, EGI and EBI colour indices.	Colour features	Weed detection	Maximum accuracy of 97%	Kazmi et al. (2015)
Tomatoes	Area, Major axis length, Minor axis length, centroid, Area/length, Compactness, Elongation, Perimeter/Broadness, Log_{10} (Height/Width) and Sum of radius of curvature.	Shape features	Weed detection	73.1% for tomatoes and 68.8% for weeds	Lee et al. (1999)
Weeds	Angular second moment, mean, variance, correlation, product moment, inverse difference moment, entropy, sum entropy, difference entropy and information measures of correlation I and II.	CCMs based textural features	Weed detection	93% with hue and saturation only.	Burks et al. (2000)
Wheat	Eccentricity, Compactness and Eight Hu invariant moments (φ_1, φ_2, φ_3, φ_4, φ_5, φ_6, φ_7, φ_8).	Shape and geometrical features	Weed detection	Not reported	Zhang and Chaisattapagon (1995)

(*continued on next page*)

Table 2 (*continued*)

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
Wheat and Barley	Short run emphasis, long run emphasis, grey level non-uniformity, run length non-uniformity, low grey level run emphasis, run percentage and high grey GLRM for angles 0°, 90°, 45° and 135°. Mean, variance and ranges of RGB and HSI, area, shape features and 20 harmonics of Fourier descriptors (FD) and features extracted from GLCM.	Colour, shape and textural features	Grain grading based on its moisture content	Up to 98% for bulk grain samples and up to 68% for the individual grain samples	Tahir et al. (2007)
White radish	Mean reflectance, semi transmittance and transmittance-spectra extracted using VINIR hyperspectral camera (400–1000 nm).	Spectral features	Hollowness detection	Overall accuracy of five class problem is 74.7%, 36.3% and 40.4% for transmittance, reflectance and semi-transmittance	Pan et al. (2017)
Wild blueberry and weed	Angular second moment, contrast, sum of squares, correlation, product moment, inverse difference moment, entropy, sum entropy, difference entropy and information measures of correlation I and II.	CCMs based textural features from HSI and NTSC luminance	Weed detection from wild blueberry and soil.	Accuracies of DF_HSI$_{SD}$, DF_SI$_{SD}$ and HSI$_{LD}$ algorithms were 94.9%, 92.7% and 91.4%, respectively.	Chang et al. (2012)
Winter wheat	Six approximation coefficients extracted at wavelengths ranging from 430 to 930 nm from continuous wavelet at scale of 4–8.	Textural features extracted at different wavelengths	Disease and pest detection	Average overall accuracy of 77%	Zhang et al. (2017a)
Winter wheat	MSR, NDVI, NRI, PRI, SIPI, PhRI, NPCI, ARI, RVSI, MCARI, HI, YRI, AI and PMI.	Spectral features	Disease and pest detection	Overall accuracies of 82.9%, 89.2%, 87.9% for three occurrence levels, i.e., slight, moderate, and severe	Shi et al. (2017)

broad leave weeds in the grasslands. The authors used the DA and achieved an accuracy ranging from 71% to 95% depending upon the phenological characteristics. The results of the study concluded that the shape features changes over time thus resulting in reduction of classification accuracy (Table 2).

DA-based techniques are among the most commonly used for fruit grading and sorting based on external defects, shape and size (Blasco et al., 2007; Gómez-Sanchis et al., 2008; Blasco et al., 2009a; Blasco et al., 2009b). A comparison between five different colour spaces (RGB, HSI, LUV, Lab and XYZ) was performed to identify the citrus peel defects by Blasco et al. (2007). The results indicated that classification criterion trained based on the LDA achieved good classification accuracy (> 80%) for all colour spaces except XYZ colour space. Blasco et al. (2009a) developed a real-time sorting facility for grading the satsuma (mandarin) segments into whole and broken one using their shape attributes. The circularity, compactness, symmetry, elongation and Fourier descriptors were used to quantify the shape, while, area and length were used to represent the size of the segments. The developed sorting facility was able to correctly classify 93.2% of complete fruit segments. In addition to the shape (circularity) and size (perimeter) feature, colour, texture and defect ratio were used to grade the apples in two and multiple classes. The LDA and several other non-linear models were compared for grading and results showed that two-class grading approach achieve higher accuracy when compared with multi-class counterpart. Sun et al. (2017) extracted mean spectral features from the VINIR hyperspectral reflectance imaging (400–1000 nm) for classifying peaches effected by chilling injury. Discriminant models for two-class (non- and chilled), three-class (non-, semi- and heavy-chilled) and four-class (non-, slight-, moderate- and heavy-chilled) were developed by using partial least square discriminant analysis (PLS-DA) and LDA. Results showed that both PLS-DA and LDA had achieved an overall accuracy of 94.37% and 92.96% for two-class scenario with six optimally selected wavelengths. However, the classification accuracies of both PLS-DA and LDA for three-class and four-class showed poor classification accuracies than two-class. Pan et al. (2017) used the similar approach for the hollowness detection in white radish, but, rather than just using the mean spectral features from reflectance imaging, they extracted features from reflectance, semi-transmittance and transmittance hyperspectral imagining (400–1000 nm). The overall classification accuracies of two-class problem were found to be 92.9%, 84.8% and 95.9% for transmittance (3 optimal wavelengths), reflectance (10 wavelengths) and semi-transmittance (14 wavelengths), respectively, using optimally selected wavelengths with PLS-DA analysis. The results of study concluded that semi-transmittance hyperspectral imaging have more potential for non-invasive hollowness identification (Table 2).

In addition to the fruit grading, DA was also used for the vegetable (Shearer and Payne, 1990; Gowen et al., 2009), grain grading (Tahir et al., 2007; Chen et al., 2010) and disease detection in plants (Zhang et al., 2017a; Shi et al., 2017). Shearer and Payne (1990) graded the bell peppers based on their colour attributes and physical defects by mapping the hue component for different primary and secondary colours. Feature selection procedure and QDA were used to develop classification criterion to achieve an overall accuracy of 96%. Gowen et al. (2009) used the red band for calculating the mean spectral reflectance from 3 × 3 region of interest (ROI) marked on the centre of the mushroom surface to identify the defects caused by thawing and freezing. Principal component analysis (PCA) along with LDA used to correctly detect the infected mushrooms with having an accuracy of 87.9% for the calibration set. The combination of colour, shape and textural features was used to quantify the effect of different moisture levels on the appearance and grain kernel morphology (Tahir et al., 2007). The images of Canada Western Red Spring (CWRS), Canada Western Amber Durum (CWAD) wheat and barley were taken in individual and bulk fashion from the conditioned grains with moisture content varying from 12% to 20% with a 2% step. The colour features

Table 3
Overview of example applications of k-Nearest Neighbour algorithm in agriculture.

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
Alfalfa, corn, soybeans, pasture and winter wheat	Mean and correlation from GLCM, large number emphasis and second moment from NGLDM extracted from multi-polarization airborne radar	Textural features	Crop classification	Highest mean classification accuracy of 83%	Treitz et al. (2000)
Apple	Mean reflectance spectra in a range of 460–1130 nm extracted using adjustable filters	Spectral features	Surface defects detection	91.9–95.7% depending upon the number of input variables	Miller et al. (1998)
Apple and Guava	3D volume reconstruction followed by the normalization, mean and standard deviation extracted from binary, RGB and HSI colour models along with colour co-occurrence matrices and run-length textural features extracted from RGB and HSI	Colour and textural features	Fruit grading based on bruise detection for apple and skin damage for guava	78–92.8% for apple and 80.2–88.5% for guava	Yimyam and Clark (2016)
Blueberry	R, G, B, H, S, I, Q, Cb, Cr and EGI	Colour features	Identification of different growth/maturity stages of fruit	85%-98% for different fruit maturity stages	Li et al. (2014)
Cereal	Major axis length, Aspect ratio, Area, (Major axis length)2/Area, Roundness, Seven Hu invariant moments (φ1, φ2, φ3, φ4, φ5, φ6, φ7)	Colour and Shape features	Weed detection	79.2% with k = 5	Perez et al. (2000)
Citrus	$NDVI_{870}$, $NDVI_{970}$, SR_{870}, SR_{970}, $MTVI_1$, $MTVI_2$, RDVI, greenness index, TVI, MCARI, SIPI	Spectral features	Disease detection in leaves	96% with five leaf samples from each tree	Mishra et al. (2011)
Citrus	Mean reflectance spectra extracted using SWIR spectrometer (350–2500 nm), pre-processed with 1stderivative, 2nd derivative, SG smoothing and derivative	Spectral features	Disease detection in leaves	Overall accuracy of 83.3%, 86.8%, and 86.8% for 1st derivative, 2nd derivative and combined spectral features	Sankaran et al. (2011)
Citrus	Mean reflectance spectra extracted using visible-near infrared spectrometer (350–2500 nm) and absorbance spectra using mid-infrared spectrometer (5.15 to 10.72 µm)	Spectral features	Diseases detection in leaves	Overall accuracy of 89.4% and 99% for visible and mid-near infrared, respectively	Sankaran and Ehsani (2013)
CWRS wheat, CWAD wheat, barley, oats and rye	Mean, variance and range of red, green, blue, hue, saturation and intensity	Colour features	Identification of two grain types based on their external characteristics	94.1, 92.3, 93.1, 95.2 and 92.5%, for CWRS wheat, CWAD wheat, barley, oats and rye, respectively	Majumdar and Jayas (2000a)
CWRS wheat, CWAD wheat, barley, oats and rye	Area, Perimeter, length, width, major axis length, minor axis length, thinness ratio, aspect ratio, rectangular aspect ratio, area ratio, maximum radius, minimum radius, standard deviation of all radii, Haralick ratio, Fourier descriptors and spatial moments	Shape features	Identification of two grain types based on their external characteristics	98.9, 93.7, 96.8, 99.9 and 81.6%, respectively for CWRS wheat, CWAD wheat, barley, oats and rye	Majumdar and Jayas (2000b)
CWRS wheat, CWAD wheat, barley, oats and rye	Mean, variance, uniformity, entropy, maximum probability, correlation, homogeneity, inertia, cluster shade, cluster prominence, short run, long run, grey level non-uniformity, run length non-uniformity, run percent, GLRM entropy	Texture features	Identification of two grain types based on their external characteristics	85.2, 98.2, 100.0, 100.0 and 76.3%, for CWRS wheat, CWAD wheat, barley, oats and rye, respectively	Majumdar and Jayas (2000c)
Land cover crops	Six reflectance spectra in visible and near infrared region extracted using Landsat TM and ETM +	Spectral features	Land use and land cover classification for 9 different crops	Not reported	Samaniego and Schulz (2009)
Mango	Mean reflectance spectra extracted using VINIR (650–1080 nm) hyperspectral imaging	Spectral features	Mechanical damage detection	94.87% to 98.08%	Rivera et al. (2014)
Rapeseed	Angular second moment, contrast, correlation variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measure of correlation I and II, maximal correlation coefficient, short run emphasis, long run emphasis, low grey-level run emphasis, high grey level run emphasis, short run low grey-level run emphasis, short run high grey-level run emphasis, long run low grey-level run emphasis, long run high grey-level run emphasis, grey level non-uniformity, run-length non-uniformity, run percentage and 59 LBP features	Textural features	Identification of seed varieties	Highest accuracy of 97.8% with k = 3 and recursive feature elimination technique	Kurtulmuş and Ünal (2015)
Sugar beet	Green, red and blue mean, green, red and blue standard deviation, area, perimeter, compactness, elongation, solidity, forma factor, six rotation scale and translation invariant moments as defined by Jian (1989) (φ1, φ2, φ3, φ4, φ5, φ6)	Colour and shape features	Weed detection	97% classification accuracy with all 19 features	Åstrand and Baerveldt (2002)
Tobacco	Features extracted using local binary pattern, local binary pattern variance, grey level local texture pattern with different radius and number of pixels	Textural features	Identification of leaves based on their maturity	Maximum accuracies of 72.75%, 75.93% and 80.91% for LBPV, LBP and GLTP, respectively	Guru et al. (2012)
Tomatoes					Xu et al. (2011)

(*continued on next page*)

Table 3 (*continued*)

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
	Percent intensity histogram, percent differential histogram, Fourier rectangular direction spectrum energy percentage of 'b' channel in CIE Lab, wavelet packet decomposition in three spaces using M-band wavelet packet (Acharyya and Kundu, 2001)	Colour and textural features	Identification of nutrient deficient leaves	92.5%, 85% and 82.5% for normal, nitrogen and potassium deficient leaves	
weed database	200 horizontal, vertical, diagonal and approximation coefficients extracted during decomposition stage	Textural features	Weed detection	94% for broad and 92% for narrow leave weeds	Ahmad et al. (2011)
Wheat and Barley grains	Area, perimeter and shape features, Mean and standard deviation of RGB, mean and variance of XY, uniformity, maximum probability, correlation homogeneity, cluster shade and cluster prominence of the GLCM for angles 0°, 90°, 45° and 135°	Colour, shape and textural features	Identification of two grain types on the basis of their external characteristics	99% by combining the properly selected shape, colour and texture features	Guevara-Hernandez and Gomez-Gil (2011)

attributed more towards the identification followed by the textural features. Zhang et al. (2017a) used spectral features extracted from UV/VINIR spectroradiometer (350–2500 nm) to identify three different diseases and their severity in wheat crop. The continuously decomposed wavelet scalograms along with overlapping strategy was used to select a total of six wavelet features extracted at wavelengths ranging from 430 to 930 nm at scale of 4–8. The training of LDA based on the selected features resulted in overall classification accuracy of 77% using k-fold cross-validation (k = 5). Fourteen different spectral vegetation indices extracted from VINIR hyperspectral images (400–1000 nm) were used to identify yellow rust, aphids and powdery mildew disease in winter wheat. Kernel discriminant analysis (KDA) trained on the selected spectral vegetation indices achieved an accuracy of greater than 87% for classifying healthy and infected leaves at canopy level (Shi et al., 2017). However, the major limitation of DA is that it is suitable for the data having unimodal Gaussian distribution for each class and therefore can be successfully applied to only these kind of scenarios (Bishop, 2007) (Table 2).

2.1.3. *K-Nearest Neighbour*

kNN as opposed to NB and DA is a discriminative, non-parametric and instance-based classifier that does not make any explicit assumption regarding the distribution of the dataset in every class (Bishop, 2007), thus can be suitable for the non-Gaussian data sets. Since kNN is an instance based, therefore it does not learn the model explicitly during training and uses the training instances as a knowledge for the prediction purposes requiring them to retain all the observations as a part of model (Mitchell, 1997). The new unknown observation is classified by computing the similarity measure (Euclidean distance) between unknown and each training observation. The predefined 'k' closest points (based on measured distance) in training data are used for calculating the conditional class probability, followed by the assignment of the observation to a class with largest probability (Mitchell, 1997; Bishop, 2007). In addition to the Euclidian distance, Manhattan, Chebyshev and Mahalanobis can be used for numerical data (depending on the properties of data), while the Hamming distance can be used for the categorical or binary data (Tibshirani et al., 2013) (Table 3).

In the literature for agriculture, the most common application of kNN algorithm was found to be more effective for classification of different grains/grain cultivars. Majumdar and Jayas (2000a) used colour features for discriminating between CWRS wheat, CWAD wheat, barley, oats and rye with kNN and LDA. The results reported that kNN achieved better accuracies for almost all the classes (k = 5), when compared with LDA. The accuracies of 94.1, 92.3, 93.1, 95.2 and 92.5% were reported for the CWRS wheat, CWAD wheat, barley, oats and rye, respectively. Majumdar and Jayas (2000b) also used the morphological features for comparing the performance of LDA against kNN for the same grain classes and reported that kNN outperformed the LDA for morphological features also. Guevara-Hernandez and Gomez-Gil (2011) extracted the textural features to classify the wheat and barley kernels with LDA and kNN. In addition to the 6 colours, 21 shape features and 72 textural features were extracted by developing grey level co-occurrence matrix (GLCM) and grey level run length matrix (GLRM) in four different orientations (0°, 45°, 90° and 135°). The authors concluded that the combination of shape, colour and texture can provide better accuracy as compared to any of these individually. The classification accuracy can be as high as 99% by careful selecting a set from these pooled features. Kurtulmuş and Ünal (2015) identified the different rapeseed varieties based on their external appearance by using kNN, SVMs and stochastic gradient descent (SGD) approaches. They used the 14, 11 and 59 textural features extracted from GLCM, GLRM and local binary pattern (LBP), respectively, as a descriptor of the grain external appearance. The results indicated that highest classification accuracy achieved by kNN algorithm was 97.8% at k = 3 and recursive feature elimination, while for SVMs and SGD, highest classification accuracies were 99.7% and 95.8%, respectively (Table 3).

Table 4
Overview of example applications of SVMs algorithm in agriculture.

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
Apple	A total of 70 features, including 14 from colour (Chroma, hue angle, mean, variance, skewness and kurtosis), 24 from wavelet (energy and entropy), and 32 from grey level co-occurrence matrix (homogeneity, measure of smoothness, third moment, measure of uniformity, entropy, energy, contrast and correlation) technique.	Colour, shape, and texture-based features	Leaf diseases classification	Coefficient of determination was achieved 96.3% with rbf kernel during testing	Omrani et al. (2014)
Banana, beans, guava, jackfruit, lemon, mango, potato and tomato	Texture features like contrast, energy, local homogeneity, cluster shade and cluster prominence	Texture features from hue-saturation-intensity colour model using colour co-occurrence matrix analysis	Detection of unhealthy region of plant leaves	Diseases detected and classified with an accuracy of 94%	Arivazhagan et al. (2013)
Chilli	Total fourteen features were extracted from three categories from digital images. colour features: red, green and blue colour components; Size independent shape features: form factor, elongatedness, convexity, solidity; moment invariants including second and third order moment invariants	Colour, shape and moment invariant features	Crop and weed classification	Accuracy achieved above 97%	Ahmed et al. (2012)
Corn	Hyperspectral images that collected from 72 narrow bands in a range from 408.73 to 947.07 nm (visible to near-infrared), with bandwidths which was varied from 4.27 to 4.41 nm	Spectral features	Weed and nitrogen stress detection	Classification results were 86 and 81%, respectively for weed and nitrogen	Karimi et al. (2006)
Cotton	Colour features: mean of red (R), green (G), and blue (B) colour component, mean of combination of R, G and B, standard deviation shape features: form factor, aspect ratio, rectangularity, solidity, eccentricity, sphericity, Euler number and texture features: mean intensity, mean contrast, roughness, third-order moment, consistency and entropy	Colour, shape and texture features	Classification of foreign fibres in cotton lint	Highest average accuracy was 93.57% using one-against-one voting based MSVM	Li et al. (2010)
Cotton	Energy, inertia, entropy, homogeneity, correlation, solidity, extent, minor axis length, eccentricity, dispersion R, G and B, Gary R, G, B, H, S and V and Grey_hist R, G, B, H, S and V	Shape, texture features, fractal dimension, lacunarity, dispersion, grey level, grey histogram discrimination and Fourier descriptor	Visual symptoms of plant diseases	The classification accuracy was 90% using all features	Camargo and Smith (2009)
Dense and sparse grasses	Co-occurrence of Binary Pattern feature extraction using Local Binary Operator (LBP) and GLCM	Texture features	Vegetation classification	Overall accuracy was 91.82% using SVMs for linear kernel function	Chowdhury et al. (2015)
Grape	Hue, saturation and Cr component from YCbCr colour space are applied to extract salient colour features by Gabor wavelet	Colour features	Leaf disease detection	Detection accuracy was 97.8%	Meunkaewjinda et al. (2008)
Maize	Intensity, Mean, Energy,				
Entropy, Standard Deviation, Smoothness and Third Moment	Texture features	Classifying crop or weed	Classification of weed or a crop resulted in an accuracy of 82% with texture features	Athani and Tejeshwar (2017)	
Pomegranate	Hue, saturation and value for colour and contrast, energy, entropy for texture features	Colour and texture features extracted after segmentation by K-means clustering	Leaf disease detection	An accuracy of 97.30% was found for detection of disease spots and classifying the leaf image	Sannakki et al. (2013)
Potato	Mean, standard deviation, skewness, kurtosis, RMS, contrast, correlation, energy, entropy and homogeneity	Colour and texture statistical feature	Classifying diseases on plant	An accuracy of 95%	Islam et al. (2017)
Rice	Shape features: rectangularity, compactness, elongation, roundness and textural features: contrast, uniformity, entropy, inverse difference and linearity correlation	Shape and colour texture features	Classifying rice bacterial leaf blight, rice sheath blight and rice		
blast	Disease spots detected with an accuracy of 97.2%	Yao et al. (2009)			
Rice	Morphological traits including seedling height, width, aspect ratio, seedling length, second-internode length, second-leaf length, third-leaf length and leaf angle and	Morphological and colour traits of the seedlings	Bakanae disease detection in rice seedlings	Distinguishing healthy and infected seedlings achieved an overall accuracy of 87.9%	Chung et al. (2016)

(continued on next page)

Table 4 (*continued*)

Crop	Image features	Type of features	Application area	Classification accuracy	Reference
	colour traits including mean of the CIE L*c*h* colour model and hue values				
Rose, bean, lemon and banana	Local homogeneity, contrast, cluster shade, energy and cluster prominence	Texture features	Plant leaf diseases detection	SVMs showed an accuracy of 95.71%	Singh and Misra (2017)
Sugar beet	Hyperspectral reflectance was measured with the range from 400 nm to 1050 nm and with a spectral resolution of 1.4 nm.	Spectral vegetation indices	Leaf diseases (i.e., Cercospora leaf spot, leaf rust and powdery mildew) detection	Classification accuracies up to 97% for healthy sugar beet leaves and diseased leaves	Rumpf et al. (2010)
Tomato	Energy, contrast, sum of squares, correlation, entropy, sum entropy, cluster shade, cluster prominence, homogeneity from GLCM	Texture features	Detection and identification of unhealthy leaves	Highest classification accuracy of 99.83% was achieved using linear kernel function	Mokhtar et al. (2015)
Wheat	Colour features were average, correlation, deviation, energy, shape features were area ratio, compactness, first, second, third and fourth rotation invariant moment and texture features were energy, entropy, moment of inertia, local smooth and correlation	Colour, shape and texture features	Recognition of leaf diseases including Powdery Mildew, leaf rust *Puccinia triticina*, leaf blight and Puccinia *striiformis*	SVMs based multiple classifier system (MCS) had better accuracy rate of 95.16%	Tian et al. (2012)

Miller et al. (1998) developed a baseline spectral reflectance data for different types of surface defects on apples using two adjustable filter systems in the range of visible (460–750 nm) and NIR region (750–1030 nm). The accuracies were reported to be dependent on the number of spectral features used for training a kNN and found to be in a range of 91.9% to 95.7%. Similarly, a method of early detection of mechanical damage on mangoes was developed by using VINIR hyperspectral (650–1080 nm) images (Rivera et al., 2014). The results indicated that kNN achieved highest classification accuracy within a range of 94.87% to 98.08% depending on the days after the mechanical injury. Yimyam and Clark (2016) reconstructed the 3D images from multi view images (one top image and 3 side images) for apple and guava grading. The newly constructed 3D image was used to extract the colour features from binary, RGB and HSI colour spaces along with textural features (GLCM, GLRM) extracted from RGB and HSI images. The accuracies were reported to be in the range of 78% to 92.8% and 80.2% to 88.5% for apple and guava, respectively. Li et al. (2014) developed a five-stage kNN classifier using the colour features to discriminate the blueberry fruit based on their maturity stage. During the first stage the kNN classified the fruit from background, followed by separation of "mature and near mature" from "young and near young" classes. Lastly, the kNN classifiers separated the mature from near mature and young from near young fruits. The reported accuracy was ranged from 85% to 98% depending on the maturity stages (Table 3).

A multiband active optical sensor was developed to identify the Huanglongbing (HLB) infected citrus trees from the healthy ones (Mishra et al., 2011). The optical sensor was designed to monitor the spectral reflectance response of tree canopies using two visible wavelengths (570 nm and 670 nm) and two NIR wavelengths (870 nm and 970 nm). The authors reported that detection of the infected trees with only one measurement from each tree resulted in poor accuracy and there, used five sets of measurements from each tree to achieve an accuracy of 96% with kNN. Sankaran et al. (2011) as opposed to Mishra et al. (2011) extracted the mean reflectance spectra by extending the effective wavelength range up to short wavelength infrared spectrum (350–2500 nm) using spectrometer for detecting the HLB infected trees. The raw spectra were pre-processed using 1st derivative, 2nd derivative and SG smoothing and derivative. The maximum overall classification accuracy of 86.8% was observed with kNN and all combined features (raw spectra, 1st derivative and 2nd derivative). Sankaran and Ehsani (2013) extracted mean reflectance spectra in short wavelength infrared (350–2500 nm) and mean absorbance spectra in mid-infrared (5.15 to 10.72 µm) for detecting the HLB using kNN. The reported accuracy (99%) was higher for the mid infrared compared to short wavelength infrared (89.4%). Other applications of kNN algorithm in agriculture involves weed detection (Perez et al., 2000; Åstrand and Baerveldt, 2002; Ahmad et al. 2011), land use land cover classification (Treitz et al., 2000; Samaniego and Schulz, 2009) nutrition deficient plant leaves (Xu et al., 2011). A major drawback of the kNN algorithm is large processing time to classify a new unknown observation based on distance calculated from every observation, which is not desirable for real-time applications (Mitchell, 1997). Furthermore, the multidimensional data tends to decrease the classification accuracy of kNN as the difference between nearest and farthest neighbour is little (Shi and Judd, 2013) (Table 3).

2.1.4. Support vector machines

SVMs are binary classifiers able to classify experimental data samples in two disassociate classes (Cortes and Vapnik, 1995; Burges, 1998; Vapnik, 1998). The principle of SVMs comes from the simplified case is which the two accounted classes are linearly separable. A hyperplane able to differentiate all data samples in two classes exists. The SVMs classifier can also be used for classifying data which are not linearly separable. The data space is transformed into a higher-dimensional space where classes become linearly separable. Kernels functions are used to implicitly apply a suitable transformation on the data space.

The nonlinearly maps the data to a high-dimensional space by using kernels and then tries to find the hyperplane that separates data with maximum margin in new space (Burges, 1998; Vapnik, 1998). The SVMs require a training phase where data samples are conveniently provided for measuring the required parameters (support vectors) and then the posterior decision phase is highly dependent of the data samples supplied, i.e., from the data which have been utilized for training the classifier (Guerrero et al., 2012) (Table 4).

The SVMs based ML technique is successful and applied different areas of agricultural machine vision systems such as plant diseases detection (Meunkaewjinda et al., 2008; Camargo and Smith, 2009; Yao et al., 2009; Arivazhagan et al., 2013; Omrani et al., 2014; Chung et al., 2016), weed detection (Karimi et al., 2006; Tellaeche et al., 2011; Guerrero et al., 2012; Athani and Tejeshwar, 2017), fiber classification (Li et al., 2010), vegetation classification (Chowdhury et al., 2015), grading and sorting (Mizushima and Lu, 2013) and vegetable inspection (Razmjooy et al., 2012). A grape leaf disease detected by Meunkaewjinda et al. (2008) using SVMs based multiple artificial intelligent techniques. Camargo and Smith (2009) utilized SVMs to identify visual symptoms of cotton diseases in images. The SVMs used a radial basis function (RBF) kernel and one-against-one method to overcome the problems of nonlinearity and classifying the multiple classes. Yao et al. (2009) detected rice diseases by using SVMs with shape and texture features. The classification was performed between three diseases such as rice bacterial leaf blight, sheath blight, and blast and achieved 97.2% of accuracy using combination of shape and texture features. The accuracy was dropped when the features were used separately. All the experiments were evaluated by using rbf kernel of SVMs classifier. Omrani et al. (2014) found radial basis function was effective for apple disease detection. The authors also compared SVMs classifier with artificial neural networks (ANNs) and suggested SVMs approach provided better results than the ANNs for disease classification. Arivazhagan et al. (2013) followed five main steps to identify plant diseases as follows: image acquisition, colour transformation, masked green pixels as segmentation process, computation of texture statistics by colour co-occurrence matrix and finally classifier is used for the features that are extracted to classify the disease. Using extracted features, the trained SVMs based ML classifier was able to identify several plants bacterial and viral diseases such as early and late scorch and fungal diseases in beans with an accuracy of 94%. Chung et al. (2016) detected Bakanae disease from rice seedlings by using SVMs based ML technique. Two SVMs classifiers were used in a cascade for distinguishing the control and inoculated seedling and the seedlings at different inoculation levels. A soft-margin SVMs classifiers with radial basis function kernel were used and achieved an overall accuracy of 87.9% (Table 4).

The SVMs based ML is not only proved as effective technique in plant diseases detection but also proved effective in weed identification. Karimi et al. (2006) evaluated SVMs based ML as a tool for classifying airborne hyperspectral images taken over a corn field. The nitrogen application rates and weed management practices were considered to evaluate classifier performance and the performance was compared with ANNs based classifier on the same data. The SVMs method resulted in very low misclassification rates as compared to the ANNs approach for all cases with high relative value of kappa of 0.79 and 0.71 for weeds and nitrogen classes. Detection of stresses in early crop growth stage using the SVMs method could aid in effective early application of site-specific remedies. Tellaeche et al. (2011) successfully applied SVMs to identify *Avena sterillis* weed grown in cereal crop fields. The images of this experiment were segmented to extract cells from images as the low-level units by combining basic suitable image processing techniques. Finally, SVMs classifier evaluated each cell and determined whether it needs to be sprayed with highest correct classification rate of 85%. Guerrero et al. (2012) addressed greenness identification problem by using a learning approach based on SVMs in maize field. The authors concluded SVMs is capable to identify plants with green spectral components masked and unmasked. The Gaussian radial basis, polynomial and sigmoid functions were experimented and achieved an average percentage of success of 93.1% (Table 4).

SVMs were also applied for fruit grading and sorting, and vegetable inspection in crop cultivation. Mizushima and Lu (2013) developed an automatic adjustable algorithm by using segmentation technique based on a linear SVMs in combination with the Otsu's thresholding method to sort and grade apples and achieved an error fewer than 2%. The proposed method automatically adjusted the classification hyperplane calculated by using linear SVMs and required minimum training and time. Unay et al. (2011) graded bi-colour apples by using SVMs based multispectral machine vision system (centred at 450, 500, 750 and 800 nm with respective bandwidths of 80, 40, 80 and 50 nm). The authors observed the recognition accuracy on healthy fruits was 94.6%, while on that of defective fruits was slightly lower (92.3%) with selected features. On the other hand, Razmjooy et al. (2012) inspected potato by using SVMs based ML and achieved 1% of false rejection rate with sequential minimal optimization (SMO) kernel. Four different kernels (linear, quadratic, poly nominal and SMO) were used and achieved 96.86% of accuracy for sizing and about 95% for the best classifier with SVM-SMO. However, a major drawback of SVMs classifier is that it has some limitations in size and speed both in training and testing and it also has very difficult structure to understand (VijayaLakshmi and Mohan, 2016) (Table 4).

2.1.5. Other supervised machine learning algorithms

Besides previously introduced supervised machine learning algorithms, other statistical machine learning algorithms that can potentially be used for agricultural machine vision systems includes decision trees, random forests, and logistic regression. Yang et al. (2003) used decision trees coupled with hyperspectral imaging to distinguish between different tillage methods and residual levels with accuracies of 89% and 98%, respectively. Random forests are ensemble learning techniques which are generally based on several decision trees (Criminisi et al., 2012). Many researches have attempted to use the random forests to classify land cover using multispectral and hyperspectral satellite imagery (Rodriguez-Galiano et al., 2012). Logistic regression was applied to the development of land use map to detect the level and degree of deforestation (Schneider and Pontius Jr., 2001). Shin et al. (2012) studied a machine vision system for estimating fruit sizes and fruit count during post-harvest processing. A logistic regression model-based pixel classification algorithm was developed for fruit detection from images of the postharvest citrus in commercial citrus orchard.

2.2. Unsupervised machine learning algorithms

2.2.1. K-Means Clustering

The K-means clustering is an unsupervised learning technique that used unlabelled data for classification. The principle of this classifier is to find groups in the data, with the number of groups represented by the variable K. The K-mean classifier works iteratively to assign each data point to one of K groups based on the features that are provided. Then, data points are clustered based on feature similarity. The K-means clustering algorithm is mostly used in agriculture to delineate the region of interest (ROI) by using geometrical distribution of the spectral features (Bishop, 2007). The spectral information of an agronomic image can be used for partitioning the vegetative biomass and soil residues (Steward and Tian, 1999; Steward et al., 2004; Carroll and Holden, 2005; Nieuwenhuizen et al., 2007; Pérez-Ortiz et al., 2016; Senthilnath et al., 2017) or for segmenting the diseased/defected regions on plants /fruits from healthy areas (Leemans and Destain, 2004; Unay and Gosselin, 2006; Al Bashish et al., 2011; Dubey and Jalal, 2014; Hu et al., 2014; Pujari et al., 2013; Deshpande et al., 2014; Kruse et al., 2014; Pham and Lee, 2015) based on pre-defined number of clusters. The segmentation results of K-means clustering algorithm are

Table 5
Overview of example applications of K-means clustering algorithm in agriculture.

Crop	ML task	Data/sensor type	Application area	Classification accuracy	Reference
Apple	Classification	Colour camera	Fruit grading	73%	Leemans and Destain (2004)
Apple	Identification	Colour camera	Automated robotic harvesting system	Approximate accuracy of 80% under lighting condition	Bulanon et al. (2004)
Apple	Identification	Multispectral camera with 450, 500, 750 and 800 nm filters	Defects detection	~85% for all defects	Unay and Gosselin (2006)
Apple	Defected pixel segmentation and classification	Colour camera	Defects and disease detection in fruits	Maximum accuracy of 93.17%	Dubey and Jalal (2014)
Apple	Identification	Thermal infrared and colour camera	Automated fruit harvesting	An accuracy of 74.37% with combination of high and low-level features	Wachs et al. (2010)
Banana	Banana Finger and Flaws Segmentation	Colour camera	Fruit grading	Visual categorization	Hu et al. (2014)
Cereal	Database of weed maps (Lutman et al., 1998)	Grey-scale scanned maps	Weed mapping	Not reported	Carroll and Holden (2005)
Cotton	Soil management zone delineation	VERIS 3100 electrical resistivity meter and Landsat 8 imagery	Variable rate irrigation	40% of soil available water content variance explained	Haghverdi et al. (2015)
Cucumbers	Disease pixel segmentation and classification	Colour Camera	Disease detection on leaves	Overall accuracy was 85.70%	Zhang et al. (2017b)
Different plants in Ghor-area	Diseased pixel segmentation and classification	Colour camera	Disease detection on leaves	Precision of around 93%	Al Bashish et al. (2011)
Different tree and vegetation species	Spectral-spatial classification	GoPro and Pi colour camera	Crop region and tree type mapping using UAV	80.7% for crop region and 45.6% for tree type	Senthilnath et al. (2017)
Grapes	Classification	Video/Colour Camera	Yield estimation based on shoot detection	Accuracy was calculated to be 86.83%	Liu et al. (2017)
Mango, Grape and Pomegranate	Diseased pixel segmentation and Classification	Colour Camera	Fungal disease detection on fruit and leaves	84.65% for normal and 76.6% for defected fruits with anthracnose	Pujari et al. (2013)
Oil palm	Classification	Web Cam	Fruit grading based on ripeness	Not determined	Jaffar et al. (2009)
Oil palm	Classification	Web cam	Fruit grading	Overall 93.53%	Makky and Soni (2013)
Orange	Fruit shape topologies estimation for different genotype classification	Colour Camera	Genotype classification	74.50%	Costa et al. (2009)
Orange and Potato	Identification	Colour camera	Defects detection	Not reported	Pham and Lee (2015)
Pomegranate	Segmentation	Colour camera	Disease detection on leaf and fruit followed by estimation of infestation level	Not determined	Deshpande et al. (2014)
Rice and Cotton	Segmentation	Colour camera	Vegetation segmentation from non-vegetative portions	88.1% for rice and 91.7% for cotton	Bai et al. (2013)
Soybeans	Segmentation	Colour camera	Weed detection	Not reported	Steward and Tian (1999)
Soybeans	Vegetation segmentation from background	Colour camera	Weed detection	89.6% and 91.9% on cloudy and sunny conditions	Steward et al. (2004)
Strawberry	Classification (long taper, square, taper, round)	Colour camera	Fruit grading with shape characteristics	Strawberry size detection accuracy was around 95%	Liming and Yanchao (2010)
Subterranean clover	Injured pixel segmentation and quantification	Colour camera	Leaf injury induced by ozone	93%	Kruse et al. (2014)
Sugar beet	Classification	Colour camera	Weed detection	Maximum accuracy of 97%	Nieuwenhuizen et al. (2007)
Sunflower and Maize	Segmentation followed by classification	Colour camera	Weed mapping using UAV	Maximum accuracy of 95.50%	Pérez-Ortiz et al. (2016)
Tomato	Identification	Colour camera	Automated fruit harvesting	Not reported	Yin et al. (2009)
Wheat	Classification	Colour camera	Grain grading	Overall accuracy of 88.33%	Olgun et al. (2016)
Wheat	Identification	Colour camera	Row detection	Detection rate was up to 90%	Jiang et al. (2016)

different for predefined manually adjusted (Mery and Pedreschi, 2005; Blasco et al., 2009a) or automated (Liming and Yanchao, 2010) histogram-based thresholding in way that it can result in more continuous regions (Ho and Lee, 2003) (Table 5).

Steward and Tian (1999) identified the weed locations by using K-means clustering algorithm to cluster the individual pixel with a class of similar pixel based on their colour attributes. Their algorithm used two clusters to represent the background and two for the vegetative regions (plants and weeds). Steward et al. (2004) developed a reduced-dimension clustering (RDC) algorithm for classifying the image pixels to vegetation and background followed by training a Bayes classifier for further segmentation of vegetation and background. The results of the algorithm showed accuracies of 89.6% and 91.9% for cloudy and sunny conditions, respectively. Nieuwenhuizen et al. (2007) used an algorithm like RDC and compared its performance with adaptive resonance theory 2 (ART2) neural network for the identification of potato in sugar beet fields. They concluded that both algorithms performed similarly with 97% and 49% correct classification in first and second field, respectively. Pérez-Ortiz et al. (2016) mapped weeds in sunflower and maize fields using Otsu thresholding algorithm (Otsu, 1979) along with K-means clustering to pool the segmented image in three different clusters (background, plant and weeds) (Table 5).

Presence of any external blemishes or disease on both fresh and processed agricultural products can be attributed to its colour or shape change (Cubero et al. 2011, Chang and Rehman, 2018). The external attributes can be used only for disease/defects detection or can be explored further with shape and/or textural features for automated online grading systems (Chang and Rehman, 2018). Leemans and Destain (2004) developed an apple grading system by segmenting the fruit surface to different blobs. An array of different features was calculated from these blobs followed by features clustering using K-means clustering algorithm. The results indicated a correct classification of 73% for apples. Unay and Gosselin (2006) compared the performance of several supervised and unsupervised classification-based techniques for pixel-wise segmentation of defects on apple surfaces. The results of their study indicated that supervised classifiers were more accurate with an approximate error range of 20–30%. Jaffar et al. (2009) graded the oil palm fresh fruit bunches based on the estimated ripeness using K-means clustering algorithm to segment the different coloured clusters. Dubey and Jalal (2014) used K-means clustering technique for defect segmentation along with a multiclass SVMs classifier to classify among different types of diseases common to apple fruit. The results showed that their system was able to correctly classify 93% of apple diseases. Olgun et al. (2016) developed an automated wheat grain classification system by using K-means clustering which was operated on Dense Scale Invariant features (DSIFT). The performance of DSIFT was evaluated by SVMs and achieved an overall accuracy of 88.33% for classification (Table 5).

Similar to the detection of diseased/defected surfaces on the fruits, K-means clustering technique was used for identifying the diseases on the leaf surfaces of different crops. Kruse et al. (2014) compared four techniques to classify individual pixel as healthy or injured by the air pollutant ozone. The results indicated the highest classification mean accuracy of 95% for LDA, whereas K-means clustering was able to achieve an accuracy of 93%. Zhang et al. (2017b) used K-means clustering algorithm for segmenting the diseased portions of the cucumber leaves followed by training five multiclass classifiers to differentiate among seven diseases. The results showed that highest classification accuracy (91.25%) was achieved for grey mould (Table 5).

Likewise, the application of K-means clustering is also being explored to identify and locate the position of fruits as a part of harvesting robot development. Bulanon et al. (2004) developed a harvesting robot for red Fuji apple by using K-means clustering. The authors used chromaticity method-based colour model for fruit detection instead of colour difference method and RGB model. The chromaticity method was represented by two chromaticity coordinates (r and g) which were calculated by $r = \frac{R}{R+G+B}$ and $g = \frac{G}{R+G+B}$ and followed by a noise removal low-pass filter. The authors reported that approximately 80% of fruit pixels were correctly classified under lighting condition. Wachs et al. (2010) used high and low-level visual features with colour and thermal infra-red images to detect the green apples within a tree canopy. The 'a' and 'b' channels of the 'Lab' colour space (Hunter, 1948) were clustered to segment the pixels for obtaining apple, background and leaf classes. Their approach showed poor performance due to unconstrained illumination in natural scenes (53.16% recognition accuracy) by using high-level visual features whereas low-level features showed 66.28% of recognition rate. However, the combination of high and low-level features reported the increased of accuracy of 74.37%. Yin et al. (2009) developed a harvesting robotic system to extract ripe tomato by using K-means clustering with Lab colour space (Table 5).

In addition to the weed mapping, disease/defects detection and sensing system for automated harvesting robots, K-means clustering technique was also used for classifying 17 different orange genotypes on the basis of their shape topologies (Costa et al., 2009), strawberry grading on basis of its shape (Liming and Yanchao, 2010), development of soil management zones on the basis electrical resistivity (Haghverdi et al., 2015), wheat row detection (Jiang et al., 2016) and early grapes yield estimation based on shoot detection (Liu et al., 2017). However, the adoption of K-means clustering algorithm for the realistic field applications in these areas still needs to be analysed in greater depth. The iterative non-convex optimization for moving the cluster's centre in K-means clustering algorithm is suitable for the data distributions supporting spherical clusters (Buchta et al., 2012). Furthermore, hard assignment of the data points is to move into a cluster as a part of centre initialization and its dependency on the cluster results are the major drawbacks of the algorithm (Bishop, 2007). These limitations can be overcome by using several enhancements like fuzzy k-means/ mixture models (soft assignment) and k-means++ as well as improved initialization (Arthur and Vassilvitskii, 2007) (Table 5).

2.2.2. Fuzzy clustering

Fuzzy clustering (also referred as soft clustering) enables each data point to have a probability of belonging to each (predefined) cluster, rather than just belonging to one cluster as it is the case in traditional k-means clustering algorithm (Bezdek et al., 1984). Fuzzy approach is suitable for the dataspaces which contains the data points that are somewhat in between the cluster centres or otherwise are ambiguous (Bishop, 2007). The distance function (Euclidean, Mahalanobis, Minkowski, etc.) between data point and cluster centre is replaced with measure of having a probability relative to the inverse of distance (Ghosh and Dubey, 2013). The data points are then assigned to a cluster that has maximum posterior probability or can be analysed as different probabilistic distributions rather than a fixed assignment (Boydell and McBratney, 2002; Ghosh and Dubey, 2013). The fuzzy clustering is commonly applied in agriculture for soil management zone delineation on the basis of soil and/or crop characteristics (Boydell and McBratney, 2002; Triantafilis and Lesch, 2005; Vrindts et al., 2005), vegetation segmentation or weed mapping (Meyer et al., 2004a and 2004b; Neto et al., 2006; Tellaeche et al., 2007; Guijarro et al., 2011; Solahudin et al., 2010; Romeo et al., 2013), disease detection (El-Helly et al., 2003; Majumdar et al., 2015; Mondal and Kole, 2016) or crop row detection (Romeo et al., 2012).

Boydell and McBratney (2002) delineated the temporally stable management zones based on cotton yield estimated from 11 consecutive years Landsat thematic mapper (TM) imager. The authors used modified fuzzy k-means algorithm (De Gruijter and McBratney, 1988) to generate yield maps. The results of study reported that 5 consecutive years of data provides temporally stable yield zones as indicated by the Kappa index of agreement (Carstensen Jr, 1987). In addition to the crop yield, soil compaction, normalized difference vegetation index (NDVI) and green index ($G - R/G + R$) along with fuzzy k-means was used to

Table 6
Overview of example applications of Fuzzy clustering algorithm in agriculture.

Crop	ML task	Data/sensor type	Application area	Classification accuracy	Reference
Barley	Classification	Colour camera	Weed detection	Highest classification accuracy of 91%	Tellaeche et al. (2007)
Barley and corn	Segmentation	Colour camera	Green vegetation identification from background and sky	Accuracy of 91.69% using combined colour approach	Guijarro et al. (2011)
Barley and Maize	Segmentation	Colour Camera	Vegetation segmentation from soil and residues	Different accuracies with different illumination conditions	Romeo et al. (2013)
Corn	Management zones delineation	Soil texture, organic matter, available phosphorus, pH, moisture content and electrical conductivity were measured using standard laboratory procedures, yield was measured manually	Site-specific crop management	Coincidence in the classification of 98% of the sampled grid points	Gili et al. (2017)
Corn and Peanut	Identification	Colour Camera	Weed density mapping	Accuracy for weed mapping was not reported	Solahudin et al. (2010)
Cotton	Temporally stable management zones delineation	Landsat TM imagery for 11 consecutive years	Yield estimation over the consecutive years	Five consecutive years of data provides temporally stable estimates of yield zones	Boydell and McBratney (2002)
Cucumber	Classification	Colour Camera	Three different disease detections on leaves	Not reported	El-Helly et al. (2003)
Grapevines	Management zones delineation	EC_a and soil depth using EM38, elevation using DGPS, NDVI using ACS-210 crop circle sensor, manually collected yield data and chemical properties using refractometer	Site-specific crop management based on yield and grapes chemical quality	EC_a and NDVI is highly related with yield and quality index of grapes	Tagarakis et al. (2013)
Maize	Identification	Colour camera	Row detection	Average accuracy of 97.32% for green vegetation identification using CRD	Romeo et al. (2012)
Soybean and Velvet leaf	Individual leaf extraction from young canopies	Colour camera	Weed detection	46% leaf extraction rate of field plants	Neto et al. (2006)
Sunflower, soybean, redroot pigweed and velvet	Classification	Colour camera	Weed detection	10 – 69% of correct classification in bare soil	Meyer et al. (2004a)
Wheat	Classification	Colour camera	4 different types of disease detection on leaves	56% for classification of different diseases	Majumdar et al. (2015)
Wheat	Identification	Colour camera	Disease detection on leaves	95% and 94% for diseased and non-diseased images	Mondal and Kole (2016)
Wheat and Cotton	Mapping clay content variation	EC_a using EM34/38 sensor	Soil resources management	Average measured clay contents were 47.38% and predicted were 47.71%	Triantafilis and Lesch (2005)
Wheat and Soybeans	Subfield management class delineation	EC_a using VERIS 3100 sensor, elevation, hardpan depth using hydraulic penetrometer and yield	Site-specific crop management	Spatial information with fuzzy k-means performed better than only fuzzy K-means	Córdoba et al. (2013)
Winter wheat	Management zones delineation	Soil compaction with sensor developed by Mouazen et al., 2003, spectral information measured with spectral line imager and monochrome camera with 480 to 850 nm range, yield measured grain mass flow sensor	Site-specific crop management based on the correlation between soil compaction, yield and crop spectral reflectance	Yield was limited for soil with dry bulk density of 1.6 Mg/m^3	Vrindts et al. (2005)

generate the management zones (Vrindts et al., 2005). The results indicated that soil compaction (dry bulk density > 1.6 Mg/m^3) resulted in reduced yield. Tagarakis et al. (2013) measured NDVI at five different stages during the grapevines growth, soil apparent electrical conductivity (EC_a), soil depth and field topography to develop the zones resulting in improved yield and grape chemical contents. The fuzzy clustering along with pixel-by-pixel comparison indicated that EC_a and NDVI are highly correlated with the yield and quality of grapes. However, the major drawback of fuzzy clustering approach is that it does not take the spatial correlation of the different data (i.e. crop, soil, yield etc.) variables into account and therefore may not show the mapping in contagious zoning (Córdoba et al., 2013). To address this issue, Dray et al., (2008) proposed a principal component analysis (PCA) based MULTISPATI-PCA algorithm to transform the input variables based on the spatial autocorrelation between them followed by fuzzy k-means on transformed data. The results indicated that proposed strategy showed smallest within class variance and highest yield difference between delineated classes (Table 6).

Meyer et al. (2004a) used fuzzy c-means and Gustafson-Kessel approach to segment sunflower, soybean, redroot pigweed and velvet leaf plants against bare clay soil, corn residue and wheat residue. The results indicated that the fuzzy based clustering approach was correctly segmented 10–69% of plants in bare soil but failed for plants with corn and wheat residue as background. Guijarro et al. (2011) classified the green plants (corn, barley and weeds) from background soil and sky (if present) using the combination of different colour indices followed by the threshold. The fuzzy clustering was used to cluster the textures with in same class i.e. differentiating the trees and their shades from the green crop pixels. The results indicated that the accuracy of segmenting green plants from background was around 91.69% using combined colour approach. Romeo et al. (2012) developed a crop row detection algorithm for the autonomous agricultural vehicles by removing the background pixels using the fuzzy approach. The crop rows were identified by using perspective projection by searching for the maximum accumulation of green pixels. The developed algorithm showed an average accuracy of around 97.32%, while the standard Hough transform based algorithm was able to correctly identify around 89% (on an average) crop lines. The other applications of fuzzy approach include plant disease detection (El-Helly et al., 2003; Majumdar et al., 2015; Mondal and Kole, 2016) (Table 6).

2.2.3. Gaussian mixture models

GMMs, Gaussian mixture models, similar to their fuzzy counterparts are suitable for the scenarios in which the data points overlap between true clusters (Bishop, 2007). However, GMMs tends to classify or cluster the spectral attributes of the objects in agronomic images by assuming that each class has its own normal distribution and the complete image is mixture of different Gaussians (Dempster et al., 1977). To cluster the scene elements, the mixture of different Gaussians can be fitted using an iterative expectation maximization (EM) algorithm (Dempster et al., 1977) thereby estimating the parameters (mean, variance and weights) that maximized the likelihood of observing the image data (Casanova et al., 2014; Naik et al., 2017; Park et al., 2017). GMM with EM expresses each data point as a weighted sum of k Gaussian distributions and assigns it to a cluster that maximizes the posterior probability (Reynolds, 2015). Alternatively, the mixture of different distributions can be separated by a predefined mean determined through the threshold-based image segmentation techniques (Casanova et al., 2014). GMM applications in agricultural sector includes vegetation segmentation/mapping (Nielsen et al., 2012; Bai et al., 2013; De Rainville et al., 2014), high throughput plant phenotyping (Kirk et al., 2009; Bauer et al., 2011; Rousseau et al., 2013; Naik et al., 2017; Park et al., 2017), variable rate irrigation (Casanova et al., 2014; Haghverdi et al., 2015) and yield estimation (Dey et al., 2012) (Table 7).

Bauer et al. (2011) differentiated infected (leaf spot pathogen and rust fungus) and healthy leaves of sugar beet by using multispectral (700–950 nm) and colour cameras. GMM followed by the adaptive Bayes classification was carried out to perform the pixelwise classification. The results indicated median of 94%, 91% and 86% for healthy leaf area, leaf spot disease and rust fungus, respectively. Rousseau et al. (2013) quantified the resistance of the green bean plants to the inoculated pathogens using chlorophyll fluorescence imaging. The maximum quantum yield of photosystem II photochemistry (F_v/F_m) was extracted followed by the GMM to cluster the pixels to 4 classes. Park et al. (2017) estimated the crop water stress index ($CWSI = T_{canopy} - T_{wet}/T_{dry} - T_{wet}$) in nectarine and peach orchards using thermal infrared camera attached to UAV. The GMM was fitted to temperature distribution clustering the temperature pixels of soil and canopy followed by calculation 'T_{wet}' and 'T_{dry}' thresholds for CWSI. The measured CWSI was cross-validated with stomatal conductance and stem water potential and coefficient of determination (R^2) were found to be 0.82 and 0.72, respectively. Naik et al. (2017) compared the performance of different supervised and unsupervised classifiers for identifying the iron deficiency chlorosis stress severity in soybean plants. The results showed that colour signatures (yellow and brown) along with GMM and Bayes classifier were able to achieve an accuracy of 99.1% on sub-set data and 99.4% on complete dataset (Table 7).

Tabb et al. (2006) modelled the background of the apple orchards (background subtraction from motion information) using global GMM to identify and locate the apple fruit from the video feed for an autonomous harvesting system. The algorithm correctly identified approximately 85–96% of yellow and red apples with a frame rate of 14–16 frames per second. Nielsen et al. (2012) used LIDAR and stereo vision for orchard tree mapping along with the tree height estimation. Individual trees in stereo feed were segmented from using the GMM and the measured height was correlated with manually measured ground truth. The study reported that R^2 for the height estimated from LIDAR was 0.83 and measured with stereo camera was 0.60. Dey et al. (2012) performed 3D reconstruction of the grapevines tree structures using the motion information. The reconstructed plant structures were classified using GMM to three semantic classes (berry, branch and leaves) and thereby used for the yield estimation. The results indicated an accuracy of 98% prior to grapes ripening and 96% during ripening (Table 7).

Bai et al. (2013) compared the performance of GMM with eight other techniques when applied on the CIE L*a*b (Bergman et al., 2005) colour space for segmenting the rice crop from background. The results indicated that the mean accuracies of segmentation were 85.8%, 79.2% and 84.4% on cloudy, overcast and sunny days, respectively. De Rainville et al. (2014) identified the intra-row weeds in soybean and corn fields using morphological characteristics of the weeds along with a NB classifier. To improve the classification accuracy, a GMM model was used to filter out the misclassified weeds as plants (soybean and corn). The authors reported that developed technique was able to achieve an accuracy of 94% for corn and soybean along with 85% correct identification of the weeds. Casanova et al. (2014) developed a wireless sensing system to detect the effect of fungus infection and water stresses using an Arduino Mega ADK board and CMUCAM4 camera. The GMM with EM was used to differentiate between soil and vegetation followed by the detection of healthy and stressed wheat plants. The results indicated that the mean hue value of infected and stressed plants is significantly different form the healthy, well-watered plants. GMM have been used and studied for the disease/ stress detection, tree/orchard mapping and vegetation segmentation. However, there is relatively less research towards the application of these methods for soil management zone delineation (Haghverdi et al., 2015) and fruit/vegetable/grain grading and sorting (Patil et al., 2016). Whilst the GMM consider the covariance between different input variables using covariance matrix, therefore it might be a good alternative to fuzzy clustering and/or MULTISPATI-PCA for generating more contagious soil management zones as indicated by Córdoba et al. (2013) (Table 7).

Table 7
Overview of example applications of Gaussians mixtures model algorithm in agriculture.

Crop	ML task	Data/sensor type	Application area	Classification accuracy	Reference
Apple	Identification	Video camera	Automated harvesting robot	Correctly classified approximately 85–96% for red and yellow apples	Tabb et al. (2006)
Arabidopsis	Quantification of plant growth, photosynthesis and lead temperature	Time-lapse video/ Colour camera, chlorophyll fluorescence, thermal imaging	High throughput phenotyping	Not determined	De Vylder et al. (2012)
Barley and Wheat	Quantification	Colour camera and LAI-2000	Leaf area index estimation	R^2 = 0.68–0.81○ for camera method	Kirk et al. (2009)
Corn and Soybean	Classification	Colour Camera	Weed detection	An accuracy of 94% for corn and soybean and 85% of weeds	De Rainville et al. (2014)
Cotton	Soil management zone delineation	VERIS 3100 electrical resistivity meter and Landsat 8 imagery	Variable rate irrigation	40% of soil available water content variance explained	Haghverdi et al. (2015)
Grapevines	Fine-scale plant structure in 3D point clouds from the motion information	Colour camera	Early crop yield estimation	98% prior to ripening and 96% during ripening	Dey et al.,
Green beans	Plant resistance quantification	Chlorophyll fluorescence imaging	High throughput phenotyping of plant resistance to pathogens	Not reported	Rousseau et al. (2013)
Nectarine and Peach	Quantification	Thermal infrared camera (Stomatal conductivity using LI-6400 sensor and handheld infrared thermometer, stem water potential using Scholander pressure chamber for cross validation)	Estimation of crop water stress	R^2 of 0.72 and 0.82 with stem water potential and stomatal conductivity, respectively	Park et al. (2017)
Peach Orchard	Height quantification	Stereo colour camera and LIDAR	Orchard and tree mapping	R^2 = 0.83 using LIDAR and 0.60 using stereo vision	Nielsen et al. (2012)
Rice	Segmentation	Colour camera	Vegetation segmentation from non-vegetative portions	85.8%, 79.2%, 84.4% for cloudy, overcast and sunny days, respectively	Bai et al. (2013)
Soybean	Classification	RAW format/ Colour camera	Iron deficiency chlorosis stress severity	~99% accuracy	Naik et al. (2017)
Sugar beet	Classification	Multi-spectral and colour stereo images	Two types of disease detection on leaves	94% for healthy, 91% for leaf spot and 86% for rust fungus	Bauer et al. (2011)
Walnut	Classification	Hyperspectral fluorescence imagery	Separation of walnut shell and meet	Overall classification accuracy of 95.6%	Jiang et al. (2007)
Winter wheat	Segmentation of diseased and diseased free crop from background	Colour Camera (CAMUCAM4 for Arduino Mega ADK)	Disease detection and variable rate irrigation	Mean vegetation hue significantly affected by water and virus stress	Casanova et al. (2014)

2.2.4. Other unsupervised machine learning algorithms

In addition to previously introduced unsupervised statistical machine learning algorithms for agricultural machine vision systems, other algorithms including principal component analysis (PCA) for feature selection, hidden Markov model (HMM), and association analysis with Apriori and FP-growth (frequent pattern growth) algorithm could also be used in agricultural machine learning. Educated terrain classifier with contact features achieved 85.1% accuracy and all the features in the expanded space was combined to achieve an accuracy of 89.1% (Reina et al., 2017). PCA can be used as a pre-processing tool for extracting linear combinations of the most relevant features extracted from the different agronomic images and using them for further ML tasks. The PCA was found to be very commonly used for selecting the specific wavelengths (features) from high dimensional hyperspectral imaging for plant phenotypic applications (Singh et al., 2016; Pandey et al., 2017; Liang et al., 2017). The HMM, on the other hand, is an extension of the Markov model by adding hidden conditions and directly observable observations (Blunsom, 2004). Leite et al. (2008) used HMM to classify the different crops by using their spectral features along the complete crop cycle satellite imagery. The results of the study indicated that the HMM method was able to achieve an overall accuracy of 86%.

Apriori is a very simple algorithm, it takes a lot of time to find a specific pattern by repeatedly scanning the data (Kumar and Rukmani, 2010). FP-growth is one of the fastest and most popular algorithms (Borgelt, 2005). It is the process of storing information in what is called the FP-tree and removing the unused data in each. The advantage of FP-growth is that it requires only two times of scanning. The first time compresses a large database into a small frequent-pattern tree and second time derives an efficient FP-tree-based frequent pattern, FP-growth (Kumar and Rukmani, 2010). Lu et al. (2018) asserted detecting immature fruit before harvesting helps to increase yield and profits. Using only texture and intensity distribution, method of finding green colour fruit from the image of the tree was proposed.

2.3. Reinforcement machine learning algorithms

Machine learning is largely categorized as supervised learning and unsupervised learning. However, reinforcement learning focused on interaction-to-goal-oriented learning that can learn the behaviour through interaction (Sutton and Barto, 1998). A very demanding promising application of the reinforcement learning could be the automated agricultural robots/intelligent machines as it can be used to teach robots to improve their behaviour according to the relationship between themselves and the surrounding environment over the time. The robot can be used in agriculture for planting the crops precisely in a row with more uniform plant size. The reinforcement learning together with the robotic technology/ intelligent machines can fit very well to trend that is decreasing numbers of farmers and increasing crop production (Bechar and Vigneault, 2016).

3. Conclusion

ML technology has the potential to become very important to the agricultural machine vision system. The use of ML technology for weed detection, plant diseases and stress detection, yield prediction and estimation, plant water content determination, grading and sorting, soil analysis and real-time field operations may become routine operation in near future agriculture. Advancement of ML with machine vision will make agricultural technologies accurate, robust and low cost. A machine vision could come up with image acquisition and processing that might need a discriminator to classify desire target with high dimensional data. Convincing ML techniques combine a suitable feature extraction and selection procedures with appropriate prediction algorithms. The potential applications of ML approaches are highly depending on appropriate application of ML algorithms to specific fields of cropping system. Based on current research dynamics in ML technologies and methods for agricultural machine vision system feature analysis, the following trends for the future of feature data analysis in precision agricultural systems are expected.

- The NB based ML algorithms is suitable for the tasks with strong prior knowledge of the data distribution and level of the correlation between different features. It was observed that the NB algorithm did not perform well for the combination of different features. This might be because this algorithm is not able to learn the interaction because of the correlation between the features. However, careful feature selection (with PCA or other feature selection procedures) could help to achieve the conditional independence of the features. In summary of reviewed literature for NB, 47% of papers were found to be related with the agricultural product grading, 29% were used for crop disease detection, 18% were used for weed detection and 6% were used for the detection of nutrient deficiency in plants (Table 1).
- The DA like NB is also suitable for known data distributions; however, these types of algorithms are able to learn the interaction between the different features. Generally, non-linear algorithms (QDA, KDA) performed better than the LDA and NB. In summary of reviewed literature for DA, 50% of papers were used for solving the problems related with the agricultural product grading, 42% were used for weed detection and 8% were used for crop disease detection (Table 2).
- For the applications with little or no knowledge of the data distribution, the kNN can be used as it doesn't make any assumption about the data. It should be noted that kNN usually take large processing time. The reviewed literature indicated that kNN perform better than the NB for the combination of different features. In summary of reviewed literature for kNN, 47% of papers were found to be used for solving the problems related with the agricultural product grading, 21% were used for crop disease detection, 16% were used for weed detection and the rest were used for nutrient deficiency, land classification etc. (Table 3).
- Despite having some limitations of SVMs in training and testing, the classifier has a good potential in plant diseases detection and weed identification. Especially in row crop cultivation, SVMs based machine learning classifiers are very effective in field conditions and the place where the data are skewed. In summary of reviewed literature for SVM, 70% of papers were used for crop disease detection, 18% were used for solving the problems related with the agricultural product grading and 12% were used for weed detection (Table 4).
- Among, Unsupervised ML algorithms, K-means clustering can be used when the data points are clearly separable and thus the points can be classified to one of the class. The K-means algorithm is not suitable if the classes have some overlap. In summary of reviewed literature for K-means clustering, 39% of papers were used for solving the problems related with the agricultural product grading, 21% were used for crop disease detection, 18% were used for weed detection, 11% were used for automated robotic applications and the rest were used for yield estimation, crop row detection, tree mapping, variable rate (VR) irrigation and crop genotypic classification (Table 5).
- The fuzzy and GMM based algorithms are suitable for the cases where the data points overlap between different classes. These are good alternative to the K-means clustering for classifying the ambiguous points. The fuzzy clustering doesn't take any advantage of the correlation between different variables, therefore not suitable for spatially correlating variables. For these situations with correlation use GMM. GMM can be used for the cases where the prior information about the data distribution is known and each class is normally distributed. Fuzzy and GMM based algorithms were not found to be very common algorithms for the agricultural product

grading. In summary of reviewed literature for fuzzy clustering, 37% were used for weed detection, 31% were used for the site-specific crop management, 19% were used for crop disease detection and the rest were used for the crop row detection and yield estimation (Table 6). In summary of reviewed literature for GMM, 29% were used for plant phenotyping technologies, 14% were used for crop disease detection 14% were used for the VR irrigation and the rest were used for yield/stress estimation, weed detection, tree mapping and automated robotic applications (Table 7).

The above suggestions are based on statistical ML algorithms for machine vision application. The future application of ML technology will be spread in agriculture. Thus, present study suggests, but are not limited to, NB, DA, kNN, SVM and K-means clustering for agricultural product grading, crop disease detection and weed detection, Fuzzy clustering for site-specific crop management and soil analysis and GMM for plant phenotyping/stress detection and water content measurement. Although some supervised and unsupervised ML algorithms including ANNs and deep learning other than statistical ML have also proved to be potential in agricultural crop production, however current study could not cover up all these techniques.

Acknowledgements

This work was supported by Nova Scotia Research and Innovation Graduate Scholarship Program and Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grants Program (RGPIN-2017-05815).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2018.12.006.

References

Acharyya, M., Kundu, M.K., 2001. Adaptive basis selection for multi texture segmentation by M-band wavelet packet frames. In: Image Processing, 2001. Proceedings. 2001 International Conference on, vol. 2. IEEE, pp. 622–625.
Ahmad, I., Siddiqi, M.H., Fatima, I., Lee, S., Lee, Y.K., 2011. Weed classification based on Haar wavelet transform via k-nearest neighbor (k-NN) for real-time automatic sprayer control system. In: Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication. ACM, pp. 17.
Ahmed, F., Al-Mamun, H.A., Bari, A.H., Hossain, E., Kwan, P., 2012. Classification of crops and weeds from digital images: a support vector machine approach. Crop Prot. 40, 98–104.
Al Bashish, D., Braik, M., Bani-Ahmad, S., 2011. Detection and classification of leaf diseases using K-means-based segmentation and Neural-network based classification. Inform. Technol. J. 10 (2), 267–275.
Arivazhagan, S., Shebiah, R.N., Ananthi, S., Varthini, S.V., 2013. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agric. Eng. Int.: CIGR J. 15 (1), 211–217.
Arthur, D., Vassilvitskii, S., 2007. k-means++: the advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp. 1027–1035.
Asadollahi, H., Kamarposhty, M.S., Teymoori, M.M., 2009. Classification and evaluation of tomato images using several classifier. In: Paper presented at the Computer Science and Information Technology-Spring Conference, 2009. IACSITSC'09. International Association of, pp. 471–474.
Åstrand, B., Baerveldt, A., 2002. An agricultural mobile robot with vision-based perception for mechanical weed control. Autonomous Robots 13 (1), 21–35.
Athani, S.S., Tejeshwar, C., 2017. Support vector machine-based classification scheme of maize crop. In: Advance Computing Conference (IACC), 2017 IEEE 7th International. IEEE, pp. 84–88.
Bai, X., Cao, Z., Wang, Y., Yu, Z., Zhang, X., Li, C., 2013. Crop segmentation from images by morphology modeling in the CIE L* a* b* color space. Comput. Electron. Agric. 99, 21–34.
Bandi, S.R., Varadharajan, A., Chinnasamy, A., 2013. Performance evaluation of various statistical classifiers in detecting the diseased citrus leaves. Int. J. Eng. Sci. Technol. 5 (2), 298–307.
Bauer, S.D., Korč, F., Förstner, W., 2011. The potential of automatic methods of classification to identify leaf diseases from multispectral images. Precis. Agric. 12 (3), 361–377.
Bechar, A., Vigneault, C., 2016. Agricultural robots for field operations: concepts and components. Biosyst. Eng. 149, 94–111.
Behmann, J., Mahlein, A.K., Rumpf, T., Römer, C., Plümer, L., 2015. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Precis. Agric. 16 (3), 239–260.
Bergman, L., Verikas, A., Bacauskiene, M., 2005. Unsupervised colour image segmentation applied to printing quality assessment. Image Vis. Comput. 23 (4), 417–425.
Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. 10 (2–3), 191–203.
Bishop, C., 2007. Pattern Recognition and Machine Learning (Information Science and Statistics). 2006. corr. 2nd printing edn.. first ed. Springer, New York.
Blasco, J., Aleixos, N., Gómez, J., Moltó, E., 2007. Citrus sorting by identification of the most common defects using multispectral computer vision. J. Food Eng. 83 (3), 384–393.
Blasco, J., Aleixos, N., Cubero, S., Gómez-Sanchís, J., Moltó, E., 2009a. Automatic sorting of satsuma (*Citrus unshiu*) segments using computer vision and morphological features. Comput. Electron. Agric. 66 (1), 1–8.
Blasco, J., Cubero, S., Gómez-Sanchís, J., Mira, P., Moltó, E., 2009b. Development of a machine for the automatic sorting of pomegranate *(Punica granatum)* arils based on computer vision. J. Food Eng. 90 (1), 27–34.
Blunsom, P. (2004). Hidden Markov Models. Lecture notes, August, 15(18-19), 48.
Borgelt, C., 2005. An Implementation of the FP-growth algorithm. In: Proceedings of the 1st international workshop on open source data mining: frequent pattern mining implementations. ACM, pp. 1–5.
Boydell, B., McBratney, A., 2002. Identifying potential within-field management zones from cotton-yield estimates. Precis. Agric. 3 (1), 9–23.
Buchta, C., Kober, M., Feinerer, I., Hornik, K., 2012. Spherical k-means clustering. J. Stat. Softw. 50 (10), 1–22.
Bulanon, D.M., Kataoka, T., Okamoto, H., Hata, S.I., 2004. Development of a real-time machine vision system for the apple harvesting robot. August In: SICE 2004 Annual Conference. IEEE, pp. 595–598.
Burges, C.J., 1998. A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. 2 (2), 121–167.
Burks, T., Shearer, S., Payne, F., 2000. Classification of weed species using color texture features and discriminant analysis. Trans. ASAE 43 (2), 441.
Caglayan, A., Guclu, O., Can, A.B., 2013. A plant recognition approach using shape and color features in leaf images. In: International Conference on Image Analysis and Processing. Springer, Berlin, Heidelberg, pp. 161–170.
Camargo, A., Smith, J.S., 2009. Image pattern classification for the identification of disease causing agents in plants. Comput. Electron. Agric. 66 (2), 121–125.
Carroll, J., Holden, N., 2005. A method to quantify weed distribution for relating to patch spraying systems. Trans. ASAE 48 (1), 27–35.
Carstensen Jr, L.W., 1987. A measure of similarity for cellular maps. Am. Cartogr. 14 (4), 345–358.
Casanova, J.J., O'Shaughnessy, S.A., Evett, S.R., Rush, C.M., 2014. Development of a wireless computer vision instrument to detect biotic stress in wheat. Sensors 14 (9), 17753–17769.
Cen, H., Lu, R., Zhu, Q., Mendoza, F., 2016. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. Postharvest Biol. Technol. 111, 352–361.
Chang, Y.K., Zaman, Q.U., Schumann, A.W., Percival, D.C., Esau, T.J., Ayalew, G., 2012. Development of color co-occurrence matrix based machine vision algorithms for wild blueberry fields. Appl. Eng. Agric. 28 (3), 315–323.
Chang, Y.K., Rehman, T., 2018. Current and future applications of Cost-Effective Smart Cameras in Agriculture. In: Robotics and Mechatronics in Agriculture. CRC Press, Taylor and Francis Group, USA, pp. 75–121.
Chen, X., Xun, Y., Li, W., Zhang, J., 2010. Combining discriminant analysis and neural networks for corn variety identification. Comput. Electron. Agric. 71, S48–S53.
Chowdhury, S., Verma, B., Stockwell, D., 2015. A novel texture feature based multiple classifier technique for roadside vegetation classification. Expert Syst. Appl. 42 (12), 5047–5055.
Chung, C.L., Huang, K.J., Chen, S.Y., Lai, M.H., Chen, Y.C., Kuo, Y.F., 2016. Detecting Bakanae disease in rice seedlings by machine vision. Comput. Electron. Agric. 121, 404–411.
Conners, R.W., Trivedi, M.M., Harlow, C.A., 1984. Segmentation of a high-resolution urban scene using texture operators. Computer Vis. Graph. Image Process. 25 (3), 273–310.
Córdoba, M., Bruno, C., Costa, J., Balzarini, M., 2013. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. Comput. Electron. Agric. 97, 6–14.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learn. 20 (3), 273–297.
Costa, C., Menesatti, P., Paglia, G., Pallottino, F., Aguzzi, J., Rimatori, V., Recupero, G.R., 2009. Quantitative evaluation of tarocco sweet orange fruit shape using optoelectronic elliptic fourier based analysis. Postharvest Biol. Technol. 54 (1), 38–47.
Criminisi, A., Shotton, J., Konukoglu, E., 2012. Decision forests: a unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. Found. Trends® Comput. Graph. Vis. 7 (2–3), 81–227.
Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., Blasco, J., 2011. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food Bioprocess Technol. 4 (4), 487–504.
De Gruijter, J. J., McBratney, A. B. (1988). A Modified Fuzzy k-Means Method for Predictive Classification.
De Rainville, F., Durand, A., Fortin, F., Tanguy, K., Maldague, X., Panneton, B., Simard, M., 2014. Bayesian classification and unsupervised learning for isolating weeds in row crops. Pattern Anal. Appl. 17 (2), 401–414.
De Vylder, J., Vandenbussche, F., Hu, Y., Philips, W., Van Der Straeten, D., 2012. Rosette tracker: an open source image analysis tool for automatic quantification of genotype effects. Plant Physiol. 160 (3), 1149–1159.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. B(methodological) 1–38.
Deshpande, T., Sengupta, S., Raghuvanshi, K., 2014. Grading & identification of disease in pomegranate leaf and fruit. Int. J. Comput. Sci. Inform. Technol. 5 (3), 4638–4645.
Dey, D., Mummert, L., Sukthankar, R., 2012. Classification of plant structures from uncalibrated image sequences. In: Paper Presented at the Applications of Computer Vision (WACV), 2012 IEEE Workshop on, pp. 329–336.
Dingle Robertson, L., King, D.J., 2011. Comparison of pixel-and object-based classification in land cover change mapping. Int. J. Remote Sens. 32 (6), 1505–1529.
Dray, S., Saïd, S., Débias, F., 2008. Spatial ordination of vegetation data using a generalization of Wartenberg's multivariate spatial correlation. J. Veg. Sci. 19 (1), 45–56.
Du, C.J., Sun, D.W., 2006. Learning techniques used in computer vision for food quality evaluation: a review. J. Food Eng. 72 (1), 39–55.
Dubey, S.R., Jalal, A.S. (2014). Adapted approach for fruit disease identification using images. ArXiv Preprint arXiv:1405.4930.
Elferink, M., Schierhorn, F. 2016. Global Demand for Food is Rising. Harvard Business Review April 07.
El-Helly, M., Onsi, H., Rafea, A., El Gamal, S., 2003. Detecting leaf spots on cucumber crop using fuzzy clustering algorithm. In: 11th International Conference on AI, pp. 102–107.
FAO, (2009). Global agriculture towards 2050: high-level expert forum on how to feed the world in 2050, 12-13 Oct 2009 by Food and Agriculture Organization of the United Nations. https://reliefweb.int/report/world/global-agriculture-towards-2050-high-level-expert-forum-how-feed-world-2050-12-13-oct.
FAO, (2017). The future of food and agriculture: trends and challenges by Food and Agriculture Organization of the United Nations. https://reliefweb.int/sites/reliefweb.int/files/resources/a-i6583e.pdf.
Frank, E., Trigg, L., Holmes, G., Witten, I.H., 2000. Naive bayes for regression. Machine Learn. 41 (1), 5–25.
Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I.H., Trigg, L., 2009. Weka-a machine learning workbench for data mining. In: Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA, pp. 1269–1277.
Gao, W., Sammes, N.M., 1999. An Introduction to Electronic and Ionic Materials. World Scientific Publishing Company.
Garner, S.R., Cunningham, S.J., Holmes, G., Nevill-Manning, C.G., Witten, I.H., 1995. Applying a machine learning workbench: experience with agricultural databases. In: Proc Machine Learning in Practice Workshop, Machine Learning Conference, Tahoe City, CA, USA, pp. 14–21.
Gebhardt, S., Schellberg, J., Lock, R., Kühbauch, W., 2006. Identification of broad-leaved dock (Rumex obtusifolius L.) on grassland by means of digital image processing. Precis. Agric. 7 (3), 165–178.
Ghosh, S., Dubey, S.K., 2013. Comparative analysis of k-means and fuzzy c-means algorithms. Int. J. Adv. Comput. Sci. Appl. 4 (4), 35–39.
Gili, A., Álvarez, C., Bagnato, R., Noellemeyer, E., 2017. Comparison of three methods for delineating management zones for site-specific crop management. Comput. Electron. Agric. 139, 213–223.
Gómez-Sanchis, J., Gómez-Chova, L., Aleixos, N., Camps-Valls, G., Montesinos-Herrero, C., Moltó, E., Blasco, J., 2008. Hyperspectral system for early detection of rottenness caused by penicilliumdigitatum in mandarins. J. Food Eng. 89 (1), 80–86.
Gowen, A.A., Taghizadeh, M., O'Donnell, C.P., 2009. Identification of mushrooms subjected to freeze damage using hyperspectral imaging. J. Food Eng. 93 (1), 7–12.
Granitto, P.M., Navone, H.D., Verdes, P.F., Ceccatto, H., 2002. Weed seeds identification by machine vision. Comput. Electron. Agric. 33 (2), 91–103.
Granitto, P.M., Verdes, P.F., Ceccatto, H.A., 2005. Large-scale investigation of weed seed identification by machine vision. Comput. Electron. Agric. 47 (1), 15–24.
Guerrero, J.M., Pajares, G., Montalvo, M., Romeo, J., Guijarro, M., 2012. Support vector machines for crop/weeds identification in maize fields. Expert Syst. Appl. 39 (12), 11149–11155.
Guevara-Hernandez, F., Gomez-Gil, J., 2011. A machine vision system for classification of wheat and barley grain kernels. Spanish J. Agric. Res. 9 (3), 672–680.
Guijarro, M., Pajares, G., Riomoros, I., Herrera, P., Burgos-Artizzu, X., Ribeiro, A., 2011. Automatic segmentation of relevant textures in agricultural images. Comput. Electron. Agric. 75 (1), 75–83.
Guru, D., Mallikarjuna, P., Manjunath, S., Shenoi, M., 2012. Machine vision based classification of tobacco leaves for automatic harvesting. Intell. Autom. Soft Comput. 18 (5), 581–590.
Haghverdi, A., Leib, B.G., Washington-Allen, R.A., Ayers, P.D., Buschermohle, M.J., 2015. Perspectives on delineating management zones for variable rate irrigation. Comput. Electron. Agric. 117, 154–167.
Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybernet. 6, 610–621.
Ho, S.Y., Lee, K.Z., 2003. Design and analysis of an efficient evolutionary image segmentation algorithm. J. VLSI Signal Process. Syst. Signal Image Video Technol. 35 (1), 29–42.
Hu, M., 1962. Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory 8 (2), 179–187.
Hu, M., Dong, Q., Liu, B., Malakar, P.K., 2014. The potential of double K-means clustering for banana image segmentation. J. Food Process Eng 37 (1), 10–18.
Hunter, R.S., 1948. Photoelectric color-difference meter. JOSA 38 (7), 661 In Proceedings of the Winter Meeting of the Optical Society of America.
Islam, M., Dinh, A., Wahid, K., Bhowmik, P., 2017. Detection of potato diseases using image segmentation and multiclass support vector machine. In: Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on. IEEE, pp. 1–4.
Jaffar, A., Jaafar, R., Jamil, N., Low, C.Y., Abdullah, B., 2009. Photogrammetric grading of oil palm fresh fruit bunches. Int. J. Mech. Mechatron. Eng. 9, 18–24.
Jiang, L., Zhu, B., Jing, H., Chen, X., Rao, X., Tao, Y., 2007. Gaussian mixture model-based walnut shell and meat classification in hyperspectral fluorescence imagery. Trans. ASABE 50 (1), 153–160.
Jiang, G., Wang, X., Wang, Z., Liu, H., 2016. Wheat rows detection at the early growth stage based on hough transform and vanishing point. Comput. Electron. Agric. 123, 211–223.
Karimi, Y., Prasher, S.O., Patel, R.M., Kim, S.H., 2006. Application of support vector machine technology for weed and nitrogen stress detection in corn. Comput. Electron. Agric. 51 (1–2), 99–109.
Kaur, K., 2016. Machine learning: applications in Indian agriculture. Int. J. Adv. Res. Comput. Commun. Eng. 5 (4).
Kazmi, W., Garcia-Ruiz, F.J., Nielsen, J., Rasmussen, J., Andersen, H.J., 2015. Detecting creeping thistle in sugar beet fields using vegetation indices. Comput. Electron. Agric. 112, 10–19.
Kirk, K., Andersen, H.J., Thomsen, A.G., Jørgensen, J.R., Jørgensen, R.N., 2009. Estimation of leaf area index in cereal crops using red–green images. Biosyst. Eng. 104 (3), 308–317.
Klecka, W.R., 1980. Discriminant Analysis. Sage Printing Press, London, UK.
Kruse, O.M.O., Prats-Montalbán, J.M., Indahl, U.G., Kvaal, K., Ferrer, A., Futsaether, C.M., 2014. Pixel classification methods for identifying and quantifying leaf surface injury from digital images. Comput. Electron. Agric. 108, 155–165.
Kumar, B.S., Rukmani, K.V., 2010. Implementation of web usage mining using APRIORI and FP growth algorithms. Int. J. Adv. Netw. Appl. 1 (06), 400–404.
Kurtulmus, F., Lee, W.S., Vardar, A., 2011. Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. Comput. Electron. Agric. 78 (2), 140–149.
Kurtulmus, F., Lee, W.S., Vardar, A., 2014. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. Precis. Agric. 15 (1), 57–79.
Kurtulmuş, F., Ünal, H., 2015. Discriminating rapeseed varieties using computer vision and machine learning. Expert Syst. Appl. 42 (4), 1880–1891.
Lachenbruch, P.A., 1975. Discriminant Analysis. Macmillan Publishing Co, New York, NY, USA.
Laursen, M.S., Midtiby, H.S., Krüger, N., Jørgensen, R.N., 2014. Statistics-based segmentation using a continuous-scale naive bayes approach. Comput. Electron. Agric. 109, 271–277.
Lee, W.S., Slaughter, D., Giles, D., 1999. Robotic weed control system for tomatoes. Precis. Agric. 1 (1), 95–113.
Leemans, V., Destain, M., 2004. A real-time grading method of apples based on features extracted from defects. J. Food Eng. 61 (1), 83–89.
Leite, P.B.C., Feitosa, R.Q., Formaggio, A.R., Da Costa, G.A.O.P., Pakzad, K., IDA, S., 2008. Hidden Markov models applied in agricultural crops classification. Proceeding of GEOBIA (GEOgraphic Object-Based Image Analysis for the 21St Century).
Li, D., Yang, W., Wang, S., 2010. Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine. Comput. Electron. Agric. 74 (2), 274–279.
Li, H., Lee, W.S., Wang, K., 2014. Identifying blueberry fruit of different growth stages using natural outdoor color images. Comput. Electron. Agric. 106, 91–101.
Liming, X., Yanchao, Z., 2010. Automated strawberry grading system based on image processing. Comput. Electron. Agric. 71, S32–S39.
Liang, Z., Pandey, P., Stoerger, V., Xu, Y., Qiu, Y., Ge, Y., Schnable, J.C., 2017. Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. GigaScience 7 (2), gix117.
Liu, S., Cossell, S., Tang, J., Dunn, G., Whitty, M., 2017. A computer vision system for early stage grape yield estimation based on shoot detection. Comput. Electron. Agric. 137, 88–101.
Lu, J., Lee, W.S., Gan, H., Hu, X., 2018. Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis. Biosyst. Eng. 171, 78–90.
Lutman, P.J.W., Rew, L.J., Cussans, G.W., Miller, P.C.H., Paice, M.E.R., Stafford, J.E. (1998). Development of a'patch spraying'system to control weeds in winter wheat. Development of a'patch spraying'system to control weeds in winter wheat. 158.
Majumdar, D., Ghosh, A., Kole, D.K., Chakraborty, A., Majumder, D.D., 2015. Application of fuzzy C-means clustering method to classify wheat leaf images based on the presence of rust disease. In: Paper presented at the Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014, pp. 277–284.
Majumdar, S., Jayas, D., 2000a. Classification of cereal grains using machine vision: II. color models. Trans. ASAE 43 (6), 1677.
Majumdar, S., Jayas, D., 2000b. Classification of cereal grains using machine vision: I. morphology models. Trans. ASAE 43 (6), 1669.
Majumdar, S., Jayas, D., 2000c. Classification of cereal grains using machine vision: III. texture models. Trans. ASAE 43 (6), 1681.
Makky, M., Soni, P., 2013. Development of an automatic grading machine for oil palm fresh fruits bunches (FFBs) based on machine vision. Comput. Electron. Agric. 93, 129–139.
Mery, D., Pedreschi, F., 2005. Segmentation of colour food images using a robust algorithm. J. Food Eng. 66 (3), 353–360.
Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K., Srikaew, A., 2008. Grape leaf disease detection from color imagery using hybrid intelligent system. In: Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on, vol. 1. IEEE, pp. 513–516.
Meyer, G., Mehta, T., Kocher, M., Mortensen, D., Samal, A., 1998. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. Trans. ASAE 41 (4), 1189.
Meyer, G.E., Neto, J.C., Jones, D.D., Hindman, T.W., 2004a. Intensified fuzzy clusters for

classifying plant, soil, and residue regions of interest from color images. Comput. Electron. Agric. 42 (3), 161–180.
Meyer, G., Hindman, T., Jones, D., Mortensen, D., 2004b. Digital camera operation and fuzzy logic classification of uniform plant, soil, and residue color images. Appl. Eng. Agric. 20 (4), 519.
Miller, W., Throop, J., Upchurch, B., 1998. Pattern recognition models for spectral reflectance evaluation of apple blemishes. Postharvest Biol. Technol. 14 (1), 11–20.
Mishra, A., Karimi, D., Ehsani, R., Albrigo, L.G., 2011. Evaluation of an active optical sensor for detection of huanglongbing (HLB) disease. Biosyst. Eng. 110 (3), 302–309.
Mitchell, T.M., 1997. Machine Learning. WCB.
Mizushima, A., Lu, R., 2013. An image segmentation method for apple sorting and grading using support vector machine and Otsu's method. Comput. Electron. Agric. 94, 29–37.
Mokhtar, U., El Bendary, N., Hassenian, A.E., Emary, E., Mahmoud, M.A., Hefny, H., Tolba, M.F., 2015. SVM-based detection of tomato leaves diseases. In: Intelligent Systems' 2014. Springer, Cham, pp. 641–652.
Mondal, D., Kole, D.K., 2016. A time efficient leaf rust disease detection technique of wheat leaf images using pearson correlation coefficient and rough fuzzy C-means. In: Information Systems Design and Intelligent Applications. Springer, pp. 609–618.
Mondal, D., Kole, D.K., Roy, K., 2017. Gradation of yellow mosaic virus disease of okra and bitter gourd based on entropy based binning and naive bayes classifier after identification of leaves. Comput. Electron. Agric. 142, 485–493.
Morrison, D.F., 1976. Multivariate Statistical Methods, second ed. McGraw-Hill, New York, NY.
Mouazen, A.M., Dumont, K., Maertens, K., Ramon, H., 2003. Two-dimensional prediction of spatial variation in topsoil compaction of a sandy loam field-based on measured horizontal force of compaction sensor, cutting depth and moisture content. Soil Tillage Res. 74 (1), 91–102.
Munera, S., Besada, C., Aleixos, N., Talens, P., Salvador, A., Sun, D.W., Blasco, J., 2017. Non-destructive assessment of the internal quality of intact persimmon using colour and VIS/NIR hyperspectral imaging. LWT-Food Sci. Technol. 77, 241–248.
Mursalin, M., Mesbah-Ul-Awal, M., 2014. Towards classification of weeds through digital image. In: Paper presented at the Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on, pp. 1–4.
Naik, H.S., Zhang, J., Lofquist, A., Assefa, T., Sarkar, S., Ackerman, D., Ganapathysubramanian, B., 2017. A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. Plant Methods 13 (1), 23.
Neto, J.C., Meyer, G.E., Jones, D.D., 2006. Individual leaf extractions from young canopy images using Gustafson-Kessel clustering and a genetic algorithm. Comput. Electron. Agric. 51 (1), 66–85.
Nielsen, M., Slaughter, D.C., Gliever, C., Upadhyaya, S., 2012. Orchard and tree mapping and description using stereo vision and lidar. Paper Presented at the International Conference of Agricultural Engineering.
Nieuwenhuizen, A., Tang, L., Hofstee, J., Müller, J., Van Henten, E., 2007. Colour based detection of volunteer potatoes as weeds in sugar beet fields using machine vision. Precis. Agric. 8 (6), 267–278.
Okamoto, H., Murata, T., Kataoka, T., Hata, S., 2007. Plant classification for weed detection using hyperspectral imaging with wavelet analysis. Weed Biol. Manage. 7 (1), 31–37.
Olgun, M., Onarcan, A.O., Özkan, K., Işik, Ş., Sezer, O., Özgişi, K., Koyuncu, O., 2016. Wheat grain classification by using dense SIFT features with SVM classifier. Comput. Electron. Agric. 122, 185–190.
Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N.B., Nasir, M.H.N.M., 2014. Potential of radial basis function-based support vector regression for apple disease detection. Measurement 55, 512–519.
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybernet. 9 (1), 62–66.
Pan, L., Sun, Y., Xiao, H., Gu, X., Hu, P., Wei, Y., Tu, K., 2017. Hyperspectral imaging with different illumination patterns for the hollowness classification of white radish. Postharvest Biol. Technol. 126, 40–49.
Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Front. Plant Sci. 8, 1348.
Park, S., Ryu, D., Fuentes, S., Chung, H., Hernández-Montes, E., O'Connell, M., 2017. Adaptive estimation of crop water stress in nectarine and peach orchards using high-resolution imagery from an unmanned aerial vehicle (UAV). Remote Sens. 9 (8), 828.
Patil, K., Kadam, S., Kale, S., Rachetti, Y., Jagtap, K., Inamdar, K. (2016). Machine Vision Based Autonomous Fruit Inspection and Sorting.
Pereira, L.A., Nakamura, R.Y., De Souza, G.F., Martins, D., Papa, J.P., 2012. Aquatic weed automatic classification using machine learning techniques. Comput. Electron. Agric. 87, 56–63.
Perez, A., Lopez, F., Benlloch, J., Christensen, S., 2000. Colour and shape analysis techniques for weed detection in cereal fields. Comput. Electron. Agric. 25 (3), 197–212.
Pérez-Ortiz, M., Peña, J.M., Gutiérrez, P.A., Torres-Sánchez, J., Hervás-Martínez, C., López-Granados, F., 2016. Selecting patterns and features for between-and within-crop-row weed mapping using UAV-imagery. Expert Syst. Appl. 47, 85–94.
Phadikar, S., Sil, J., Das, A.K., 2013. Rice diseases classification using feature selection and rule generation techniques. Comput. Electron. Agric. 90, 76–85.
Pham, V.H., Lee, B.R., 2015. An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm. Vietnam J. Comput. Sci. 2 (1), 25–33.
Piron, A., Leemans, V., Kleynen, O., Lebeau, F., Destain, M., 2008. Selection of the most efficient wavelength bands for discriminating weeds from crop. Comput. Electron. Agric. 62 (2), 141–148.
Piron, A., Leemans, V., Lebeau, F., Destain, M., 2009. Improving in-row weed detection in multispectral stereoscopic images. Comput. Electron. Agric. 69 (1), 73–79.
Puerto, D.A., Gila, D.M.M., García, J.G., Ortega, J.G., 2015. Sorting olive batches for the milling process using image processing. Sensors 15 (7), 15738–15754.
Pujari, J.D., Yakkundimath, R., Byadgi, A.S., 2013. Grading and classification of anthracnose fungal disease of fruits based on statistical texture features. Int. J. Adv. Sci. Technol. 52 (1), 121–132.
Qureshi, W.S., Payne, A., Walsh, K.B., Linker, R., Cohen, O., Dailey, M.N., 2017. Machine vision for counting fruit on mango tree canopies. Precis. Agric. 18 (2), 224–244.
Raschka, S. (2014). Naive bayes and text classification i-introduction and theory. ArXiv Preprint arXiv:1410.5329.
Ravikanth, L., Singh, C.B., Jayas, D.S., White, N.D., 2015. Classification of contaminants from wheat using near-infrared hyperspectral imaging. Biosyst. Eng. 135, 73–86.
Razmjooy, N., Mousavi, B.S., Soleymani, F., 2012. A real-time mathematical computer method for potato inspection using machine vision. Comput. Math. Appl. 63 (1), 268–279.
Rehman, T., 2017. Development of a Machine Vision Based Weed (Goldenrod) Detection System for Spot-Application of Herbicides in Wild Blueberry Cropping System. M.Sc. thesis. Dalhousie university, Halifax, Canada.
Rehman, T.U., Zaman, Q.U., Chang, Y.K., Schumann, A.W., Corscadden, K.W., Esau, T.J., 2018. Optimising the parameters influencing performance and weed (goldenrod) identification accuracy of colour co-occurrence matrices. Biosyst. Eng. 170, 85–95.
Reina, G., Milella, A., Galati, R., 2017. Terrain assessment for precision agriculture using vehicle dynamic modelling. Biosyst. Eng. 162, 124–139.
Rennie, J.D., Shih, L., Teevan, J., Karger, D.R., 2003. Tackling the poor assumptions of naive Bayes text classifiers. In: Proceedings of the 20th international conference on machine learning (ICML-03), pp. 616–623.
Reynolds, D., 2015. Gaussian mixture models. Encyclopedia Biometr. 827–832.
Rivera, N.V., Gómez-Sanchis, J., Chanona-Pérez, J., Carrasco, J.J., Millán-Giraldo, M., Lorente, D., Blasco, J., 2014. Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. Biosyst. Eng. 122, 91–98.
Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104.
Romeo, J., Pajares, G., Montalvo, M., Guerrero, J.M., Guijarro, M., de la Cruz, Manuel, Jesús, 2013. A new expert system for greenness identification in agricultural images. Expert Syst. Appl. 40 (6), 2275–2286.
Romeo, J., Pajares, G., Montalvo, M., Guerrero, J., Guijarro, M., Ribeiro, A., 2012. Crop row detection in maize fields inspired on the human visual perception. Sci. World J.
Romualdo, L., Luz Pedro Henrique de Cerqueira, Devechio, F., Marin, M., Zúñiga, A., Bruno, O.M., Herling, V.R., 2014. Use of artificial vision techniques for diagnostic of nitrogen nutritional status in maize plants. Comput. Electron. Agric. 104, 63–70.
Rousseau, C., Belin, E., Bove, E., Rousseau, D., Fabre, F., Berruyer, R., Boureau, T., 2013. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. Plant Methods 9 (1), 17.
Rumpf, T., Mahlein, A.K., Steiner, U., Oerke, E.C., Dehne, H.W., Plümer, L., 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. Comput. Electron. Agric. 74 (1), 91–99.
Samaniego, L., Schulz, K., 2009. Supervised classification of agricultural land cover using a modified k-NN technique (MNN) and Landsat remote sensing imagery. Remote Sens. 1 (4), 875–895.
Sankaran, S., Mishra, A., Maja, J.M., Ehsani, R., 2011. Visible-near infrared spectroscopy for detection of huanglongbing in citrus orchards. Comput. Electron. Agric. 77 (2), 127–134.
Sankaran, S., Ehsani, R., 2012. Detection of huanglongbing disease in citrus using fluorescence spectroscopy. Trans. ASABE 55 (1), 313–320.
Sankaran, S., Ehsani, R., 2013. Comparison of visible-near infrared and mid-infrared spectroscopy for classification of huanglongbing and citrus canker infected leaves. Agric. Eng. Int.: CIGR J. 15 (3), 75–79.
Sannakki, S.S., Rajpurohit, V.S., Nargund, V.B., 2013. SVM-DSD: SVM Based diagnostic system for the detection of pomegranate leaf diseases. In: Proceedings of International Conference on Advances in Computing. Springer, New Delhi, pp. 715–720.
Schneider, L.C., Pontius Jr, R.G., 2001. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agric. Ecosyst. Environ. 85 (1–3), 83–94.
Senthilnath, J., Kandukuri, M., Dokania, A., Ramesh, K., 2017. Application of UAV imaging platform for vegetation analysis based on spectral-spatial methods. Comput. Electron. Agric. 140, 8–24.
Shearer, S.A., Holmes, R., 1990. Plant identification using color co-occurrence matrices. Trans. ASAE 33 (6), 1237–1244.
Shearer, S.A., Payne, F.A., 1990. Color and defect sorting of bell peppers using machine vision. Tran. ASAE 33 (6), 1245–1250.
Shi, Y., Judd, M., 2013. Finding nearest neighbors for multi-dimensional data. In: Proceedings of the 5th International Conference on Advances in Databases, Knowledge, and Data Applications, pp. 52–55.
Shi, Y., Huang, W., Luo, J., Huang, L., Zhou, X., 2017. Detection and discrimination of pests and diseases in winter wheat based on spectral indices and kernel discriminant analysis. Comput. Electron. Agric. 141, 171–180.
Shin, J.S., Lee, W.S., Ehsani, R., 2012. Postharvest citrus mass and size estimation using a logistic classification model and a watershed algorithm. Biosyst. Eng. 113 (1), 42–53.
Siedliska, A., Baranowski, P., Mazurek, W., 2014. Classification models of bruise and cultivar detection on the basis of hyperspectral imaging data. Comput. Electron. Agric. 106, 66–74.
Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 21 (2), 110–124.
Singh, V., Misra, A.K., 2017. Detection of plant leaf diseases using image segmentation and soft computing techniques. Inform. Process. Agric. 4 (1), 41–49.
Solahudin, M., Boro Seminar, K., Astika, I. W., Buono, A. (2010). Weeds and Plants Recognition Using Fuzzy Clustering and Fractal Dimension Methods for Automatic

Weed Control.
Stegmayer, G., Milone, D.H., Garran, S., Burdyn, L., 2013. Automatic recognition of quarantine citrus diseases. Expert Syst. Appl. 40 (9), 3512–3517.
Steward, B.L., Tian, L.F., 1999. Real-time weed detection in outdoor field conditions. In: Precision Agriculture and Biological Quality, vol. 3543. International Society for Optics and Photonics, pp. 266–279.
Steward, B.L., Tian, L.F., Nettleton, D., Tang, L., 2004. Reduced-dimension clustering for vegetation segmentation. Trans. ASAE 47 (2), 609.
Sun, Y., Gu, X., Sun, K., Hu, H., Xu, M., Wang, Z., Pan, L., 2017. Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches. LWT-Food Sci. Technol. 75, 557–564.
Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT press.
Tabb, A.L., Peterson, D.L., Park, J., 2006. Segmentation of apple fruit from video via background modeling. Paper Presented at the 2006 ASAE Annual Meeting, 1.
Tagarakis, A., Liakos, V., Fountas, S., Koundouras, S., Gemtos, T., 2013. Management zones delineation using fuzzy clustering techniques in grapevines. Precis. Agric. 14 (1), 18–39.
Tahir, A., Neethirajan, S., Jayas, D., Shahin, M., Symons, S., White, N., 2007. Evaluation of the effect of moisture content on cereal grains by digital image analysis. Food Res. Int. 40 (9), 1140–1145.
Tellaeche, A., BurgosArtizzu, X., Pajares, G., Ribeiro, A., 2007. A vision-based hybrid classifier for weeds detection in precision agriculture through the bayesian and fuzzy k-means paradigms. Innovat. Hybrid Intell. Syst. 72–79.
Tellaeche, A., Pajares, G., Burgos-Artizzu, X.P., Ribeiro, A., 2011. A computer vision approach for weeds identification through Support Vector Machines. Appl. Soft Comput. 11 (1), 908–915.
Tian, Y., Zhao, C., Lu, S., Guo, X., 2012. SVM-based multiple classifier system for recognition of wheat leaf diseases. In: World Automation Congress (WAC), 2012. IEEE, pp. 189–193.
Tibshirani, R., James, G., Witten, D., Hastie, T., 2013. An Introduction to Statistical Learning-with Applications in R. Springer publishing Co, New York, NY, USA.
Treitz, P., Howarth, P., Filho, O., Soulis, E., 2000. Agricultural crop classification using SAR tone and texture statistics. Canad. J. Remote Sens. 26 (1), 18–29.
Triantafilis, J., Lesch, S., 2005. Mapping clay content variation using electromagnetic induction techniques. Comput. Electron. Agric. 46 (1), 203–237.
Unay, D., Gosselin, B., 2006. Automatic defect segmentation of 'Jonagold'apples on multi-spectral images: a comparative study. Postharvest Biol. Technol. 42 (3), 271–279.
Unay, D., Gosselin, B., Kleynen, O., Leemans, V., Destain, M., Debeir, O., 2011. Automatic grading of bi-colored apples by multispectral machine vision. Comput. Electron. Agric. 75 (1), 204–212.
Vapnik, V., 1998. Statistical Learning Theory. Wiley, New York.
Veernagouda Ganganagowder, N., Kamath, P., 2017. Intelligent classification models for food products basis on morphological, colour and texture features. Acta Agronómica 66 (4).
VijayaLakshmi, B., Mohan, V., 2016. Kernel-based PSO and FRVM: an automatic plant leaf type detection using texture, shape, and color features. Comput. Electron. Agric. 125, 99–112.
Vrindts, E., Mouazen, A.M., Reyniers, M., Maertens, K., Maleki, M., Ramon, H., De Baerdemaeker, J., 2005. Management zones based on correlation between soil compaction, yield and crop data. Biosyst. Eng. 92 (4), 419–428.
Wachs, J.P., Stern, H.I., Burks, T., Alchanatis, V., 2010. Low and high-level visual feature-based apple detection from multi-modal images. Precis. Agric. 11 (6), 717–735.
World Bank (2015). The World Bank: Agricultural Land (% of Land Area). https://data.worldbank.org/indicator/AG.LND.AGRI.ZS.
Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F., Xiang, Q.L., 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. In: Signal Processing and Information Technology, 2007 IEEE International Symposium on. IEEE, pp. 11–16.
Xu, G., Zhang, F., Shah, S.G., Ye, Y., Mao, H., 2011. Use of leaf color images to identify nitrogen and potassium deficient tomatoes. Pattern Recogn. Lett. 32 (11), 1584–1590.
Yang, C.C., Prasher, S.O., Enright, P., Madramootoo, C., Burgess, M., Goel, P.K., Callum, I., 2003. Application of decision tree technology for image classification using remote sensing data. Agric. Syst. 76 (3), 1101–1117.
Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y., Yang, B., 2009. Application of support vector machine for detecting rice diseases using shape and color texture features. In: Engineering Computation, 2009. ICEC'09. International Conference on. IEEE, pp. 79–83.
Yimyam, P., Clark, A.F., 2016. 3D reconstruction and feature extraction for agricultural produce grading. In: Paper Presented at the Knowledge and Smart Technology (KST), 2016 8th International Conference on, pp. 136–141.
Yin, H., Chai, Y., Yang, S.X., Mittal, G.S., 2009. Ripe tomato extraction for a harvesting robotic system. In: Paper presented at the Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on, pp. 2984–2989.
Zhang, N., Chaisattapagon, C., 1995. Effective criteria for weed identification in wheat fields using machine vision. Trans. ASAE 38 (3), 965–974.
Zhang, J., Wang, N., Yuan, L., Chen, F., Wu, K., 2017a. Discrimination of winter wheat disease and insect stresses using continuous wavelet features extracted from foliar spectral measurements. Biosyst. Eng. 162, 20–29.
Zhang, S., Wu, X., You, Z., Zhang, L., 2017b. Leaf image based cucumber disease recognition using sparse representation classification. Comput. Electron. Agric. 134, 135–141.