IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 20, 2021, accepted May 28, 2021, date of publication June 4, 2021, date of current version June 14, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3086269

Cascading Feature Filtering and Boosting Algorithm for Plant Type Classification Based on Image Features

ADEL BAKHSHIPOUR
Department of Biosystems Engineering, Faculty of Agricultural Sciences, University of Guilan, Rasht 4199613776, Iran
e-mail: abakhshipour@guilan.ac.ir

ABSTRACT Crop and weeds identification is of important steps towards the development of efficient automotive weed control systems. The higher the accuracy of plant detection and classification, the higher the performance of the weeding machine. In this study, the capability of two popular boosting methods including Adaboost.M1 and LogitBoost algorithms was evaluated to enhance the plant classification performance of four classifiers, namely Multi-Layer Perceptron (MLP), k-Nearest Neighbors (kNN), Random Forest (RF), and Support Vector Machine (SVM). Four feature filtering techniques including Correlation-based Feature Selection (CFS), Information Gain (IG), Gain Ratio (GR), and OneR were applied to the image-extracted features and 10 of the most significant features were selected and fed into single and boosted classifiers. The RF model trained by IG selected features (IG-RF) was the most appropriate classifier among the evaluated models whether in single or boosted modes. It was also found that boosting of IG-RF by using Adaboost.M1 and LogitBoost algorithms improved the classification accuracy. Regarding the performance values, the LogitBoost-IG-RF structure, which provided a classification accuracy of 99.58%, a kappa (k) of 0.9948, and a Root Mean Squared Error (RMSE) of 0.0688 on training dataset, was selected as the most appropriate classifier for plant discrimination in peanut fields. The accuracy, k, and RMSE criteria of this combination on test dataset were 95.00%, 0.9375, and 0.1591, respectively. It was concluded that combination of boosting algorithms and feature selection methods can promote plant type discrimination accuracy, which is a crucial factor in the development of precision weed control systems.

INDEX TERMS Ensemble learning, feature selection, image processing, plant identification, precision agriculture.

I. INTRODUCTION

Presence of weeds in fields and their competition with the main plant for water, light, nutrients, and space can cause irreparable damage to crop performance if the weeds are not appropriately treated.

Yield losses from 37% to 61% were reported by Dille *et al.* [1] in grain sorghum with weed interference in different regions of the United States. The results from field studies showed that growers in different regions of the United States and Canada would potentially lose an average of 31% to 94% of their dry bean yield [2] and 61% to 83% of the sugar beet yield [3] due to weeds.

Timely and effective weed management is very important in peanut, as it increases the yield performance and economic return of peanut compared to the non-treated strategy [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Senthil Kumar.

Precision weed control such as selective spraying or accurate mechanical removal of weeds, is a challenging task that aims to reduce the amount of herbicides without compromising the quality of crops [5]. Accurate weed detection in croplands, as a prerequisite for applying any precision weed management technology [6], is still a challenging step toward the development of site-specific weed control machines, especially when there are intra-row weeds that are highly overlapped with the main plant. Efficient weed removal weather using variable-rate sprayers, or precise mechanical, electrical, or thermal hoeing systems, preliminary requires to detect and segregate weeds from main crop [7]. Computer vision is a well-known approach that has shown to be successful in object recognition and detection in a variety of applications including agriculture automation and monitoring. The computer vision system uses image processing techniques to process and analyze colour images and achieve the required classification [8]. The critical procedure for precise weed detection is digital image processing, through which weed can be segmented and extracted from the acquired images [9]. Several image processing algorithms can be used to extract distinguishing features from images or videos obtained from a scene. Numerous types of image-based features including colour [10]–[12], shape [13]–[15], texture [16]–[18], wavelet transform [19]–[21], and fusion of different features [22]–[24] have been applied to plant type identification with acceptable accuracies encouraging the further application of these image-based features for crop-weed discrimination.

There are large number of features that can be extracted from images and introduced into the classification models to discriminate invasive weeds from the main plant. The problem here is that having too many input features doesn't always guarantee obtaining higher classification performance. In addition to useful relevant features, there are often some features that have collinearity and, maybe, there are some of the features that have very little correlation with the plant type. These redundant and irrelevant features can significantly decrease the accuracy of the developed model and can increase the training time. The existent classification techniques are inadequate to handle a high number of attributes in terms of training time and/or effectiveness in selecting the relevant set of features [25]. Therefore, dimensionality reduction and removal of non-informative or redundant data from a high dimensional training dataset is an important preprocessing step in machine learning that enhances the performance and simplifies the complexity of the classifier model. The feature selection can be performed in a supervised mode that takes into account the class information, or an unsupervised mode where the class label information is unavailable or ignored [26], [27]. One popular type of supervised feature selection method is feature filtering which is accomplished by selecting a feature subset from the original feature set that is the most relevant and pertinent to the target classes [28], [29].

One of the most usual applications of feature selection is in classifiers that use the selected features to discriminate objects into different classes. Different supervised and unsupervised classifier algorithms including Artificial Neural Networks (ANN), Support Vector Machine (SVM), Decision Tree (DT), Principal Component Analysis (PCA), Bayesian Classifier (BC), Linear Discriminant Analysis (LDA), K-Means, k-Nearest Neighbors (kNN), etc. have been applied for distinguishing the crops from weeds [7], [13], [30]–[38].

A technique to increase the classification accuracy is combining the decisions of multiple independent base classifiers to achieve a booster classifier, which is called ensemble classifier. Ensemble learning is a machine learning paradigm in which multiple learners are trained to solve the same problem [39]. A form of ensemble learning is known as boosting, in which a set of simple classifiers that are also called weaker learners, are combined to construct a relatively stronger classifier [40], [41]. Two of the most common boosting algorithms are Adaboost.M1 and LogitBoost. AdaBoost.M1, which was proposed by Freund and Schapire [42], is a simple generalization of Adaboost to be applicable to problems with more than two classes [43]. According to Cortes *et al.* [44]: "Adaboost.M1 is based on building consecutive classifiers on modified versions of the training set generated according to the error rate of the previous classifier, while focusing on the hardest examples of the training set".

LogitBoost, which was formulated by Friedman *et al.* [45], is another expansion of Adaboost that uses a combination of the boosting method and the logistic regression for classification [46], [47].

Although several researches have been conducted for weed detection using different classification techniques, however, to the best of our knowledge, there is almost no research that focuses on applying boosting learning methods for image-based plant discrimination. This study is conducted to fulfill this research gap. Several image features are extracted from the images of different plants in the peanut field. Feature selection algorithms are applied to extract the most significant features which are fed into boosted classifiers to distinguish different types of plants. A comparative study is performed to find the most effective strategy.

II. MATERIAL AND METHODS

A. IMAGE ACQUISITION

The data acquisition was executed in May 2020. In order to obtain the required dataset for this study, the colour images were acquired from a peanut field in Astaneh-ye Ashrafieh county of Guilan province, Iran. Image capturing was performed at after three weeks from seed sowing (7 days after first emergence of the peanut seedlings), when the weeds also emerged. A metal frame was constructed having a platform at the elevation of 40 cm above the crop row for placing the image capturing device. In this study, the images were captured using an affordable smartphone with a resolution of 1344 × 2240 pixels. A fabric shade was applied to avoid the effect of direct sunlight on the capturing scene. A total number of 150 images were captured from fields containing multiple plants.

B. IMAGE PREPARATION

Image processing and feature extraction operations were implemented in MATLAB programming software (MATLAB 2018b, The MathWorks Inc., MA, USA). In addition to the peanut plant (*Arachis hypogaea*), four other plants that were the most common weeds in peanut farms, namely; thorn apple (*Datura stramonium*), morning glory (*Ipomoea purpurea*), purslane (*Portulaca oleracea*), and velvetleaf (*Abutilon theophrasti*) were investigated in this study (Figure 1). In order to extract image features from plants, the regions of the studied plants were manually selected in the images. The corner coordinates of the rectangles around the mentioned plants were determined using simple marking in the "*imtool*" function of MATLAB software and then the desired regions were cropped using the "*imcrop*" function.

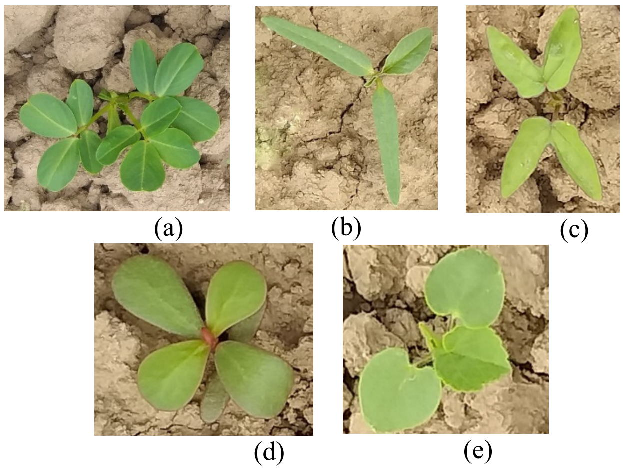

FIGURE 1. Images of studied plants: a) peanut, b) thorn apple, c) morning glory, d) purslane, and e) velvetleaf.

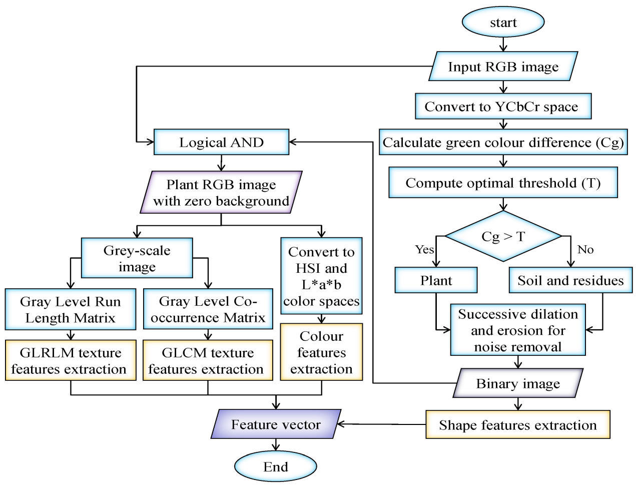

FIGURE 2. Schematic diagram of image processing for feature extraction.

Eighty samples were selected by this method for each plant type.

The flowchart of the image processing algorithm for colour, shape, and texture feature extraction is shown in Figure 2. Since the RGB images of plants contained field soil and little residues in the background, it was necessary to separate the plants from the image background before performing the feature extraction processes. The Red (R), Green (G), and Blue (B) colour components were firstly extracted from the RGB image and the luminance component (Y) was calculated using equation 1 [48], [49], which was used to calculate the green colour difference image (Cg) by equation 2 [50], [51].

$$Y = 0.3R + 0.6G + 0.1B \tag{1}$$

$$C_g = G - Y \tag{2}$$

Image segmentation was performed by applying optimal threshold on the Cg image which resulted in the binary image showing plant. Possible noises in the binary image were omitted by applying successive dilation and erosion using the "*imopen*" function in MATLAB software. The resulted binary images of plants were used for shape feature extraction. However, in order to extract colour and texture features, one more operation was performed. Logical AND was applied between RGB images of plants and the binary image to obtain colour images of plants with zero pixel value for background regions. This allowed the colour and texture features to be extracted from only the plant regions in the images. The images obtained at different steps of such described algorithm are shown in figure 3.

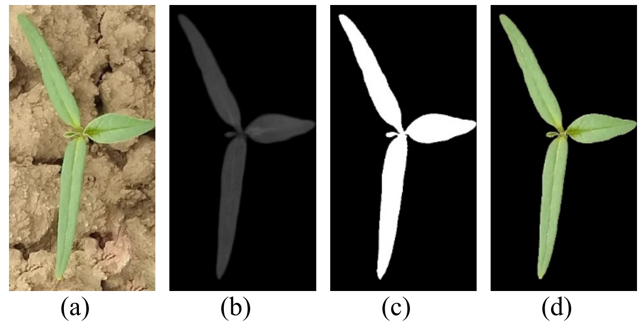

FIGURE 3. Gallery of image preparation steps: a) RGB image of thorn apple weed, b) enhanced green colour difference, c) segmented image, d) colour image of thorn apple weed with zero intensity for background pixels.

C. SHAPE FEATURES

In this study, in order to obtain features invariant to position, orientation and size, to be more generalized in different images, two types of features were extracted, including shape factors and moment invariant shape descriptors. After image segmentation, geometric shape features including area, perimeter, major axis length, minor axis length, and equivalent diameter (diameter of a circle with the same area as the region) values were extracted from binary images of plants (figure 3c) using the "*regionprops*" function in MATLAB software. These values were used to calculate four shape factors using the equations (3) to (6) which are also described in the literature [13], [50], [52], [53].

$$shape\ factor\ 1 = 4\pi \frac{area}{perimeter^2} \tag{3}$$

$$shape\ factor\ 2 = \frac{major\ axis\ length}{area} \tag{4}$$

$$shape\ factor\ 3 = \frac{area}{major\ axis\ length^3} \tag{5}$$

$$shape\ factor\ 4 = \frac{4\ area}{\pi .major\ axis\ length.minor\ axis\ length} \tag{6}$$

In addition to shape factors, 7 moment invariant shape descriptors (M1 to M7) which are called Hu moments [54] were also extracted from binary images of plants. The Hu moment invariants are independent of geometric translation, scaling, and rotation, providing a high discrimination power to discriminate different morphological classes of objects [13]. These features and their calculation formulas are described by Rhouma *et al.* [55], Fatma and Dash [15] and Sabzi *et al.* [56].

D. COLOUR FEATURES

The colour values of plant images in RGB, HSI and L*a*b colour spaces were considered as colour features. Compared to RGB colour space, the HSI and L*a*b colour spaces are closer to human visual perception to colour. Moreover, the I component in HSI space and the L component in L*a*b space are representations of image luminance information. This helps to extract the image brightness as a separate component in these two colour spaces, leading to better consideration of the effects of possible lightness variations in the images. In order to extract the colour values in this study, the images of plant regions that obtained from the image preparation section (figure 3d) were converted from RGB colour space to HSI, and L*a*b colour spaces and the average and standard deviation measures of different colour components in these three spaces, namely; Red (R), Green (G), Blue (B), Hue (H), Saturation (S), Intensity (I), Lightness (L), a* and b* colour components, were determined. The corresponding colour space transformation equations are presented by researchers [7], [57], [58]. A total number of 18 colour features including nine colour averages and nine colour standard deviations were extracted in this case.

E. GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) TEXTURE FEATURES

In this section, the colour images of plant regions obtained from *image preparation* section (figure 3d) were converted from RGB colour space to gray-level images. The Gray-Level Co-occurrence Matrices (GLCM) were constructed for obtained gray-level images in four directions from 0° to 135° with 45° interval with one-pixel distance between the compared pixels to obtain four GLCMs for each image, which were then averaged to be used for extracting the texture features. From so many GLCM-based texture features that have been described and used in several studies [17], [59]–[61], 17 texture features were calculated and used in this study for plant type detection. These GLCM features, which are also called Haralick feature, were autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, and inverse difference moment.

F. GRAY LEVEL RUN LENGTH MATRIX (GLRLM) TEXTURE FEATURES

In order to acquire more insight into the plant leaf texture information, the GLRLM based texture features were also extracted and analyzed in this study. The GLRLM which was introduced by Galloway [62] is a method of extracting higher-order statistical texture features. Each element P(i, j) in run length matrix (P) is equal to the number of runs with pixels of gray level intensity equal to i and length of run equal to j along a specific direction [63]. The GLRLMs were constructed for gray-level images of plants in four different directions of 0, 45, 90, and 135° and averaged to be used for feature extraction. In this study, 11 GLRLM based texture features including Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray-Level Non-Uniformity (GLN), Run Length Non-Uniformity (RLN), Run Percentage (RP), Low Gray-Level Run Emphasis (LGRE), High Gray-Level Run Emphasis (HGRE), Short Run Low Gray-Level Emphasis (SRLGE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE), and Long Run High Gray-Level Emphasis (LRHGE) were extracted from images. These features and their formulas have been previously described in several research articles [64]–[66] and the related MATLAB codes are revealed by Wei [67].

G. FEATURE SELECTION

In this study four feature filtering techniques including Correlation-based Feature Selection (CFS), Information Gain (IG), Gain Ratio (GR), and OneR were used to determine the most significant feature vectors for discriminating different plant types.

CFS is a simple feature filtering method that selects a subset of features that is highly correlated to the class label and minimum relevance to each other [68], [69]. IG is a feature evaluation method based on the reduction in entropy of dataset features [70]. It ranks the features based on the amount of information that they provide for the target feature [71] and it ignores the feature correlation [72]. Gain ratio (GR) is a modification of the information gain method that reduces its bias by taking intrinsic information from each attribute and eliminating the bias value of each attribute [73], [74]. OneR feature selection which was firstly introduced by Holte [75], is based on the One Rule theory and creates association rules by identifying the correlation between a particular feature and its impact on the output class [76], to be used for feature ranking.

The total number of image extracted features in this study was 59 (18 colour features, 13 shape features, and 28 texture features). After applying the mentioned feature filters on the original features, the first ten features were selected and fed into the classifiers to discriminate different plants.

H. SINGLE AND BOOSTED CLASSIFIERS

In order to classify the plants using the image features that selected by feature filtering methods, two strategies were applied. First, the selected features were fed into five classifiers including Multilayer Perceptron (MLP), k-Nearest Neighbor (kNN), Random Forest (RF), and Support Vector Machine (SVM) to evaluate and compare the capability of these single classifiers for plant type classification.

Second, two of the most popular meta-learning algorithms including Adaptive Boosting M1 (AdaBoost.M1) and Logistic Boosting Regression (LogitBoost) were utilized as boosting algorithms for plant type classification, in which MLP, kNN, RF, and SVM were used as base (or weaker) classifiers to construct boosted models.

I. MODEL EVALUATION

In this study, 10 fold cross-validation strategy was applied to train multiclass (5 classes) classifiers. In order to evaluate the performance of the classifiers, three statistical criteria including accuracy (ACC), Cohen kappa statistics (k), Root Mean Squared Error (RMSE), were determined for the developed models. These criteria are detailedly described [7], [16], [64], [77]. Higher values for ACC and k, and lower value of RMSE correspond to better classification performance.

III. RESULTS

A. CORRELATIONS BETWEEN IMAGE FEATURES AND PLANT TYPE

The absolute values of the correlation coefficients among the averages of colour features, and plant type are illustrated as colourmaps in figure 4. Figure 4 shows that the averages of Hue, Red, and b* colour components (ave_H, ave_b* and ave_R) have higher correlation coefficients with the plant type. The absolute values of correlation coefficients between plant type and ave_H, ave_b* and ave_R were 0.682, 0.558 and 0.426, respectively. The lowest correlation coefficient was observed between average Blue values and the plant type (0.209).

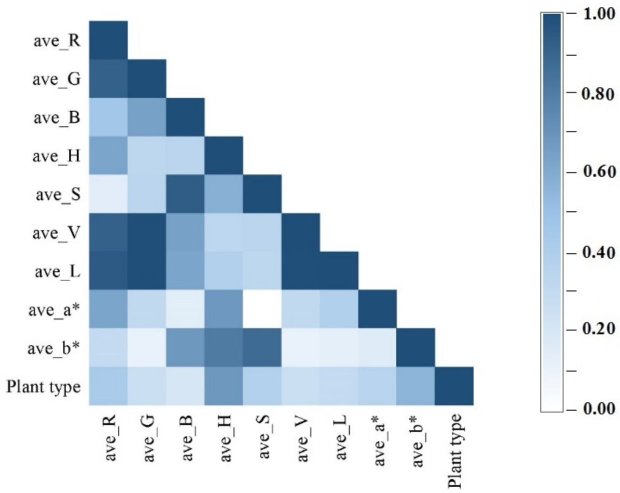

FIGURE 4. Colourmap of correlation matrix between colour feature averages and plant type.

The figure 4 also shows that there are high values of intercorrelation between the averages of several colour components (for example, between ave_R, ave_G, ave_I, and ave_L). This is due to the fact that colour spaces could be converted by mathematical transformation. Also it justifies the use of feature selection methods to extract the most significant features and to ignore the redundant features or those with low correlation to plant type.

The absolute correlation coefficients between standard deviations of colour components and plant type is presented graphically in figure 5. It can be seen from this figure that there are very low correlation coefficients between the standard deviations of colour features and plant type (bottom row in figure 5). This indicates that the standard deviations of the colour values of the leaf surfaces do not provide useful information for discrimination of studied plants. The highest absolute value of correlation, in this case, was 0.201 which was between the standard deviation of Hue component (std_H) and plant type. Also, high intercorrelations among standard deviations of Red, Green, Blue, Intensity, and Lightness colour components can be observed.

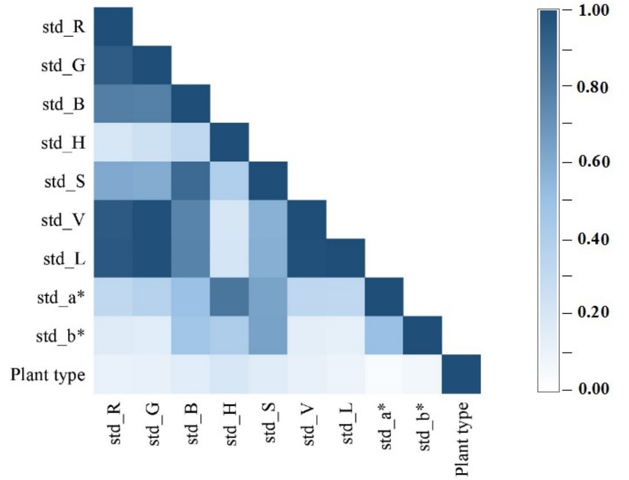

FIGURE 5. Colourmap of correlation matrix between colour feature standard deviations and plant type.

The correlation values between shape features and plat type are indicated in figure 6. The shape factor 1 has a high correlation with the plant type (0.742). The next highest correlated shape feature to plant type is shape factor 4 (0.650). Among the moment invariants, the highest correlation to plant type belongs to M1 (0.526). The moment invariants of M5 to M9 has very small correlations with plant type (<0.03), indicating the inappropriacy of these features in the classification of studied plants. The intercorrelation among the shape features was not as strong as those in the colour features.

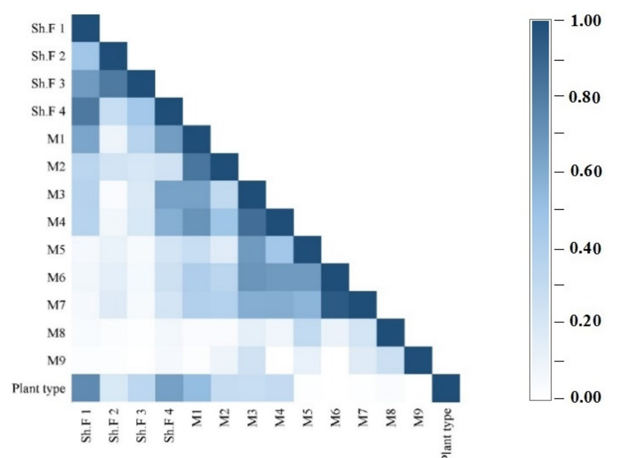

FIGURE 6. Colourmap of correlation matrix between shape features and plant type.

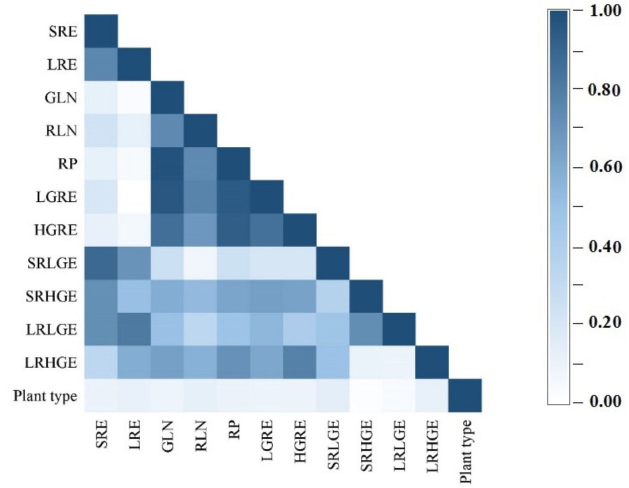

FIGURE 7. Colourmap of correlation matrix between GLRLM texture features and plant type.

Figure 7 illustrates the colourmap of the absolute correlation coefficients among GLRLM texture features and plant type. There were high intercorrelations among the GLRLM features themselves and low correlations between these features and plant type were obtained, which can be seen from figure 7. Besides, by observing the colourmap of correlation coefficient values between GLCM texture features and the plant type in figure 8, the highest correlation coefficient, in this case, was obtained 0.248 that observed between autocorrelation and plant type. There are also high intercorrelations among most GLCM texture features.

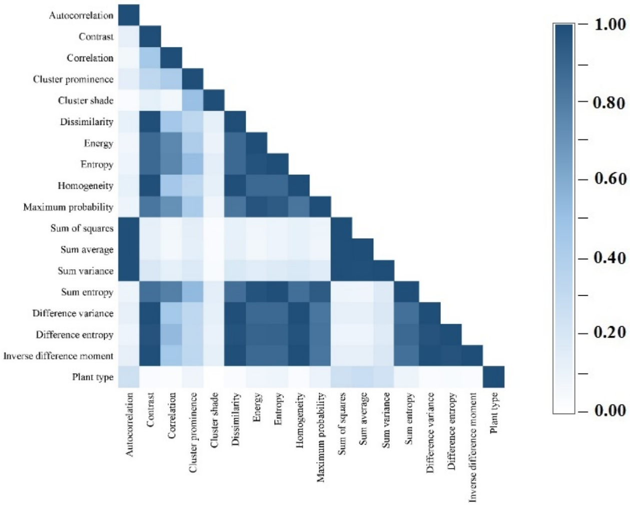

FIGURE 8. Colourmap of correlation matrix between GLCM texture features and plant type.

B. RESULTS OF FEATURE SELECTION METHODS

The top ten most informative features are listed in table 1. These features are ranked based on their correlation to plant type which ranged from 0.298 for M4, to 0.742 for shape factor 1. It can be seen from this table that half of the top ten features that have the highest correlations to plant type are shape features and the other half of this set are colour features, while none of the texture features are in this set which shows the importance of colour and shape features for plant classification at early growth stages of plants.

In order to include all of the extracted features in the plant type classification process to avoid any information loss, a fusion of different feature types of colour, shape, and texture data was used. However, to get rid of redundant and non-useful features, feature selection methods were applied to extract the most significant features. The selected features by different feature selection methods are presented in table 2.

TABLE 1. Top ten most correlated image-extracted features to plant type.

Image Feature	Correlation coefficient
shape factor 1	0.742
average Hue	0.682
shape factor 4	0.650
average b*	0.558
M1	0.526
average Red	0.426
average Saturation	0.378
average a*	0.353
shape factor 3	0.334
M4	0.298

TABLE 2. The selected features using CFS, IG, GR, and OneR feature selection methods.

Feature importance rank	Feature selection method			
	CFS	IG	GR	OneR
1	average Red	M1	average Hue	shape factor 4
2	std of Red	shape factor 4	shape factor 1	shape factor 1
3	average Blue	shape factor 1	average Red	average Hue
4	average Hue	M3	shape factor 2	shape factor 3
5	std of Hue	shape factor 3	shape factor 4	M3
6	average a*	average Hue	M3	shape factor 2
7	average b*	shape factor 2	shape factor 3	M1
8	shape factor 1	M4	M4	M4
9	shape factor 2	average a*	M1	average a*
10	shape factor 3	average b*	M2	average Saturation

Comparing the results of different feature selection methods in table 2, it can be seen that different subsets of relevant attributes are selected by applying different feature selection algorithms. For example, The CFS method selected 10 attributes, including 7 colour attributes and 3 shape attributes, as the most important input data for the plant classification, while looking at the first 10 ranked attributes in the IG method, 7 shape features, and 3 colour features were extracted as the most important features.

Moreover, the two other feature selection methods introduced different ranked features in their selected subsets. This is because different feature selection methods use different statistical criteria to calculate the importance of each feature and its relevance to the class label [78], [79]. Further, from table 2, only colour and morphological characteristics are selected as the top 10 most informative attributes. Meanwhile, none of GLRLM and GLCM texture features are selected as important features for plant type classification.

C. RESULTS OF SINGLE AND ENSEMBLE CLASSIFIERS

The 10 significant CFS-selected features and the first 10 ranked features by IG, GR, OneR feature selection methods were fed into the single and boosted classifiers to differentiate different plants. Table 3 shows the results of the MLP, kNN, RF, and SVM classifiers used in this study for plant type classification. All of these applied single classifiers achieved high classification performances. Regarding the performance criteria on the training dataset, the most accurate single classifier for plant type identification was the RF classifier when used the attributes selected by the IG feature selection method as the input data. This classifier obtained an accuracy of 98.75%, k of 0.9844, and RMSE of 0.0734. The IG-RF model was also evaluated on a separated test dataset which was not included in the training procedure where the resulted accuracy, k, and RMSE values of this model were 91.67%, 0.8958, and 0.1605, respectively.

TABLE 3. Classification performance of single classifiers with different input selected features.

Feature selection	Classifier	ACC (%)	k	RMSE
CFS	MLP	97.92	0.9740	0.8920
	kNN	94.17	0.9271	0.3203
	RF	97.92	0.9740	0.9000
	SVM	94.58	0.9323	0.1472
IG	MLP	96.67	0.9583	0.1155
	kNN	95.83	0.9479	0.1183
	RF	98.75	0.9844	0.0734
	SVM	93.75	0.9219	0.1572
GR	MLP	97.92	0.9740	0.0871
	kNN	86.25	0.8281	0.2320
	RF	94.17	0.9271	0.3211
	SVM	90.00	0.875	0.1759
OneR	MLP	92.50	0.9158	0.1513
	kNN	95.83	0.9479	0.1171
	RF	98.33	0.9792	0.0843
	SVM	93.33	0.9167	0.1402

In addition to single models, two boosting algorithms were also evaluated in this study for plant type classification, and the results are revealed as follows. Performance criteria of the Adaboost.M1 ensemble learning algorithm when fed by features selected by CFS, IG, GR, and OneR feature selection methods, and constructed from 4 different base classifiers, are presented in table 4. Regarding the performance statistics of the models on the training dataset, the developed Adaboost.M1 classifiers resulted in satisfactory performances.

TABLE 4. Classification performance of Adaboost.M1 classifiers with different input selected features and base classifies.

Feature selection	Base classifier	ACC (%)	k	RMSE
CFS	MLP	98.75	0.9844	0.0723
	kNN	95.00	0.9375	0.1401
	RF	99.58	0.9948	0.0713
	SVM	96.67	0.9583	0.1157
IG	MLP	97.50	0.9687	0.0867
	kNN	96.67	0.9583	0.1145
	RF	99.58	0.9948	0.0688
	SVM	95.83	0.9479	0.1187
GR	MLP	97.97	0.974	0.0827
	kNN	97.08	0.9635	0.1072
	RF	99.17	0.9896	0.0680
	SVM	95.83	0.9479	0.117
OneR	MLP	97.92	0.974	0.0874
	kNN	96.67	0.9583	0.1145
	RF	99.58	0.9948	0.0730
	SVM	95.00	0.9375	0.1285

The highest classification accuracy obtained by Adaboost.M1 learning method was 99.58% which was obtained by the Adaboost.M1 classifier with the base classifier of RF, when the IG selected feature subsets were used as input data (Adaboost.M1-IG-RF). The k, and RMSE values of the Adaboost.M1-IG-RF model were 0.9948 and 0.0688, respectively. The accuracy, k, and RMSE values of the Adaboost.M1-IG-RF model on test data were 93.33%, 0.9167, and 0.1627 respectively. Regarding table 4, the Adaboost.M1–RF structure also resulted in the highest accuracies among other boosted classifiers while fed by CFS, GR and OneR feature filters.

Results of the LogitBoost classifier with different base classifiers and different feature selection techniques are summarized in table 5. Generally, the MLP and SVM classifiers resulted in lower performances than kNN, and RF classifiers when used as the weaker classifiers in the LogitBoost ensemble algorithm. The most successful LogitBoost structure for plant type classification had the base classifier of RF and trained using the IG feature selection method (called LogitBoost-IG-RF).

The accuracy, k, and RMSE values of this model on the training dataset were obtained as 99.58%, 0.9948, and 0.0408, respectively. The accuracy, k, and RMSE results of the LogitBoost-IG-RF were respectively 95.00%, 0.9375, and 0.1591 on test data. The LogitBoost-RF structure had also resulted in the same classification accuracy when trained by other feature subsets generated by CFS, GR, and OneR feature selection methods.

IV. DISCUSSION

The idea behind this study was to enhance the plant classification accuracy by employing feature selection algorithms and boosting ensemble techniques. For better comparison, the most accurate single and boosted models of this study, as well as the results of some other related studies are presented in table 6. From this table, and also by tables 3 to 5, the ensemble models are more accurate than single classifier, showing the advantage of boosting algorithms over single classifiers for plant classification.

TABLE 5. Classification performance of LogitBoost classifiers with different input selected features and base classifies.

model	Base classifier	ACC (%)	k	RMSE
CFS	MLP	73.33	0.6667	0.3234
	kNN	95.00	0.9375	0.1401
	RF	99.58	0.9948	0.0413
	SVM	85.42	0.8177	0.2058
IG	MLP	65.42	0.5677	0.3573
	kNN	96.67	0.9583	0.1155
	RF	99.58	0.9948	0.0408
	SVM	95.83	0.9479	0.1102
GR	MLP	36.67	0.2083	0.4916
	kNN	97.08	0.9635	0.1080
	RF	99.58	0.9948	0.0409
	SVM	86.25	0.8281	0.1900
OneR	MLP	47.08	0.3385	0.4447
	kNN	96.67	0.9583	0.1155
	RF	99.58	0.9948	0.0409
	SVM	96.67	0.9583	0.1015

TABLE 6. Classification performances of the selected classifiers of this study and some other related articles.

Method	ACC (%)	Reference
LogitBoost-IG-RF	99.58	this study
Adaboost.M1-IG-RF	99.58	this study
IG-RF	98.75	this study
Hybrid of ANN and Harmony Search algorithm (ANN-HS)	98.38	[31]
Image processing	98.23	[12]
Neighborhood Component Feature Selection (NCFS) with kNN classifier (NCFS-kNN)	98.00	[59]
CFS feature selection with J48 classifier (J48-CFS)	95.56	[7]
Hybrid of ANN and imperialist competitive algorithm (ANN-ICA)	95.21	[38]
SVM	95.00	[13]

Moreover, between ensemble classifiers, although the accuracy and k values of the most successful LogitBoost model were the same as those of the most successful Adaboost.M1 model, but the obtained RMSE value of the LogitBoost-IG-RF model was less than the RMSE of the Adaboost.M1-IG-RF model. Regarding these descriptions, the LogitBoost-IG-RF model, having the least error rate and the highest accuracy criteria, is the most effective classifier over the other models for the classification of plants based on image-extracted features.

It is also notable that the RF model, either alone or when boosted by the Adaboost.M1 and LogitBoost algorithms, yielded the best results compared to the other classifiers. RF itself is an ensemble classifier composed of several decision trees and aggregates the predictions of separately trained decision trees to make a final decision [80], [81] making it more robust to overfitting and noises. Integration of random forest with boosting methods combines the advantages of both adaptive evaluation of boosting and diversity of random forest [82] to enhance the classification accuracy. Also it is seen that among the feature selection methods, the IG algorithm extracted the most significant features toward plant classification. The ten selected features by IG algorithm were M1, shape factor 4, shape factor 1, M3, shape factor 3, average Hue, shape factor 2, M4, average a*, average b*, respectively (table 2). It can be observed that seven of the ten selected significant features were the shape features, which shows the significance of shape characteristics in discriminating between plants in the early stages of growth. The remaining three features, of the ten significant features, were average values of Hue, a* and b* colour components. This indicates that the colour space transformation from the RGB to other colour spaces such as HSI and L*a*b, can help distinguish plants from each other. Furthermore, texture features are not included in the IG selected features, which shows that the gray level spatial distribution of the leaf surface does not give useful information about the plant type at early growth stages, while the colour and shape of plants are good factors for plant type identification in this period.

Comparing the results of this study with some related studies in table 6 shows that the boosted models of this study have almost higher accuracies than the reported accuracies in previous researches. These results emphasize the efficiency of the employed boosting methods for promoting the plant classification accuracy.

V. CONCLUSION

A comparison between different single and boosted classifiers was performed in this study for the classification of plants based on image extracted features filtered via feature selection algorithms. It was observed that the performance of the RF classifier, when fed by features selected through the IG algorithm, was better than other evaluated combinations of feature selection methods and classifier models. Also, the classification performance of the IG-RF classifier was enhanced when it was boosted by Adaboost.M1 and

LogitBoost algorithms. Considering the performance criteria, the LogitBoost-IG-RF is introduced as the best classification model for plant type discrimination. It is concluded that integration of such boosted classifier into a computer vision system can enhance the performance of crop and weed detection, toward the development of a robotic vision-based system for weed control at early emergence stages in peanut fields.

ACKNOWLEDGMENT
The author would like to thank the University of Guilan for supporting this research.

REFERENCES
[1] J. A. Dille, P. W. Stahlman, C. R. Thompson, B. W. Bean, N. Soltani, and P. H. Sikkema, ''Potential yield loss in grain sorghum (Sorghum bicolor) with weed interference in the United States,'' *Weed Technol.*, vol. 34, no. 4, pp. 624–629, Aug. 2020.
[2] N. Soltani, J. A. Dille, R. H. Gulden, C. L. Sprague, R. K. Zollinger, D. W. Morishita, N. C. Lawrence, G. M. Sbatella, A. R. Kniss, P. Jha, and P. H. Sikkema, ''Potential yield loss in dry bean crops due to weeds in the United States and Canada,'' *Weed Technol.*, vol. 32, no. 3, pp. 342–346, Jun. 2018.
[3] N. Soltani, J. A. Dille, D. E. Robinson, C. L. Sprague, D. W. Morishita, N. C. Lawrence, A. R. Kniss, P. Jha, J. Felix, R. E. Nurse, and P. H. Sikkema, ''Potential yield loss in sugar beet due to weed interference in the United States and Canada,'' *Weed Technol.*, vol. 32, no. 6, pp. 749–753, Dec. 2018.
[4] A. T. Hare, D. L. Jordan, R. G. Leon, K. L. Edmisten, A. R. Post, D. J. Mahoney, and D. Washburn, ''Impact of weed management on peanut yield and weed populations the following year,'' *Peanut Sci.*, vol. 46, no. 2, pp. 182–190, Jul. 2019.
[5] M. Fawakherji, A. Youssef, D. D. Bloisi, A. Pretto, and D. Nardi, ''Crop and weed classification using pixel-wise segmentation on ground and aerial images,'' *Int. J. Robot. Comput.*, vol. 2, no. 1, pp. 39–57, Jan. 2020.
[6] B. Liu and R. Bruch, ''Weed detection for selective spraying: A review,'' *Current Robot. Rep.*, vol. 1, no. 1, pp. 19–26, Mar. 2020.
[7] A. Bakhshipour and H. Zareiforoush, ''Development of a fuzzy model for differentiating peanut plant from broadleaf weeds using image features,'' *Plant Methods*, vol. 16, no. 1, pp. 1–16, Dec. 2020.
[8] D. P. Cavallo, M. Cefola, B. Pace, A. F. Logrieco, and G. Attolico, ''Non-destructive and contactless quality evaluation of table grapes by a computer vision system,'' *Comput. Electron. Agricult.*, vol. 156, pp. 558–564, Jan. 2019.
[9] A. Wang, W. Zhang, and X. Wei, ''A review on weed detection using ground-based machine vision and image processing techniques,'' *Comput. Electron. Agricult.*, vol. 158, pp. 226–240, Mar. 2019.
[10] J.-L. Tang, X.-Q. Chen, R.-H. Miao, and D. Wang, ''Weed detection using image processing under different illumination for site-specific areas spraying,'' *Comput. Electron. Agricult.*, vol. 122, pp. 103–111, Mar. 2016.
[11] Y. Zheng, Q. Zhu, M. Huang, Y. Guo, and J. Qin, ''Maize and weed classification using color indices with support vector data description in outdoor fields,'' *Comput. Electron. Agricult.*, vol. 141, pp. 215–222, Sep. 2017.
[12] M. G. Forero, S. Herrera-Rivera, J. Ávila-Navarro, C. A. Franco, J. Rasmussen, and J. Nielsen, ''Color classification methods for perennial weed detection in cereal crops,'' in *Proc. Iberoamerican Congr. Pattern Recognit.*, Madrid, Spain. Cham, Switzerland: Springer, 2018, pp. 117–123.
[13] A. Bakhshipour and A. Jafari, ''Evaluation of support vector machine and artificial neural networks in weed detection using shape features,'' *Comput. Electron. Agricult.*, vol. 145, pp. 153–160, Feb. 2018.
[14] W. Kazmi, F. Garcia-Ruiz, J. Nielsen, J. Rasmussen, and H. J. Andersen, ''Exploiting affine invariant regions and leaf edge shapes for weed detection,'' *Comput. Electron. Agricult.*, vol. 118, pp. 290–299, Oct. 2015.
[15] S. Fatma and P. Parimita Dash, ''Moment invariant based weed/crop discrimination for smart farming,'' in *Proc. Int. Conf. Comput., Electr. Commun. Eng. (ICCECE)*, Jan. 2019, pp. 1–5.
[16] S. Chowdhury, B. Verma, and D. Stockwell, ''A novel texture feature based multiple classifier technique for roadside vegetation classification,'' *Expert Syst. Appl.*, vol. 42, no. 12, pp. 5047–5055, Jul. 2015.
[17] C. P. Rojas, L. S. Guzmán, and N. V. Toledo, ''Weed recognition by SVM texture feature classification in outdoor vegetable crops images,'' *Ingeniería e Investigación*, vol. 37, no. 1, pp. 68–74, 2017.
[18] R. Kamath, M. Balachandra, and S. Prabhu, ''Crop and weed discrimination using laws' texture masks,'' *Int. J. Agricult. Biol. Eng.*, vol. 13, no. 1, pp. 191–197, 2020.
[19] A. Bakhshipour, A. Jafari, S. M. Nassiri, and D. Zare, ''Weed segmentation using texture features extracted from wavelet sub-images,'' *Biosyst. Eng.*, vol. 157, pp. 1–12, May 2017.
[20] A. A. Aware and K. Joshi, ''Wavelet based crop detection and automatic spraying of herbicides,'' *Int. J. Innov. Adv. Comput. Sci.*, vol. 4, no. 2, Feb. 2015.
[21] A. A. Omeer and R. R. Deshmukh, ''Improving the classification of invasive plant species by using continuous wavelet analysis and feature reduction techniques,'' *Ecol. Informat.*, vol. 61, Mar. 2021, Art. no. 101181.
[22] Y. Chen, Z. Wu, B. Zhao, C. Fan, and S. Shi, ''Weed and corn seedling detection in field based on multi feature fusion and support vector machine,'' *Sensors*, vol. 21, no. 1, p. 212, Dec. 2020.
[23] R. Kamath, M. Balachandra, and S. Prabhu, ''Paddy crop and weed discrimination: A multiple classifier system approach,'' *Int. J. Agronomy*, vol. 2020, pp. 1–14, Jun. 2020.
[24] F. Lin, D. Zhang, Y. Huang, X. Wang, and X. Chen, ''Detection of corn and weed species by the combination of spectral, shape and textural features,'' *Sustainability*, vol. 9, no. 8, p. 1335, Aug. 2017.
[25] Z. C. Dagdia and C. Zarges, ''A detailed study of the distributed rough set based locality sensitive hashing feature selection technique,'' *Fundamenta Informaticae*, vol. 2020, pp. 1–70, Jun. 2020.
[26] P. Zhu, W. Zuo, L. Zhang, Q. Hu, and S. C. K. Shiu, ''Unsupervised feature selection by regularized self-representation,'' *Pattern Recognit.*, vol. 48, no. 2, pp. 438–446, Feb. 2015.
[27] E. Bijanzadeh, Y. Emam, and E. Ebrahimie, ''Determining the most important features contributing to wheat grain yield using supervised feature selection model,'' *Austral. J. Crop Sci.*, vol. 4, no. 6, pp. 402–407, 2010.
[28] S. Khalid, T. Khalil, and S. Nasreen, ''A survey of feature selection and feature extraction techniques in machine learning,'' in *Proc. Sci. Inf. Conf.*, Aug. 2014, pp. 372–378.
[29] B. Ghojogh, M. N. Samad, S. A. Mashhadi, T. Kapoor, W. Ali, F. Karray, and M. Crowley, ''Feature selection and feature extraction in pattern analysis: A literature review,'' 2019, *arXiv:1905.02845*. [Online]. Available: https://arxiv.org/abs/1905.02845
[30] F. Ahmed, A. Bari, E. Hossain, H. A. Al-Mamun, and P. Kwan, ''Performance analysis of support vector machine and Bayesian classifier for crop and weed classification from digital images,'' *World Appl. Sci. J.*, vol. 12, no. 4, pp. 432–440, 2011.
[31] S. Sabzi, Y. Abbaspour-Gilandeh, and G. García-Mateos, ''A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms,'' *Comput. Ind.*, vol. 98, pp. 80–89, Jun. 2018.
[32] M. H. Siddiqi, S.-W. Lee, and A. M. Khan, ''Weed image classification using wavelet transform, stepwise linear discriminant analysis, and support vector machines for an automatic spray control system,'' *J. Inf. Sci. Eng.*, vol. 30, no. 4, pp. 1–18, 2014.
[33] S. Zhang, J. Guo, and Z. Wang, ''Combing K-means clustering and local weighted maximum discriminant projections for weed species recognition,'' *Frontiers Comput. Sci.*, vol. 1, p. 4, Sep. 2019.
[34] J. Tang, Z. Zhang, D. Wang, J. Xin, and L. He, ''Research on weeds identification based on K-means feature learning,'' *Soft Comput.*, vol. 22, no. 22, pp. 7649–7658, Nov. 2018.
[35] G. Khurana and N. K. Bawa, ''Weed detection approach using feature extraction and KNN classification,'' in *Advances in Electromechanical Technologies*. Singapore: Springer, 2021, pp. 671–679.
[36] D. Saha, G. Hamer, and J. Y. Lee, ''Development of inter-leaves weed and plant regions identification algorithm using histogram of oriented gradient and K-means clustering,'' in *Proc. Int. Conf. Res. Adapt. Convergent Syst.*, 2017, pp. 136–139.
[37] L. Voorhoeve, ''Machine learning for crop and weed classification,'' Wageningen Univ., Final Internship Rep., Jan. 2018.
[38] S. Sabzi, Y. Abbaspour-Gilandeh, and H. Javadikia, ''The use of soft computing to classification of some weeds based on video processing,'' *Appl. Soft Comput.*, vol. 56, pp. 107–123, Jul. 2017.
[39] Y.-A. Chen, W.-H. Hsieh, Y.-S. Ko, and N.-F. Huang, ''An ensemble learning model for agricultural irrigation prediction,'' in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2021, pp. 311–316.
[40] X. Fan, C.-H. Lung, and S. A. Ajila, ''An adaptive diversity-based ensemble method for binary classification,'' in *Proc. IEEE 41st Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, Jul. 2017, pp. 822–827.

[41] A. Narassiguin, M. Bibimoune, H. Elghazel, and A. Aussem, "An extensive empirical comparison of ensemble learning methods for binary classification," *Pattern Anal. Appl.*, vol. 19, no. 4, pp. 1093–1128, Nov. 2016.
[42] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
[43] S. G. T. de Carvalho Santos and R. S. M. de Barros, "Online adaboost-based methods for multiclass problems," *Artif. Intell. Rev.*, vol. 53, no. 2, pp. 1293–1322, Feb. 2020.
[44] E. A. Cortés, M. G. Martínez, and N. G. Rubio, "Multiclass corporate failure prediction by AdaBoost. M1," *Int. Adv. Econ. Res.*, vol. 13, no. 3, pp. 301–312, Jul. 2007.
[45] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting (With discussion and a rejoinder by the authors)," *Ann. Statist.*, vol. 28, no. 2, pp. 337–407, Apr. 2000.
[46] C. Xing, X. Geng, and H. Xue, "Logistic boosting regression for label distribution learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4489–4497.
[47] P. Kadavi, C.-W. Lee, and S. Lee, "Application of ensemble-based machine learning models to landslide susceptibility mapping," *Remote Sens.*, vol. 10, no. 8, p. 1252, Aug. 2018.
[48] H. M. Lalabadi, M. Sadeghi, and S. A. Mireei, "Fish freshness categorization from eyes and gills color features using multi-class artificial neural network and support vector machines," *Aquacultural Eng.*, vol. 90, Aug. 2020, Art. no. 102076.
[49] S. H. Payman, A. Bakhshipour, and H. Zareiforoush, "Development of an expert vision-based system for inspecting rice quality indices," *Qual. Assurance Saf. Crops Foods*, vol. 10, no. 1, pp. 103–114, Mar. 2018.
[50] A. Arefi, A. M. Motlagh, and R. F. Teimourlou, "Wheat class identification using computer vision system and artificial neural networks," *Int. Agrophys.*, vol. 25, no. 4, pp. 1–7, 2011.
[51] K. H. B. Ghazali, J. Ma, R. Xiao, and S. A. lubis, "An innovative face detection based on YCgCr color space," *Phys. Procedia*, vol. 25, pp. 2116–2124, Apr. 2012.
[52] A. R. Pazoki, F. Farokhi, and Z. Pazoki, "Classification of rice grain varieties using two artificial neural networks (MLP and neuro-fuzzy)," *J. Animal Plant Sci.*, vol. 24, no. 1, pp. 336–343, 2014.
[53] S. P. Shouche, R. Rastogi, S. G. Bhagwat, and J. K. Sainis, "Shape analysis of grains of Indian wheat varieties," *Comput. Electron. Agricult.*, vol. 33, no. 1, pp. 55–76, Dec. 2001.
[54] M.-K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inf. Theory*, vol. IT-8, no. 2, pp. 179–187, Feb. 1962.
[55] M. B. H. Rhouma, J. Žunić, and M. C. Younis, "Moment invariants for multi-component shapes with applications to leaf classification," *Comput. Electron. Agricult.*, vol. 142, pp. 326–337, Nov. 2017.
[56] S. Sabzi, Y. Abbaspour-Gilandeh, and J. I. Arribas, "An automatic visible-range video weed detection, segmentation and classification prototype in potato field," *Heliyon*, vol. 6, no. 5, May 2020, Art. no. e03685.
[57] F. Rahimi-Ajdadi, Y. Abbaspour-Gilandeh, K. Mollazade, and R. P. R. Hasanzadeh, "Development of a novel machine vision procedure for rapid and non-contact measurement of soil moisture content," *Measurement*, vol. 121, pp. 179–189, Jun. 2018.
[58] A. Jafari, H. E. Jahromi, S. Mohtasebi, and M. Omid, "Color segmentation scheme for classifying weeds from sugar beet using machine vision," *Int. J. Inf. Sci. Manage.*, vol. 4, no. 1, pp. 1–12, 2012.
[59] H. A. Elnemr, "Feature selection for texture-based plant leaves classification," in *Proc. Int. Conf. Adv. Control Circuits Syst. (ACCS) Syst. Int. Conf. Paradigms Electron. Inf. Technol. (PEIT)*, Nov. 2017, pp. 91–97.
[60] W. Yuan, N. K. Wijewardane, S. Jenkins, G. Bai, Y. Ge, and G. L. Graef, "Early prediction of soybean traits through color and texture features of canopy RGB imagery," *Sci. Rep.*, vol. 9, no. 1, pp. 1–17, Dec. 2019.
[61] Y. Adhitya, S. W. Prakosa, M. Köppen, and J.-S. Leu, "Feature extraction for cocoa bean digital image classification prediction for smart farming application," *Agronomy*, vol. 10, no. 11, p. 1642, Oct. 2020.
[62] M. M. Galloway, "Texture analysis using grey level run lengths," *NASA STI/Recon Tech. Rep. N*, vol. 75, Jul. 1974, Art. no. 18555.
[63] B. E. Park, W. S. Jang, and S. K. Yoo, "Texture analysis of supraspinatus ultrasound image for computer aided diagnostic system," *Healthcare Informat. Res.*, vol. 22, no. 4, p. 299, 2016.
[64] F. R. Ajdadi, Y. A. Gilandeh, K. Mollazade, and R. P. Hasanzadeh, "Application of machine vision for classification of soil aggregate size," *Soil Tillage Res.*, vol. 162, pp. 8–17, Sep. 2016.
[65] M. Khojastehnazhand and H. Ramezani, "Machine vision system for classification of bulk raisins using texture features," *J. Food Eng.*, vol. 271, Apr. 2020, Art. no. 109864.
[66] S. Naeem, A. Ali, C. Chesneau, M. H. Tahir, F. Jamal, R. A. K. Sherwani, and M. U. Hassan, "The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach," *Agronomy*, vol. 11, no. 2, p. 263, 2021.
[67] X. Wei. (2021). *Gray Level run Length Matrix Toolbox*. MATLAB Central File Exchange. (Mar. 11, 2021). [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/17482-gray-level-run-length-matrix-toolbox
[68] P. Yildirim, "Filter based feature selection methods for prediction of risks in hepatitis disease," *Int. J. Mach. Learn. Comput.*, vol. 5, no. 4, p. 258, 2015.
[69] M. B. Shahbaz, X. Wang, A. Behnad, and J. Samarabandu, "On efficiency enhancement of the correlation-based feature selection for intrusion detection systems," in *Proc. IEEE 7th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Oct. 2016, pp. 1–7.
[70] M. A. Thanoon, M. J. M. Zedan, and A. N. Hameed, "Feature selection based on wrapper and information gain," in *Proc. 1st AL-Noor Int. Conf. Sci. Technol. (NICST)*, Oct. 2019, pp. 32–37.
[71] A. Abdollahpouri, S. Rahimi, F. Zamani, and P. Moradi, "A text classification method based on combination of Information gain and graph clustering," *Int. J. Inf. Commun. Technol. Res.*, vol. 11, no. 4, pp. 57–65, 2019.
[72] X. Wang, B. Guo, Y. Shen, C. Zhou, and X. Duan, "Input feature selection method based on feature set equivalence and mutual information gain maximization," *IEEE Access*, vol. 7, pp. 151525–151538, 2019.
[73] R. D. Priyatna and M. Ramli, "K-means algorithm and modification using gain ratio," *IOP Conf. Mater. Sci. Eng.*, vol. 420, no. 1, Sep. 2018, Art. no. 012133.
[74] S. Ravichandran and C. Ramasamy, "Performance comparison of dimensionality reduction methods using MCDR," *Training*, vol. 3, no. 7, pp. 1–6, 2016.
[75] R. C. Holte, "Very simple classification rules perform well on most commonly used datasets," *Mach. Learn.*, vol. 11, no. 1, pp. 63–90, Apr. 1993.
[76] Q. Yanghua, H. Ng, and N. Kapre, "Boosting convergence of timing closure using feature selection in a learning-driven approach," in *Proc. 26th Int. Conf. Field Program. Log. Appl. (FPL)*, Aug. 2016, pp. 1–9.
[77] P. Yariyan, S. Janizadeh, T. Van Phong, H. D. Nguyen, R. Costache, H. Van Le, B. T. Pham, B. Pradhan, and J. P. Tiefenbacher, "Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping," *Water Resour. Manage.*, vol. 34, no. 9, pp. 3037–3053, Jul. 2020.
[78] W. Wang, B. Jiang, S. Ye, and L. Qian, "Risk factor analysis of bone mineral density based on feature selection in type 2 diabetes," in *Proc. IEEE Int. Conf. Big Knowl. (ICBK)*, Nov. 2018, pp. 221–226.
[79] H. K. Darshan, A. R. Shankar, B. S. Harish, and K. K. HM, "Exploiting RLPI for sentiment analysis on movie reviews," *J. Adv. Inf. Technol.*, vol. 10, no. 1, pp. 14–19, 2019.
[80] J. Ma, A. Ovalle, and D. M.-K. Woodbridge, "Medhere: A smartwatch-based medication adherence monitoring system using machine learning and distributed computing," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 4945–4948.
[81] L. Li, W. Jing, and H. Wang, "Extracting the forest type from remote sensing images by random forest," *IEEE Sensors J.*, early access, Dec. 17, 2020, doi: 10.1109/JSEN.2020.3045501.
[82] A. K. Tiwari, A. Nath, K. Subbiah, and K. K. Shukla, "Effect of varying degree of resampling on prediction accuracy for observed peptide count in protein mass spectrometry data," in *Proc. 11th Int. Conf. Natural Comput. (ICNC)*, Aug. 2015, pp. 691–695.

ADEL BAKHSHIPOUR received the Ph.D. degree in biosystems engineering from Shiraz University, in 2016. He is currently working in the Department of Biosystems Engineering, University of Guilan. His main research interests include applications of computer vision, image processing, data mining in agriculture, and non-destructive techniques for food quality assessment.