Automatic Detection and Severity Assessment of Crop Diseases Using Image Pattern Recognition

Liangxiu Han, Muhammad Salman Haleem and Moray Taylor

Abstract Disease diagnosis and severity assessment are necessary and critical for predicting the likely crop yield losses, evaluating the economic impact of the disease, and determining whether preventive treatments are worthwhile or particular control strategies could be taken. In this work, we propose to make advances in the field of automatic detection and diagnosis and severity assessment of crop diseases using image pattern recognition. We have developed a two-stage crop disease pattern recognition system which can automatically identify crop diseases and assess sevrity based on combination of marker-controlled watershed segmentation, superpixel based feature analysis and classification. We have conducted experimental evaluation using different feature selection and classification methods. The experimental result shows that the proposed approach can accurately detect crop diseases (i.e. Septoria and Yellow rust, which are the two most important and major types of wheat diseases in UK and across the world) and assess the disease severity with efficient processing speed.

Keywords Image processing · Machine learning/pattern recognition · Computer vision · Crop disease

L. Han (✉) · M.S. Haleem
School of Computing, Mathematics and Digital Technology,
Manchester Metropolitan University, Manchester, UK
e-mail: l.han@mmu.ac.uk

M.S. Haleem
e-mail: m.haleem@mmu.ac.uk

M. Taylor
The Food and Environment Research Agency, York, UK
e-mail: Moray.Taylor@fera.gsi.gov.uk

L. Chen et al. (eds.), *Emerging Trends and Advanced Technologies for Computational Intelligence*, Studies in Computational Intelligence 647,
DOI 10.1007/978-3-319-33353-3_15

1 Introduction

Food is of key importance in a sustainable society. Plant diseases are a major foe to global food security, which have the potential to significantly decrease crop yields and increase postharvest losses, especially in the face of climate change. It is estimated that almost 40 % of worldwide crops are lost to diseases [1], which may cause devastating economical, social and ecological losses. For instance, yield losses from Septoria diseases of wheat, most prevalent diseases across UK, can range from 30 to 50 % [2].

Disease prevention and management are essential for the sustainability of our society. Effective control measures often rely on the symptom recognition and assessment of the severity of diseases. In the context of plants, a disease refers to any impairment of normal physiological function of a plant. Most plant diseases produce characteristic symptoms or some kind of manifestation in the visible spectrum, which are usually used for disease diagnosis and severity assessment (the proportion of the plant area affected by a disease) through visual inspection. Disease diagnosis and severity assessment are necessary and critical for predicting the likely crop yield loss using yield loss models [3], evaluating the economic impact of a disease, and determining whether preventive treatments are worthwhile or particular control strategies could be taken. There is no point in implementing a control measure if it will cost more than the economic return of the increased crop yield. Currently, the diagnosis and severity assessment are performed visually by trained surveyors, which can lead to inaccuracy and is subject to individual bias as well as being costly and time consuming. This is because the level of detail is variable and inevitably, errors arise from the subjectivity of individuals and the tedious nature of the task. The surveyors have to be regularly trained to maintain quality [4], which is costly. With large areas or fields to be inspected, the scarcity of the trained surveyors makes the monitor of disease a challenging task. Additionally, visual inspection can be destructive if samples are collected in a field for assessment later in the laboratory. Therefore, automatic, accurate and timely diagnosis and quantification of crop diseases are very important.

In this work, we have developed a novel image pattern recognition approach to automatically detect and identify crop diseases (i.e. Septoria and Yellow rust. Two types of most important and major wheat diseases in UK). Some research efforts have been made on automating plant disease identification using image processing approaches, for example, identification of banana leaf diseases [5], disease related coloration in circuit [6], detection of rice leaf disease [7] and quantification of symptoms present in unhealthy plants (e.g. pumpkin, pepper, bean) [8] using image color analysis, which tried to discriminate the diseases based on color difference. The method in [9] tried to discriminate a given disease from other pathologies of rubber tree leaves based on PCA and Neural networks, which directly applied to the RGB values of the pixels of a low resolution (15×15 pixels) and does not employ any image segmentation techniques. The system [10] was proposed to monitor the health vineyard, which mainly used thresholding to separate diseased leaves and ground from healthy leaves using both RGB and HSV color representation of the image

1 Introduction

Food is of key importance in a sustainable society. Plant diseases are a major foe to global food security, which have the potential to significantly decrease crop yields and increase postharvest losses, especially in the face of climate change. It is estimated that almost 40 % of worldwide crops are lost to diseases [1], which may cause devastating economical, social and ecological losses. For instance, yield losses from Septoria diseases of wheat, most prevalent diseases across UK, can range from 30 to 50 % [2].

Disease prevention and management are essential for the sustainability of our society. Effective control measures often rely on the symptom recognition and assessment of the severity of diseases. In the context of plants, a disease refers to any impairment of normal physiological function of a plant. Most plant diseases produce characteristic symptoms or some kind of manifestation in the visible spectrum, which are usually used for disease diagnosis and severity assessment (the proportion of the plant area affected by a disease) through visual inspection. Disease diagnosis and severity assessment are necessary and critical for predicting the likely crop yield loss using yield loss models [3], evaluating the economic impact of a disease, and determining whether preventive treatments are worthwhile or particular control strategies could be taken. There is no point in implementing a control measure if it will cost more than the economic return of the increased crop yield. Currently, the diagnosis and severity assessment are performed visually by trained surveyors, which can lead to inaccuracy and is subject to individual bias as well as being costly and time consuming. This is because the level of detail is variable and inevitably, errors arise from the subjectivity of individuals and the tedious nature of the task. The surveyors have to be regularly trained to maintain quality [4], which is costly. With large areas or fields to be inspected, the scarcity of the trained surveyors makes the monitor of disease a challenging task. Additionally, visual inspection can be destructive if samples are collected in a field for assessment later in the laboratory. Therefore, automatic, accurate and timely diagnosis and quantification of crop diseases are very important.

In this work, we have developed a novel image pattern recognition approach to automatically detect and identify crop diseases (i.e. Septoria and Yellow rust. Two types of most important and major wheat diseases in UK). Some research efforts have been made on automating plant disease identification using image processing approaches, for example, identification of banana leaf diseases [5], disease related coloration in circuit [6], detection of rice leaf disease [7] and quantification of symptoms present in unhealthy plants (e.g. pumpkin, pepper, bean) [8] using image color analysis, which tried to discriminate the diseases based on color difference. The method in [9] tried to discriminate a given disease from other pathologies of rubber tree leaves based on PCA and Neural networks, which directly applied to the RGB values of the pixels of a low resolution (15×15 pixels) and does not employ any image segmentation techniques. The system [10] was proposed to monitor the health vineyard, which mainly used thresholding to separate diseased leaves and ground from healthy leaves using both RGB and HSV color representation of the image

and morphological operations. The method [11] was to detect and differentiate two diseases that affect rice crops (i.e. blast and brown spots). It first used an entropy-based thresholding and the edge detection to segment images, and then the intensity of green components was used to detect spots. The spots were resized to 80×100 pixels and the pixel values are extracted as features for final classification using self organising map (SOM).

For the purpose of automatic disease recognition from images, it is very important to extract the features and find the most discriminative features for efficient disease recognition by developing appropriate image processing and machine learning approaches.

To the best of our knowledge, there is currently no existing work for solving our problem i.e. automatically detecting crop diseases of Septoria and Yellow rust. Additionally, the existing methods are not suitable and can not be directly applied to our case. Therefore, we have developed a new approach to automatic disease detection using marker-controlled watershed segmentation, superpixel based feature analysis and classification. The proposed system can take an image as input and automatically classify it into different categories of diseases and also assess how severe that type of disease is. Our work provides added value and complement the existing research efforts. Our contributions lie in:

- Development of a new way to extract features by combining marker controlled watershed segmentation and superpixels to avoid over segmentation, reduce the complexity of subsequent image processing tasks, thus producing high quality segmentation;
- A comprehensive comparison study using different feature selection and classification algorithms;
- A prototype of two-stage crop disease pattern recognition system. The first stage determines the severity level whereas the second stage detects the disease type in a crop leaf.

The remainder of this paper is organised as follows. Section 2 describes the dataset used, Sect. 3 describes our proposed methods, which includes algorithms and methods for feature extraction, feature selection and classification. In Sect. 4, we have conducted experimental on real datasets. Section 5 concludes the proposed work and highlights the future work.

2 Ground Truth

The data were collected by domain experts in the area of food and agriculture, which were captured in real fields using the camera phone. Each image has a dimension of 3264×2448. We have two types of dataset i.e., the training set and the testing set. The training dataset is composed of 20 leaf images each diseased with ‘Septoria’ and ‘Yellow Rust’. These images have been annotated around diseased regions. In

this way, we have 40 images in our training set. For the testing dataset, we have the images from 57 leaves diseased with 'Septoria' and 60 leaves diseased with 'Yellow Rust'. The test images have been annotated with their disease severity level.

3 The Proposed Approach

A high-level overview of our classification system for disease diagnosis is shown in Fig. 1. Our case is two-stage classification framework in which the first stage determines the severity level by classification between healthy and unhealthy regions in a wheat leaf. The second stage is to determine the type of disease in a wheat leaf. During the training procedure, we need to train the classification model based on a training set of images with human annotations. At the testing procedure the performance of the classification model has to be evaluated in terms of accuracy. Finally, when the performance is satisfactory, at the deployment procedure the model has to be deployed to perform the classification of images without annotation. This work mainly focuses on the training, testing and evaluation procedures.

In most cases, crop diseases exhibit a range of visual symptoms on the leaves such as colored spots, streaks, etc. These visual symptoms continuously change their color, shape and size as disease progress. It is very important to identify features of a certain disease, extract the most discriminative features and then build a classification model using suitable image processing and machine learning approaches. The feature extraction procedure has been briefed as follows which has been explained in detail in our previous paper [12].

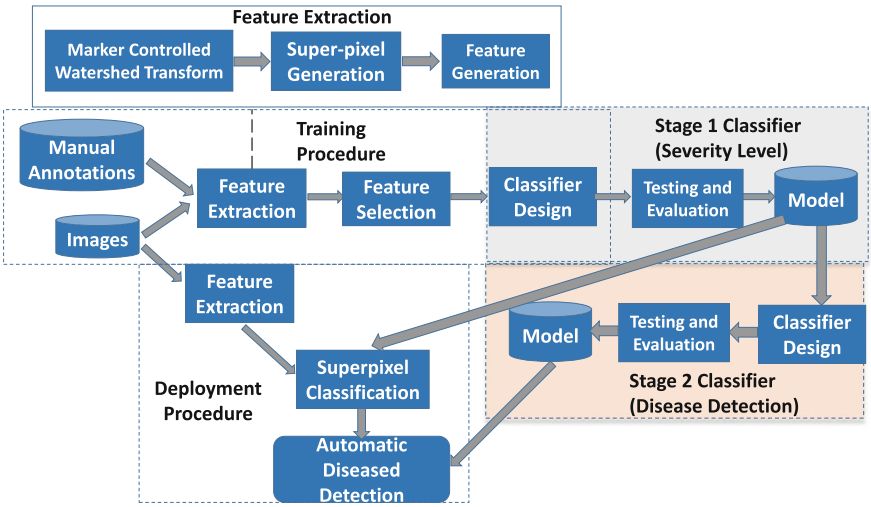

Fig. 1 A high-level overview of classification for disease diagnosis

this way, we have 40 images in our training set. For the testing dataset, we have the images from 57 leaves diseased with 'Septoria' and 60 leaves diseased with 'Yellow Rust'. The test images have been annotated with their disease severity level.

3 The Proposed Approach

A high-level overview of our classification system for disease diagnosis is shown in Fig. 1. Our case is two-stage classification framework in which the first stage determines the severity level by classification between healthy and unhealthy regions in a wheat leaf. The second stage is to determine the type of disease in a wheat leaf. During the training procedure, we need to train the classification model based on a training set of images with human annotations. At the testing procedure the performance of the classification model has to be evaluated in terms of accuracy. Finally, when the performance is satisfactory, at the deployment procedure the model has to be deployed to perform the classification of images without annotation. This work mainly focuses on the training, testing and evaluation procedures.

In most cases, crop diseases exhibit a range of visual symptoms on the leaves such as colored spots, streaks, etc. These visual symptoms continuously change their color, shape and size as disease progress. It is very important to identify features of a certain disease, extract the most discriminative features and then build a classification model using suitable image processing and machine learning approaches. The feature extraction procedure has been briefed as follows which has been explained in detail in our previous paper [12].

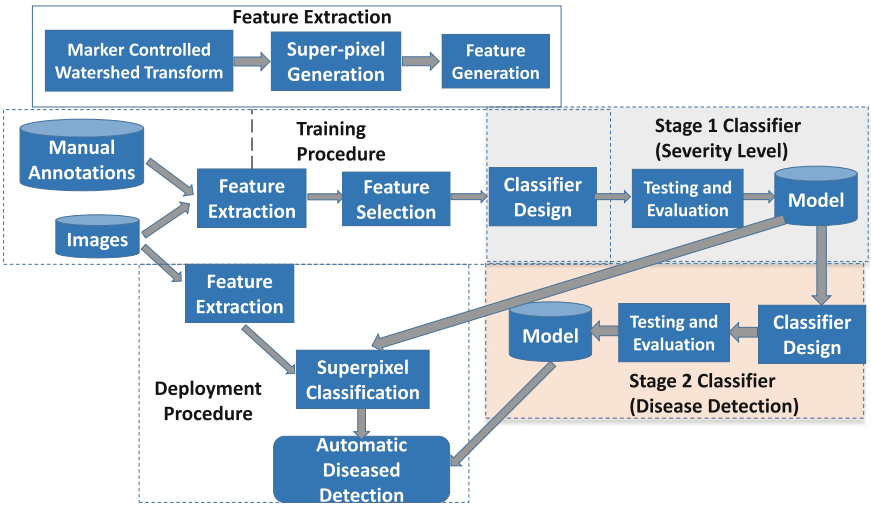

Fig. 1 A high-level overview of classification for disease diagnosis

3.1 Leaf Extraction from the Complicated Background

Our image samples are taken from real fields. Each image usually contains more than one object including foreground objects (e.g. target leaves) and background objects (e.g. non-target leaves, soil, racemes and other interference), as shown in Fig. 3a. It is critical to extract a target plant leaf first and then identify the disease patterns from that leaf using image processing approaches. Most of existing approaches for leaf extraction require manual operations or prior information and images taken in controlled environments and cannot deal with complicated background. We have first used marker-controlled watershed segmentation [13, 14] to separate the target leaf from the background. The marker-controlled watershed algorithm uses operations of mathematical morphology to place foreground markers in the blob and background markers for the areas without blobs. The details have been summarized in [12] but a block diagram representing the leaf extraction from a complicated background has been shown in Fig. 2. Figure 3b shows the result of marker-controlled watershed segmentation.

3.2 Superpixel Generation

After the target leaf is extracted, we need to further extract disease features from that target leaf. We have applied superpixels method to the target leaf for computing local image features. Superpixels [15] group pixels of an image into small meaningful regions and each region is equivalent to a pixel (A superpixel refers to a group of pixels which have similar characteristics.). The feature vector is calculated for

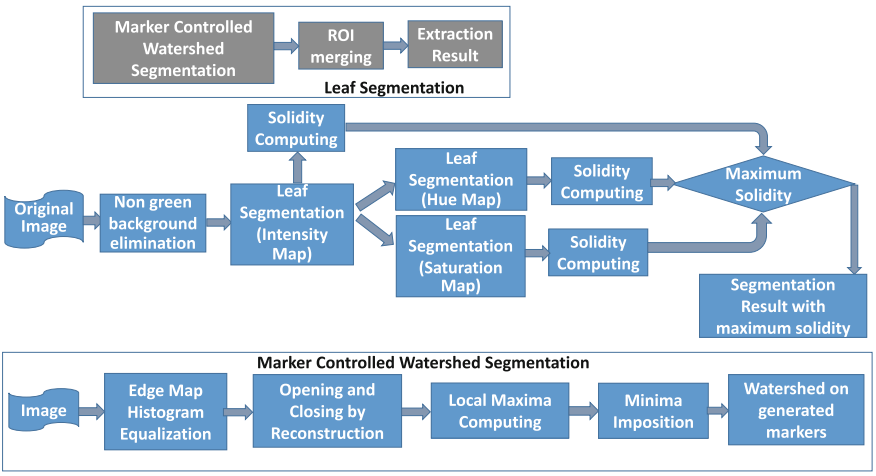

Fig. 2 Leaf extraction from complicated background

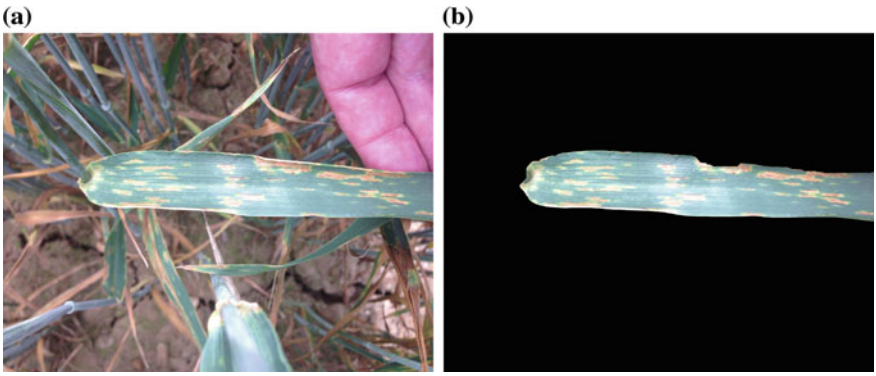

Fig. 3 **a** An image sample from a real crop field and **b** extraction of the target leaf from the background using marker controlled watershed transformation

each superpixel rather than pixel for high computational efficiency. It can capture redundancy in the image, reduce the complexity of subsequent image processing tasks, and produce high quality segmentations. There are various ways for superpixel grouping [16–18]. In this work, we have adopted Simple Linear Iterative Clustering (SLIC) superpixels [19]. The basic idea of SLIC superpixels is that it decomposes an image in visually homogeneous regions based on a spatially localized version of k-means clustering. Each pixel is associated to a feature vector (grayscale values in our case):

$$\Psi(x, y) = \begin{bmatrix} \lambda x \\ \lambda y \\ I(x, y) \end{bmatrix}$$

and then k-means clustering is applied. Where the coefficient λ is defined as

$$\lambda = \frac{\texttt{regularizer}}{\texttt{regionSize}}$$

where regularizer represents a nominal size of the superpixel regions and regionSize represents the strength of the spatial regularization. The image is first divided into a grid with step regionSize. The center of each grid tile is then used to initialise a corresponding k-means. The distance vector is calculated in terms of intensity values and pixel positions. After assigning each pixel to the nearest centre, an updating step adjusts the cluster centres to the mean of pixels in this group. The residual error is then calculated between new cluster centre and previous cluster centre. Such iterations continue until the error convergences. Figure 4 shows the image after using superpixels. By using superpixels, the original features from 3264×2448 can be reduced to 10,000, which dramatically increases speed and improves the quality of the results.

Fig. 4 The resulting image based on superpixels

3.3 Feature Extraction

Based on the resulting image of superpixels, we have extracted textural, gradient, gabor, and biologically inspired features for each of the superpixels in the images. The details of those features are as follows:

3.3.1 Textural Features

We have focused on textural features based on Grey Level Co-occurrence Matrix (GLCM) [20] (also called Haralick features). It is used for determining how often a pixel of a grey scale value i occurs adjacent to a pixel of the value j. Four angles for observing the pixel adjacency i.e. $\theta = 0°, 45°, 90°, 135°$ are used. GLCM also needs an offset value **D** which defines pixel adjacency by certain distance. In our case, offset value is set to 1. The features, which are calculated using GLCM matrix are summarised in [12]. The mean value in each direction was taken for each Haralick feature and they were calculated from both red and green channels.

3.3.2 Gradient Features

The reason for including gradient features was illumination non-uniformity of the artefacts. In order to calculate these features, the response from Gaussian filter bank is calculated. The equation of Gaussian filter bank is given as:

$$\mathcal{N}(\sigma, i, j) = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}}$$

The Gaussian filter bank includes Gaussian $\mathcal{N}(\sigma)$, its two first order derivatives $\mathcal{N}_x(\sigma)$ and $\mathcal{N}_y(\sigma)$ and three second order derivatives $\mathcal{N}_{xx}(\sigma)$, $\mathcal{N}_{xy}(\sigma)$ and $\mathcal{N}_{yy}(\sigma)$ in horizontal (x) and vertical (y) directions at different scales (σ). After convolving the image with the filter bank at a particular channel, the mean value is taken over of each filter response over all pixels of each superpixel.

For determining features at different smoothing scales, both red and green channels of images are convolved with the Gaussian at scales $\sigma = 1, 2, 4, 8, 16$.

3.3.3 Gabor Features

We have also applied Gabor filter [21, 22] for extracting features based on the following equation:

$$Gb(x, y, \theta, f, \sigma_x, \sigma_y) = \exp(-\frac{1}{2}(\frac{x^{'2}}{\sigma_x^2} + \frac{y^{'2}}{\sigma_y^2})) * \exp(i2\pi f x) \tag{1}$$

$$x^{'} = xcos\theta + ysin\theta \tag{2}$$

$$y^{'} = ycos\theta - xsin\theta \tag{3}$$

x and y are image pixel coordinates. Here we have varied $\sigma x = \sigma y = [1,2,4,8,16]$, $f = [0.333,0.5,1,0.5,0.333]$, $\theta = [0, 45, 90, 135]$.

3.3.4 Biologically Inspired Features

The Biologically Inspired Features reflect the process of the human visual cortex in image recognition tasks [23–25]. Here the crop leaf image is filtered in a number of low-level visual 'feature channels' at multiple spatial scales, for features of colour, intensity, flicker etc. It consists of 6 *Intensity Units I* and 12 *Colour Units* representing neuronal working of mammals [26]. The *Colour Units* has 4 different colour channels representing excitation as well as inhibition of colours. The excitation and inhibition of different colours make two pairs. They are formed as red-green (R-G) and blue-yellow (B-Y) channels. By using dyadic Gaussian pyramids [27] convolved on the intensity channel of an input colour image, nine spatial scales are generated with a ratio from 1:1 (scale 0) to 1:256 (scale 8). To get intensity feature maps, the centre-surround [28] operation is performed between centre levels ($c = 2, 3, 4$) and surround levels ($s = c + d$ with $d = 3, 4$ i.e., six feature maps are computed at levels of 2–5, 2–6, 3–6, 3–7, 4–7, and 4–8. Because scales are different between centre levels and surround levels, maps of surround levels are interpolated to the same size as the corresponding centre levels and are subtracted point-by-point from the corresponding centre levels to generate the relevant feature maps. For each region, we have generated the mean response of these features as:

$$I^{reg}(c,s) = \sum_{i}^{N} \frac{I(c,s,n)}{N}$$
$$RG^{reg}(c,s) = \sum_{i}^{N} \frac{RG(c,s,n)}{N} \tag{4}$$
$$BY^{reg}(c,s) = \sum_{i}^{N} \frac{BY(c,s,n)}{N}$$

where N is number of pixels in the region and,

$$\begin{aligned} I(c,s) &= |I(c) - Interp_{s-c}I(s)| \\ RG(c,s) &= |(R(c) - G(c)) - Interp_{s-c}(R(s) - G(s))| \\ BY(c,s) &= |(B(c) - Y(c)) - Interp_{s-c}(B(s) - Y(s))| \end{aligned} \tag{5}$$

where $Interp_{s-c}$ represent interpolation to $s - c$ level.

3.4 Feature Selection

After determination of features, we have performed feature selection on the features calculated in Sect. 3.3 so as to determine features most relevant towards classification. We have selected features for two-stage classification. For the first stage, we have selected features for classification between the diseased area and non-diseased area of the crop leaf extracted in Sect. 3.1 so as to determine severity level of the disease. For the second stage, we have performed classification between diseased region for 'Septoria' and 'Yellow Rust' so as to determine the type of disease. In our case, we have applied wrapper feature selection [29] which is an iterative procedure of feature selection until the certain task performance is maximized. For quantification of task performance, we have area under the Receiving Operating Characteristics (ROC), also known as Area Under the Curve (AUC). ROC is a graphical plot that illustrates the performance of a binary classifier system by area under it as it is created by plotting the true positive rate against the false positive rate at various threshold settings [30]. In feature selection procedure, the feature with the highest AUC is selected in the first iteration. During the next iterations, the feature which in together with previously selected features result in highest AUC. This process is repeated until there is little or no improvement. In this way there are 40 features selected for the feature set.

We have compared the performance of the features selected under the proposed criteria with other feature selection methods. i.e. Fisher Ratio [31], Quadratic Discriminant Analysis (QDA) [31], Maximum Relevancy Minimum Redundancy (mRMR) [32] and Fast Correlation-Based Filter (FCBF) [33].

Fisher Ratio is the linear discriminatory analysis which maximizes the distance between classes while minimizing the variance within each class.

Quadratic Discriminant Analysis (QDA) is same as Fisher ratio except that covariance matrix may not be identical for each class.

Maximum Relevancy Minimum Redundancy (mRMR) selects the features with highest mutual information yet, minimum redundancy among them.

Fast Correlation-Based Filter (FCBF) selects the features based on heuristics that good feature subsets contain features highly correlated with the class, yet uncorrelated with each other.

The comparison has been performed both in terms of severity level as well as disease detection procedures. The feature selection procedures such as wrapper-AUC and mRMR have been forced to stop as there has been little or no difference in their classification performance on the training set. Table 1 summarizes the total number of features selected by each feature selection procedure for both classifier stages as well as their attributes if the feature selection algorithm was converged or forced to stop. The table as well as the feature selection procedure shown in Fig. 5 shows that wrapper-AUC can perform well on least numbers of features selected compared to other feature selection procedures. The other feature selection procedure such as mRMR can perform well but their progress with number of selected features is slow. Although other feature selection procedures have been selected after improvement in their respective classification performance (e.g. Fisher Ratio in case of wrapper-LDA); their feature selection performance on Fig. 5 has been compared in terms of AUC so as to keep on one platform.

The performance has been compared with other feature selection methods in terms of ROC curves as shown in Fig. 6. We have also performed comparison with the feature set with having all features in the training set (340 features altogether).

Table 1 Number of features selected by each feature selection method for both stage 1 and stage 2 classifiers

Feature selection procedures	Stage 1 severity level		Stage 2 disease detection	
	Number of features selected	Attributes	Number of features selected	Attributes
Wrapper-AUC	40	Forced to stop	32	Forced to stop
Wrapper-LDA	23	Algorithm converged	28	Algorithm converged
Wrapper-QDA	19	Algorithm converged	17	Algorithm converged
mRMR	40	Forced to stop	40	Forced to stop
FCBF	3	Algorithm converged	4	Algorithm converged

The attributes reflect either forced stopping or the algorithm convergence of the feature selection procedure. The procedures which are forced to stop have little or no improvement while increase in feature set

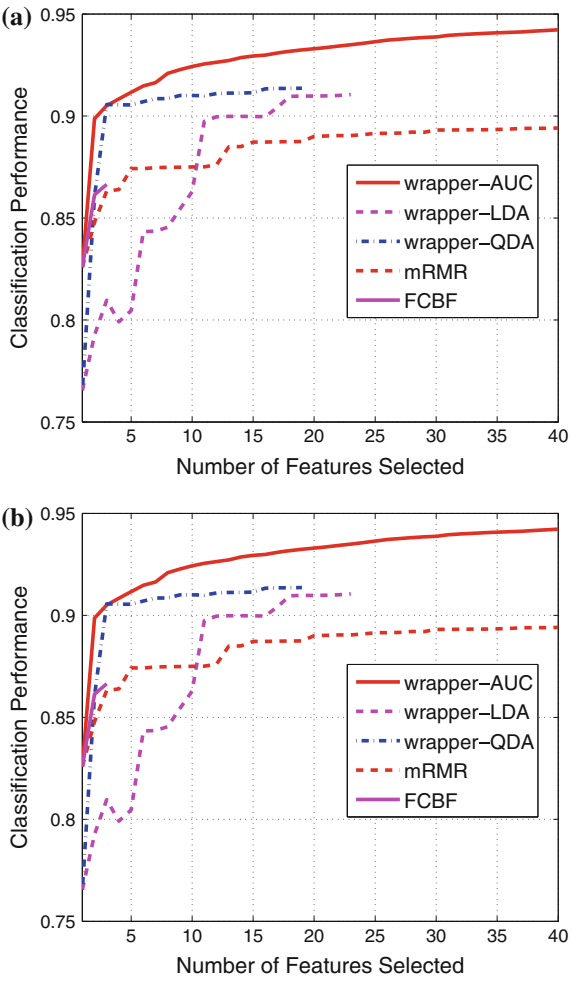

Fig. 5 Comparison of feature selection procedures by increase in classification performance with number of selected features. **a** Stage-1 classifier of disease severity detection and **b** stage-2 classifier of disease detection. The classification performance have been calculated in terms of AUC

The ROC curves have been generated by 5-fold cross validation on the training set. The results show significant dominance of feature set selected by wrapper feature selection maximizing AUC. In both the cases, the 'wrapper-AUC' approach has not only outperformed its counterparts but also its performance has been near to the feature set which consists of all features. This can be quite useful in determination of classifier as classifiers with large number of features have significantly reduced computational efficiency both in terms of training as well as testing.

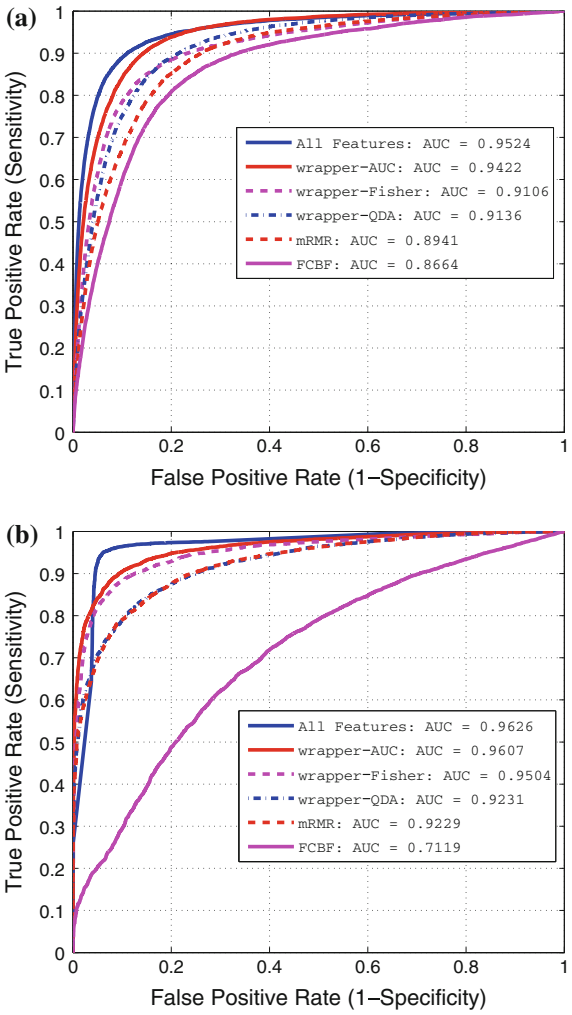

Fig. 6 Comparison of feature selection procedures in terms of ROC curves with **a** comparison of *green* area in the leaf with disease area and **b** comparison of septoria versus *yellow* rust. The result show that wrapper-AUC have performed significantly better compared to other two counterparts while performing nearest to the feature set with having all features

3.5 *Classifier Design for Severity Level and Disease Diagnosis*

Based on the feature set obtained from the above, we have developed the classifier for both stage 1 and stage 2 based on support vector machines (SVM) [34–36]. An SVM function can be defined as:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) - b$$

where

- x_i is the vectors of the training set.
- N is the number of training samples used.
- α_i is a scalar, a real number that takes values between 0 and C.
- y_i is either 1 or -1, indicating the class to which the point $\mathbf{x}_i$ belongs.
- b is a scalar that shifts the output of the SVM by a constant
- $K(\mathbf{x}, \mathbf{x}_i)$ is the kernel function. A kernel is a function that takes two vectors as inputs and produces a single scalar value which is positive.

There are different types of SVM depending upon the kernel function. We decide the kernel function based upon type of data as well as with few numerical difficulties. In our case, the obvious choice was Linear SVM (LSVM) where value of kernel is set to 1. The LSVM training is based on separating the two classes by linear hyperplane. If $\mathbf{x}$ belongs to class +1, $f(x)$ is larger or equal than 1. When $\mathbf{x}$ belongs to class -1, $f(x)$ is smaller or equal than -1. The α_i values and the b value are selected to match these requirements. As for the parameters C and γ, these are meta parameters which are chosen based on past performance on a training window. We need to replicate this procedure carefully to validate the model without incurring forward-looking bias.

For the comparative study, we have different types of SVMs such as Radial Based Function (RBF)-SVM, polynomial SVM and sigmoid SVM [37]. We have also performed comparison with other classifiers such as Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Artificial Neural Networks (ANN) [38]. For RBF kernel, we have chosen the kernel value equal to $k(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2 / M)$. For ANN, we have one hidden layer with 20 neurons as there has been maximum classification accuracy compared to other ANN specifications.

The comparative results of LSVM with other classifiers have been presented in Fig. 7. For both stage 1 and stage 2, the results have been generated by performing 5-fold cross validation on the training set across different feature sets generated in Sect. 3.4 and different classifiers. The results show that feature set selected by wrapper-AUC approach has performed significantly better compared to other feature selection procedures for all of the classifiers except QDA. This shows that wrapper-AUC was an obvious choice for selecting features for both stage 1 as well as stage

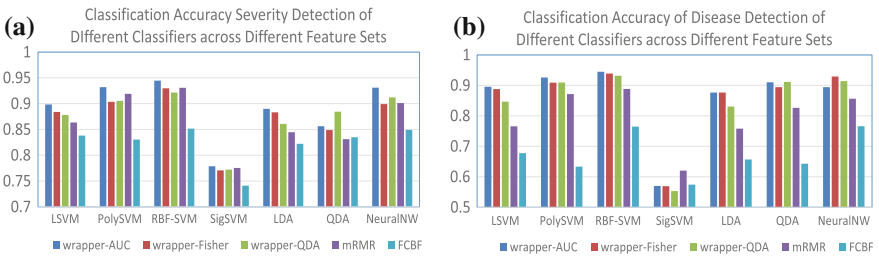

Fig. 7 Comparative results of different classifiers in stage 1 and stage 2 across different feature sets

2 classifier. Among different classifiers, all of the classifiers except sigmoid SVM have performed exceptionally well not only for severity detection (stage 1) but also for disease diagnosis (stage 2). Therefore, we omitted Sigmoid SVM for testing on our test set. As far as other classifiers are concerned, we have selected the feature set in the testing stage for each classifier which has performed the best in the respective category e.g. wrapper-AUC has performed the best for all of the classfiers except QDA, therefore, we have selected wrapper-AUC feature set for these classifiers in the testing stage. For QDA, we have selected the feature set selected by wrapper-QDA for both stage 1 and stage 2.

4 Experimental Evaluation

We have performed experimental evaluation on our test set which is composed 117 images in which 57 wheat leaf images have been taken from 'Septoria' while 60 images have been taken from 'Yellow Rust'. As stated in Sect. 3.5, we have constructed the LSVM classifier on the feature set selected by wrapper-AUC approach. For the purpose of comparison study on the test set, we have developed other classifiers on the basis of feature sets which have performed the best in the cross-validation stage in Fig. 7. We have calculated the error between the estimated severity level and annotated severity level. The severity level can be estimated by calculating the degree of overlap of the disease region with the crop leaf. We have Dice Coefficient to measure the disease severity which is the degree of overlap between the automatically determined diseased regions (D_1) and the automatically extracted leaf region (D_2) using our proposed approach. The mathematical representation of Estimated Severity (ES) is described as follows:

$$ES = \frac{2|D_1 \cap D_2|}{|D_1 + D_2|} \tag{6}$$

and therefore, the absolute difference of the particular image i can be represented as

$$E_i = |AS_i - ES_i| \tag{7}$$

We have also calculated Mean Absolute Difference (MAD) for estimating severity levels in both 'Septoria' and 'Yellow Rust' which can be represented as:

$$MAD = \frac{\sum_i E_i}{n} \tag{8}$$

where n is total number of images.

According to our collaborator i.e. Fera, the absolute difference should not be more than 10 units between the estimated severity level and the annotated severity level.

Table 2 Severity assessment across different classifiers in terms of MAD

	LSVM	RBF-SVM	PolySVM	LDA	QDA	ANN
MAD-Septoria	8.04	13.42	11.79	9.43	7.42	11.29
MAD-Yellow Rust	10.28	20.37	20.14	11.65	10.75	19.12

Table 2 represent the severity assessment across different classifiers with MAD for both 'Septoria' and 'Yellow Rust'. By observing Table 2 and Fig. 7, we come to know that our LSVM has performed significantly better compared to its other counterparts on the test set despite of being little bit inferior performance compared to RBF-SVM and PolySVM on the training set. This shows that determining the classifier model by over estimating on the training set can actually depreciate the classifier performance on the test set. The MAD for LSVM in 'Septoria' and 'Yellow Rust' is 8.04 and 10.28 which is at or below 10 units. Although the performance of QDA in severity level assessment was comparable to LSVM however, its performance with respect to disease detection was not as good as LSVM.

The performance of stage-2 classifier has been evaluated in Table 3. The stage-2 classifier has been applied on the diseased area extracted by stage-1 classifier. The disease area has been detected as either 'Septoria' or 'Yellow Rust' if the higher number of pixels have been classified as the respective class. Like stage-1, the stage-2 classifier also shows significantly better performance for LSVM compared to other classifiers.

Some visual results of automatic disease recognition are shown in Fig. 8. As shown in Fig. 8, our system can automatically detect this image which has Septoria as well as Yellow Rust diseases with severity level around comparable to the manual estimation from experts. We have also calculated computational time for our proposed approach as shown in Fig. 9. The software was run on a normal working computer (Hardware specification: CPU Intel Core i7-3630 QM, 2.4G 4 cores/8 logical processers; Memory: 8 GB (Physical)/16 GB (Virtual)). The total execution time is around 1.5 min.

Table 3 Accuracy of disease detection

	LSVM (%)	RBF-SVM (%)	PolySVM (%)	LDA (%)	QDA (%)	ANN (%)
Septoria	78.94	82.45	80.71	61.41	64.91	61.41
Yellow Rust	91.67	61.67	70	86.67	85	91.67
Overall	85.47	71.79	75.21	74.35	75.21	76.92

The second stage classifier has been applied after extracting disease area in the crop leaf

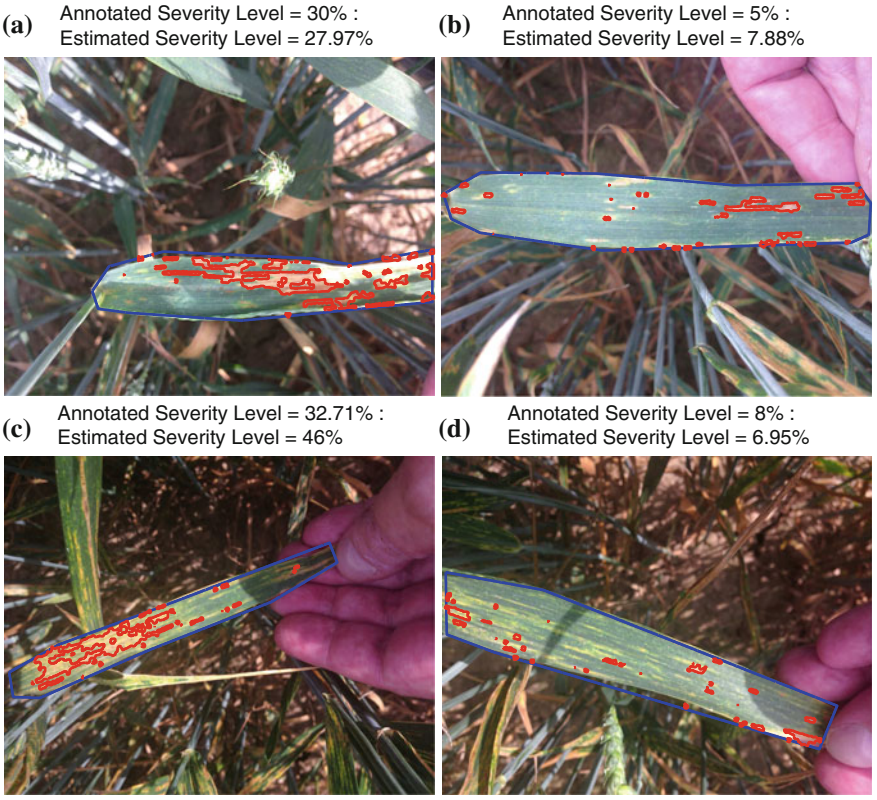

Fig. 8 Visual results of automatic septoria disease and yellow rust severity detection of crops with **a** and **b** examples from Septoria disease leaves whereas **c** and **d** are examples from Yellow Rust

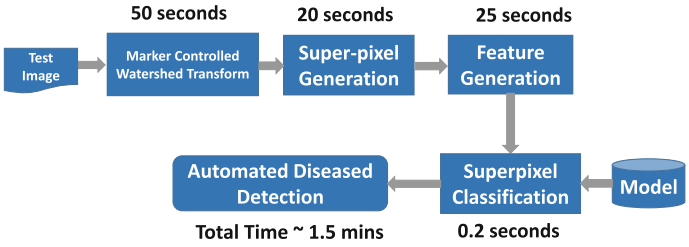

Fig. 9 The computational time for each process

5 Conclusion and Future Work

Accurate detection and identification of crop diseases plays an important role in effectively controlling and preventing diseases for sustainable agriculture and food security. Effective control measures rely on the symptom recognition and assessment

of severity of diseases. In vast majority of cases, diseases exhibit a range of visual symptoms such as colored spots or streaks. Currently, the identification of diseases and their severity assessment are mainly achieved by visual observation and estimation when crop walking, which is costly and time consuming and may lead to inaccuracy and bias as well. In this paper, we have developed a novel image pattern recognition system for automatic diagnosis and severity assessment of crop diseases. There are four major contributions:

- We have developed new feature extraction methods based on combination of marker controlled watershed segmentation and superpixels methods for avoiding over segmentation, reducing the complexity of subsequent image processing tasks, and improving the quality of the results;
- We have conducted a comprehensive study on different feature selection methods (i.e. Fisher Ratio, Quadratic Discriminant Analysis (QDA), Maximum Relevancy Minimum Redundancy (mRMR) and Fast Correlation-based Filter (FCBF) and classification approaches (i.e. LSVM, RBF-SVM, PolySVM, LDA, QDA, and ANN);
- We have prototyped a two-stage crop disease pattern recognition system. The first stage determines the severity level whereas the second stage detects the disease type in a crop leaf.

The data used in the experimental evaluation were collected and captured by domain experts in the crop fields. The images are 157 in total. 40 images were used for training models and the rest of 117 images for testing. The experimental result demonstrates that our approach performs well on disease recognition and severity assessment. The preliminary result from this work is promising.

The future will be to further look into the feature extraction of diseases and collect more data from different seasons for improving robustness of algorithms. Additionally, to enable to process more images in real time, we will also develop new algorithms for parallel processing images.

Acknowledgments The work reported in this paper has formed part of the COPE project, which is funded by Sustainable Society Network+ (RCUK digital programme) and Manchester Metropolitan University in UK. This project is in collaboration with Fera (Food and Environment Research Agency). The authors acknowledge Ms. Sharon Elcock at the Fera for her support.

References

1. Strange, R.: Food security **2**(2), 111 (2010). doi:10.1007/s12571-010-0064-5
2. Basf, A.: Managing septoria in wheat (2014). http://www.agricentre.basf.co.uk/
3. Madden, I.M., Hughes, L.V.: Phytopathology **90**, 788 (2000)
4. Domiciano, G.P., Duarte, H.S.S., Moreira, E.N., Rodrigues, F.A.: Plant Pathol. **63**, 922 (2014)
5. Camargo, A., Smith, J.S.: Comput. Electr. Agric. **66**(2), 121 (2009)
6. Pydipati, R., Burks, T., Lee, W.: Comput. Electr. Agric. **52**(1–2), 49 (2006)
7. Pugoy, R.A.D.L., Mariano, V.Y.: In: Third International Conference on Digital Image Processing (ICDIP 2011), vol. 8009, pp. F1–F7 (2011)

8. Contreras-Medina, L.M., Osornio-Rios, R.A., Torres-Pacheco, I., de J. Romero-Troncoso, R., Guevara-González, R.G., Millan-Almaraz, J.R.: Sensors **12**(1), 784 (2012)
9. Abdullah, N., Rahim, A., Hashim, H., Kamal, M.: In: 2007 5th Student Conference on Research and Developmen, pp. 1–6 (2007)
10. Lloret, J., Bosch, I., Sendra, S., Serrano, A.: Sensors **11**(6), 6165 (2011)
11. Phadikar, S., Sil, J.: In: 11th International Conference on Computer and Information Technology, pp. 420–423. IEEE (2008)
12. Han, L., Haleem, M.S., Taylor, M.: In: Science and Information Conference (SAI) (2015)
13. Wang, X., Huang, D., Du, J., Xu, H., Heutte, L.: J. Appl. Math. Comput. **205**, 916 (2008)
14. Tang, X., Liu, M., Zhao, H., Tao, W.: In: Proceedings of IEEE International Congress on Image and Signal Processing, pp. 1–5 (2009)
15. Ren, X., Malik, J.: In: 9th IEEE International Conference on Computer Vision, pp. 10–17. IEEE (2003)
16. Felzenszwalb, P., Huttenlocher, D.: Int. J. Comput. Vis. **59**(2), 167 (2004)
17. Shi, J., Malik, J.: IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 888 (2000)
18. Veksler, O., Boykov, Y., Mehrani, P.: In: European Conference on Computer Vision (ECCV) (2010)
19. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274 (2012)
20. Haralick, R.M., Shanmugam, K., Dinstein, I.: IEEE Trans. Syst. Man Cybern. **3** (1973)
21. Gabor, D.: J. Inst. Electr. Eng. **93**(26), 429 (1946)
22. Daugman, J.G.: IEEE Trans. ASSP **36**(7), 1169 (1988)
23. Tian, Y., Kanade, T., Cohn, J.: IEEE Trans. Pattern Anal. Mach. Intell. **23**(2), 97 (2001)
24. Siagian, C., Itti, L.: IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 300 (2007)
25. Song, D., Tao, D.: IEEE Trans. Image Process. **19**(1), 174 (2010)
26. Itti, L., Koch, C., Niebur, E.: IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1254 (1998)
27. Adelson, E.H., Anderson, C.H., Bergen, J.R., Burt, P.J., Ogden, J.M.: RCA Eng. **29**, 33 (1984)
28. Sun, S.G., Kwak, D.M.: J. Multimed. **1**, 16 (2006)
29. Serrano, A.J., Soria, E., Martin, J.D., Magdalena, R., Gomez, J.: In: The International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (2010)
30. Alberg, A.J., Park, J.W., Hager, B.W., Brock, M.V., Diener-West, M.: J. General Intern. Med. **19**, 460 (2004)
31. Duda, R.O., Hart, P.E., Stork, D.G. (eds.): Pattern Classification. Wiley-Interscience, New York (2000)
32. Peng, H., Long, F., Ding, C.: IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1226 (2005)
33. Yu, L., Liu, H.: In: Proceedings of the Twentieth International Conference on Machine Learning, pp. 856–863 (2003)
34. Vapnik, V.N.: The nature of statistical learning theory. Springer, New York (1995). ISBN 0-387-94559-8
35. Shevade, S., Keerthi, S., Bhattacharyya, C., Murthy, K.: IEEE Trans. Neural Netw. **11**(5), 1188 (2000)
36. Muller, K., Mika, S., Ratsch, G., Tsuda, K., Scholkopf, B.: IEEE Trans. Neural Netw. **12**(2), 181 (2001)
37. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classifcation (2010)
38. Smola, A., Vishwanathan, S.: Introduction to Machine Learning (Cambridge University Press, 2008)