Computers and Electronics in Agriculture 138 (2017) 200–209

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case

Alexander Johannes [a,1,*], Artzai Picon [b,1], Aitor Alvarez-Gila [b,1], Jone Echazarra [b], Sergio Rodriguez-Vaamonde [b], Ana Díez Navajas [c], Amaia Ortiz-Barredo [c]

[a] BASF SE, Speyererstrasse 2, 67117 Limburgerhof, Germany
[b] Computer Vision, TECNALIA, Parque Tecnológico de Bizkaia, C/ Geldo. Edificio 700, E-48160 Derio, Bizkaia, Spain
[c] NEIKER, Plant Health Dp, Arkaute Agrifood Campus, E-01080 Vitoria-Gasteiz, Araba, Spain

ARTICLE INFO

Article history:
Received 25 November 2016
Received in revised form 20 April 2017
Accepted 21 April 2017

ABSTRACT

Disease diagnosis based on the detection of early symptoms is a usual threshold taken into account for integrated pest management strategies. Early phytosanitary treatment minimizes yield losses and increases the efficacy and efficiency of the treatments. However, the appearance of new diseases associated to new resistant crop variants complicates their early identification delaying the application of the appropriate corrective actions. The use of image based automated identification systems can leverage early detection of diseases among farmers and technicians but they perform poorly under real field conditions using mobile devices. A novel image processing algorithm based on candidate hot-spot detection in combination with statistical inference methods is proposed to tackle disease identification in wild conditions. This work analyses the performance of early identification of three European endemic wheat diseases – septoria, rust and tan spot. The analysis was done using 7 mobile devices and more than 3500 images captured in two pilot sites in Spain and Germany during 2014, 2015 and 2016. Obtained results reveal AuC (Area under the Receiver Operating Characteristic–ROC–Curve) metrics higher than 0.80 for all the analyzed diseases on the pilot tests under real conditions.

1. Introduction

The identification of early symptoms for diseases is one of the major milestones for the crop protection industry but also for two social and environmental challenges: increasing the yield or minimizing yield losses to ensure food security for 9 billion people in 2050 and increasing the efficacy of the phytosanitary treatments to reduce their use.

Agrios (2006) divides the measurement of plant diseases into three parts: Measuring of the incidence, severity and yield loss. Incidence is described as the proportion of the plant which is diseased. The severity is specified as proportion of area which is infected. And the yield loss is the share of the harvest which is destroyed or which has an effect to the quality of the produce. Even though severity and yield loss are of a much bigger importance for the grower, the incidence of a disease is much more difficult to measure and in some cases not possible until it is too late due to incipient incidence that is not detected by the farmers. It is necessary to take into account that the detection of early symptoms is a usual threshold considered for integrated pest management strategies of diseases in cereals. Consequently, research needs to focus on activities that can identify diseases in an early stage, so that targeted activities can be triggered, symptoms can be treated and even epidemics can be prevented. If solutions can be advanced to achieve this, the corresponding yield losses could be minimized. The annual yield loss of 20–40% of agricultural productivity is therefore due to impacts of pathogens, animal and plant weeds (Oerke, 2006). Studies have shown that yield losses per year accumulate up to $ 5 billion in the US (Savary et al., 2012). Another example for leaf rust epidemics and their respective impact on losses during 2001–2003 in Mexico equated $ 32 million in durum wheat (Herrera-Foessel et al., 2006). Therefore, the use of resistant varieties or crop protection applications can minimize the loss risk significantly as it is shown by a survey of Chai et al. (2015) where a decrease of yield losses could be verified between 1985 and 1999 due to the usage of resistant varieties and fungicides. Nevertheless, an increase on the yield loss can be observed since 2000 due to new rust races. In Europe new multi-virulent rust strains appeared in 2011. Several wheat varieties became affected and the new strains seems to be very aggressive based on field observations. These recently appeared races named as Warrior and Warrior (-) (data from Global Rust Reference Center, 2017) have changed

* Corresponding author.
E-mail address: alexander.a.johannes@basf.com (A. Johannes).
[1] Equally contributing authors.

http://dx.doi.org/10.1016/j.compag.2017.04.013
0168-1699/

European situation and urged research to elaborate on tools to help farmers detect the early presence of rust in the field avoiding to increase losses. Unexpected annual weather variation and the effects of climate change for diseases management tend to add uncertainty to decision making and inducing high incidence and severe outbreaks of endemic diseases unknown before (Coakley et al., 1999; Garrett et al., 2007).

As new pathotypes are emerging, it is essential to detect upcoming diseases in an early stage to prevent or minimize yield losses. Especially an early detection is almost impossible and requires very specific knowledge, as symptoms are not yet well developed. There is a need to provide image processing based plant disease identification, to diagnose diseases in their early development stages to be able to react in time with crop protection applications. At the same time, it is necessary to increase the reliability of disease identification and validate it on real environment. Other sensor devices for direct or indirect color variation detection could provide useful information as severity and spreading in the plant or crop but they do not produce enough information to diagnose a specific damage in the plant, including biotic (disease, pest and weeds) and abiotic hazards.

This paper gives an insight into a trainable system to identify plant diseases which has been validated in three European endemic wheat diseases by image processing based techniques in combination with statistical inference methods to solve the above mentioned technical problem. Plant disease identification as used herein includes the determination of a probability that a particular disease is present. Typically, plant diseases cause characteristic damage on the surface of plant elements (e.g. leaf, root, blossom, fruit, flower, stem, etc.). Therefore, characteristic symptoms become visible on some elements of an infected plant.

2. Related work

Different research has addressed automated plant disease identification by computerized visual diagnostic methods. Satellite or Airborne remote sensing have been proposed for disease identification. Huang et al. (2007) performed a comparative study of in-situ and airborne hyperspectral images in order to evaluate tailored spectral indexes to detect yellow rust in wheat. Mahlein et al. (2013) developed spectral disease indexes to detect the presence of sugar beet diseases by selecting the two most discriminant wavelengths per disease whereas Moshou et al. (2005) extracted suitable spectral indices by an unsupervised approach. Although these methods are appropriate to evaluate and identify the extension of a disease over a region, they fail on the detection of early symptoms as they require the presence of the disease on its later stage to be easily recognizable by airborne imagery. Other authors (Sankaran et al., 2010) make an exhaustive review on advanced techniques for plant disease diagnosis analyzing the volatiles produced by a diseased plant in order to develop specific electronic sensors.

In order to identify the diseases based on early symptoms, automatic picture analysis has been proposed by different authors. Sannakki et al. (2011) introduced an approach to automatically grade diseases on plant leaves. The proposed system uses image processing techniques to analyze color specific information in an image of the infected plant. A k-means clustering method is performed for every pixel in the image to extract clusters with infected spots. The segmented image is saved and the total leaf area is calculated. Finally, the disease spreading on plant leaves is graded by employing Fuzzy Logic to determine a particular disease. A high computational effort is required for such an image processing based method although some initiatives have reduced the computational cost (Xie et al., 2016). Other authors (Siricharoen et al., 2016) satisfactorily used an approach combining texture, color and shape features to detect the presence of a disease. However, they focus on leafs containing one disease at the same time and not focusing on early symptoms.

Other initiatives such as PlantVillage (Hughes and Salathé, 2015) have released more than 50,000 expertly curated images of healthy and infected leaves of different 14 different crops (apple, blueberry, corn, grape...) 12 of them having also healthy leaves and a total number of 26 different diseases. The use of deep convolutional neural networks have been already proposed (Mohanty et al., 2016; Sladojevic et al., 2016), obtaining accuracies greater than 99% when the testing images belong to the same dataset. However, when the model is tested against images collected from online trusted sources (Sladojevic et al., 2016), the accuracy is degraded down to 31.4%. The fact that the database is taken under controlled conditions and the presence of only late stage diseases on the database precludes its use as a real digital farming application. Besides this, the existing databases (Mohanty et al., 2016) does not consider the case where more than one disease are present in the same plant; therefore, models trained and tested on it will only be able to detect the most visible disease, which is not necessarily the one of most importance for the crop.

3. Materials and methods

In order to develop an algorithm capable of distinguishing diseases at early stages, an extensive image database has been created for wheat crop. Although wheat crop is used for validation purposes, the algorithm has been designed to be used on different crops and diseases and can be applied to other crops such as barley or corn. An acquisition campaign took place during the years 2014, 2015 and 2016 in Germany and Spain coincident with the crop season. Images were acquired at field conditions on the field with seven different devices: iPad, iPhone4, Dell-tablet, Samsung Galaxy Note, Windows Phone, Samsung S3, iPhone5. The acquisition campaign was designed to last the whole crop season, so that pictures

Fig. 1. Examples of pictures of infected images in the dataset.

of the disease at different phenological growth stages could be acquired. In order to assure variability, the pictures were taken in at least 36 different wheat varieties searching for the symptoms in the most common or recently commercialized ones with different resistance disease level and different life cycle. Examples of pictures acquired during these campaigns can be seen in Fig. 1.

The three diseases considered in this work were septoria, rust and tan spot (Clark et al., 2010). Besides them pictures containing other diseases such as powdery mildew or abiotic damages were included to assure the algorithm is able to generalize the absence of disease appropriately.

The early symptoms of septoria are characterized by water-soaked (chlorotic or yellowed) patches which quickly turn to ashen gray-brown oval lesions surrounded by a chlorotic halo, randomly distributed on leaves. Later in the season, the disease severity increased, these lesions frequently coalesce to produce large areas of brown-dry tissue and can appear in the whole plant, from the lower to upper leaves. Depending on the climate and the fungus implied, the patches contain the visible black pycnidia (small black points) which are the most characteristic sign of septoria in mature lesions.

Symptoms of brown-red and yellow rust at early stages seen as individual pustules yellow to orange-brown in color and about 0.5–1.0 mm. A Chlorotic halo surrounded the pustules is quite common which appear even before de pustule formation but not always appears. Later in the season, the brown pustules (brown-red rust) tend to be scattered at random compared with the more striped symptoms of yellow rust.

The first foliar symptoms caused by tan spot appear as small (1–2 mm), light brown blotches that develop into oval-shaped, light brown, necrotic lesions, sometimes bordered with a thin yellow halo. Later in the season, these lesions coalesce to produce large areas of brown tissue, similar to septoria symptoms in late stages.

A total number of 3637 images of wheat diseases have been taken under natural conditions during 2014 and 2015 on both pilot sites and they have been used as a training and development database. The details of these dataset are summarized in Table 1.

Table 1
Generated database for training purposes.

Database name (acquisition time)	Rust	Septoria	Tan spot
Wheat 2014 (W-2014)	471	700	183
Wheat 2015 (W-2015)	516	1805	457

Taking into account the symptoms and signals for the above mentioned diseases, the pictures were taken from an expanded leaf, form the upper leaf surface, avoiding pictures of symptoms or signals located in the margins or in the tip of the leaves. The image for these databases should be representative of the disease symptoms which appear in a cluster of plants or in a field. The picture showed damaged and not damaged tissue and were performed avoiding direct light.

In order to design and validate an algorithm capable of identify crop diseases at their early stages, all dataset images were accurately segmented by expert technicians. All the spots presenting each disease as well as the image region comprising the leaf were segmented. A specific segmentation was done for each disease (see Fig. 2).

To validate the classification results under real conditions, a group of 20 farmers and technicians belonging to Neiker-Tecnalia performed a pilot validation test with wheat crop in Germany and Spain. A smartphone application was given to them and they acquired images containing healthy leaves, rust, septoria, tan spot, other common diseases as Powdery mildew (*Blumeria graminis* f. sp. *tritici*) and abiotic diseases on the field under natural conditions. A gray cardboard was given to the farmers in order to optionally help them taking the picture under windy conditions. However, no information from the cardboard was used by the developed algorithm. The taken images were diagnosed online by the application and the results were stored to calculate performance of the system at real conditions. These images were validated by a technician at Neiker and this groundtruth was checked against previous algorithm results already stored on the database.

During the validation of the system, a total of 179 images were taken (examples of these images can be seen in Fig. 3). From those images, 96 presented at least rust, 88 septoria and 18 contained tan spot (Table 2).

A screenshot of the application is shown at Fig. 4.

4. Disease identification algorithm

Although several algorithms have been proposed on the literature for disease identification, they have not demonstrated their capabilities under natural conditions as a step decrease on their accuracy was seen when the images are taken in real field conditions. To overcome this inconvenience, an algorithm pipeline that is able to work under image acquisition variable conditions is pro-

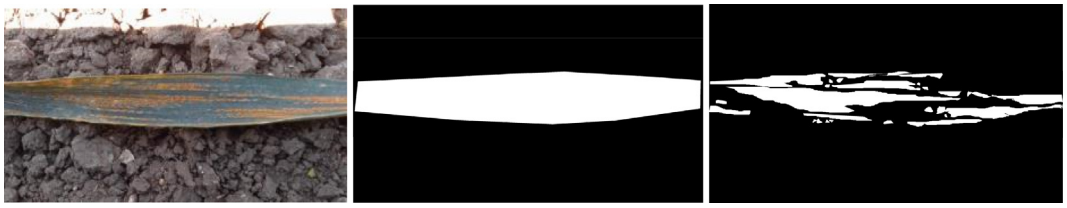

Fig. 2. Manual segmentation process (left) original image, (center) leaf region, (right) segmented rust hot-spots.

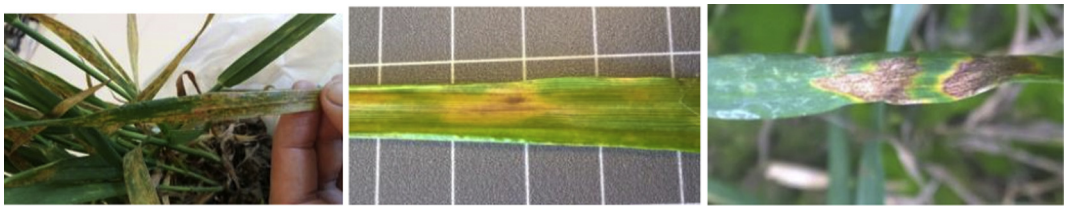

Fig. 3. Validation dataset images (left) image taken in a green-house, (center) image taken in the field using the accessory gray-board, (right) image taken in the field.

Table 2
Distribution of diseases for the pilot validation database.

Database name (acquisition time)	Rust	Septoria	Tan spot	Healthy or other diseases	Total images
Pilot database (P-2016)	96	88	18	27	179

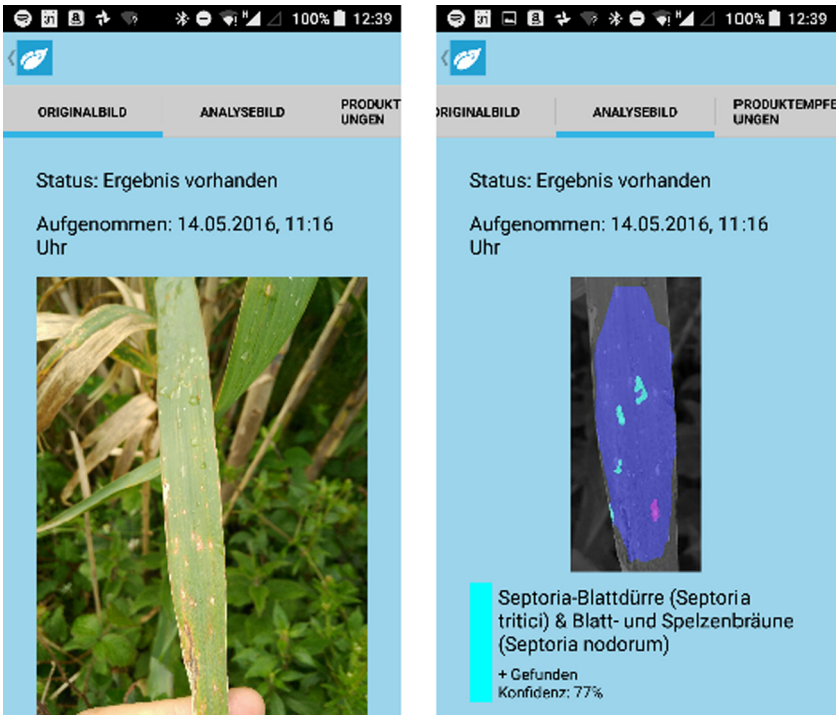

Fig. 4. Left image, the leaf photo taken by the farmer is shown; Right image, the output result from the algorithms is visualized.

posed. These variabilities can be caused, among others, by differences on the acquisition sensor, image scale, illumination, background and picture orientation. We propose a generic algorithm being able to be trained for different crops and diseases. To validate this versatility, the algorithm has been tested with three wheat diseases that present different visual characteristics. The proposed algorithm minimizes the accuracy loss under these variability conditions while coping with both early and late stage diseases.

In order to achieve this, a hierarchical approach based on the following stages is proposed:

(a) Image preprocessing: The acquired image is processed by means of color constancy algorithms to minimize the natural illumination variability effects. After that, leaf is segmented and isolated from the rest of the image. Different approaches have been taken depending on the nature of the image (natural image, existing draft mask from the user input...).

(b) Disease identification algorithm that comprises the following steps:
   a. Extraction of disease candidate regions: The segmented leaf isolated image is corrected to achieve color constancy, normalized and candidate sub-regions susceptible of containing diseases, called Hot-Spots, are extracted.
   b. Each extracted Hot-Spot region is analyzed in detail by local descriptors that extract and categorize each region in terms of its visual characteristics. Next, each Hot-Spot is checked against different disease detection inference models.
   c. All this information is gathered and processed by a high-level classifier, called meta-classifier, that is able to extract the complementarities of the different inherent features that are embedded within an image.

This global layout of the algorithm is described in Fig. 5. Each stage is detailed in the following sub-sections.

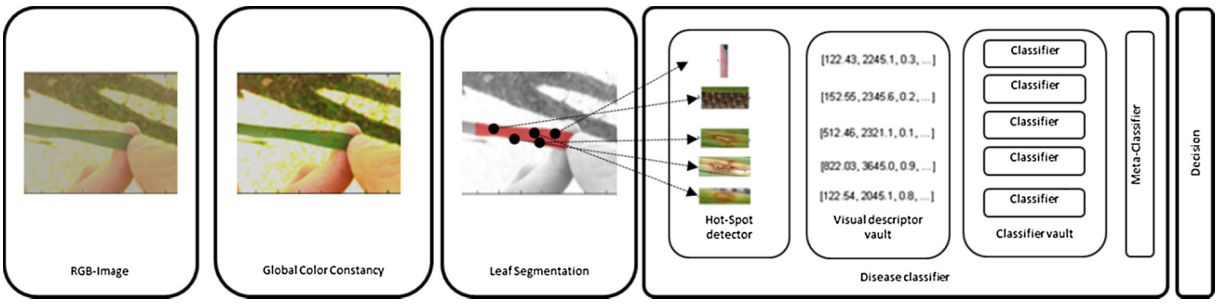

Fig. 5. Global algorithm layout.

4.1. Image preprocessing module

This module performs color constancy normalization on the acquired images to reduce the effects for natural image variability. After that, the portion of the image that contains the leaf is segmented and isolated to obtain a color normalized leaf image. In order to achieve this normalization several steps are involved (Fig. 6) and described in the following paragraphs.

4.1.1. Global color constancy

The differences in color appearance of the images in the database are due to various factors, such as changes in the physical lighting conditions (caused by weather, location, date and time of capture, orientation, etc.), differences in terms of color response and reproducible gamut among the diverse capture device models and individual units, or variability in the in-device image preprocessing and compression pipeline on a per-device and per-image basis. Such lack of control in the image acquisition workflow implies that a full color normalization process is not feasible, unless a relatively large set of color patches (printed under a color-managed workflow and measured with a colorimeter) are included with each and every capture and used to compute and apply a per-image color normalizing transformation (Finlayson et al., 2015). The cost and non-convenience of deploying such a solution at scale makes this approach inadvisable.

However, a color constancy transformation that aims at reducing the variance of the images in terms of white balance (while leaving out the color-gamut diversity) can be computed and applied per image in a feasible manner. Therefore, as a first preprocessing step, a global color constancy method is proposed in order to try to make the system more robust to changing illumination conditions during the recording of the images, and thus reach better classification results. The proposed method uses image statistics information to model the lighting conditions that were present while the image was captured, estimates such illuminant's color, and applies an inverse transform that maps the recorded colors to a common white point, yielding a white-balanced image. Shades of gray (Finlayson and Trezzi, 2004), Gray world

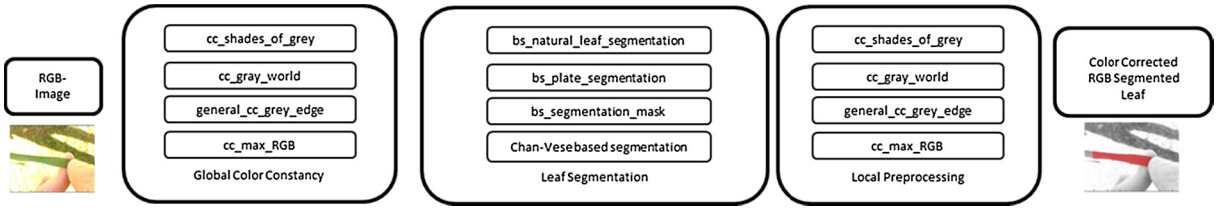

Fig. 6. Diagram for color constancy and image leaf segmentation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

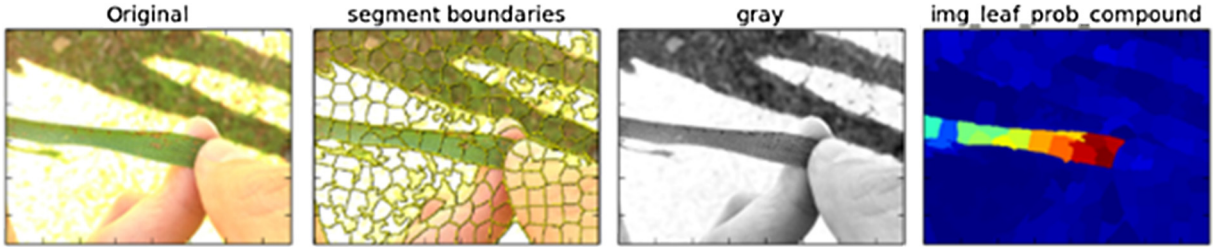

Fig. 7. Natural image segmentation algorithm.

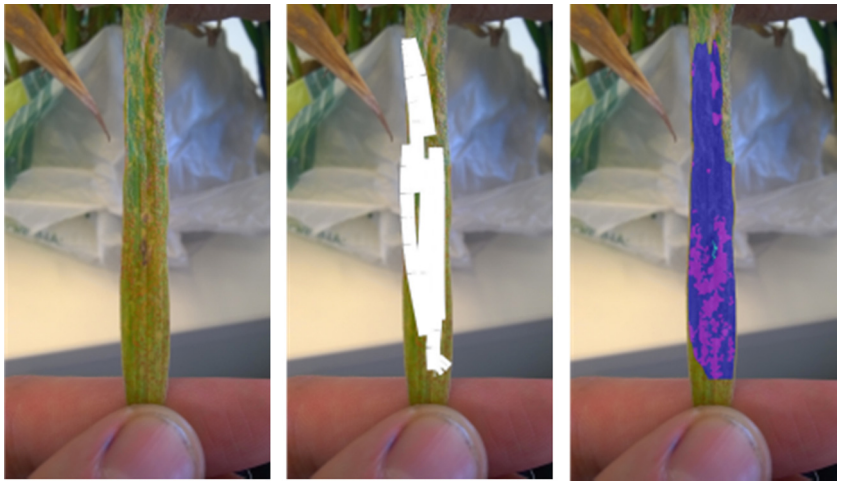

Fig. 8. Leaf mask refinement algorithm (left) original image, (center) input mask given by a user through the smartphone app right) in blue, the refined mask (blue) with the detected diseases (other colors). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Buchsbaum, 1980), Gray edge (van de Weijer et al., 2007) and Max-RGB (Land and McCann, 1977) color constancy algorithms have been proposed and evaluated to this respect.

4.1.2. Leaf segmentation

Leaf segmentation is a module to extract one or more portions from the prior color-constant image. Plant element extraction is set up in such a way that the portions extracted from the color-normalized image correspond to a leaf. Therefore, the extractor performs image processing operations which segment the portions of the image associated with plant elements from other portions in the image which do not provide any information on the disease (e.g. image background).

Two different alternatives for leaf extraction were considered.

4.1.2.1. Automated natural image leaf segmentation. The automated natural image leaf segmentation algorithm computes a binary leaf mask without any user intervention. The segmentation pipeline in this step relies on the Simple Linear Iterative Clustering (SLIC) Achanta et al., 2012 superpixel extraction algorithm, although other approaches such as Quickshift (Vedaldi and Soatto, 2008) or Felsenszwalb's efficient graph-based image segmentation (Felzenszwalb and Huttenlocher, 2004) could be also used. SLIC adapts a k-means clustering-based approach in the CIELAB color space while keeping a configurable spatial coherence as to enforce compactness. Next is the analysis of each of the generated superpixels (segments) to distinguish between leaf and non-leaf segments, by means of the extraction of different visual features. Each of such features maps a value to each superpixel, representing a weighting factor. The used features are the following: (1) mean, maximum or variance of a Gaussian weighting function, (2) mean, maximum or variance of the magnitude of the Local Binary Patterns (LBP) Ojala et al., 2002 flatness of the superpixel (3) maximum of the average of the b color channel and average of the inverted a color channel or (4) intra-superpixel perceptual color variance.

In a third step, all the probabilities available for each superpixel are combined either by means of the product or the sum of the individual probabilities. This yields a unique real value in the range [0, 1] for each superpixel, representing its probability of being part of the leaf. In the last step, a threshold value is selected and the image is binarized so that all the pixels pertaining to a superpixel with combined probability greater or equal such threshold are considered as being part of the leaf. The result can be seen in Fig. 7.

4.1.2.2. User input mask refinement method. Under real use cases it can happen that the taken image can be quite blurred or with strong artifacts or that the user wants to check a specific region of the leaf. In this case, a manually generated mask can be used to delimitate the region of interest (Fig. 8 – center image). However, this input mask generated by the user is not systematically generated and it is just a rough estimation of the region where the identification algorithm should focus. Because of this, an additional step of user input mask refinement is required.

In order to perform the leaf mask refinement, an initial segmentation is defined with its starting point as the user's given input mask. After that, Chan and Vese (2001) segmentation algorithm over the saturation(s) image channel is performed. The use of active contours without edges allows us to divide the image on two differentiated regions that minimize their internal variability. In practical terms, each interaction of the active contour minimization flows into a more homogeneous and accurate leaf mask estimation.

4.2. Disease identification classifier

The disease identification classifier performs the identification of one or several diseases over the previously segmented leaf (Fig. 9). First, a primary segmentation is done in order to detect and select any suspicious sub-region on the leaf as a disease candidate (Hot-Spots). Each disease candidate is then analyzed by the extraction of specific visual features of the candidate region and the use of a specific statistical inference model for the specific disease. It is noteworthy to indicate that this sub-segmenting approach smartly increases the accuracy of the system when working with early diseases. The hypothesis is that this training mechanism helps to deal with early stage diseases, as the classifiers are trained to distinguish among the candidate detected hot-spots the segmented candidate regions are trained for disease/not disease characterization, allowing a better training of the discriminative features. Once each Hot-Spot is properly identified, a meta-classifier module weights all individual decision to make a global assessment.

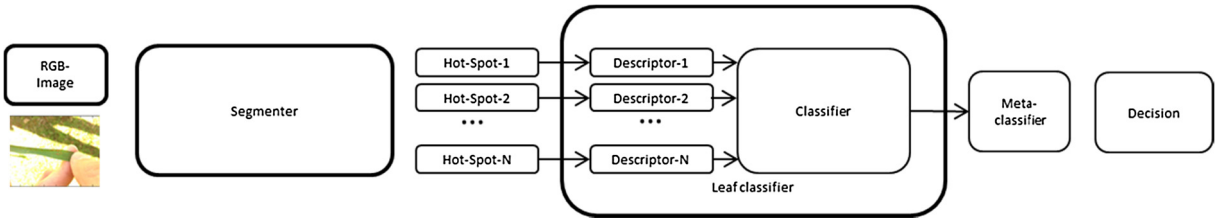

Fig. 9. Diagram of the disease identification classifier.

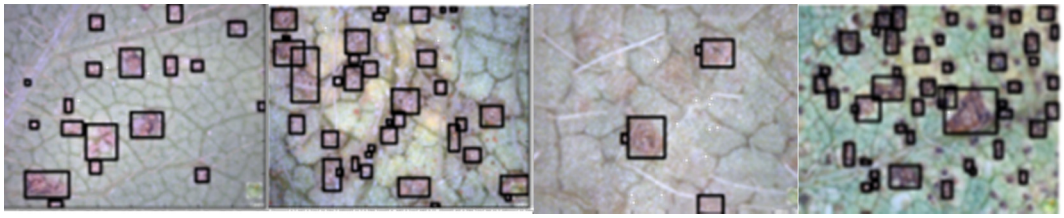

Fig. 10. Primary segmentation based on Lab clustering and Bayes classifier showing the detected candidate hotspots.

4.2.1. Primary segmentation module

The developed primary segmentation algorithm consists of a Naïve Bayes classifier (Lewis, 1998) that robustly models the color features that are always present on an infected image. This initial segmentation is performed over selected color channels including Lab and HSL. Based on these features, the image is segmented into different feature clusters (Computer Vision - A Modern Approach) where a cluster represents a group of pixels having similar visual feature values (e.g., color values or textural values). Each image blob feeds a Naïve Bayesian segmentation model. The Bayesian model acceptance threshold is selected to assure that every blob in the image feasible for presenting the desired disease is classified as a candidate.

In view of the full image content the identified candidate regions typically cover only a relatively small number of pixels when compared to the number of pixels of the entire image. Consequently, the amount of data that is used as the basis for the next stages of image processing is significantly reduced by the Naïve-Bayes filter mechanism and thus, the computational cost of the algorithm reduced. The visual feature statistics of each identified cluster is then confronted with this Naïve-Bayes classification model. Each identified image cluster is analyzed and its disease likelihood probability is calculated and is segmented by using the Bayes segmentation model. This model is biased according to the predefined threshold to ensure a low percentage (e.g. 1–5%) of discarded positives while maintaining a low rate of false positives. That is, after the application of the Bayesian filter to the clusters only such clusters above the predefined threshold are kept as candidate region for further analysis. In other words, the predefined threshold ensures that every candidate region in the image will be segmented for further analysis. Those candidate regions can then again be processed in the cluster function to receive more precise results as the infected area can be further isolated. Detected candidate hot-spot regions are depicted in Fig. 10.

One of the biggest advantages of this approach is that it makes a first segmentation that eliminates the regions in the images that have useless visual information.

4.2.2. Hot-Spot identification module

Each obtained Hot-Spot region is described by means of two visual descriptors. A first descriptor designed to sample the color of the disease contains information about the mean and standard deviation of the Lab channels over each segmented Hot-Spot blob. In order to also capture the color of the healthy plant as reference, the descriptor concatenates the mean and variance of the Lab channels on the pixels from the Hot-Spot blob bounding boxes that do not belong to the Hot-Spot region. A second descriptor, designed to map the Hot-Spot texture is also calculated for each blob. Concretely, a Uniform LBP descriptor over the RGB channels is used in order to describe blob texture assuring rotation invariance.

For each disease and descriptor, a Random-Forest based classifier is trained with the extracted descriptors that allow tagging each Hot-Spot blob with a disease feasibility value. In a second step a meta-classifier is used, to compute a confidence score for the particular disease by evaluating all determined probabilities of the candidate regions. A classification decision for one or more diseases is processed based on the entire picture. Information considered within the meta-classifier are for example, the probability given by the different descriptors, location, size or shape of the candidate regions, weighting of each classifier and the confidence necessary for an assessment to be taken into account.

This process can be seen in Fig. 11 that shows a leaf that presents only rust disease. Each row depicts an attempt of the algorithm of detecting one specific disease. The primary segmentation algorithm performs a first attempt to detect candidate spots. In the case of this leaf, candidate spots are found for septoria and rust and none for tan spot as it is depicted on the central column. The candidate spots for septoria and rust were further analyzed and only rust spots were detected as real spots (depicted in green on the right column) whereas the candidate hot-spots for septoria were detected as not real septoria and depicted in red. This is caused by the ability of the hot-spot classifier to combine the textural information from the LBP descriptor that is able to distinguish between the high frequency textures of rust and the lower level frequencies of septoria.

5. Results

The presented algorithm was developed on Python programming language and deployed as a service on a Linux based processing server. The deployed service was connected to a middleware server that managed the connections from Android and Windows-phone applications. Average processing time of the algorithm was 5.2 s with a standard deviation of 2.6 s depending on the number suspicious hot-spots found.

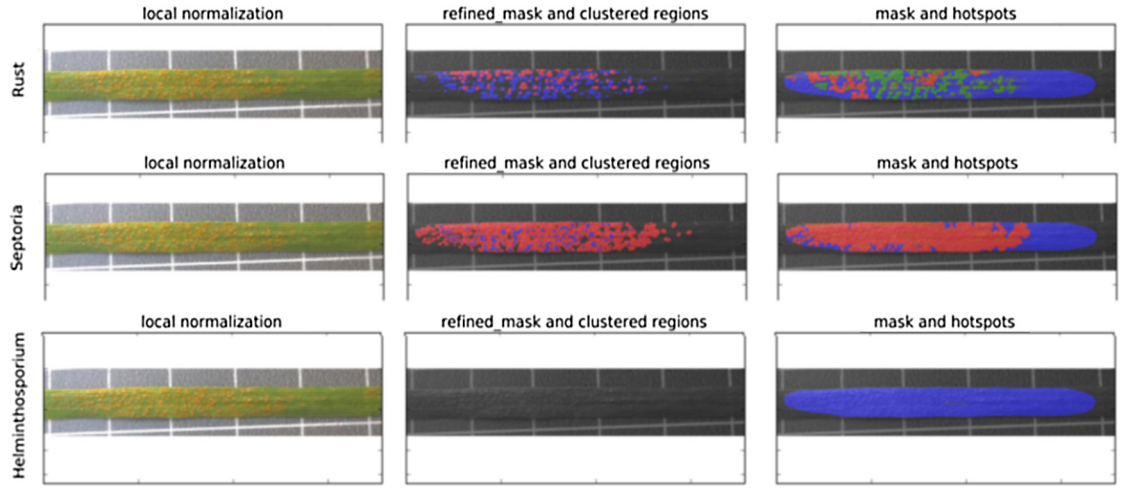

Fig. 11. (Left) RGB image after local color constancy; (center) located rust Hot-Spots candidates for each disease extracted by the primary segmentation; Right) green: HotSpots identified as the disease, red: candidate hot-spots refused by the hotspot identification module. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
classification AuC metrics for color normalized images and images without normalization.

Preprocessing color algorithm	Rust	Septoria	Tan spot
Gray world (van de Weijer et al., 2007)	0.83	0.87	0.82
Max-RGB (Achanta et al., 2012)	0.81	0.85	0.83
Shades of gray (Fawcett, 2006)	0.85	0.90	0.89
Gray edge	0.83	0.81	0.75
Without color constancy algorithm	0.77	0.73	0.72

Table 4
Identification results on the K-fold validation dataset.

Disease	AuC	Accuracy	Sensitivity	Specificity
Rust	0.85	0.82	0.7	0.95
Septoria	0.90	0.85	0.91	0.79
Tan spot	0.89	0.73	0.69	0.78

Table 5
Identification results on the May-2016 pilot.

Disease	AuC	Accuracy	Sensitivity	Specificity
Rust (Early)	0.81	0.78	0.80	0.76
Septoria (Early)	0.81	0.76	0.75	0.77
Tan spot (Early)	0.83	0.73	0.76	0.70
Rust (medium-late)	0.83	0.81	0.80	0.82
Septoria (medium-late)	0.82	0.79	0.80	0.78
Tan spot (medium-late)	0.81	0.82	0.96	0.69

In order to validate the results of the proposed method, a database was created using the images acquired in 2014 and 2015 (named W-2014 and W-2015) with 987 images containing rust, 2505 containing septoria and 657 containing tan spot on a total of 3637 images.

This training database was divided into training and validation sets. In order to avoid over-fitting and biasing, the dataset was divided into ten folds where the picture acquisition date was used as set divider. This means that, at each fold, the pictures belonging to the acquisition dates selected for training will be selected for the training set whereas the rest were selected for validation. At each fold, 90% of the acquisition dates were set as train and the remaining 10% were set as validation.

The Area under the Receiver Operating Characteristic (ROC) Curve (AuC Bradley, 1997; Fawcett, 2006) was selected as the most suitable algorithm performance metric, in order to account for the class imbalance present in the dataset (in such cases, the use of accuracy is discouraged). The AuC for a binary classification problem is constructed by first sorting all the samples by the disease presence probability predicted by the model for each of them. The classification threshold value is then moved all the way from 0 to 1, and the result at each threshold value is mapped into the plot representing False Positive (x-axis) vs. True Positive rates (y-axis), and measuring the resulting area (in the $[0, 1]$ range, higher is better) under such curve. AuC values over 0.85 were obtained for all diseases on the k-fold validation sets. It is important to remark that even diseases at very early stages were identified. An additional test was performed in order to quantify the effects of color constancy normalization on the identification metrics. It can be seen in Table 3, the use of color constancy normalization increases the overall accuracy from 0.75 up to 0.81.

Detailed metrics for each disease are depicted in Table 4.

In order to validate the results under real conditions, a pilot study was set in Germany and Spain in 2016. The pilot users employed a specific developed identification smartphone application as described in Section 3. The results of the pilot validation

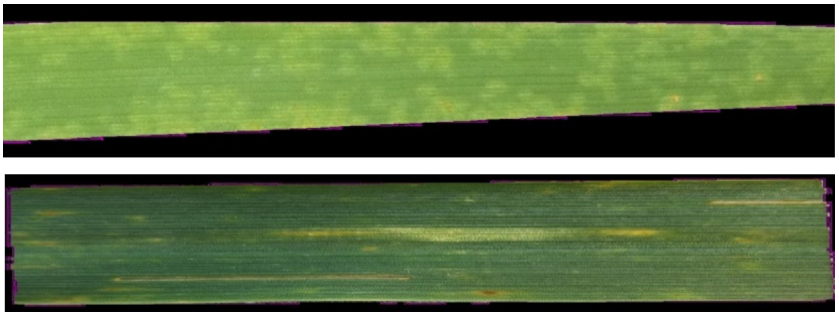

Fig. 12. Average AuC, accuracy and FAR/FRR curves for rust, septoria and tan spot diseases obtained at real pilot conditions.

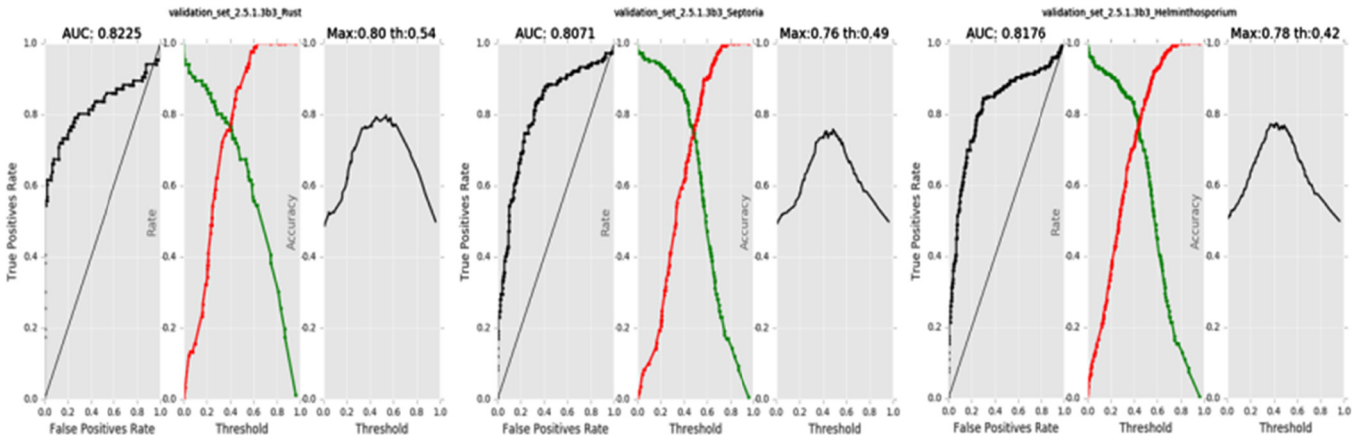

Fig. 13. Example of images with early diseases from the validation set.

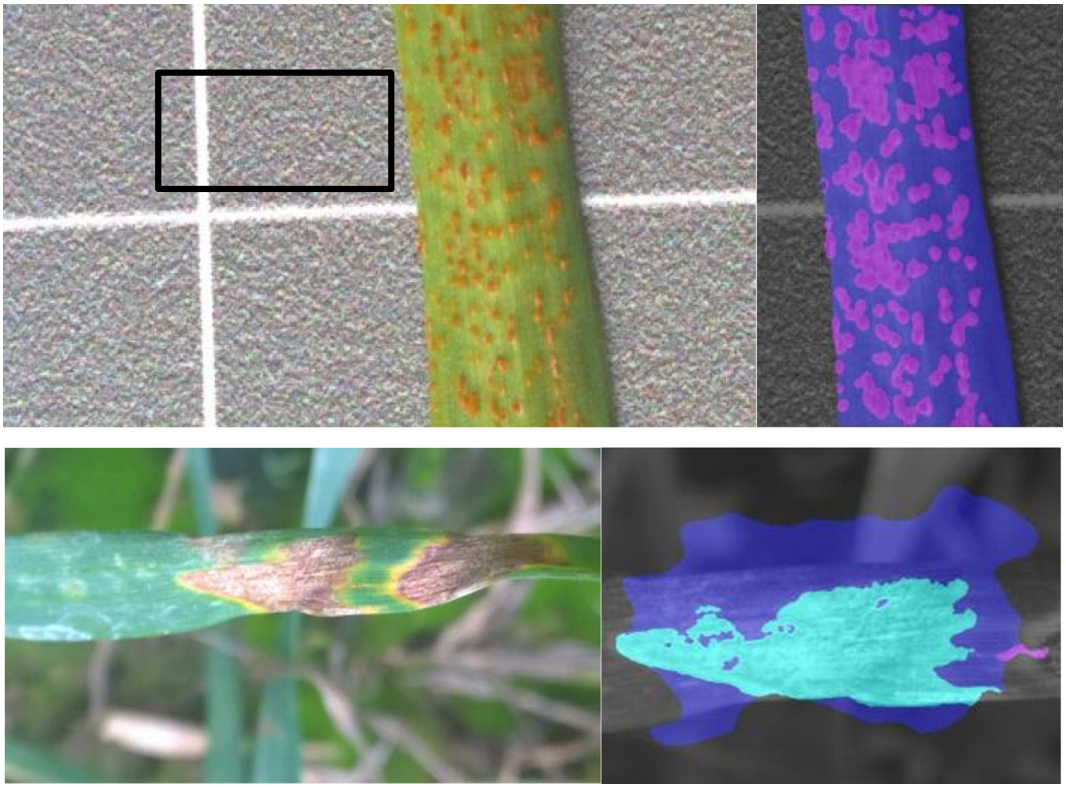

Fig. 14. Algorithm results under pilot conditions (cyan: detected septoria, purple: detected rust).

can be seen in Table 5. Captures leaves were divided into early stage diseased plants and medium-late stage diseased plants to validate early disease detection. Fig. 13 shows an example of early diseases leaves included on the early set database.

We observed a small decrease on the AuC from an average 0.88–0.81 when moving into real field conditions. Compared with other results that fail when moving into real conditions (Sladojevic et al., 2016), the developed algorithm can cope better with the variability of real light illumination, different acquisition devices and multi-located users under real use conditions. Besides this, performance is not degraded too much when dealing with early diseases. Detailed numbers are depicted in Table 5 and in Fig. 12.

Detailed algorithm results are shown on the next pictures (Fig. 14).

6. Conclusion

In this proposal a general use multi-disease identification algorithm has been presented. The algorithm was validated over three different kinds of diseases (rust, septoria and tan spot) on wheat images. The algorithm has been deployed on a real smartphone application and validated under real field conditions in a pilot study located in Spain and Germany over more than 36 wheat varieties. The results on real field tests obtained AuCs (Area under the Receiver Operating Characteristic–ROC–Curve) higher than 0.8 when assuring global color constancy on the image by means of the developed algorithm succeeding on real time conditions and being able to cope with different diseases simultaneously. The AuC decreased to 0.7 when no color constancy algorithm is applied on the processing workflow. However, the use of shades of gray color constancy algorithm assures AuC values higher than 0.8. We have also validated algorithm performance when dealing with early diseases. Although there is a small degradation on the performance of the algorithm when dealing with early diseases, the algorithm can obtain almost the similar performance on early and late diseases.

The preliminary hot-spot detection and its ulterior description by color and textural descriptors allow real time performance as only the suspicious regions are trained and described by the higher level classifiers and descriptors.

The presented image processing technology provides new possibilities for the detection of weeds and diseases in earlier stages. Next steps will be focused on measuring how this early stage detection can help the user to react in time and plan for some preventive activities, e.g. crop protection application. Being able to react in an early stage could minimize the yield losses and therefore guarantee the food security in the upcoming years.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34 (11), 2274–2282.

Agrios, G.N., 2006. Plant Pathology, fifth ed. Academic Press, p. 952. ISBN: 9780120445653.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn. 30 (7), 1145–1159.

Buchsbaum, G., 1980. A spatial processor model for object colour perception. J. Franklin Inst. 310 (1), 1–2.

Chai, Y., Pardey, P., Beddow, J., Hurley, T., Kriticos, D., Joachim-Braun, H., 2015. The global occurrence and economic consequences of stripe rust in wheat. In: Advancing Pest and Disease Modeling Workshop. University of Minnesota, CSIRO, and CIMMYT. <http://www.globalrust.org/sites/default/files/2014%20BGRI%20Pardey.pdf> (accessed 14th October 2016).

Chan, T.F., Vese, L.A., 2001. Active contours without edges. IEEE Trans. Image Process. 10 (2), 266–277.

Clark, B., Bryson, R., Tonguc, L., Kelly, C., Jellis, G., 2010. The encyclopaedia of cereal diseases. HGCA and BASF plc, Crop Protection.

Coakley, S.M., Scherm, H., Chakraborti, S., 1999. Climate change and plant disease management. Annu. Rev. Phytopathol. 37, 399–426.

"Computer Vision - A Modern Approach", by Forsyth, Ponce, Pearson Education ISBN 0-13-085198-1, pp. 301–307.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874.

Felzenszwalb, P.F., Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. Int. J. Comput. Vision 59, 167. http://dx.doi.org/10.1023/B:VISI.0000022288.19776.77.

Finlayson, G.D., Trezzi, E., 2004. Shades of gray and colour constancy. In: Color and Imaging Conference Jan (1), pp. 37–41.

Finlayson, G.D., Mackiewicz, M., Hurlbert, A., 2015. Color correction using root-polynomial regression. IEEE Trans. Image Process. 24 (5), 1460–1470.

Garrett, K.A., Dendy, S.P., Frank, E.E., Rouse, M.N., Travers, S.E., 2007. Climate change effects on plant disease: genomes to ecosystems. Annu. Rev. Phytopathol. 44, 489–509.

Herrera-Foessel, S.A., Singh, R.P., Huerta-Espino, J., Crossa, J., Yuen, J., Djurle, A., 2006. Effect of leaf rust on grain yield and yield traits of durum wheats with race-specific and slow-rusting resistance to leaf rust. Plant Dis. 90 (8), 1065–1072.

Huang, W., Lamb, D.W., Niu, Z., Zhang, Y., Liu, L., Wang, J., 2007. Identification of yellow rust in wheat using *in situ* spectral reflectance measurements and airborne hyperspectral imaging. Precision Agric. 8 (4–5), 187–197.

Hughes, D.P., Salathé, M., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. Computers and Society. In Computer Science. Cornell University Library. arXiv:1511.08060v2 [cs.CY].

Land, E.H., McCann, J.J., 1977. Lightness and retinex theory. JOSA 61 (1), 1–11.

Lewis, D.D., 1998. Naive (Bayes) at forty: the independence assumption in information retrieval. In: European Conference on Machine Learning. Springer, Berlin Heidelberg, pp. 4–15. April.

Mahlein, A.K., Rumpf, T., Welke, P., Dehne, H.W., Plümer, L., Steiner, U., Oerke, E.C., 2013. Development of spectral indices for detecting and identifying plant diseases. Remote Sens. Environ. 128, 21–30.

Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7.

Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, L., McCartney, A., Ramón, H., 2005. Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. Real Time Imaging 11 (2), 75–83.

Oerke, E.C., 2006. Crop losses to pests. J. Agric. Sci. 144, 31–43.

Ojala, T., Pietikainen, M., Maenpaa, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. 24 (7), 971–987.

Sankaran, S., Mishra, A., Ehsani, R., Davis, C., 2010. A review of advanced techniques for detecting plant diseases. Comput. Electr. Agric. 72 (1), 1–13.

Sannakki, Sanjeev S., Rajpurohit, Vijay S., Nargund, V.B., Kumar, Arun, Yallur, Prema S., 2011. Leaf disease grading by machine vision and fuzzy logic. Int. J. Comput. Technol. Appl. 2 (5), 1709–1716.

Savary, S., Ficke, A., Aubertot, J.N., Hollier, C., 2012. Crop losses due to diseases and their implications for global food production losses and food security. Food Secur. http://dx.doi.org/10.1007/s12571-012-0200-5.

Siricharoen, P., Scotney, B., Morrow, P., Parr, G., 2016. A lightweight mobile system for crop disease diagnosis. In: Campilho, A., Karray, F. (Eds.), Image Analysis and Recognition. ICIAR 2016. Lecture Notes in Computer Science, vol. 9730. Springer, Cham.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. Comput. Intelligence Neurosci. http://dx.doi.org/10.1155/2016/328980.

van de Weijer, J., Gevers, T., Gijsenij, A., 2007. Edge-based color constancy. IEEE Trans. Image Process. 16 (9), 2207–2214.

Vedaldi, A., Soatto, S., 2008. Quick shift and kernel methods for mode seeking. In: European Conference on Computer Vision, Marseille, France.

Xie, X., Zhang, X., He, B., Liang, D., Zhang, D., Huang, L., 2016. A system for diagnosis of wheat leaf diseases based on Android smartphone. Opt. Measur. Technol. Instr. http://dx.doi.org/10.1117/12.2246919.