Smart Agricultural Technology 9 (2024) 100532

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Proximally sensed RGB images and colour indices for distinguishing rice blast and brown spot diseases by k-means clustering: Towards a mobile application solution

Suvanthini Terensan [a,*], Arachchige Surantha Ashan Salgadoe [b], Nisha Sulari Kottearachchi [c], O.V.D.S. Jagathpriya Weerasena [d]

[a] Institute of Biochemistry Molecular Biology and Biotechnology, University of Colombo, Sri Lanka and Senior Lecturer at Department of Agricultural Biology, University of Jaffna, Sri Lanka
[b] Department of Horticulture and Landscape Gardening, Faculty of Agriculture and Plantation Management, Wayamba University of Sri Lanka, Makandura, Gonawilla (NWP), Sri Lanka
[c] Department of Biotechnology, Faculty of Agriculture and Plantation Management, Wayamba University of Sri Lanka, Makandura, Gonawilla (NWP), Sri Lanka
[d] Institute of Biochemistry Molecular Biology and Biotechnology, University of Colombo, Sri Lanka

ARTICLE INFO

Keywords:
Rice blast
Brown spot
Digital imaging
K-means clustering
RGB images
Proximal sensing

ABSTRACT

Rice blast (RB) and Brown spot (BS) are economically important diseases in rice that cause greater yield losses annually. Both share the same host and produce quite similar lesions, which leads to confusion in identification by farmers. Proper identification is essential for better management of the diseases. Visual identification needs trained experts and the laboratory-based experiments using molecular techniques are costly and time-consuming even though they are accurate. This study investigated the differentiation of the lesions from these two diseases based on proximally sensed digital RGB images and derived colour indices. Digital images of lesions were acquired using a smartphone camera. Thirty-six colour indices were evaluated by k-means clustering to distinguish the diseases using three colour channel components; RGB, HSV, and La*b*. Briefly, the background of the images was masked to target the leaf spot lesion, and colour indices were derived as features from the centre region across the lesion, coinciding with the common identification practice of plant pathologists. The results revealed that 36 indices delineated both diseases with 84.3 % accuracy. However, it was also found that the accuracy was mostly governed by indices associated with the R, G and B profiles, excluding the others. G/R, NGRDI, $(R + G + B)$/R, VARI, $(G + B)$/R, R/G, Nor_r, G-R, Mean_A, and Logsig indices were identified to contribute more in distinguishing the diseases. Therefore, these RGB-based colour indices can be used to distinguish blast and brown spot diseases using the k-means algorithm. The results from this study present an alternative, and non-destructive, objective method for identifying RB and BS disease symptoms. Based on the findings, a mobile application, Blast O spot is developed to differentiate the diseases in fields.

1. Introduction

Rice is an important source of carbohydrate as well as a staple food for more than half of the world's population International Rice Research Institute (IRRI) [1]. Rice production has been limited by many rice diseases in which rice blast (RB) and rice brown spot (BS) are considered important [2–4]. Both RB and BS are caused by fungus. RB disease is caused by ascomycete fungus; *Magnporthe oryzae* [5], and BS disease is mainly associated with *Bipolaris oryzae* [6]. Once infected, leaves develop elongated spot lesions and they gradually spread over the leaves, limiting the photosynthetic ability in the green tissues in both the diseases [4,7,8]. This can result in significant yield loss if not managed, where RB can cause 30 % yield loss [5] and BS can cause 52 % yield loss [9], on average, in rice.

It has been a challenge for the farmers in Sri Lanka and other countries to manage RB and BS diseases due to the confusion in discriminating the disease symptoms. According to expert knowledge and information available from IRRI, the lesions of RB are described as elliptical or spindle-shaped with whitish to grey centres and lateral necrotic edges [10], while BS lesions are described as circular to oval

* Corresponding author.
E-mail address: suvanthinit@univ.jfn.ac.lk (S. Terensan).

https://doi.org/10.1016/j.atech.2024.100532
Received 7 April 2024; Received in revised form 5 August 2024; Accepted 7 August 2024
Available online 9 August 2024

with a light brown/ grey centre, surrounded by reddish-brown lateral edges [11]. The stage wise differences exist between the two diseases are illustrated in Fig. 1.

There are existing colour gradient differences from the centre to the edges of the lesions that can be used to identify both RB and BS diseases. However, farmers and agriculture extension officers in Sri Lanka often find it difficult to differentiate them. Because the symptoms of the two diseases are often similar and appear on the same leaf, differentiation is challenging. Traditional identification methods and detailed symptom descriptions are ineffective when the diseases present identically due to various reasons; climatic pattern, agronomic practices, selection of cultivars and competitive interaction between pathogens. We conducted a survey involving 60 randomly selected farmers to differentiate the symptoms, but none could accurately identify them. Additionally, during sample collection, agriculture extension officers (5) were unable to distinguish the symptoms correctly and often confused them with blast, which is a prevalent disease in the Northern Province of Sri Lanka. This confusion emphasizes the necessity for precise diagnostic tools to ensure accurate disease identification for appropriate treatment. One of the most promising and environmentally friendly methods is the selection of resistant cultivars against specific diseases in the region. By accurately diagnosing the disease, farmers can implement targeted management practices, reducing the reliance on chemical treatments and promoting sustainable agriculture. On the other hand, visual identification with the naked eye requires trained expertise which is always not readily available with farmers and agriculture officers. Moreover, the results of any qualitative judgment also vary according to an individual's experience, where the decision can be affected by temporal variation of the symptom, as well as those incurred from different assessors [12,13]. Thus, a quantitative approach to identifying the lesions is an immediate requirement for timely disease management and avoiding serious outbreaks.

Out of the existing methods for disease identification, laboratory-based molecular-level genetic methods are accurate. For instance, a specific primer set PfH2A/PfH2B was used to selectively screen a fungus causing RB; *M. oryzae*, in PCR [14]. However, such methods require sophisticated laboratory facilities and specialized training, along with destructive sampling and off-site analysis. Therefore, an objective way of distinguishing RB and BS leaf lesions using a less time-consuming, onsite, non-destructive, and cost-effective method would be highly preferable.

Remotely sensed digital colour imagery and associated image analysis techniques present an alternative to human vision [15] and has been widely used in disease identification, symptom detection, and quantification [16]. It is a popular method of acquiring reflected wavebands/-colour information from plant leaves without physically contacting them to estimate the status of the tissues [17]. Proximal sensing tools like a camera on smartphones offer greater advantages than unmanned aerial vehicles and satellites, as the smartphone is a convenient and affordable tool that can be used by any farmer on the ground for the observation of local disease symptoms like RB and BS lesions.

1.1. Contribution summary

- An innovative approach for distinguishing between RB and BS, two closely resembling diseases, by utilizing laboratory findings and translating them into image sensing techniques. A comparative and comprehensive analysis of various color indices was conducted to identify the most suitable indices for accurate disease symptom diagnostics.
- Application of k-means clustering for RB and BS classification: The research applies k-means clustering algorithm to effectively classify and distinguish RB and BS. This methodological advancement demonstrates the potential of unsupervised learning techniques in disease diagnostics in agriculture.
- Development of a simple and user-friendly mobile application: A simple user-friendly mobile application named "Blast O Spot" was developed based on this research. This will be highly helpful for the farmers and agriculture field officers to effectively diagnose RB and BS.

1.2. Section-wise outline

The manuscript consists of following sections;

- Introduction section including the literature review which outlines the rice production, challenges experienced by the sector, impact of RB and BS, and methods employed in the disease diagnosis in agriculture which leads to the justification of the present study.
- The Method section describes the sample collection, identification of RB and BS in the laboratory and the detailed procedures used in choosing the best model for the classification of RB and BS.
- The results and discussion section outlines the results obtained based on our present study and the comparison of the performance of our model with already established models.

1.3. Literature review

1.3.1. Challenges in rice production

Rice is being cultivated in >150 million hectares around the world and consumed by 3 billion people [18]. The productivity of main cereals such as maize, wheat, rice, and soy should be increased by 87 % to feed

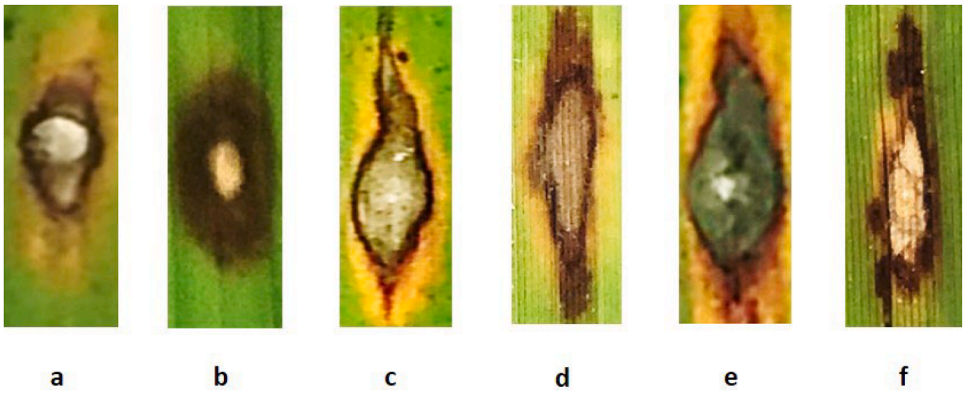

Fig. 1. Visual representation of RB and BS lesions at different disease stages: (a-b) initial stages of disease; (a) the RB lesion water-soaked in appearance, spindle shape with brown edges, (b) BS lesion oval shape in appearance with thick purplish black edges, (c-d) middle stages of disease; (c) RB lesion enlarges with the white centre because of fungal mat, (d) BS lesion often appears as a thin brown surface in the centre region most probably without mycelium, (e-f) severe stages of disease; (e) RB lesions with grey colour fungal mat appearance in the centre, and (f) BS lesion with necrotic region visible adjacent to the lesion with a reddish-brown edge.

the exponentially increasing human population [19].

Rice production is affected by various factors such as global climate change and natural disasters, and plant diseases [20]. The fungal diseases like RB, sheath blight, BS and sheath-rot, along with bacterial diseases like bacterial blight and the viral disease such as rice tungro disease known as RTD, are considered as the major diseases affecting rice [21].

1.3.2. Rice blast and brown spot disease identification

In most countries including Sri Lanka, rice production has been limited by two economically important diseases, namely, BS and RB. Beyond their ability to decrease the production significantly, they can also reduce the quality of seeds produced, which become unfit for consumption when the panicle is severely infected. Hence, they were responsible for major famines in Asian countries [22]. In Sri Lankan context, RB and BS are common and frequent fungal diseases in both *Maha* and *Yala* seasonal cultivations [23]. A severe outbreak of blast was reported, especially in the Mannar district in 2016, where significant yield loss was encountered. The studies about these diseases in the Northern Province are lacking and the literature in other parts of the country is also scarce.

Plant pathologists often rely on the symptoms to diagnose the disease. Proper diagnosis often results in a better management practice with less or no waste of resources. Accurate description of symptoms for a particular disease is not straight forward and can be problematic. The symptoms often change according to the time of the infection, plant varieties, and climatic conditions. Proper identification of them considering all these factors is a challenging task even for experienced pathologists. This is even more challenging in cases where the symptoms produced by two or more diseases are similar, as in the case of BS and RB diseases (Fig. 1). It is hard for a farmer to properly identify the symptoms in this case, and he/she often relies on expert advice, which may not be readily available in most parts of the world. Considering the fact that improper management of these diseases due to misidentification can result in complete loss of the production, it is essential to identify the diseases accurately at the early stages. The specific chemicals used to manage BS and RB are different as well [24].

1.3.3. Methods employed in disease diagnosis

The most common way of identifying the disease is based on observations made in the field or laboratory. The process involves discrimination of healthy and diseased plants, correct symptom identification, and differentiation of the symptoms based on time, and affected plant parts. The method is based on visual cues and already available identification guides. There are various guides available for this purpose to help farmers properly identify the symptoms [25].

The use of classical microbiology involves the use of millions of microbial cells as inoculum and culture them in culture media such as agar for microscopic and macroscopic analysis. The problem with the classical technique is that often the generalized traits of millions of microbial cells observed under the microscope are not the actual reflection of the traits found in each cell [26]. Another problem with microbes is that they develop into various strains within a very short period of time. Even within the same culture plate, there may be various sub-strains, serotypes, and phage types present. Many pathogens may also have the same morphological characters, which leads to improper identification. Generally, identification of plant pathogens includes the observation of the disease symptoms, along with laboratory tests such as culturing the pathogen on selective media, which is usually followed by physiological, biochemical, and/or pathogenicity tests. All these techniques are laborious and time-consuming [27].

Antibody testing has been used in plant-pathogen detection since the 1980s [28]. Even though, this is a rapid and sensitive method for detecting plant pathogens, it is expensive and related strains can react similarly to the same antibody. Molecular tools provide rapid alternatives to the above-said setbacks. PCR assays targeting polymorphism analysis in nuclear, mitochondrial, and ribosomal DNA were used for species characterization purposes [29]. Metagenomic next-generation sequencing (mNGS) is a relatively recent technique and has been widely used in pathogen detection in samples with unknown pathogenic diversity. Unlike molecular assays, mNGS captures the sequences of all DNA or RNA present in the diseased sample, providing a high-throughput technique to sequence not only an entire microbiome, but also the host genome. It is more convenient in the selection of primers as it does not require specific primers and probes, which narrows down the outcome of the data. This technique is more reliable in causal-agent identification.

Remote sensing is another tool where the characteristics of a material are analyzed without direct contact with them. The technique is widely used to characterize biophysical features of plants by using remote sensing-based techniques, either ground-based, airborne, or satellite-based. The technique has been used to study yield forecast [30], nutrient needs [31], pest attack [32], lesions and pathogenic damage [33].

Usually, in plant disease studies, the reflectance wavebands recorded in digital imagery acquired by a smartphone are depicted as colour indices which are ratios of colour channels associated with leaf pigment (i.e. chlorophyll content) in the leaf tissues [34]. As diseases progress, it alters the tissue's pigment resulting in changes in pigment composition to produce the lesions specific to the disease [35]. Whereas, the response to reflected colour by a specific lesion can be indicated as a quantitative value of the colour index. On the other hand, instead of using a single colour index (CI), a combination of CIs allows the extraction of multiple colour features from lesions and can be advantageous in image analysis.

Red, green, blue (RGB) photographs acquired by digital cameras or smartphone cameras have been successfully used as a proximal tool for a range of leaf or canopy level disease symptom studies. Salgadoe et al. [36] extracted canopy features associated with *Phytophthora* root rot disease in avocado trees using gap analysis. Saberioon et al. [37] studied 27 colour indices to determine the nitrogen content of rice plants in paddy fields at six growth stages where the RGB was the primary colour channel used to convert the images. Green (G) and Kawashima index (I_{KAW}) were reported as the best indices to determine the nitrogen content. The study also suggested that low-cost digital cameras are good enough to take digital images and it would be a useful smart tool for small-scale farmers. A study by Mohan and Gupta [38] illustrated the suitability of the RGB channel in estimating the leaf chlorophyll content of rice leaves from digitized images. The study concluded that brightness ratio (R, G, B) alone or in combination with a dark green colour index (DGCI-rgb) show a good correlation between the predicted chlorophyll content and relative chlorophyll content using artificial neuron network (ANN) model. In addition, RGB component analysis was used to obtain a prediction of the true value of foliar chlorophyll content in leaves of Quinoa and Amaranth [39]. In Barley [40], the analysis of relationships between chlorophyll content and colour indices in the RGB and Lab* colour profiles showed that the RGB colour indices such as R, G, $(R + G + B)$, R-B, $R + B$, $R + G$, and the La*b* colour space parameters such as L, a*, and b* had a significant relationship with chlorophyll content. The chlorophyll in citrus leaf is also predicted using Linear Regression and ANN successfully by using RGB based colour indices such as R, G, B, (R-B)/$(R + B)$, $R + G + B$, $(G + B)$/R, G/R, $(R + G + B)$/R, G-R)/$(G + R)$, (G-R), $(G + R)$, G/$(R + G + B)$, R/$(R + G + B)$, B/$(R + G + B)$, R-B, (G-B)/$(R + G + B)$, and $(R + B)$ [41].

K-means clustering is employed at the initial stages of discrimination of disease symptoms of plants [42]. This is used in segmentation of disease images and then more advanced methods such as ANN are used to identify the diseases accurately in many plants [43]. This algorithm reduces the noise in images substantially and used in further steps related to disease discrimination and is an effective and easy tool to classify the images [44]. There are many studies reported the use of this along with more complexed machine learning approaches to detect disease in plants such as corn, brinjal, and grape [45].

Considering the difficulty in discrimination of RB and BS in the field and in many laboratories with limited resources in Sri Lanka, it is essential to produce a digital tool which can be easy to use and also accurately identify the disease in the field with less resources. It is hypothesized that RB and BS lesions can be distinguished based on the colour indices derived from digital RGB imagery considering the existing subtle differences exist between RB and BS which cannot be discriminated by naked eye. This study investigates the applicability of proximally sensed RGB imagery collected from the field with a smartphone camera to distinguish RB and BS diseases in rice plants as a surrogate for traditional, human vision-based methods. The approach described here is resulted in a mobile application which can be easily used by a farmer in the field to discriminate the diseases. The accurate identification may also enable the farmer to identify the disease at the early stages and control the damage or even selecting appropriate cultivar for the next season.

2. Materials and methods

2.1. Overall system architecture

The "Blast O Spot" mobile application helps farmers identify Rice Blast (RB) and Brown Spot (BS) diseases in rice using smartphone-captured images. This system architecture (Fig. 2) outlines the structure and workflow toward the development of the "Blast O Spot" mobile application, illustrating how different parts of the system interact to provide accurate and efficient identification of rice diseases using digital images.

2.2. Disease symptom sampling

For this study, rice leaves showing RB and BS lesions were randomly sampled with the help of pathologists from the farmer fields. One such field is shown in Fig. 3, where the necrotic plant parts are visible as brown patches. These fields are located at five districts (Jaffna- 9°39′41.38″N, 80° 1′31.97″E, Kilinochchi- 9°22′49.04″N, 80°22′37.20″E, Mullaitheevu- 9°16′1.53″N, 80°48′51.29″E, Mannar- 8°58′51.51″N, 79°54′15.91″E and Vavuniya- 8°45′37.03″N, 80°29′40.21″E) in the Northern Province (NP) of Sri Lanka.

Symptoms associated with different severity stages from both RB and BS were sampled, labeled, and subjected to disease symptom verification using visual characteristics (as shown in Fig. 1) and molecular tests (i.e. PCR) conducted at the laboratory.

2.3. Microscopic and PCR based symptom verification

Field-collected rice leaf samples exhibiting only RB and BS symptoms were examined under a microscope at ×100 magnification to confirm the morphological characteristics of spores and causal agents. Genomic DNA was then extracted from the lesions using the Mycospin™ fungal DNA extraction kit (Ceygen Biotech, Sri Lanka) with modifications to the initial lysis step as described by Liu et al. (2000). An exclusion-based PCR method was performed using RB-specific primers (pfH2A: 5′-CGTCACACGTTCTTCAACC-3′ and pfH2B: 5′-CGTTTCACGCTTCTCCG-3′) designed from the Pot2 region [14]. PCR amplification was conducted using Promega's Taq DNA Polymerase (Cat.#M7660). DNA isolates that produced positive amplicons were confirmed as RB isolates, corroborated by microscopic observations of typical pyriform spores.

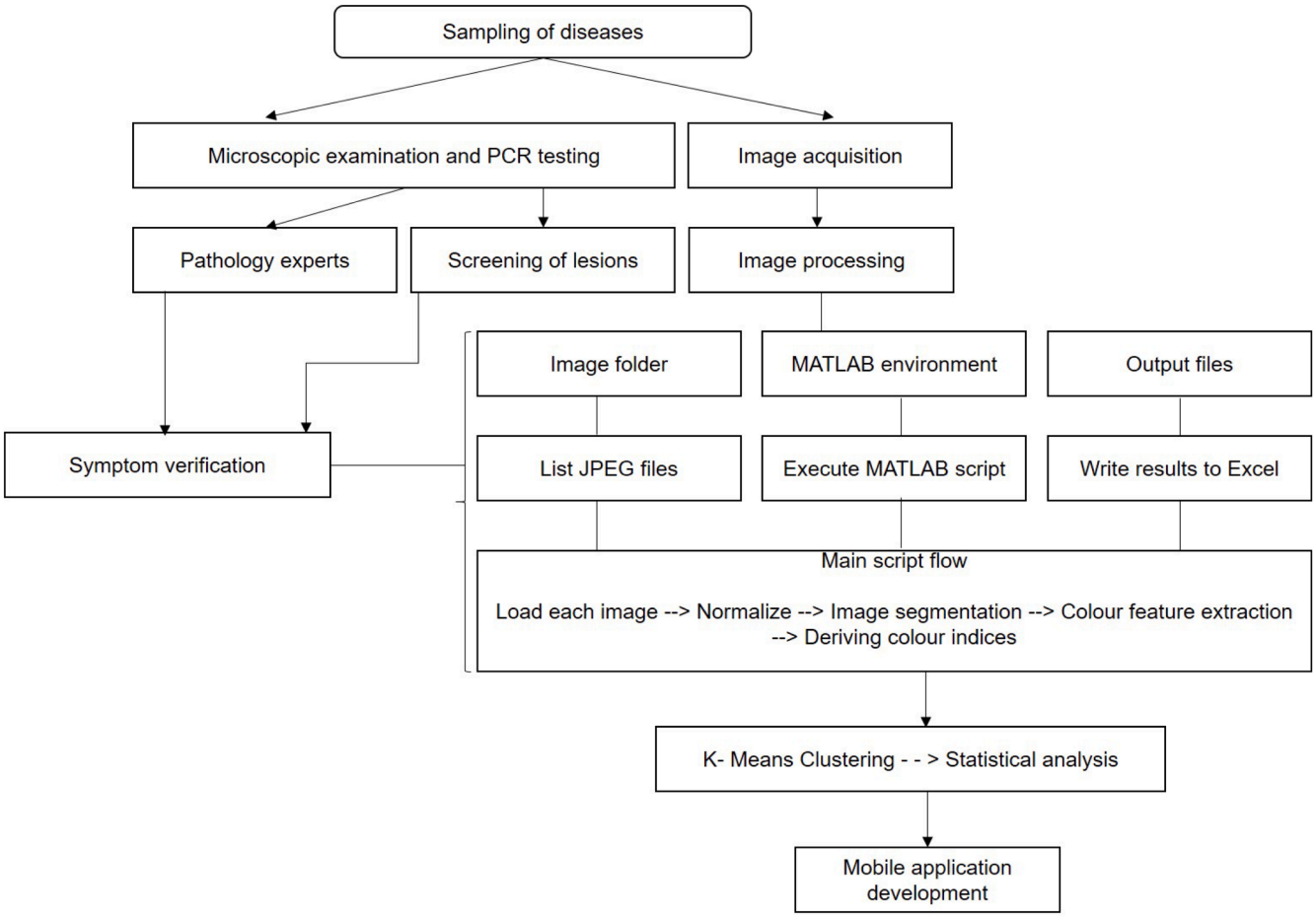

Fig. 2. Overall system architecture for disease detection. This diagram illustrates the comprehensive workflow for detecting RB and BS diseases in rice leaves.

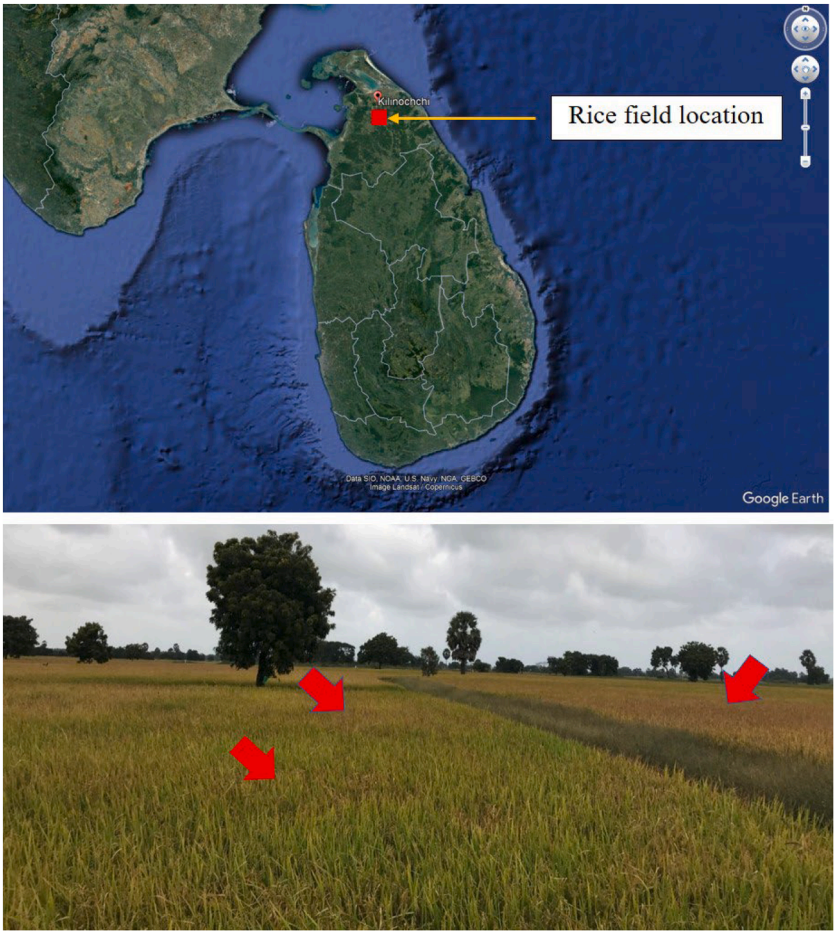

Fig. 3. A rice field infected with RB and BS diseases, located in the Kilinochchi district of the Northern Province, Sri Lanka. The necrotic parts are visible as brown patches.

Isolates that did not produce any amplicons with Pot2 blast-specific primers and exhibited BS symptoms were confirmed as BS-infected leaves.

2.4. Protocol for image acquisition through smartphone application

2.4.1. Through mobile application; 'Blast O Spot'

The "Blast O Spot" application, available in the Apple Store, helps identify RB and BS diseases. Users start by downloading the application. They then log in by entering their email address and password, which

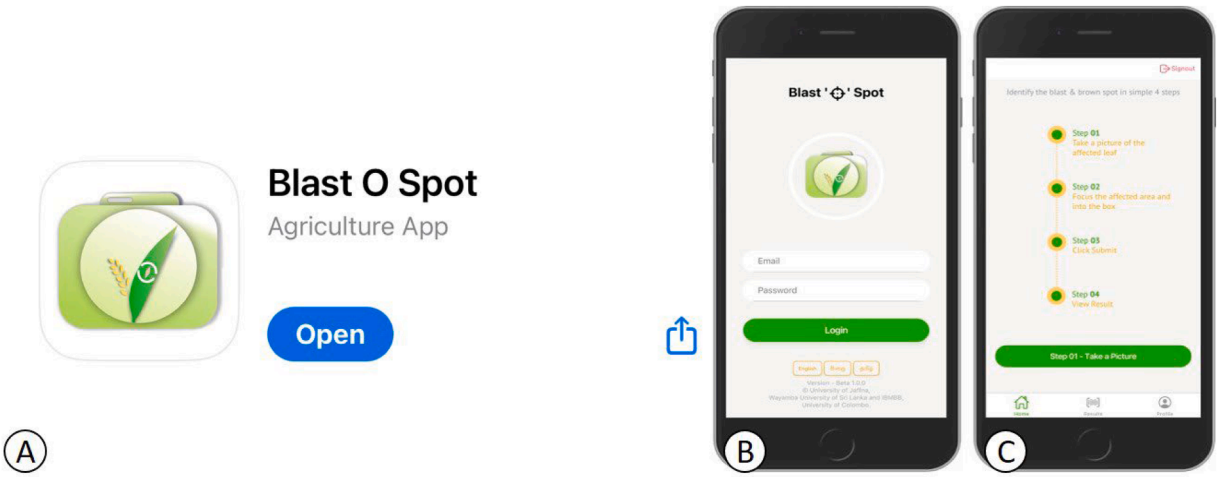

Fig. 4. This figure illustrates the process for users to access the "Blast O Spot" application. Initially, users download the application from the Apple Store (A). Upon opening the app, they are greeted with the home screen (B), which prominently features login options available in English, Sinhala, and Tamil languages. Additionally, the screen provides a succinct outline of the application process, highlighting its progression through four sequential steps (C).

activates the login function. Once authenticated, users are taken to the home screen (Fig. 4). The app also allows users to customize their language preferences, offering options such as Tamil, Sinhala, and English.

The application follows a structured four-step process, each with specific functions. Users begin by capturing images of infected leaves with RB and BS. They can then zoom in on the affected area to locate single lesion, ensuring precise positioning. Upon submitting the image, users promptly receive real-time results categorizing the infection as either RB or BS. The results screen (Fig. 5) presents classification outcomes and offers insights into the underlying business logic guiding the application's decision-making process.

2.4.2. Through smart phone camera

RB and BS lesions were captured in colour (RGB) with a smartphone camera (iPhone seven plus, Apple Inc., California). Using a wide-angle camera with a focal length of $f = 28$ mm and a telephoto lens with a focal length of $f = 56$ mm, the images were captured in open fields (between 1100 and 1400 h) to produce an array image with a resolution of 1080×1920 pixels at 401 pixels per inch (ppi). The camera was held parallel to the flat-leaf surface and a fixed distance of 50 cm was maintained between the leaf surface and the camera. Any shadows over the leaf were avoided during image capturing and the leaves were placed over a flat white background while attached to the plant to avoid the inclusion of non-target elements (i.e. other leaves and soil particles) in the imagery (Fig. 6a). Around 300 colour images containing a single disease lesion were extracted in JPEG format and labeled with the respective disease lesions (i.e. RB or BS).

2.5. RGB image processing and analysis

The analysis of RGB images was carried out using MATLAB 9.2 software (The MathWorks Inc., Natick, MA, USA) using customized scripts and in-built functions from the image processing toolbox. Here, a single lesion per single image approach was adopted for easy image analysis. In order to analyse the colour profiles from the RB or BS lesions of each single image, predominant pre-processing steps including image cropping, image normalizing, image segmentation, and colour feature extraction were executed.

2.5.1. Image cropping

The images were manually cropped to confine to targeted lesion (Fig. 6b) inclusive of a small part of surrounding live tissues (green area) [46,47]. Furthermore, a preliminary test conducted with 30 images proved that successful image segmentation of lesions was achievable when images were cropped before the segmentation process.

2.5.2. Image normalizing

Each cropped image of lesion was then normalized for varying light conditions that may have been attributed during image capturing in the open field [48]. For this, function *rescale*() was applied over each colour image to obtain pixel values between the scale of 0 to 1.

2.5.3. Image segmentation

In order to extract colour features from the targeted lesion area, a binary image was required (lesion pixels coded as 1 and other pixels as 0). Here, the area of interest was considered as the lesion area enclosed by the thick dark brown halo edge (Fig. 6b) and any pixels beyond that were segmented (green or yellow pixels) by colour thresholding [49]. Manual colour thresholding was performed using the CIE La*b* colour profile of the "Colour Thresholder App" in MATLAB, where threshold limits were set by interactively placing the image and its histogram side-by-side, visually inspecting the correct segmentation level of the desired lesion area (Fig. 7). A sensitivity test was performed using 30 images in the "Colour Threshoder App" to select the best colour profile from available RGB, HSV, YCbCr, La*b* profiles, where La*b was found to be the best performing for both RB and BS images. The resultant segmented binary image for each lesion was then applied as a mask over its normalized RGB image to obtain a masked RGB image (lesion pixels containing original RGB colour values and other pixels coded as 0).

2.5.4. Colour feature extraction

Here, the details of colour profiles recorded across the bright centre to lateral edges of the lesions were extracted coinciding with the distinctive RB and BS lesion pattern as described in Fig. 1. The following custom-developed script "extractColour.mat" was used to obtain a rectangular-shaped pointer (with a fixed height, ~40 rows) to be manually drawn over the masked RGB image connecting the opposite brown colour halo edges (as shown in Fig. 6d), to demarcate the feature extraction area.

```
Set imageFolder to the directory path containing images
List all JPEG files in the specified directory
```

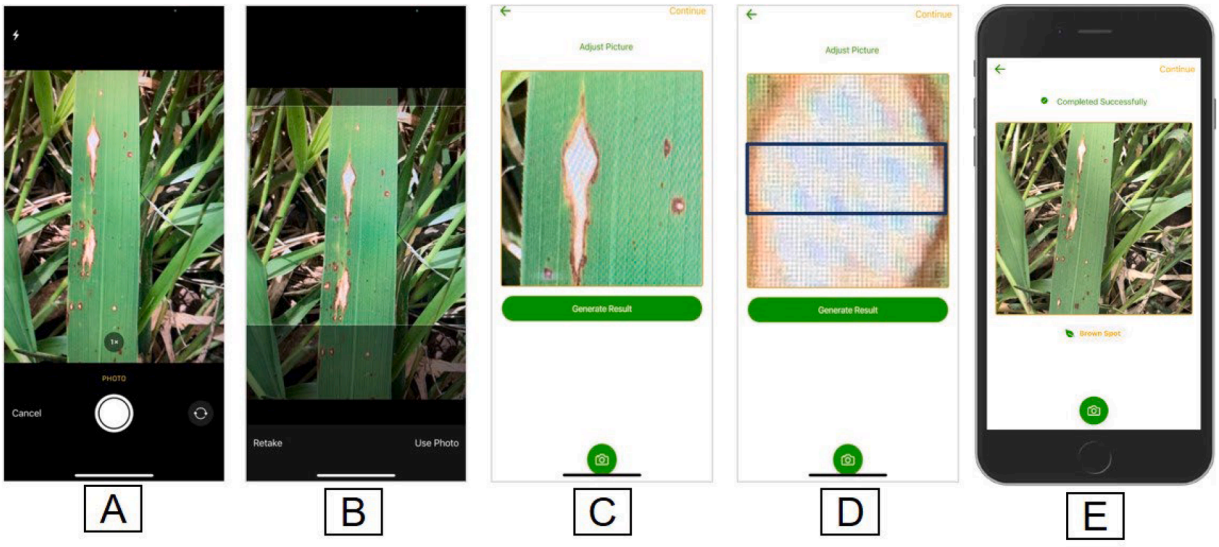

Fig. 5. The series of images illustrates the sequential steps following the capture of pictures of the lesion. A) Depicts the image immediately after capture, B) Shows the lesions magnified to a certain extent, C) Displays the entire lesion zoomed in for closer examination, and D) Demonstrates the selection of a box to analyze the area of interest, with the results displayed at the bottom, E) results page.

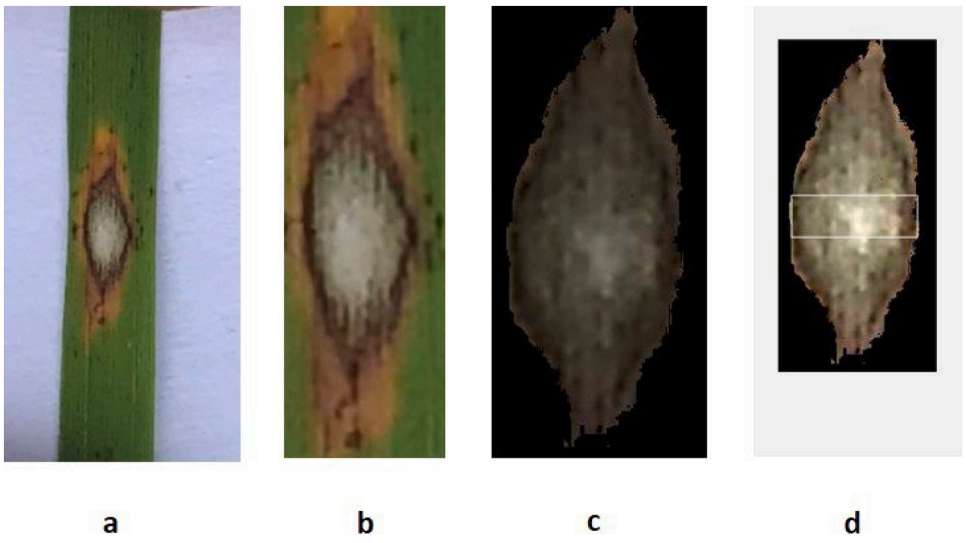

Fig. 6. Pictorial representation of the steps conducted in RGB image analysis; (a) original image of the lesion with a white background, (b) image after cropping to the targeted lesion of interest, (c) masked image after thresholding for lesion pixels (surrounding black pixels are excluded by masking), and (d) rectangular shaped pointer drawn over the lesion across the centre to lateral edges by the user for feature extraction.

```
results = {'Pxl', 'Mean_R', 'Mean_G', 'Mean_B', 'Mean_H', 'Mean_S', 'Mean_V', 'Mean_L', 'Mean_A', 'Mean_B'}
Display "started"
For each file in srcFiles
Get the path and name of the current image
Load the image
%(1) Load image
im = imread(imPath);
% (2) Normalize the image for light variations
imNorm = rescale(im);
% (3) Show where to extract information in the spot
% Prompt user to draw a line across touching widest edges horizontally
% in the spot. Line connects the widest edges across the bright center area
Display imNorm and prompt user to draw a line
Get position of the drawn line
% (4) Getting additional position details
leftX, leftY = position(1, 1), position(1, 2)
rightX, rightY = position(2, 1), position(2, 2)
centreX = min(leftX, rightX) + round(sqrt((rightX - leftX)^2) / 2)
centreY = min(leftY, rightY) + round(sqrt((rightY - leftY)^2) / 2)
numRowSample = 20
pinLeftX = leftX
pinLeftY = centreY - numRowSample
pinRightX = rightX
pinRightY = centreY + numRowSample
w = pinRightX - pinLeftX
h = pinRightY - pinLeftY
% Analysis:
% Convert image to HSV and LAB color spaces
% Extract mean values from the specified region
% Create the image with sampled location and save
Create a figure and display the image
Draw a rectangle around the sampled region
Save the image
% Prepare data for output
Store extracted pixel values and means in data variable
% Write data to Excel file
Write data to Excel file with image name
Display "end"
```

The script then derived the colour profile values from each pixel in the extraction area associated with the R, G, and B colour model, H, S, and V colour model, and L, a* and b* colour model to produce average colour values for each profile.

2.5.5. Deriving colour indices from extracted colour profiles

Colour indices listed in Table 1 were derived from the respective extracted averaged colour profile values for each RB and BS lesion.

2.6. Statistical analysis

The power of discriminating the two lesions (RB and BS) based on derived colour indices was statistically evaluated using k-means clustering in Python 3.9 package (Python Software Foundation, Virginia). The most outstanding colour indices which distinctively clustered the two lesions were selected based on the Euclidean distance and their accuracies were also recorded.

2.7. Confirmation of the results

To confirm the results, the rectangular-shaped pointer used for feature extraction was drawn freshly again over the imagery and the subsequent procedures (image processing, feature extraction, derive colour indices, and analysis) were also followed. This was repeated two times and classification accuracies were compared with the calibrated work. The data set used for the analysis is described as follows;

	Blast	Brown spot
Validation (Test)	70	70
Calibration 1	109	107
Calibration 2	100	100

The accuracy, precision, recall, and F1-score of the tested model were compared with some state-of-the art models as illustrated by Ritharson et al. [57].

2.8. Development of mobile app

The methodology for developing mobile application was based on the results obtained from laboratory-based analysis of BS and RB lesions,

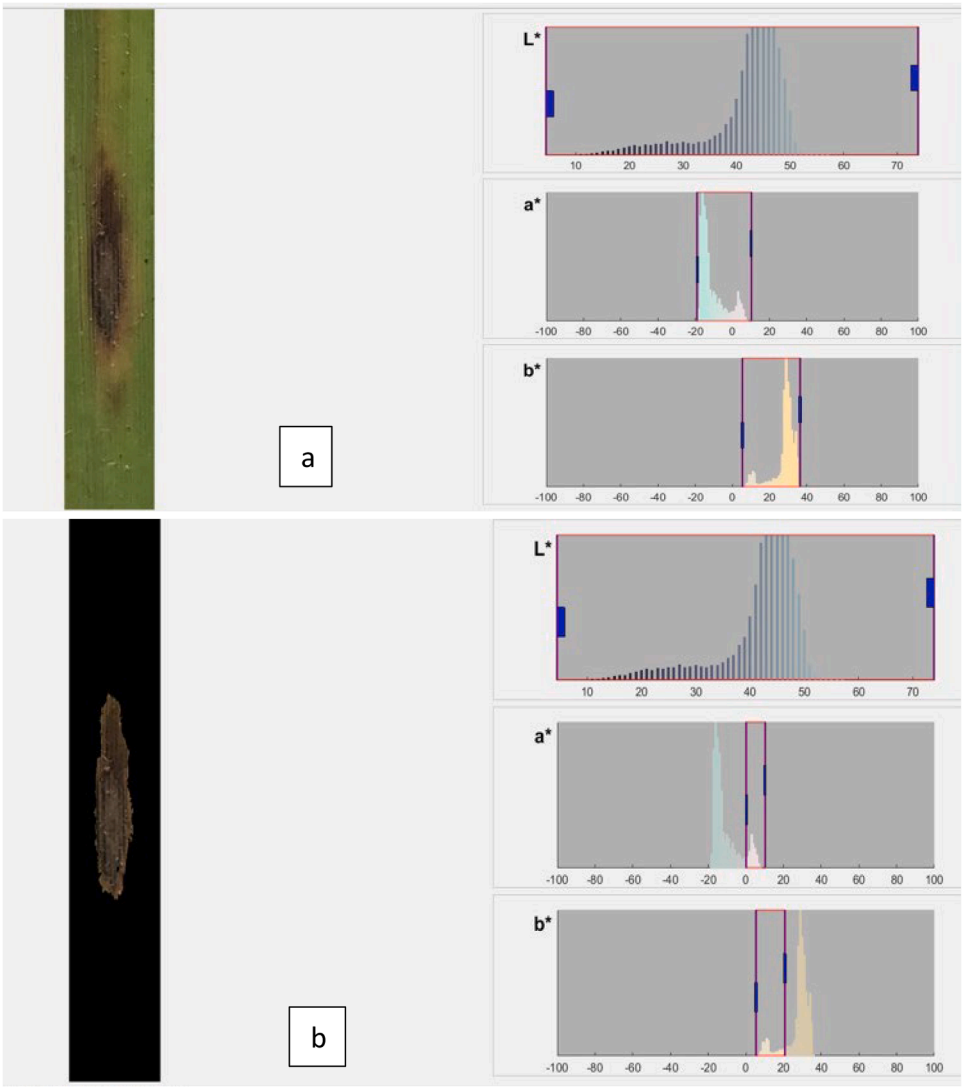

Fig. 7. Performing colour thresholding to segment lesion area from green leaf tissue pixels; a) original RGB lesion image and corresponding L*a*b histogram limits and b) threshold RGB lesion image after excluding green leaf tissue pixels (black) and resulted L*a*b histogram limits, where the lesion pixels were corresponding to the small peak in the a* and b* histograms.

along with a supervised classification method utilizing image sensing. Cerexio Pvt Ltd, a software development company based in Singapore and Sri Lanka, was tasked with the development process. The application was developed using the IonicTM framework, chosen for its hybrid application development capabilities. The system was designed to utilize smart algorithms and machine learning protocols to process images captured by users. The backend processing consisted of several critical steps. Initially, the Request Processor managed all incoming requests, translating the payload into a standard format for further processing. The Image Decoder then applied PNG and RGB decoding algorithms to extract the image from the payload, converting the byte stream into a processable format using the following:

```
file_bytes = numpy.asarray(bytearray(img_stream.read()), dtype=numpy.uint8)
img_data_ndarray = cv2.imdecode(file_bytes, cv2.IMREAD_UNCHANGED)
```

Following this, Image Validation and Enhancement involved improving the input image through geometric transformation, color space manipulation, analysis, filtering, and morphology to standardize it. Shadow correction was a key part of this step, where shadows caused by illumination changes were removed using a color-based shadow detection technique that compared chromatic and brightness components against background components. The pseudocode for this included operations such as dilation, median blurring, and normalization:

```
rgb_planes = cv2.split(img)
result_planes = []
result_norm_planes = []
for plane in rgb_planes:
dilated_img = cv2.dilate(plane, np.ones((7, 7), np.uint8))
bg_img = cv2.medianBlur(dilated_img, 21)
diff_img = 255 - cv2.absdiff(plane, bg_img)
norm_img = cv2.normalize(diff_img, None, alpha=0, beta=255, norm_type=cv2.NORM_MINMAX, dtype=cv2.CV_8UC1)
result_planes.append(diff_img)
result_norm_planes.append(norm_img)
result = cv2.merge(result_planes)
result_norm = cv2.merge(result_norm_planes)
```

Finally, Feature Extraction isolated specific image features for the trained model, using algorithms to create a 2D map array. Edge features, in particular, were extracted using the Canny edge detection algorithm, identifying areas with significant pixel value changes. The pseudocode

Table 1
Colour indices used for the analysis.

No	Name	Equation/ Definition	References
1.	R	Mean of red	[38]
2.	G	Mean of green	[38]
3.	B	Mean of blue	[38]
4.	H	Mean of Hue	[38]
5.	S	Mean of Saturation	[38]
6.	V	Mean of brightness	[38]
7.	L	Mean of lightness (L)	[50]
8.	a*	Mean of a*	[50]
9.	b*	Mean of b*	[50]
10.	Nor_r	Normalised red $=R/(R+G+B)$	[41]
11.	Nor_g	Normalised green $=G/(R+G+B)$	[41]
12.	Nor_b	Normalised blue $=B/(R+G+B)$	[41]
13.	Intensity (INT)	$INT=(R+G+B)/3$	[51]
14.	$VI_{green\_}$ Vegetation index	$VI_{Green} = (G-R)/(G+R)$	[52]
15.	Difference between green and blue	G-B	[53]
16.	Difference between green and red	G-R	[53]
17.	Sum of Green and Red	$G+R$	[41]
18.	Sum of Red and Blue	$R+B$	[41]
19.	Difference between normalized green and blue	g-b	[37]
20.	Normalized green red difference index	$NGRDI=(g-r)/(g+r)$	Tucker, 1979
21.	Excess red vegetation index	ExR = 1.4r-g	[54]
22.	Excess blue vegetation index	ExB = 1.4b-g	[54]
23.	Excess green vegetation index	ExG = 2 g-r-b	[54]
24.	Excess green minus excess red	ExGR = ExG-ExR	[54]
25.	Kawashima Index	$I_{KAW} = (R-B)/(R+B)$	[55]
26.	Dark green colour index	DGCI = [((Hue)−60)/60+(1-(Saturation))+(1-(Brightness))]/3	[56]
27.	Colour feature index	G/B	[53]
28.	Colour feature index	G/R	[53]
29.	Red green ration index	$RGRI = R/G$	[37]
30.	Green leaf index	$GLI = (2G-R-B)/(2G+R+B)$	Louhaichi et al., 2001
31.	Visible atmospherically resistance index	$VARI = (G-R)/(G+R-B)$	Gitelson, 2002
32.	Colour feature value	Logsig = [G-R/3-G/3]/255	Ali et al., 2012
33.	Colour feature index	$(G-B)/(R+G+B)$	[41]
34.	Colour feature index	$(G+B)/R$	[41]
35.	Colour feature index	$(R+G+B)/R$	[41]
36.	Colour feature index	b*/a*	[51]

for edge detection was as follows:

```
edges = feature.canny(decoded_image)
```

The trained model evaluated the input map array to select a trained label based on the actual confidence level. The result was passed back to the Request Processor component for dispatching to the mobile application. This comprehensive backend processing pipeline ensured that images were validated, enhanced, and analyzed effectively for accurate plant disease detection.

3. Results

3.1. Disease symptom sampling and verification

The collected samples were successfully identified as blast or brown spot using characteristic microscopic features of colony morphology and positive amplification of pot2 region. The identification of the lesions was also verified by the experts before capturing the images in the fields.

3.2. RB and BS lesions imagery

From acquired RB and BS lesion images, 200 (100 per each symptom) were selected which had clear lesion without any distortion caused by sunlight.

3.3. Evaluating the spatial pixel colour distribution pattern in lesions

In order to get an initial understanding of the colour distribution pattern across the area of interest (AOI; rectangular pointer area at the horizontal widest edges) of each RB and BS lesions (Fig. 7a and 7b), the spatial pixel distribution of the means of R, G, B, H, S, V and L, a*, b* colour profiles of the images after normalizing were plotted by scanning the AOI from left to right pixels or across the lesion (Fig. 8).

According to Fig. 8, each colour profile across the widest horizontal area (AOI) in RB and BS lesions resulted in a unique pixel colour distribution pattern. Both RB and BS showed a peak in colour distribution associated with the centre area of pixels of the lesion and a low mean colour at the pixels from the edges of the lesion, but the resulted peaks of RB were smooth than the irregular-shaped peak from BS (Fig. 8a and 8b). When comparing the colour models, for RGB, all means of R, G, and B showed an alteration across the lesions (Fig. 8c and 8d) while for HSV, mean of V and for La*b*, means of a* and b* remained unchanged (Fig. 8e-8h). This suggests that all available profiles (R, G, and B) in the RGB colour model are capable of providing useful colour information associated with the RB and BS lesions than the HSV and La*b* colour models.

3.4. Performance of derived colour indices on discriminating RB and BS

To distinguish the RB and BS lesions, the extracted means of R, G, B, H, S, V and L, a*, b* from AOI were computed as colour indices, where 36 colour indices were derived for a single lesion image as shown in Table 1.

Colour indices are derivable features from RGB colour digital imagery predominantly used for characterising visible changes in leaves or vegetation. Li et al. [58] reported that RGB-based colour indices such as intensity (INT) and VI Green were widely adopted in estimating the nitrogen content of the leaf. In addition, the colour indices calculated from HSV and La*b* colour channels have shown significant correlations with leaf chlorophyll concentration of rice leaves [51].

3.4.1. Performance of unsupervised clustering

A series of k-means cluster analyses (k = 0 to 10) were performed over the pool of derived colour indices to test the optimum number of clusters that were possible to produce by unsupervised method (Fig. 8) from all the images. The square of Euclidean distance resulted in each clustering attempt was plot in the elbow graph as shown in Fig. 9. The elbow graph showed two clusters as the optimum stage which were found to be associated with the two lesion groups (i.e. RB and BS). Therefore, this proved that derived 36 colour indices were strongly indicative of dissimilarities between the two major lesion groups. The silhouette coefficients for the clusters are all positive and ranged from 0.9 to 1, which validated the clusters.

3.4.2. Performance of supervised cluster analysis

A fresh k-means cluster analysis (k = 2) was performed over the ground truth labelled (RB and BS) image derived colour indices (36 colour indices as variables) to test the classification accuracy of discriminating each data point into known RB or BS group clusters. This analysis reported a high accuracy of 84.6 % for discriminating RB and BS by the selected 36 colour indices.

Furthermore, resulted Euclidean distances between the clustered groups by each colour index during the cluster analysis were plotted visually over a spider plot (Fig. 10). In the spider plot, all the colour indices depicted varying distances between the two clusters where for

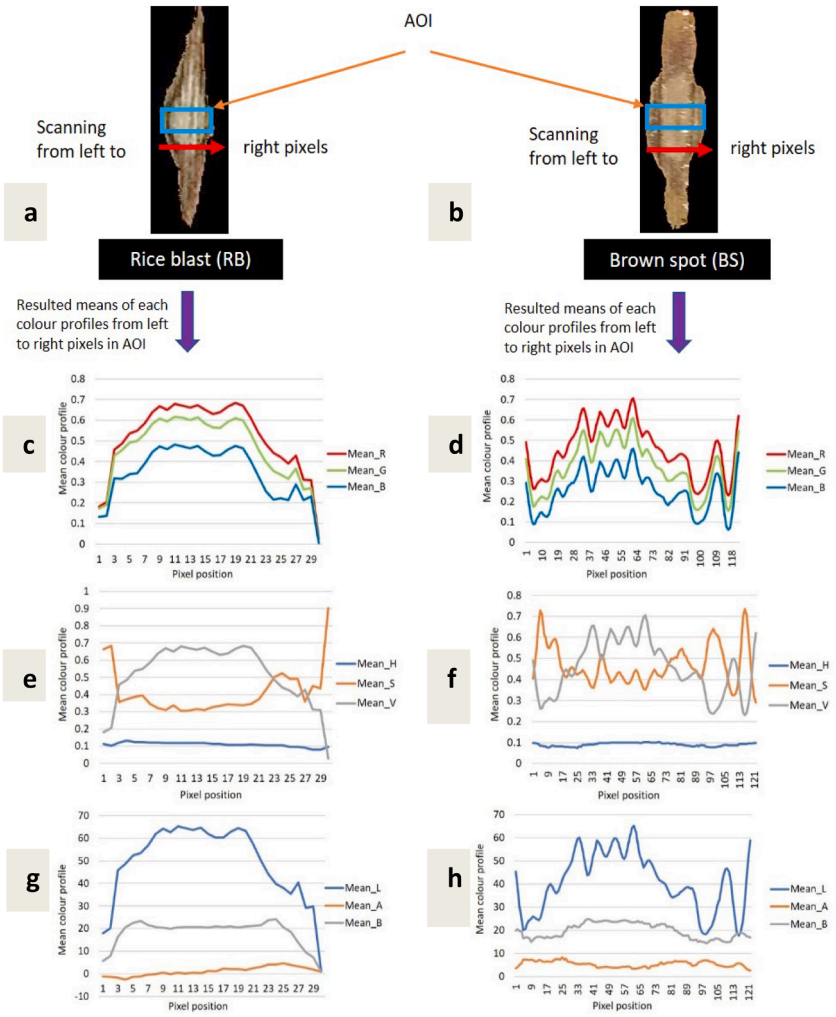

Fig. 8. Graphical interpretation of the spatial pixel colour distribution pattern in RB and BS lesions from normalized images; a) normalized RGB image of rice blast lesion, b) normalized RGB image of brown spot lesion, c-h) plots indicating the spatial pixel distribution of mean colour profiles, c & d) mean R, G, B, e & f) mean H, S, V, and g & h) mean L, a*, b*.

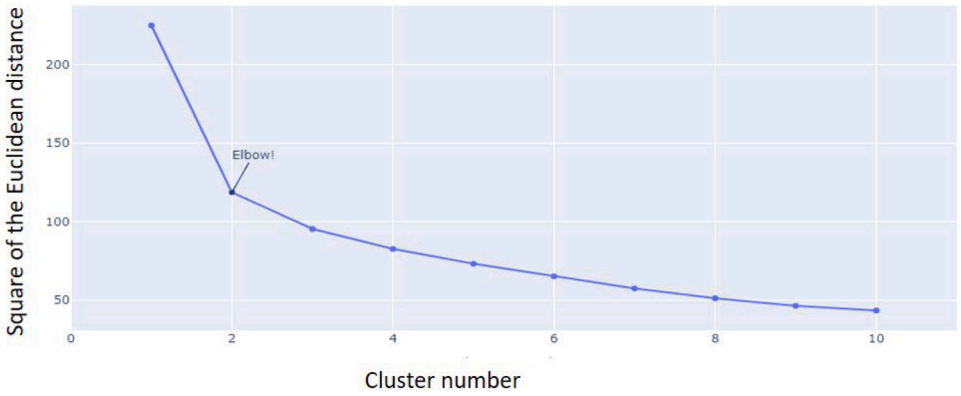

Fig. 9. The elbow graph showing the optimum number of clusters possible to produce from the pool of derived colour indices, supervised.

instance the distance between the red and blue lines for colour index 'Nor_r' showed a higher distance (> 0.3 Euclidean) than the distance between the respective lines in b*/a* (<0.1). For further understanding, the exact Euclidean distances resulted from each colour index were listed in Table 2. According to Table 2, the highest distance (0.38 Euclidean) between RB and BS clusters was obtained by colour index G/R and the lowest (0.00) separation was obtained by b*/a*. On the other hand, the results also suggest that discrimination of RB and BS lesions at an 84.6 % accuracy might be due to the contribution of several colour indices, and not necessarily by all 36 indices.

To further investigate this, scatter plots of data points showing the classification performance by the individual colour index were generated using the distance from the centroid of each respective clustered group to its member data points (Fig. 10). According to Table 2 and

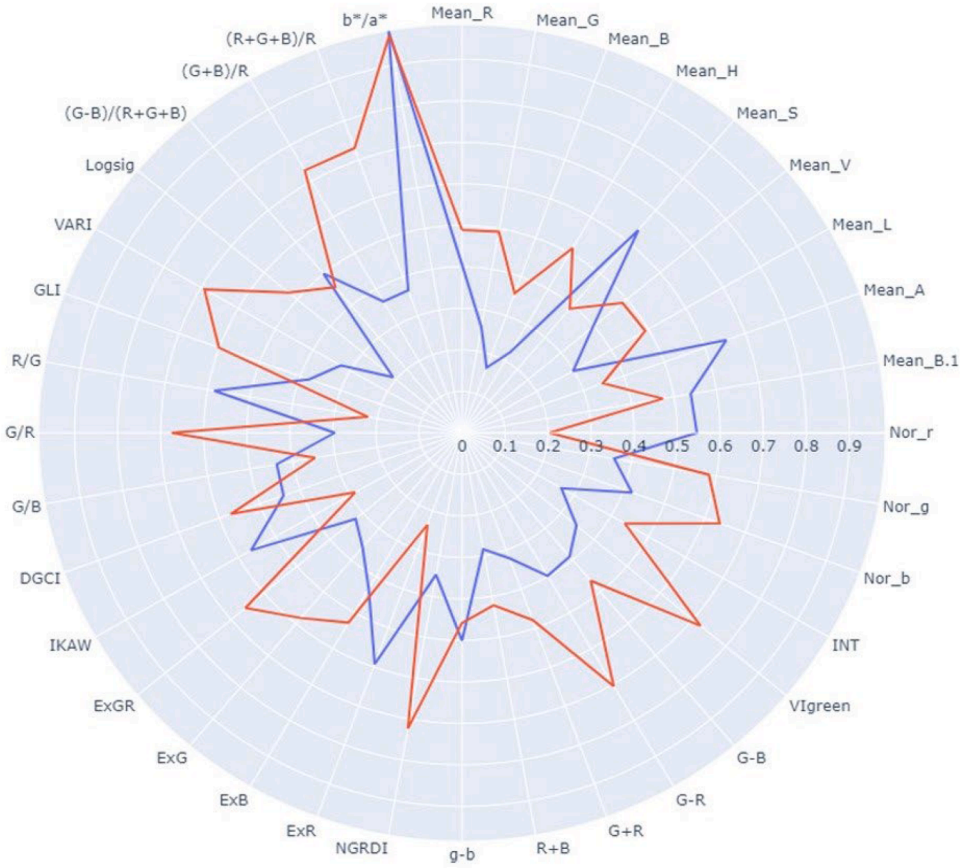

Fig. 10. Graphical interpretation of the cluster distances resulted by k-mean cluster ($k = 2$) analysis over derived colour indices; cluster corresponding with rice blast lesion (red line) and brown spot lesion (blue line), x-axis is showing the Euclidean distance.

Fig. 11, it was clearly visible that colour indices associated with a higher distance (> 0.27 Euclidean) between the clusters showed distinct separation (i.e. G/R, VIgreen, NGRDI, etc.) between the groups but colour indices associated with less distance (< 0.27 Euclidean) showed overlapping between the two groups (i.e.; b*/a*, g-b, (G-B)/($R + G + B$), etc.). This must be due to the insensitivity of such colour indices to the changes in colour in the AOI of the respective lesions.

3.5. Selecting the optimum number of colour indices for discriminating RB and BS lesions

In order to investigate the optimum number of colour indices, among the 36 selected, that are able to produce more or less the same accuracy (84.6 %), the clustering process was iterated several times by systematically excluding colour indices based on the minimum distance between produced clusters. Major stages where changes in accuracy occurred in this optimization process were listed in Table 3. Interestingly, it was found that the same accuracy (84.6 %) of discriminating RB and BS lesions when using the 36 indices, was also obtained by four colour indices (Table 3: Calibration analysis round). When the number of indices was further reduced down to three, accuracy lowered to 82.7 %, where then four colour indices stage can be suggested as the optimum number that is useful for discriminating the RB and BS lesions rather all 36 indices. Further those optimum four colour indices were G/R, VIgreen, NGRDI, $R + G + B$/R (illustrated in Table 4) which were also listed at the top in Table 2 for the indices created maximum distance (>0.36 Euclidean) between clusters.

3.6. Confirmation of the clustering results

In order to confirm the obtained results received at the calibration stage, the whole process was repeated two times by freshly drawing the AOI rectangular pointer over each image. The derived colour indices were then analyzed in the same manner and resulted accuracies are listed in Table 3. In both instances (‘repetition 1′ and ‘repetition 2′) accuracies reported in discriminating RB and BS lesions with all the indices (36) were similar to the calibration results (~84 %). Even though after optimization, the reported accuracies remained consistent with the calibration results (~84 %) (Table 3). Though the accuracies were the same, colour indices that contributed to the separation slightly changed in each repetition attempt (Tables 3 and Table 4). In ‘repetition 1′ VARI had been selected in addition to the indices selected during calibration and in ‘repetition 2′ except $R + G + B$/R the rest of the indices were different from the calibration ones (Table 4). This alteration must be due to the change in colour information derived at each repetition attempt which is attributed to the newly drawn AOI over images. However, all those colour indices were found to be listed at the top 10 in the calibration work as shown in Table 2, which had the maximum Euclidean distance (>0.28) between clusters.

The Table 5 illustrates the performance of the model tested in this study based on accuracy, precision, recall, and F1-scorewhen compared to some of the state of the art models.

4. Discussion

Digital colour images derived from smartphones can be used to generate colour indices that are sensitive to RB and BS lesions. The results proved that the proposed method of using colour indices to assess

Table 2
The resulted Euclidean distance for blast (BL) and brown spot (BS) for 36 colour indices.

No	Colour Indices	Euclidean distance of Rice Blast	Euclidean distance of Brown spot	Distance between clustered groups
1	G/R	0.66	0.30	0.36
2	VIgreen	0.71	0.35	0.36
3	NGRDI	0.71	0.35	0.36
4	$(R+G+B)/R$	0.73	0.37	0.36
5	$(G+B)/R$	0.73	0.37	0.36
6	VARI	0.68	0.32	0.35
7	R/G	0.22	0.57	0.35
8	ExR	0.23	0.58	0.34
9	Nor_r	0.19	0.53	0.33
10	ExGR	0.65	0.32	0.32
11	G-R	0.70	0.39	0.30
12	Mean_A	0.35	0.65	0.29
13	Logsig	0.54	0.25	0.28
14	IKAW	0.28	0.55	0.27
15	Nor_b	0.64	0.42	0.21
16	GLI	0.59	0.38	0.21
17	Nor_g	0.58	0.36	0.21
18	ExG	0.58	0.36	0.21
19	Mean_S	0.27	0.47	0.20
20	Mean_G	0.62	0.44	0.18
21	Mean_B	0.49	0.33	0.16
22	Mean_H	0.27	0.12	0.15
23	Mean_L	0.66	0.51	0.14
24	DGCI	0.56	0.42	0.13
25	INT	0.59	0.46	0.13
26	$G+R$	0.64	0.52	0.11
27	$R+B$	0.57	0.46	0.10
28	ExB	0.52	0.42	0.09
29	G/B	0.35	0.44	0.09
30	Mean_B.1	0.50	0.58	0.07
31	Mean_V	0.64	0.58	0.06
32	Mean_R	0.64	0.58	0.05
33	G-B	0.47	0.41	0.05
34	g-b	0.46	0.51	0.04
35	$(G-B)/(R+G+B)$	0.46	0.51	0.04
36	b*/a*	0.13	0.12	0.00

the horizontal widest edges of respective lesions is successful in delineating the lesions into two separate groups. Although this adopted method required the user to manually draw a rectangular pointer area at the horizontal widest edges (centre area) of the lesions the resulted accuracies (84.6 %) were consistent throughout the repeats even when the rectangular pointers were freshly drawn on the same lesion again and again. Thus, the findings suggest that colour indices do qualify to extract the colour behaviour across the lesion centre area at different severity stages irrespective of the user who draws the rectangular pointer. Those in the field who must distinguish the disease in order to devise efficient control strategies may find this to be useful.

In this proposed method, a combination of 36 colour indices derived from several colour profiles (R, G, B, H, S, V and L, a*, b*) produced the top accuracy (84.6 %). Similar accuracy was also produced when indices were reduced down to <10 that was coincidently associated with the R, G, and B colour profiles. This is consistent with the previous research like the work of Mohan and Gupta [38], who suggested that R, G, B colour indices are predictive of the chlorophyll content associated with leaves. Further, chlorophyll content is also used as a popular biomarker for many disease symptom diagnoses due to the property of these diseases in altering reflected spectrum from lesion areas [52,59,60]. Usually, these colour models had been optimized in plant diseases identification work [61,62] where Naik et al. [63] used RGB colour channel coupled with k-means clustering technique to classify three diseases in cabbage; black rot disease, frog eye disease and fungal disease from the healthy cabbage. Kulkarni and Patil [64] worked on La*b* colour channel to identify three major diseases that affect pomegranate fruit; bacterial blight, anthracnose and wilt complex, and the HSV colour channel was effectively used to differentiate the diseases in wheat crop and segmentation usage of the k-means algorithm [62].

The results are also consistent with similar work previously conducted to identify rice diseases based on different classification techniques [46,60,65,66]. The diseases; rice bacterial blight, RB, BS, and rice sheath rot have been classified using Minimum Distance Classifier (MDC) and k-Nearest Neighbor classifier (kNN) with the accuracy of 87.02 and 89.23 percent, respectively [46]. Shrivastava and Pradhan [67] used support vector machine (SVM) classifier to differentiate bacterial leaf blight, RB, sheath blight and achieved 94.65 % accuracy. Guchait et al. [48] detected RB and BS diseases using texture, shape, colour, and intensity distribution features and achieved 81.81 % of accuracy for IBK classifier which is a k-nearest neighborhood algorithm used for data clustering [68,69]. Shah et al. [70] reported in the work that k-mean clustering analysis is more suitable for classifying leaf images. Anthonys and Wickramarachchi [71] have discriminated RB, rice sheath blight, and BS using a classification method of membership function with the accuracy of 70 %. Yao et al. [72] used SVM to classify RB, rice sheath blight, and bacterial leaf blight with an accuracy of 97.2 %. There are few state of the art techniques available to discriminate rice diseases like Region-based Convolutional Neural Network (R-CNN) which identified healthy leaves with an accuracy of 99.25 %. A Residual Network model (ResNet19) was used to classify some rice diseases (RB, BS and Hispa) with 99.58 % of accuracy. Some researchers used a combination of methods such as FCM-KM and Faster R-CNN and obtained an accuracy of 69 to 99 % to discriminate rice diseases [57]. Considering the above accuracy levels obtained using unsupervised methods, an accuracy of 84 % (on average) produced from this supervised method-based study is an encouraging result. Additionally, the other performance indicators such as precision, F1 score and recall values are also suggested the effectiveness of this model in discriminating RB and BS.

When a rectangular box with 40 rows of height in the centre of the lesion is included in the analysis, the accuracy level stated above will be obtained. In addition, if the lesions have several spots, every single spot should be taken for analysis separately.

Image segmentation is a key step in image processing [13]. It is a process to divide an image into significant regions or segments and locate the interested target used for further analysis [12]. In other words, it is extracting the leaf spot from the background. Here, the segmentation process was done at three levels. The first one is image smoothing where the noise was eliminated as much as possible to make the true edges visible. Secondly, detection of edge points of leaf spot was done to characterize the point from where to draw the rectangular pointer. This is a significant step, as it determines the discriminating feature for both diseases. This rectangular pointer was drawn manually through the wider region of the lesion close to the middle area of the lesion so that a user can easily detect the start and stop positions. Thirdly, the selection of pixel rows was determined to calculate the mean value of each colour component in the colour model used. For this purpose, 40-pixel rows were automatically selected and the average was taken for the calculation. This might minimize the error results from different users when drawing the rectangular pointer within the center area.

The image analysis procedure could be further improved with an automated approach of drawing the rectangular pointer area across the widest horizontal edges in the lesion. An option could be adopting a combination of 'Canny edge detection' and 'Otsu's' methods as suggested by Salgadoe et al. [36]. Further, as features, additional layers of dimensions such as texture information [48] and wavelet-based features [39] can be combined with colour indices. The proposed method can also be further improved to analyse multiple lesions per single image rather than a single lesion per image which will require complex thresholding and segmentation procedures and that will help fast monitoring the number of images at a given time.

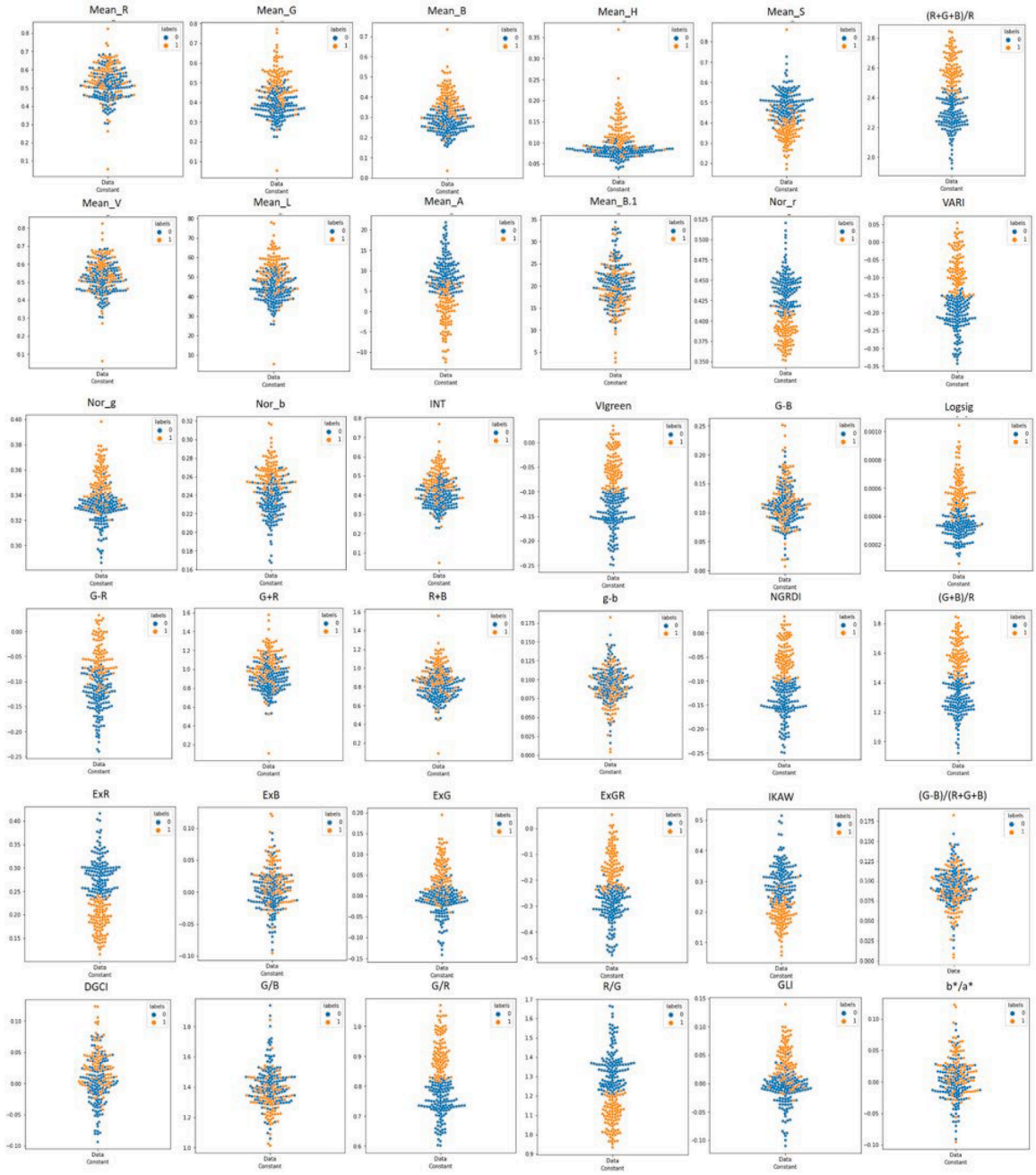

Fig. 11. Scatter plots showing the distribution of data points clustered into respective groups by each colour index; orange dots (RB) and blue dots (BS).

Table 3
The changes in accuracies resulted during optimization of colour indices at each round.

Major stages, changes in accuracy occurred	Analysis Round					
	Calibration analysis		Repetition 1		Repetition 2	
	Number of colour Indices	Accuracy (%)	Number of colour Indices	Accuracy (%)	Number of colour Indices	Accuracy (%)
All indices	36	84.6	36	84.4	36	84.0
Optimum indices	4	84.6	5	84.4	7	84.0
Lesser indices	3	82.7	4	82.5	6	82.1

The developed mobile application, "Blast O Spot" has demonstrated efficacy during trial mode with test images obtained during research. It has now been released as a Beta version in Apple store, representing a new application that may require further refinements based on user feedback and results. Notably, "Blast O Spot" exhibited a remarkable 100 % accuracy rate in identifying lesions. Its simplicity and user-friendly

Table 4
Summary of optimum colour indices contributed for discriminating RB and BS lesions for each analysis round.

Analysis Round	G/R	VIgreen	NGRDI	(R+G+B)/R	VARI	(G+B)/R	R/G	Nor_r	G-R	Mean_A	Logsig	*Acc. %
Calibration	■	■	■	■								84.6
Repetition 1	■	■	■	■	■							84.4
Repetition 2				■		■	■	■	■	■	■	84.0

*Acc- Accuracy.

Table 5
Comparison of the model used in the present study with already established state of the art models (from [57]).

Model Name	Precision	Recall	F1_score	Accuracy
Xception	0.977	0.977	0.977	97.7 %
DenseNet121	0.938	0.938	0.938	93.8 %
InceptionResnetV2	0.933	0.933	0.933	93.3 %
Resnet 50	0.811	0.811	0.811	81.1 %
InceptionV3	0.955	0.955	0.955	95.5 %
VGG 16	0.988	0.988	0.988	98.8 %
The present model used in "Blast O Spot"	0.941	0.991	0.965	84.0 %

interface, coupled with multilingual support, render it highly advantageous for farmers worldwide. Additionally, the application provides summary results, including the location of access via GPS integration in smartphones, along with the presence of RB/BS. This continuous data generation fosters a comprehensive database for future reference and enables developers to validate the application's functionality. Comparatively, RiceDoctor, an application developed by the IRRI to identify various rice-related issues such as diseases and nutrient deficiencies, presents a more intricate and cumbersome operational interface. In contrast, "Blast O Spot" boasts enhanced user-friendliness and specificity towards RB and BS lesions, thereby ensuring higher accuracy in outcomes. Notably, "Blast O Spot" "Blast O Spot" marks a pioneering effort in the development of an application specifically tailored for distinguishing between RB and BS lesions.

The device's influence as well as external variables like light intensity, shadows, and backdrop color that can affect the image quality could be the method's main limitations. Identification may also be impacted if the symptoms are not apparent in the early stages. Furthermore, there may be a problem if a user needs to be properly trained in order to capture the images appropriately. But the tool can be used in real time which makes the decision making much more easy and it can be used effectively with little practice which can be the highlights of the findings.

5. Conclusions

Our study addresses the challenge of disease identification by developing a non-destructive, cost-effective method using digital RGB images acquired from a smartphone camera. This approach leverages the accessibility of smartphones and the simplicity of image acquisition, making it a practical solution for farmers and agricultural professionals in the field. Based on our findings, we developed a mobile application, "Blast O Spot," designed to assist in accurately identifying and differentiating between RB and BS diseases. This tool enables the selection of resistant cultivars and other specific management practices. The application is user-friendly, enhancing disease management by allowing farmers to take images of symptoms and adjust the spot within a movable square to prevent false identification when multiple spots are present. This reduces dependency on expert visual identification and minimizes the need for expensive laboratory tests. By facilitating early and accurate disease detection, the application supports effective and sustainable agricultural practices.

Ethics statement

Not applicable: This manuscript does not include human or animal research.

If this manuscript involves research on animals or humans, it is imperative to disclose all approval details.

If Yes, please provide your text here:

CRediT authorship contribution statement

Suvanthini Terensan: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. **Arachchige Surantha Ashan Salgadoe:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Nisha Sulari Kottearachchi:** Writing – review & editing, Validation, Conceptualization. **O.V.D.S. Jagathpriya Weerasena:** Writing – review & editing, Validation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Suvanthini Terensan reports financial support was provided by University Grants commission, Sri Lanka. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors gratefully acknowledge the support provided by the Rice Research Station, Paranthan, Sri Lanka and the Pathology Division, Rice Research and Development Center, Bombuwala, Sri Lanka in collecting diseased samples. Also, authors would like to offer sincere gratitude to Mr. S. Muraleetharan, Environmental officer, Central Environmental

Authority, Sri Lanka for guiding in statistical analysis methods in Python software. Finally, authors would like to thank the funding agency, the University Grants Commissions, Sri Lanka.

References

[1] International Rice Research Institute. 2024 https://www.irri.org/world-food-day-2019-rice-zero-hunger.

[2] P.B. Magar, Screening of Rice Varieties against Brown Leaf Spot Disease at Jyotinagar, Chitwan, Nepal, Int. J. Appl. Sci. Biotechnol. 3 (1) (2015) 56–60, https://doi.org/10.3126/ijasbt.v3i1.12014.

[3] T.W. Mew, H. Leung, S. Savary, C.M.V. Vera Cruz, J.E. Leach, Looking ahead in rice disease research and management, CRC Crit. Rev. Plant. Sci. 23 (2) (2004) 103–127, https://doi.org/10.1080/07352680490433231.

[4] T.R. Sharma, A.K. Rai, S.K. Gupta, J. Vijayan, B.N. Devanna, Rice Blast Management through host-plant resistance: retrospect and prospects, Agricult. Res. 1 (1) (2012) 37–52, https://doi.org/10.1007/s40003-011-0003-5.

[5] L. Nalley, F. Tsiboe, A. Durand-Morat, A. Shew, G. Thoma, Economic and environmental impact of rice Blast Pathogen (Magnaporthe oryzae) alleviation in the United States, PLoS One 11 (12) (2016) e0167295, https://doi.org/10.1371/journal.pone.0167295.

[6] Y.M. Shabana, G.M. Abdel-Fattah, A.E. Ismail, Y.M. Rashad, Control of brown spot pathogen of rice (Bipolaris oryzae) using some phenolic antioxidants, Braz. J. Microbiol. 39 (3) (2008) 438–444, https://doi.org/10.1590/S1517-83822008000300006.

[7] N.K. Chakrabarti, Epidemiology and disease management of brown spot of rice in India. Major Fungal Diseases of Rice, Springer, Dordrecht, 2001, pp. 293–306, https://doi.org/10.1007/978-94-017-2157-8_21.

[8] D.E. Groth, Rice diseases and disorders in Louisiana, LSU Agricultural Experiment Station (1991). Report number: 668.

[9] M.K. Barnwal, A. Kotasthane, N. Magculia, P.K. Mukherjee, S. Savary, A.K. Sharma, H.B. Singh, U.S. Singh, A.H. Sparks, M. Variar, N. Zaidi, A review on crop losses, epidemiology and disease management of rice brown spot to identify research priorities and knowledge gaps, Eur. J. Plant Pathol. 136 (3) (2013) 443–457, https://doi.org/10.1007/s10658-013-0195-6.

[10] H. Kato, Rice blast disease, Pestic. Outlook 12 (1) (2001) 23–25.

[11] S. Sunder, R.A.M. Singh, R. Agarwal, Brown spot of rice: an overview, Indian Phytopathol 67 (3) (2014) 201–215.

[12] M.A. Bakar, A.H. Abdullah, N.A. Rahim, H. Yazid, F.S.A. Saad, K. Ahmad, Development of ripeness indicator for quality assessment of Harumanis mango by using image processing technique, in: *Materials Science and Engineering Part C, IOP Conference Series*, 932, IOP Publishing, 2020 012087.

[13] S. Savant, A review on edge detection techniques for image segmentation, Int. J. Comput. Sci. Informat. Technol. 5 (4) (2014) 5898–5900.

[14] P. Kachroo, S.A. Leong, B.B. Chattoo, Pot2, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*, Mol. Gen. Genet. 245 (3) (1994) 339–348, https://doi.org/10.1007/BF00290114.

[15] P. Mishra, G. Polder, N. Vilfan, Close range Spectral Imaging for disease detection in plants using autonomous platforms: a review on recent studies, Current Robotics Reports 1 (2) (2020) 43–48, https://doi.org/10.1007/s43154-020-00004-7.

[16] J.H. Everitt, D.E. Escobar, D.N. Appel, W.G. Riggs, M.R. Davis, Using airborne digital imagery for detecting oak wilt disease, Plant Dis. 83 (6) (1999) 502–505, https://doi.org/10.1094/PDIS.1999.83.6.502.

[17] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L.R. Goulart, C.E. Davis, Advanced methods of plant disease detection. A review, Agron. Sustainable Dev. 35 (2015) 1–25.

[18] V. Dhanushkodi, T.B. Priyadharshini, M. Baskar, S. Meena, K. Senthil, T. U. Maheshwari, Slow and controlled release nitrogen fertilizers: options for improving rice productivity: a review, Int. J. Plant Soil Sci. 34 (24) (2022) 970–981.

[19] J. Kromdijk, S.P. Long, One crop breeding cycle from starvation? How engineering crop photosynthesis for rising CO2 and temperature could be one important route to alleviation, in: Proceedings of the Royal Society B: Biological Sciences 283, 2016 20152578.

[20] W.Y. Wubneh, F.A. Bayu, Assessment of diseases on rice (Oriza sativa L.) in major growing fields of Pawe district, Northwestern Ethiopia, World Scientific News (42) (2016) 13–23.

[21] S.S. Gnanamanickam, Biological Control of Rice Diseases, 8, Springer Science & Business Media, 2009.

[22] M. Surendhar, Y. Anbuselvam, J. Ivin, Status of rice brown spot (Helminthosporium oryz) management in India: a review, Agricult. Rev. 43 (2) (2022) 217–222.

[23] S.D.S. Seneviratne, P. Jeyanandarajah, Rice diseases-problems and progress, Tropical Agricult. Res. Extension 7 (2004) 30–48.

[24] I. Yamaguchi, Overview on the chemical control of rice blast disease, in: Rice Blast: Interaction with rice and control: Proceedings of the 3rd International Rice Blast Conference, Springer Netherlands, Dordrecht, 2004, pp. 1–13.

[25] R.K. Horst, Westcott's Plant Disease Handbook, Springer Science & Business Media, 2013.

[26] E.Y. Bridson, G.W. Gould, Quantal microbiology, Lett. Appl. Microbiol. 30 (2) (2000) 95–98.

[27] F.C. Hsieh, M.C. Li, T.C. Lin, S.S. Kao, Rapid detection and characterization of surfactin-producing Bacillus subtilis and closely related species based on PCR, Curr. Microbiol. 49 (2004) 186–191.

[28] H.Y. Lau, J.R. Botella, Advanced DNA-based point-of-care diagnostic methods for plant diseases detection, Front. Plant Sci. 8 (2017) 2016.

[29] W. Photita, P.W. Taylor, R. Ford, K.D. Hyde, S. Lumyong, Morphological and molecular characterization of Colletotrichum species from herbaceous plants in Thailand, Fungal Divers. 18 (2) (2005) 117–133.

[30] P.S. Thenkabail, R.B. Smith, E. De Pauw, Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization, Photogramm Eng. Remote Sens. 68 (6) (2002) 607–622.

[31] D. Stroppiana, M. Boschetti, P.A. Brivio, S. Bocchi, Plant nitrogen concentration in paddy rice from field canopy hyperspectral radiometry, Field Crops Res. 111 (1–2) (2009) 119–129.

[32] G. Ranjitha, M.R. Srinivasan, A. Rajesh, Detection and estimation of damage caused by thrips Thrips tabaci (Lind) of cotton using hyperspectral radiometer, Agrotechnology 3 (1) (2014) 123.

[33] M. Zhang, Z. Qin, X. Liu, S.L. Ustin, Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing, Int. J. Appl. Earth Obs. Geoinf. 4 (4) (2003) 295–310.

[34] K. Gröll, S. Graeff, W. Claupein, Use of vegetation indices to detect plant diseases. Agrarinformatik im Spannungsfeld zwischen Regionalisierung und globalen Wertschöpfungsketten–Referate der 27, GIL Jahrestagung, 2007.

[35] R. Devadas, D.W. Lamb, S. Simpfendorfer, D. Backhouse, Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves, Precis. Agric. 10 (2009) 459–470.

[36] A.S.A. Salgadoe, A.J. Robson, D.W. Lamb, E.K. Dann, C. Searle, Quantifying the severity of *Phytophthora* root rot disease in avocado trees using image analysis, Remote Sens (Basel) 10 (2) (2018) 226, https://doi.org/10.3390/rs10020226.

[37] M.M. Saberioon, M.S.M. Amin, W. Aimrun, A. Gholizadeh, A.A.R. Anuar, Assessment of colour indices derived from conventional digital camera for determining nitrogen status in rice plants, J. Food Agric. Environ. 11 (2013) 655–662.

[38] P.J. Mohan, S.D. Gupta, Intelligent image analysis for retrieval of leaf chlorophyll content of rice from digital images of smartphone under natural light, Photosynthetica 57 (2) (2019) 388–398, https://doi.org/10.32615/ps.2019.046.

[39] M. Riccardi, G. Mele, C. Pulvento, A. Lavini, R. D'andria, S.E Jacobsen, Non-destructive evaluation of chlorophyll content in quinoa and amaranth leaves by simple and multiple regression analysis of RGB image components, Photosyn. Res. 120 (3) (2014) 263–272, https://doi.org/10.1007/s11120-014-9970-2.

[40] H. Hu, H.Q. Liu, H. Zhang, J.H. Zhu, X.G. Yao, X.B. Zhang, K.F. Zheng, Assessment of chlorophyll content based on image colour analysis, comparison with SPAD-502. 2nd International Conference On Information Engineering and Computer Science, IEEE, 2010, pp. 1–3.

[41] U. Barman, R.D. Choudhury, Smartphone image based digital chlorophyll meter to estimate the value of citrus leaves chlorophyll using Linear Regression, LMBP-ANN and SCGBP-ANN, J. King Saud Univ. – Comput. Informat. Sci. 34 (6) (2022) 2938–2950, https://doi.org/10.1016/j.jksuci.2020.01.005.

[42] S.M. Javidan, A. Banakar, K.A. Vakilian, Y. Ampatzidis, Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning, Smart Agricult. Technol. 3 (2023) 100081, https://doi.org/10.1016/j.atech.2022.100081.

[43] H. Yu, J. Liu, C. Chen, A.A. Heidari, Q. Zhang, H. Chen, H. Turabieh, Corn leaf diseases diagnosis based on K-means clustering and deep learning, IEEE Access 9 (2021) 143824–143835, h.

[44] R. Anand, S. Veni, J. Aravinth, An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method, in: 2016 international conference on recent trends in information technology (ICRTIT), IEEE, 2016, pp. 1–6.

[45] K.P. Swain, S.R. Nayak, V. Ravi, S. Mishra, T.J. Alahmadi, P. Singh, M. Diwakar, Empowering Crop Selection with Ensemble Learning and K-means Clustering: a Modern Agricultural Perspective, Open Agricult. J. 18 (1) (2024).

[46] A.A. Joshi, B.D. Jadhav, Monitoring and controlling rice diseases using Image processing techniques, in: International conference on computing, analytics and security trends (CAST), IEEE, 2016, pp. 471–476.

[47] S. Mutalib, M.H. Abdullah, S. Abdul-Rahman, Z.A. Aziz, A brief study on paddy applications with image processing and proposed architecture, in: 2016 IEEE Conference on Systems, Process and Control (ICSPC), IEEE, 2016, pp. 124–129.

[48] N. Guchait, I. Bhakta, S. Phadikar, K. Majumder, Visual computing for blast and brown spot disease detection in rice leaves, in: Proceedings of the 2nd International Conference on Communication, Devices and Computing, ICCDC, Springer, Singapore, 2020, pp. 595–606.

[49] R. Chakraborty, R. Sushil, M.L. Garg, ICQPSO-based multilevel thresholding scheme applied on colour image segmentation, IET Signal Proc. 13 (3) (2019) 387–395, https://doi.org/10.1049/iet-spr.2018.5073.

[50] H.J. Trussell, M.J. Vrhel, E. Saber, Colour image processing, IEEE Signal Process. Mag. 22 (2015) 14–22.

[51] Y. Wang, D. Wang, P. Shi, K. Omasa, Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light, Plant Methods 10 (1) (2014) 36, https://doi.org/10.1186/1746-4811-10-36.

[52] F. Vesali, M. Omid, H. Mobli, A. Kaleita, Feasibility of using smart phones to estimate chlorophyll content in corn plants, Photosynthetica 55 (4) (2017) 603–610, https://doi.org/10.1007/s11099-016-0677-9.

[53] Y.F. Xu, X.M. Wang, H. Sun, H.H. Wang, Y. Zhan, Study of monitoring maize leaf nutrition based on image processing and spectral analysis, World Autom. Congr. (2010) 465–468.

[54] W. Mao, Y. Wang, Y. Wang, Real time detection of between row weeds using machine vision. Paper No. 031004. ASAE Annual Meeting, 2003. St Joseph, MI, USA.

[55] S Kawashima, M Nakatani, An algorithm for estimating chlorophyll content in leaves using a video camera, Ann. Bot. 81 (1998) 49–54.
[56] M.D. Richardson, D.E. Karcher, Quantifying turfgrass colour using digital image analysis, Crop Sci. 43 (2003) 943–951.
[57] P.I. Ritharson, K. Raimond, X.A. Mary, J.E. Robert, J. Andrew, DeepRice: a deep learning and deep feature based classification of Rice leaf disease subtypes, Artificial Intellig. Agricult. 11 (2024) 34–49, https://doi.org/10.1016/j.aiia.2023.11.001.
[58] Y. Li, D. Chen, C.N. Walker, J.F. Angus, Estimating the nitrogen status of crops using a digital camera, Field Crops. Res. 118 (3) (2010) 221–227, https://doi.org/10.1016/j.fcr.2010.05.011.
[59] T.S. Sazzad, A. Anwar, M. Hasan, M.I. Hossain, An image processing framework to identify rice blast. 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE, 2020, pp. 1–5.
[60] M. Xiao, Y. Ma, Z. Feng, Z. Deng, S. Hou, L. Shu, Z. Lu, Rice blast recognition based on principal component analysis and neural network, Comput. Electron. Agric. 154 (2018) 482–490, https://doi.org/10.1016/j.compag.2018.08.028.
[61] K.R. Gavhale, U. Gawande, K.O. Hajari, Unhealthy region of citrus leaf detection using image processing techniques, in: International conference for convergence for technology, IEEE, 2014, pp. 1–6, https://doi.org/10.1109/I2CT.2014.7092035.
[62] K.V. Kumar, T. Jayasankar, An identification of crop disease using image segmentation, Int. J. Pharm. Sci. Res. 10 (2019) 1054–1064.
[63] D. Naik, R. Shaikh, S. Shetti, Detection and quantification of disease in cabbage using clustering and RGB colour, Int. J. Emerg. Technol. Comput. Sci. Electr. 14 (2015) 194–199.
[64] A.H. Kulkarni, A. Patil, Applying image processing technique to detect plant diseases, Int. J. Modern Eng. Res. 2 (5) (2012) 3661–3664.
[65] Nidhis, A.D., Pardhu, C.N.V., Reddy, K.C. & Deepa, K. (2019). Cluster based paddy leaf disease detection, classification and diagnosis in crop health monitoring unit. In: Peter, J., Fernandes, S., Eduardo Thomaz, C. & Viriri, S. (Ed). *Computer Aided Intervention and Diagnostics in Clinical and Medical Images. Lecture Notes in Computational Vision and Biomechanics* (p. 281–291). Springer. doi:10.1007/978-3-030-04061-1_29.
[66] S. Shrivastava, D.S. Hooda, Automatic brown spot and frog eye detection from the image captured in the field, Am. J. Intellig. Systems 4 (2014) 131–134.
[67] V.K. Shrivastava, M.K. Pradhan, Rice plant disease classification using color features: a machine learning paradigm, J. Plant Pathol. 103 (1) (2021) 17–26, https://doi.org/10.1007/s42161-020-00683-3.
[68] K. Alsabti, S. Ranka, V. Singh, An efficient k-means clustering algorithm, Electr. Eng. Comput. Sci. (1997) 43.
[69] D.T. Pham, S.S. Dimov, C.D. Nguyen, Selection of K in K-means clustering, in: Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 219, 2005, pp. 103–119, https://doi.org/10.1243/095440605x8298.
[70] J.P. Shah, H.B. Prajapati, V.K. Dabhi, A survey on detection and classification of rice plant diseases, in: International Conference on Current Trends in Advanced Computing (ICCTAC), IEEE, 2016, pp. 1–8, https://doi.org/10.1109/ICCTAC.2016.7567333.
[71] G. Anthonys, N. Wickramarachchi, An image recognition system for crop disease identification of paddy fields in Sri Lanka, in: 2009 International Conference on Industrial and Information Systems (ICIIS), IEEE, 2009, p. 430, https://doi.org/10.1109/ICIINFS.2009.5429828. -407.
[72] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu, B. Yang, Application of support vector machine for detecting rice diseases using shape and color texture features, in: 2009 international conference on engineering computation, IEEE, 2009, pp. 79–83.