Image classifiers and image deep learning classifiers evolved in detection of *Oryza sativa* diseases: survey

N. V. Raja Reddy Goluguri[1,2] · **K. Suganya Devi[1]** · **Nagesh Vadaparthi[2]**

Abstract

Growth in consumption of *Oryza sativa* (rice) has led the farmers across Asian countries to cultivate *Oryza sativa*, with an impact of 2.5 percent increase in the cultivation of the crop every year. Along with the growth in *Oryza sativa* cultivation, there are new challenges that are faced by the farmers in terms of diseases. The absence of information to recognize what sort of infection the plant is influenced with during the harvest cycle drives the farmers over the globe to lose 37 percent of the production. Involving technology to identify these diseases during the harvest cycle will help the farmers to get benefitted by attaining better yields. Deep learning being a latest technology playing a vital role in helping human in many aspects. A thorough review of the research papers on the various classifiers used in the identification of *Oryza sativa* diseases was carried out and the survey was tabulated and presented.

Keywords Deep learning · *Oryza sativa* · Rice diseases · Support vector machine · Multiclass SVM · Convolutional neural network · Back propagation neural network

1 Introduction

Oryza sativa or *Oryza glaberrima* is a grass species from which rice is produced. *Oryza sativa* is cultivated in all Asian countries whereas *Oryza glaberrima* is confined to be cultivated only in West Africa (Sano et al. 1984).

In Asia, *Oryza sativa* (Rice) is regarded to be the first cultivated crop. According to some archaeological evidence, it was believed that on the banks of the Yangtze River

✉ N. V. Raja Reddy Goluguri
rajareddygnv@mvgrce.edu.in

K. Suganya Devi
suganyanits@gmail.com

Nagesh Vadaparthi
itsnageshv@gmail.com

[1]   Computer Science and Engineering, National Institute of Technology-Silchar, Silchar 788010, Assam, India

[2]   Information Technology, MVGR College of Engineering, Chintalavalasa, Vizianagaram, Andhra Pradesh 535005, India

valley in China rice was first cultivated (Duncan et al. 2008) from where it has spread its roots across India, Sri Lanka and slowly to the rest of the world.

Oryza sativa has 6–7% of protein, 2–2.5% fat, B group of vitamins, Vitamin D, iron and calcium (Kennedy et al. 2002). Being having many nutritional benefits, consumption of *Oryza sativa* has increased which led the farmers across the world to take up its cultivation. If the world indices of farming are considered, the highest cultivated crop is *Oryza sativa* after the maize and sugarcane cultivation. *Oryza sativa* stands first if considered in terms of the human consumable crop, as 85% of the produced crop is consumed by humans (https:// archive.gramene.org/species/oryza/rice_intro.html) and the rest is used for cattle feed etc.
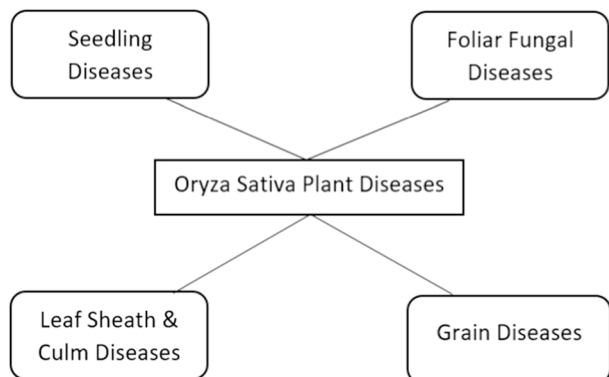
According to the United States Department of Agriculture (USDA), the world rice production (http://www.fao.org/statistics/en/) 2019–2020 evaluated it to be 497.82 million metric tons. To more readily express the significant-top 10 positioning Countries producing rice on the planet is placed in an order then the arrangement would resemble CIIVTBBPBJ which implies China, India, Indonesia, Vietnam, Thailand, Bangladesh, Burma, Philippines, Brazil, Japan. About the portion of the total populace depends for life on rice as it is staple nourishment. The nourishment security of the world relies upon this brilliant cereal. In India, 50-60% of the populace relies upon rice.

According to the data published by the United Nations Food and Agriculture Organisation (http://www.fao.org/faostat/en/#home), the total area under *Oryza sativa* cultivation is 160.8 mH. In India, the total area under *Oryza sativa* cultivation is 43.39 mH, which accounts to be 27% approximately when compared with the world statistics. Though India is holding 27% of land for *Oryza sativa* cultivation, it is unable to hold records when it comes to production. There is a huge difference i.e., other countries are producing nearly 4645 kg/ha whereas in India the farmers are only producing 2404 kg/ha. There are enormous reasons why farmers in India are unable to compete in terms of production with the other *Oryza sativa* cultivating countries. Out of those reasons the majority of the challenging issues that farmers are facing in India are with the *Oryza sativa* plant diseases.

International Rice Research Institute (IRRI) (https://www.irri.org/our-solutions/irri-education) has categorized the *Oryza sativa* fungal pathogens into four types which can occur during the plant life cycle. Those four types are Seeding diseases, Foliar fungal diseases, Leaf sheath and culm diseases, and Grain diseases. Shown in Fig. 1: Taxonomy of *Oryza sativa* Plant Diseases.

These diseases are further classified into multiple categories (http://www.knowledgeb ank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases) based on their

Fig. 1 Taxonomy of *Oryza sativa* plant diseases

type and severity of affecting the plant life cycle and the same has been listed in the Table 1 as given below.

To overcome the effects of these diseases the farmers have to take up the next generation farming i.e. Precision agriculture along with applied nanoscience (Shakeel et al. 2019a, b). Precision agriculture is a management concept in which farmers apply inputs concerning soil, crop and weather conditions, taking up these inputs it will help the farmers to identify and eradicate the plant diseases in time which would help farmers in producing a better yield. Precision agriculture includes innovation to distinguish these illnesses, it implies that the innovation ought to be information enough for recommending the farmers to make up any fundamental move when a malady is recognized. So on this inclination, there is a lot of research going on Artificial Intelligence and Image Processing.

We were paying special mind to a continuous challenging issue that artificial intelligence can be utilized to discover the solution. So everything began with the undertaking to pay special mind to something truly fascinating and furthermore it ought to be useful to the general public. From now on the thought started and we were talking about a ton of numerous themes and arrived at the resolution on *Oryza sativa* diseases (Rice Plant Diseases).

Collection of research papers In any case, to address an issue one ought to be clear about the issue and the piece of arrangement that is as of now existing. Thus we began investigating all the research articles that have been as of now distributed. Be that as it may, it is again a central issue characteristic of the timespan on the grounds that artificial intelligence is an immense region and has roots from 1970s. We choose to consider articles that are not over 10 years old and have included image processing and machine learning embodiment in them. We have gathered articles which have the catchphrases like Rice diseases or *Oryza sativa* diseases alongside watchwords like Multilayer perceptron, Artificial intelligence, Deep convolutional Neural network, Backpropagation Neural Network and so on., and the article ought to be distributed in Elsevier, Springer, Nature, and IEEE. We even need to consider a few articles distributed in IEEE conferences and other international conferences because of nature of work in it.

Table 1 Taxonomy of various *Oryza sativa* fungal pathogens

Seedling diseases	Foliar fungal diseases	Leaf sheath and culm diseases	Grain diseases
Bakanae (Bak)	Brown Spot (BSp)	Stem Rot (SR)	Kernel Smut (KSm)
Feeder Root Necrosis and Root Rot (FrT)	Downy Mildew (DM)	Sheath Blight (ShB)	Udbatta Disease (UbD)
Seedling Blight (SdB)	Narrow Brown Leaf Spot (NBLs)	Aggregate Sheath Spot (AsS)	Glume Blight (GlB)
	White Leaf Streak (WLS)	Sheath Spot (ShS)	Grain DiscolorationI (GrD
	Stackburn Disease (StD)	Sheath Net-Blotch (SnB)	Scab (ScB)
	Leaf Scald (LSc)	Sheath Rot (ShR)	Ear Blight (EB)
	Collar Rot (CR)	Crown Sheath Rot (CsR)	False Smut (FSm)
	Eyespot (ES)	Sheath Blotch (ShBl)	Bakanae (Bak)
	Red Stripe (RS)	Fusarium Sheath Rot (FuR)	
	Leaf Smut(LSm)		
	Rusts (Ru)		
	Rice Blast (RBl)		

Significance and contribution of this survey There is a great deal of research in identifying and classifying plant infections with the assistance of computational intelligence (Ning et al. 2017) however this article is limited to *Oryza sativa* diseases. The primary importance of this article is that we have overviewed 20 articles that are identified with *Oryza sativa* disease identification and classification just over the time of the most recent 10 years. Perusing this article will give an obvious thought of the considerable number of works did in *Oryza sativa* disease identification and classification using computational intelligence.

This article has eight sections including the introduction. Section 2 presents the taxonomy of *Oryza sativa* diseases. In Sect. 3, we have tended to the most generally utilized methodology in *Oryza sativa* disease identification and classification. To put it plainly, we have given a rundown of the 20 articles reviewed in Sect. 4 which have been coursed somewhere in the range of 2008 and 2019. Section 5 presents an outline table and in Sect. 6 we have addressed different challenges. In Sect. 7, we have addressed the future scope in the identification and classification of *Oryza sativa* diseases. We have concluded this article in Sect. 8 (Table 2).

Table 2 List of abbreviations

Name	Abbreviation
RGB	Red, Green, Blue
HSV	Hue, Saturation, Value
HIS	Hue, Intensity, Saturation
RGL	Red, Green, Light
SOM	Self-Organizing Map
MLP	Multilayer Perceptron
PCA	Principal Component Analysis
SVM	Support Vector Machine
GLCM	Grey Level Co-occurrence Matrix
CNN	Convolutional Neural Network
ANN	Artificial Neural Network
BPNN	Back Propagation Neural Network
KNN	K- Nearest neighbours
PSO	Particle Swarm Optimization
DWT	Discrete Wavelet Transform
RBF	Radial Basis Function
HE	Histogram Equalization
BoW	Bag of Words
HaarWT	Haar Wavelet Transform
LBPH	Local Binary Pattern Histogram
LBP	Local Binary Pattern
SIFT	Scale-invariant feature transform

2 Taxonomy of *Oryza sativa* diseases

This section has a concise note on different *Oryza sativa* fungal pathogens, shown in Fig. 2. The fundamental saying is to let the readers comprehend the sorts of fungal pathogens and their effect on the *Oryza sativa* harvest cycle.

Only through a better knowledge of these fungal pathogens can the readers realize the role that image processing has been doing to assist the farmers in classifying and identifying *Oryza sativa* fungal pathogens during the harvest period.
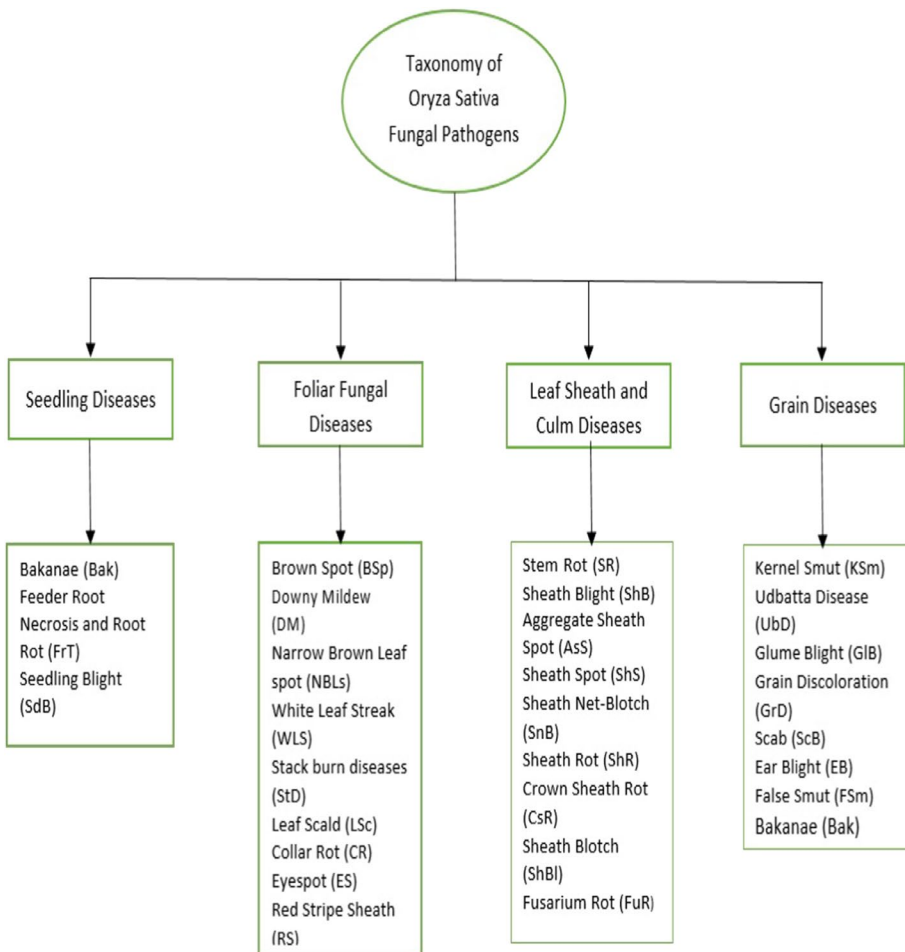
Fig. 2 Taxonomy of *Oryza sativa* fungal pathogens

Fig. 3 Rice blast affected leaves

Fig. 4 Rice blast affected collars

2.1 Rice blast

Rice blast in *Oryza sativa* is caused due to the fungus "*Magnaporthe Oryzae*" (ZIegler et al. (1994)). A Rice blast has different names geographically, such as Rice blast fungus, Rice rotten neck, Rice seedling blight, and Blast of rice, etc., (Adam et al. 2013).

It is treated to be the most highly recorded fungal pathogen which causes significant damage to rice production. This disease is feared by the rice farmers because of its nature that it can cause lesions at roots, stems, panicles, leaves, and even seeds.

The most common symptom of this disease is that it creates diamond-shaped lesions on the affected areas of *Oryza sativa* plant. Rice Blast affected leaves, collars, seeds and neck & panicles are shown in Figs. 3, 4, 5, and 6 respectively.

Fig. 5 Rice blast affected seeds

Fig. 6 RB affected neck and Panicle's

2.2 Brown spot

Brown Spot in *Oryza sativa* is caused by the *Cochliobolus miyabeans* fungus that used to be called *Helminthosporium Oryzae* (http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases). According to reports, this disease had a lot of impact on the" Bengal Famine of 1943" during World War-II where almost 2 million Indians died due to starvation. Nations fear Brown Spot as it can be used as a biological anti crop weapon which helps to attack staple crops of specific places.

Brown spot is categorized by a wide range of symptoms. A most visible sign of the brown spot is the presence of circular or oval brown spots with a yellow halo during the stage.

Fig. 7 Brown spot affected leaf

Fig. 8 Brown spot affected seeds

These symptoms may be found at leaves, stems, and seeds. Shown in Figs. 7, 8, and 9 respectively. Stem discoloration can also lead to a brown spot.

2.3 Sheath blight

Sheath blight in *Oryza sativa* is caused by the fungus *Rhizoctonia Solanihe*, fungus belongs to the phylum *Basidiomycota, Ceratobasidiaceae* family (http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases). This disease has been discovered more than 100 years ago. This disease causes significant loss in quality and grain yield.

Fig. 9 Brown spot affected stems

The sheath blight is a soil-borne fungal pathogen. Sheath blight attacks the plant at a very early stage of development like at seeds and seedlings stage. Initially, grey-green lesions develop around the collar region on both the leaf and its sheath. Subsequently, the lesions turn brown and the flag leaf dries up. Sheath Blight affected to collars, panicles, roots, and stem are shown in Figs. 10, 11, 12, and 13.

2.4 Leaf scald

Leaf Scald in *Oryza sativa* is caused by the fungus *Microdochium Oryzae* (http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases). It creates a scalded appearance on leaves of *Oryza sativa* plant. The disease occurs on mature leaves. This fungus creates zonate lesions on leaf tip or edges of alternating light tan and dark brown spots. The main sources of Leaf scald are seeds and crop stubbles. The leaves of an infected plant split near the midrib especially when there is a strong wind. Leaf Scald affected leaf tip is shown in Fig. 14.

Fig. 10 Sheath blight affected collar

Fig. 11 Sheath blight affected
Panicles

Fig. 12 Sheath blight affected
roots

Fig. 13 Sheath blight affected
stem

Fig. 14 Leaf scald affected leaf tip

2.5 Leaf smut

Leaf smut in *Oryza sativa* is caused by the fungus *Enthloma Oryzae* (http://www.knowl edgebank.irri.org/training/fact-sheets/pest-management/diseases). Due to the airborne spores on diseased leaf debris in the soil, it is spreads to the plant.

In the growing season this develops late and causes minimal damage. This fungus produces either sides of the leaves angular black spots. The leaves of *Oryza sativa* turn yellow when the leaf smut is heavily infected and the tips of the leaf die and turn grey. Leaf smut affected leaf tip is shown in Figs. 15 and 16.

3 Methodology evolved in *Oryza sativa* plant disease identification & classification

In this section, the general methodology that evolved in identifying the *Oryza sativa* plant diseases will be explained. The methodology carries a six-stage procedure. Shown in Fig. 17.
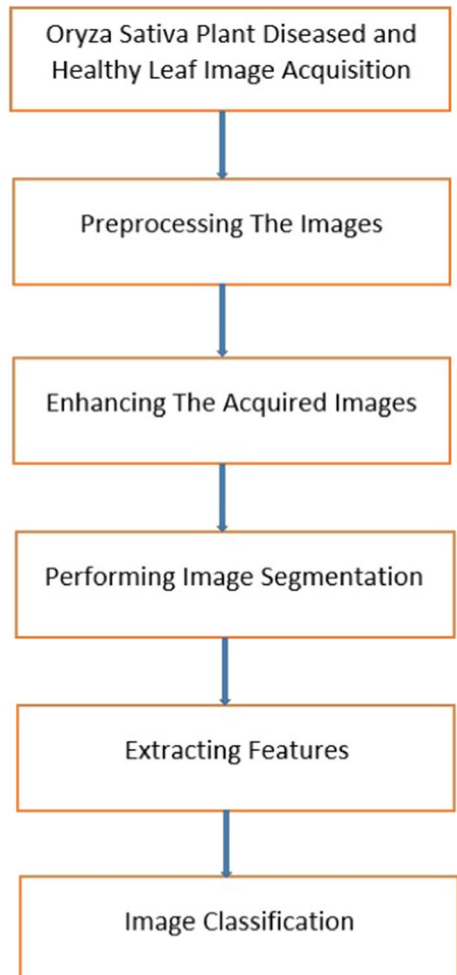
Fig. 15 Leaf smut affected leaf

Fig. 16 Leaf smut dead leaf
turned grey

Fig. 17 Methodology in identify-
ing *Oryza sativa* plant diseases

Oryza Sativa Plant Diseased and
Healthy Leaf Image Acquisition

Preprocessing The Images

Enhancing The Acquired Images

Performing Image Segmentation

Extracting Features

Image Classification

- Image Acquisition.
- Image Preprocessing.
- Image Enhancement.
- Image Segmentation.
- Feature Extraction.
- Image Classification.

3.1 Image acquisition

Image acquisition is collecting images related to an area of interest from some external source for performing any kind of image processing. The acquisition of images of the *Oryza sativa* plant diseases is the major tedious task. International Rice Research Institute (IRRI) provides *Oryza sativa* disease pictures (http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases). To work on any particular disease, we have got to prepare our own image archive. It is always better to visit farms and take pictures of the diseased plants using any digital cameras.

The acquisition of image datasets is a significant challenge for forward-looking research on plant diseases of *Oryza sativa*. As datasets are not accessible and one needs a lot of time to procure and make a dataset. The explanation is that, for example, in the event that we need to compile a dataset of Rice Blast, we can only compile it in the fields affected by rice blasts and, as a rule, rice blasts occur during the summer season. It is not that easy to build a dataset for conducting some research, as there are multiple factors such as water, soil, climate, etc. that play a major role.

We have discovered a portion of the online accessible datasets that can be used at first to examine and to achieve some great results, we can make our own image dataset and check the research work done on the dataset. Some of the currently available datasets are (Huy 2010; Shah 2019).

In Table 3, we tabled the types of *Oryza sativa* diseases, the place of image acquisition and different methods of assessment that the researchers adopted by evaluating 20 papers that are confined exclusively to *Oryza sativa* diseases.

3.2 Image preprocessing

Image without any distortion gives better information. It is always necessary for an image to be clear without any noise or distortion to give better results while processing it. Here is where the image preprocessing plays a vital role in improving the image information by eliminating unnecessary defects and by improving image characteristics. While capturing the images of *Oryza sativa* plant diseases, there can be distortion like water dew dust, drops, insects, and effects of shadow. This kind of distortion can be removed using noise filters and by using some better enhancement algorithms (shown in Fig. 18). Sometimes the captured image file size might be very large, in such cases file size should be reduced to improve computing time complexity.

It is hard to eliminate noise from images without the requisite knowledge on filtering techniques. The filters can indeed be chosen by analyzing the behavior of the noise. Filtering image details is a common technique used in almost any image pre-processing application. They eliminate noise from the images by maintaining the finer points of the equivalent. Given the fact that there are various types of filters, this article will consider a filter that is one of the most commonly used filters for pre-processing images.

Table 3 A Table presenting place of image acquisition and respective evaluation patterns

Reference article	Diseases	Image acquisition	Image evaluation patterns
Phadikar and Sil (2008)	Rice Blast and Brown spot	East Midnapur and South Bengal, India	Contrast and Brightness enhanced, RGB Color space to HSI color space
Sanyal and Patel (2008)	Rice leaf blast and Brown spot	Not mentioned	RGB
Libo and Guomin (2009)	Rice Brown spot	Northern part of Ningxia Hui, China	RGB Color space to Lab color space
>Yao et al. (2009)	Rice leaf blast, Rice sheath Blight, and Rice Bacterial Leaf Blight	China National Rice Research Institute-Fuyang, China	Median filter
Orillo et al. (2014)	Brown Spot, Rice blast, and Bacterial leaf blight	International Rice research institute, Los banos, Laguna, Philippines	RGB Color space to HSV color space, Histogram equalization
Lu et al. (2017)	Brown Spot, Rice blast, false smut, bakanae, sheath blight, sheath rot, bacterial sheath, Seedling Blight, Bacterial wilt and Bacterial leaf blight	Heilongjiang Academy of land reclamation sciences, China	ZCA whitening method
Prajapati et al. (2017)	Leaf smut, Brown Spot and Bacterial Leaf Blight	Shertha, Gandhi nagar, Gujarat, India	RGB color space to HSV color space, Background removal
Bhagyashri and Gajanan (2017)	Brown Spot, and Leaf Blast	Rice Research center Karjat, Maharastra, India	RGB Color space to Lab color space
Narmadha and Arulvadivu (2017)	Narrow brown spot, Brown Spot, and Blast	Internet	RGB Color space to gray scale color space
Archana and Arun (2018)	Brown Spot and Bacterial Leaf Blight	Not mentioned	RGB Color space to HSV color space
Gayathridevi and Neelamegam (2018)	Leaf streak, Leaf blast, False smut, Brown Spot and Bacterial Leaf Blight	Tamil Nadu Rice research Institute, Aduthurai, Thanjavur, India	Median Filter
Maohua et al. (2018)	Rice blast lesions	Nanjing Agricultural University	RGB Color space to HSV and YCbCr color space
Takuya and Yutaka (2018)	Rice blast	Not mentioned	RGB
Ramesh and Vydeki (2018)	Rice blast	Panpoli, Tirunelveli, India	RGB Color space to HSV color space
Taohodul et al. (2018)	Rice blast, Rice bacterial blight, and Rice brown spot	International Rice Research Institute	RGB
Abu Bakar et al. (2018)	Rice leaf blast	Malaysian Agricultural Research and Development Institute, Malaysia	Histogram Equalization, RGB to HSV

Table 3 (continued)

| Reference article	Diseases	Image acquisition	Image evaluation patterns
Komal et al. (2019)	Bacterial Leaf Blight, False Smut, and Brown spot	Rice Research Institute, Lahore, Pakistan & Punjab, India	RGB Color space to gray scale color space
Wan-Jie et al. (2019)	Rice Blast	Institute of Plant protection, Jiangsu, China	RGB
Chowdhury et al. (2019)	Brown spot, Neck blast, Bacterial leaf Blight, Sheath Rot, False Smut, Sheath blight, Brown Plant hopper, Stem borer, and Hispa	Bangladesh Rice Research Institute, Bangladesh	RGB
Neha et al. (2019)	Rice blast, Sheath Blight, and Brown Spot	Not mentioned	Not mentioned
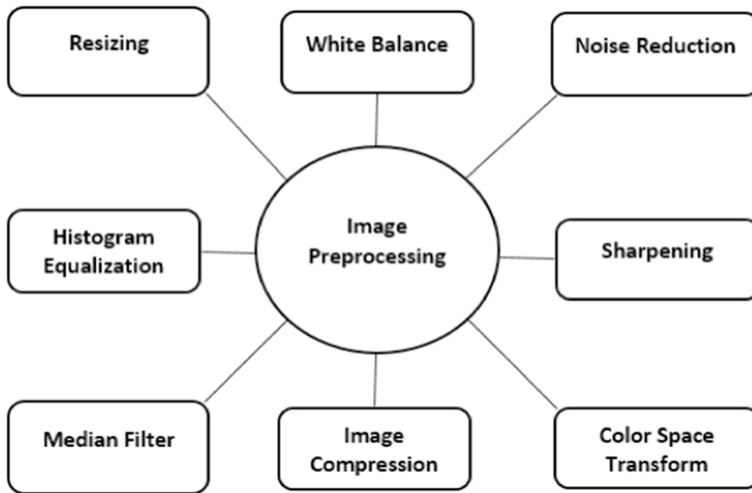
Fig. 18 Various preprocessing techniques

3.2.1 Median filter

Median filter is a nonlinear filter. It is best suited for removing the most common noise i.e. the salt and pepper noise. Median filter working process is very simple, the median is determined by just sorting all of the pixel values in sequential order from the window, and afterwards replacing the pixel found with just the middle (median) pixel.

3.2.2 Histogram equalization

To improve the contrast of an image by spreading the intensity values over the full range is the role of Histogram equalization (Dorothy et al. 2015). If the image suffers from non-uniform illumination, histogram equalization cannot be applied. Applying histogram equalization on an image suffering from non-uniform illumination would only remove extra pixels from the dark regions and add extra pixels to the light region of the image. The result would be high dynamic range in the output image.

The main idea behind histogram equalization is to manipulate the probability density function (pdf) (Pankaj Kumar et al. 2014). In simple words, it changes the pdf of a given image into that of the uniform probability density function that spreads out from the lowest pixel value (0 in this case) to the highest pixel value (L − 1). This can only be achieved if the pdf is a continuous function.

Probability density function of a digital image will be a discrete function. Let us consider an image x, and let the dynamic range for the intensity r_k varies from 0 (black) to L1 (white). The pdf can be calculated by using the probability based on the histogram $p(r_k)$ as follow:

$$pdf(x) = p(r_k) = \frac{total\ pixels\ with\ intensity r_k}{total\ pixels\ I\ image\ x} \tag{1}$$

Cumulative density function (cdf) can be obtained as follow:

$$cdf(x) = \sum_{k=0}^{L-1} p(r_k) \tag{2}$$

where $p(r_k)$ is the probability for pixel of intensity. By applying the histogram equalization, the output of a pixel would be equal to the cdf of the image. Shown in Fig. 19.

3.2.3 Resizing

Image resizing (Xia et al. 2014) plays a vital role at whatever point the general number of pixels in an image should be increased or decreased. At the point when an individual need to see more detail in a picture, they have to expand the number of pixels which is additionally alluded to as zooming. The procedure of resizing the image starting with one-pixel network then onto the next is called Image Interpolation.

3.2.4 White balancing

The method of expelling artificial color casts, with the result that objects that appear white face to face are made white in your image, is called white balancing.

The image taken under common light conditions would have issues identified with temperature, shadow, dust, and so forth. For instance, an item that is white in shading will seem reddish at a low shading temperature and will seem blueish in shading at a high shading temperature. Human eyes are superb at perceiving white color under various light sources, yet with regards to advanced cameras, there is still some disarray in deciding white color under various light sources.

There are some advanced features in the present day digital cameras like "Auto white balancing" (Sujung et al. 2008). Digital cameras with auto white balancing(AWB) features can naturally make up for the color shifts because of the distinctive shading temperatures of the light sources by changing the RGB of the comparing channel utilizing the data of the image captured.
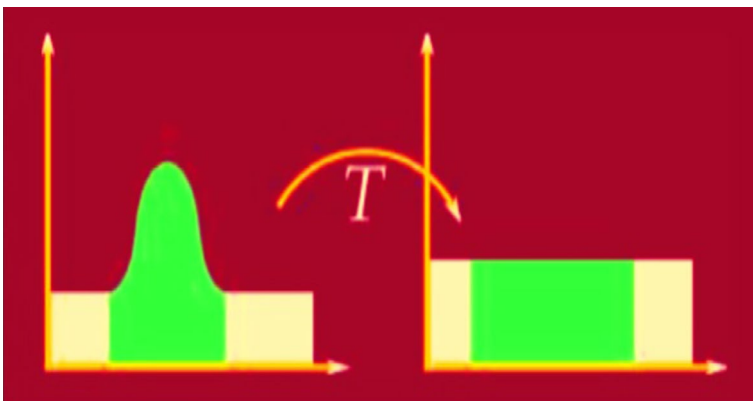

Fig. 19 Histogram equalization of an image

3.2.5 Sharpening

Image sharpening (Konstantinides et al. 1999) increases the contrast between the light and the dark regions to bring out the details. Acutance and resolution are the key segments of sharpness. Acutance can be represented as a subjective indicator of contrast at the edge. Human eyes can learn even more from images whose edges are more contrasting. Resolution characterizes the size of an image in pixels. In the event that the image is of higher resolution that implies it has an increasing number of pixels. At the point when an image has more pixels, it implies that the picture will be sharper.

3.2.6 Color space transform

A color space is a means of defining, generating and visualizing light. As humans, by its properties of brightness, hue, and texture, we can describe a color. A computer can define a color using the quantities of red, green and blue required to match a color. Therefore, a color is typically defined using three coordinates, or parameters. These parameters define the color position inside the used color space.

There are many color space models like RGB color space, The CMY and CMYK color space, YIQ color space, HSV color space, and YCbCr color space. Each color space model has its own importance. YCbCr color space model (Kamlesh et al. 2015) provides better performance as compared to other color space models available.

3.2.7 Image compression

The compression of images is defined as the process of decreasing the amount of data required to display a digital image. It is accomplished by deleting unnecessary data. The goal of image compression is to reduce the number of bits needed for storage and transmission without any noticeable loss of information.

Image compression (Jiang 1999) is a must at the pre-processing stage as all images of *Oryza sativa* are captured in live fields with the help of digital cameras and the output JPEG file would be very large in size. So the images are downsampled to small size. The reason behind the downsampling of the original image is that it should be forwarded to the neural network and large image files take longer.

Image compression can be differentiated into two categories (Mehwish et al. 2014), i.e. on the basis of data interpretation or on the basis of no data interpretation. Image compression based on data interpretation requires segmentation of the image only when high compression ratios can be achieved. But here the concern of image compression is to reduce the file size by removing unnecessary data to help reduce processing time.

Image compression without interpretation of data, i.e. removal of redundant and irrelevant information. This type of compression is further classified into two categories, i.e. lossless compression or lossy compression. lossy compression is widely used in the *Oryza sativa* preprocessing phase.

3.3 Image enhancement

In *Oryza sativa* farms, while capturing the images there will be many factors that affect the quality of the image. Image enhancement aims to improve a bad quality image to
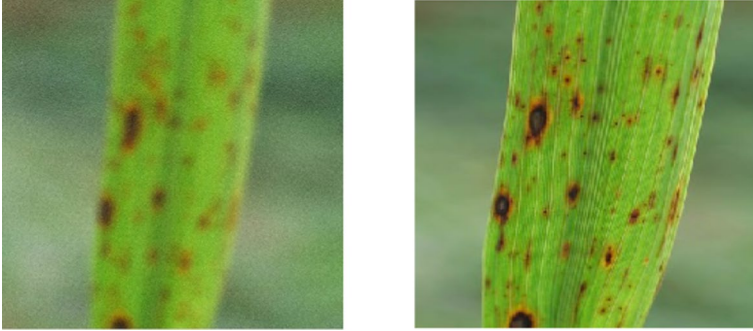
Fig. 20 Blurred and enhanced image of *Oryza sativa* infected by brown spot

a good quality image (shown in Fig. 20). The factors that demand image are contrast, brightness, spatial resolution, and noise. Human perception demands to improve the information in the images to be better perceive the intended meaning.

3.4 Image segmentation

Image segmentation separates a digital image into several sections inorder to gather some meaningful information for some research (shown in Fig. 21). The main objective is to segment the pixels of a given image into two classes where one class would be the *Oryza sativa* plant diseased area and the other class would be soil and other background factors.

Segmentation of images is a crucial role in distinguishing the area of interest i.e. diseased plant which is an essential factor and also the removal of non-area of interest i.e. background. If the removal of background is not done in an appropriate way, it might lead to misclassification.

There are various image segmentation techniques (shown in Fig. 22) which help in segmentation of the images of *Oryza sativa* Plant. Some of them are edge detection technique, region- based technique, and artificial neural network-based technique, etc.

These segmentation techniques are discussed in short and tabulated, refer Table 4: Summary of various Segmentation techniques Table 4.
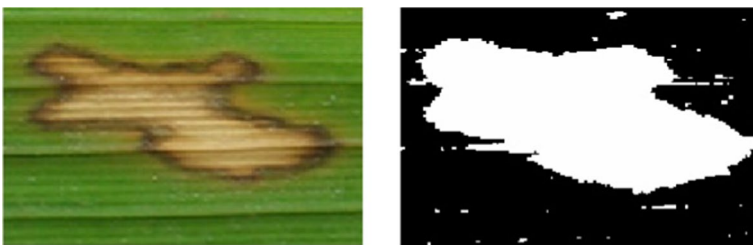

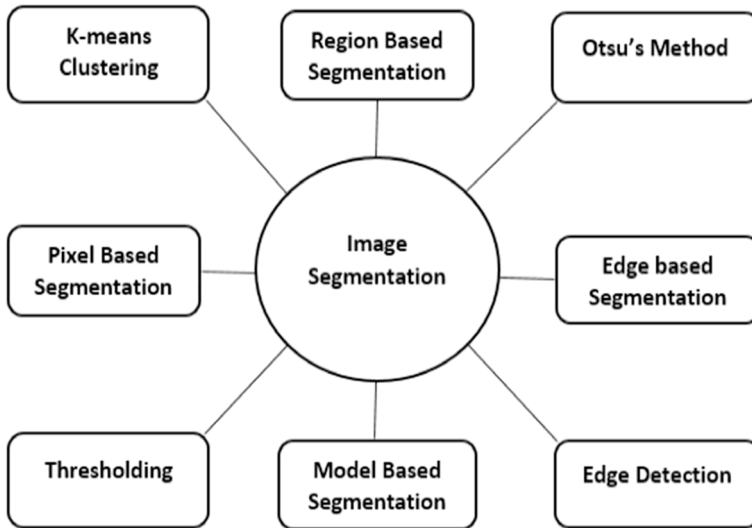**Fig. 21** Segmented image of *Oryza sativa* infected leaf enhancement

Fig. 22 Various segmentation techniques

3.5 Feature extraction

Every image will carry a unique feature or property that defines the purpose of the image captured. So to differentiate between the area of interest and the rest part of the image is where the extraction of feature plays an important role. The area of interest in an image can be boundaries of the image, and texture of the image, etc. Feature extraction does a meaningful job in gathering informative or relevant information of an image to train the machine. Some of the feature extraction techniques can be observed in the Fig. 23.

These feature extraction techniques are discussed in short and tabulated, refer Table 5.

3.6 Image classification

The classification of images is a two-stage process where initially the classification model is developed to describe the predefined classes available and then the classification algorithm is developed by learning from the data with unique category. This stage is called a stage of learning. In the second stage, classification is done with the help of the model which was generated in the first step. There are multiple classification models (Fatemeh and Mohammad 2017) (shown in Fig. 24), such as the KNN Classifier (Suganya Devi et al. 2019), rule-based classifier, support vector machine, and neural network.

These various image classifiers techniques are discussed in short and tabulated, refer Table 6.

Table 4 Summary of various segmentation techniques

| Techniques	Description	Pros	Cons
Thresholding (Capms et al.2019)	Relies upon the histogram of an Image.	A straightforward methodology where there is no any necessity of earlier information on image.	It doesn't function admirably if such a large number of edges are available or not fit for level valleys.
Edge Based Segmentation (Capms et al. 2019)	In light of discontinuity of pixels or pixels having various intensities.	Detecting discontinuities in brightness.	This technique is less invulnerable to noise and not work if the edges are not characterized flawlessly.
Region Based Segmentation (Capms et al. 2019)	Categorizing of pixels with specific properties and forming a region.	Function well when it is easy to define the region's homogeneity criteria and be more immune to noise.	This strategy comprises of two step segmentation which requires some serious time and memory.
K - Means Clustering (Naveen et al. 2016)	Gathering the pixels having comparative properties and characterizes the bunch values as indicated by their visible intensities.	It works well when large number of dataset of images are taken for segmentation.	It is Sensitive to initial outliers and centres.
Otsu's Method (Vaibhav 2015)	Automatic filtering process for region selection based on segmentation	As it operated on histograms it is very fast. It is preferred when two classes are performed in an image, such as foreground and background.	It demands bimodal histograms.
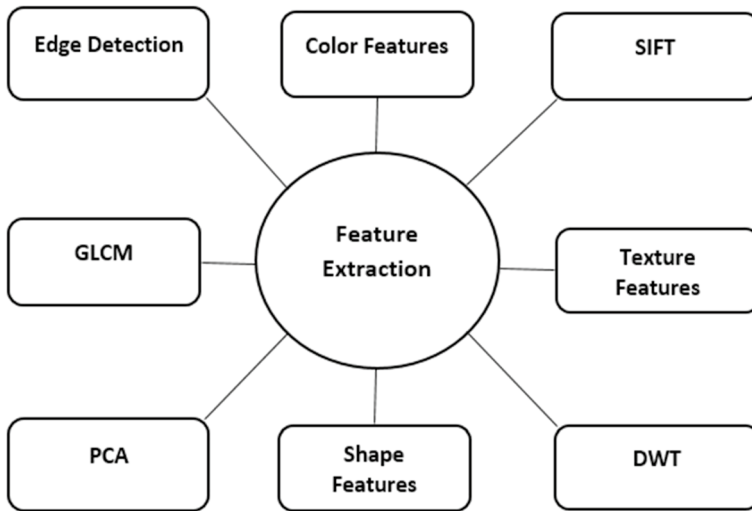
Fig. 23 Various feature extraction techniques

4 Literature survey

Throughout this section, a summary of the numerous work completed by researchers on the identification and classification of *Oryza sativa* diseases is described by segregating them according to the learning style of the algorithms.

4.1 *Oryza sativa* diseases classified by Instance Based algorithms

Phadikar and Sil (2008) have acquired images from Rice fields in East Midnapur and South Bengal, India and presented a study to identify rice diseases. In the study, the images of the diseased rice plant are captured using a digital camera and the captured images are processed by using image growing, image segmentation techniques to detect the infected parts of the plant. The images are preprocessed by increasing the brightness and contrast and then transformed to Hue Intensity Saturation (HIS) model for performing segmentation. To identify the infected part of the leaf, Entropy-based bi-level thresholding segmentation method is been used and then the boundary detection algorithm using 8 connectivity method is applied to identify the boundaries of the infected portion on the leaf. The boundaries of the segmented images vary between 300 * 300 to 1500 * 1500. Spot detection is used to identify the spots on the infected leaf. Self-Organizing Map (SOM) neural network is used for classification. By using the RGB of the spots for classification the study has achieved 92% accuracy.

Yao et al. (2009) have presented an article on detecting rice diseases with the application of support vector machine (SVM). China National Rice Research Institute Fu yang, China supported the author in acquiring the rice disease images. The study opted on a Rice leaf blast, Rice Sheath blight, and Rice bacterial leaf blight as the study object. In this study, the acquired images were preprocessed to remove noise from the images by applying the

Table 5 Summary of feature extraction techniques

| Techniques	Description	Pros	Cons
Color features (Divya et al. 2015)	This embodies the color bins' frequency distribution inside an image	It's close to human sight	Scalability, Computational time
SIFT (Wu et al. 2013)	Digital image recognition and description of local features	Rotation Invariant, Scale invariant and more accurate.	It is important to measure the gradients of each Pixel throughout the patch, and these calculations cost time
Texture features (Humeau and Anne 2019)	A methodology to collect features such as size, shape, density through texture analysis process	Salient features can be extracted	Computationally expensive
DWT (Dipalee and Siddhartha 2015)	Discrete sampling of wavelets	Flexibility	Number of Decomposition levels
Shape features (Humeau and Anne 2019)	These are generally utilized for finding and matching shapes, recognizing objects or making estimation of shapes	Reduced randomness	Dependent on CPU memory and time
PCA (Naresh 2019)	It is a statistical method used to reduce the size of the data	Improves Visualization and reduces overfitting	Loss of information, Data Standardization
GLCM (Humeau and Anne 2019)	It is an arrangement of how regularly various mixes of grey levels co-happen in an image	Works with lowest possible length of features vector	Need very high Computation
Edge Detection (Ghassan Mahmoud and Ahmed Mohamed 2015)	In light of discontinuity of pixels or pixels having various intensities	Detecting discontinuities in brightness	This technique is less invulnerable to noise and not work if the edges are not characterized flawlessly
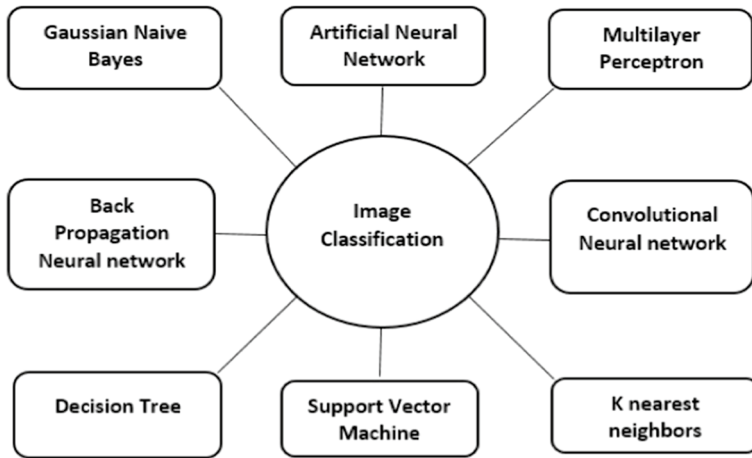
Fig. 24 Various image classifiers

median filter with a 3*3 rectangular filter window. The RGB color space was converted to a new algorithm developed by the authors. The working principle of the new algorithm was like initial it converts the RGB values into y1 and y2 representation where $y1 = 2g - r - b$ and $y2 = 2r - g - b$ and then maximum class square error (Otsu's method) was used to segment the diseased spots from the rice leaf. Shape and color features of disease spot were extracted and from Gray level co-occurrence matrix (GLCM) the texture features are adopted. Support vector machine (SVM) was used to classify the three diseases.

Takuya and Yutaka (2018) have presented a study on the implementation of a categorization system of rice disease using artificial intelligence. The study opted on Rice blast as the study object. In this study, both the healthy and infected rice leaf images were taken of 200*200 pixels. The images taken were of 40,000 Red, Green, and Blue (RGB) data. Since the data would take much learning time, the data was reduced to two dimensions by using the principal component analysis (PCA). Two-dimensional feature vectors helped for the learning and classification. To decide which learning information can be taught, a precision test has been carried out using the leave-one-out cross-validation process by enforcing a nonlinear Support vector machine (SVM) using Radial basis function (RBF) kernel. In this study by extracting the pixel values the SVM classifier attained 95% accuracy in identifying rice blast infected leaves.

Komal et al. (2019) have presented an article on the rice diseases Detection and classification. The study collected images from Rice Research Institute, Lahore, Pakistan & Some parts Punjab, India. The study opted on Bacterial Leaf Blight, False Smut, and Brown spot as the study object. In this study, the acquired images are preprocessed and converted to greyscale images using the EM GUI function. Scale-Invariant Feature Transform (Thao et al. 2013) (SIFT) was used for feature detection and extraction of the number of regions in an image. Vocabulary building for image classification was done using the Bag of Words (BoW) by converting the features identified by SIFT to words. K-means clustering was applied to create clusters and the cluster count was set to 80. For matching the features, they have applied the Brute-Force matcher so that it tries to find the closest descriptor in the set by trying each feature. Training normalized histograms were provided to train the Support vector machine (SVM) to get the maximum efficiency and accuracy. Out of the

Table 6 Summary of various image classifiers

Techniques	Description	Pros	Cons
Artificial Neural Network (Gonzalez Fernandez et al. 2019)	A model which processes information the way biological nervous system do.	Easily can handle Noisy data	Difficult to interpret, High processing time is required
Multilayer perceptron classifier (Rana et al. 2018)	It is a connected series of nodes where each node represents a model.	Would be of better model when non-linear data with large number of input features	Hard to train
Convolutional neural network (Jurgen 2014)	A model to detect complex features in data	Easy to handle noisy data	Over fitting
K nearest neighbours (Zahid et al. 2018)	A model to solve both classification and regression problems	Easy to implement on small datasets	Demands large storage space
Support vector machine (Jurgen 2014)	A model which helps to work like a discriminative classifier by separating hyperplane	Over and under fitting problems are lower	Training consumes more time
Decision tree (Jurgen 2014)	Based on feature values instances are classified by sorting	Numerical and categorical data can be handled	Numeric data set will generate complex decision tree
Back propagation neural network (Jurgen 2014)	Backward propagation of errors	Accuracy and versatility	Time consuming
Gaussian naïve Bayes (Zahid et al. 2018)	Works on Bayes theorem and gives probability of an event occurrence	Discriminative models can be converged quickly	suffer multicollinearity

three parameters of SVM (C, Gamma, and Nu), the study has performed parameter optimization on Gamma and Nu parameters with multiple values to generate more effective results. With the help of precision and recall, the system was evaluated and found that the system was 94.16% accurate in classifying the 3 rice diseases.

Neha et al. (2019) have presented a study on detection of paddy leaf disease with image processing and machine learning. The study opted on Rice blast, Sheath blight, and brown spot as the study object. In this study, the acquired images are preprocessed to enhance the quality of the image. To segment the images and convert the images into binary form, Otsu's method is applied. Image attributes are extracted using attribute analysis like shape feature extraction and color feature extraction. The research identified all three diseases with the aid of the support vector machine (SVM).

4.2 *Oryza sativa* diseases classified by artificial neural network algorithms

Sanyal and Patel (2008) have proposed a work to identify the rice life disease by opting Rice leaf blast and Brown spot as the study object. In this article, 400 images were acquired out of which 80% images have been used for training whereas the rest 20% was used for testing. 7*7 neighborhood was considered for each pixel of an image to compute the feature vector. 70 hidden nodes of color features and 40 hidden nodes texture features were given to the Multilayer perceptron classifier and the classification results correspond to both color and texture feature MLP's and quantitative analysis of the classification resulted in 89.26% accuracy.

Libo and Guomin (2009) have acquired rice leaf diseases from Northern part of Ningxia Hui, China. They have opted the Brown Spot disease as the study object. In this paper, the collected images are first transformed from RGB color space to XYZ reference color space and then XYZ color space was transformed to Lab color space. From the image of R, G, B component and Lab component, the R, G, L component is identified to have the larger difference between the lesion and background and even the L component suppressed the noise much effectively. R, G, L values are given as the parameter to input layer neurons and one output layer. With the method" Calculating the upper limit on multilayer Perceptron's Hidden Neuron number", three hidden layer neurons were drawn. On processing the image of the rice leaf Brown spot was identified. The author worked on 400 images and attained 90% accuracy in classifying the disease.

Orillo et al. (2014) have acquired rice disease images from International Rice research institute, Los banos, Laguna, Philippines and presented a study using Backpropagation Artificial Neural Network. The study opted for the Brown Spot, Rice blast, and Bacterial leaf blight diseases as the study object. In this study, the images are enhanced by performing noise reduction and contrast adjustment and they are transformed from RGB color space to HSV (Hue Saturation-Value) color space. Otsu's method is applied so that the intensity of the images is adjusted. Using the process of thresholding masking the processed images are converted to binary level images and then the Blob Extraction is used to remove noise. The segmented leaf image is converted to LAB color space and the A-color plane is used for extracting the disease. The processed image is further converted to the binary-level image with the help of histogram equalization using 2 bins (black and white) so that the image serves as the mask to identify and remove the healthy part of the leaf. The diseased part of the leaf is used for feature extraction. Features extracted helped to train the backpropagation neural network and to identify the diseases with 100% accuracy.

Ramesh and Vydeki (2018) have presented a research article acquiring Rice blast disease images from Panpoli, Tirunelveli, India. Pathogen detection and categorization has been conducted using the machine learning algorithm. The study opted on Rice blast as the study object. In this study, the collected images are processed and transformed from RGB color space to HSV color space. K-means clustering is used to segment the images to compute the mean value of each cluster and calculate the distance of each cluster to the corresponding mean. Mean value, standard deviation and GLCM are calculated in the feature extraction to further train the Artificial Neural Network (ANN). In this study, they have taken 3 hidden layers in the ANN architecture with N input layers and 1 output layer. 180 Images were used for training the ANN and 120 for testing the ANN. The results in the study show that ANN achieved 99% accuracy while training and 90% accuracy while testing in classifying the rice blast disease.

4.3 *Oryza sativa* diseases classified by clustering algorithms

Prajapati et al. (2017) have presented an article on the rice plant diseases detection and classification. The images were acquired from Shertha, Gandhi Nagar, Gujarat, India. The study opted for the Leaf smut, Brown Spot and bacterial leaf blight as the study object. In this study, the collected images were converted from RGB color space to HSV color space and only a Saturation component of the HSV image is extracted. A Mask is generated based on the original image and is applied to hue component values, value component values, saturation component values of the image in HSV color space to remove the background of the RGB color space image. Once again the image generated from the removal of background is converted to HSV color space and applied K-means clustering on the hue component of the HSV image. The cluster generated is applied thresholding to remove unwanted green regions and the infected image is used for the three-color, shape, texture feature extraction. Min–Max normalization is applied and further, it is used for training the support vector machine (SVM) classifier with features extracted and tested to check whether the classifier is identifying the diseased or infected leaf. The study has compared three segmentation techniques namely HSV color space based K-means clustering, LAB color space based K-means clustering, and Otsu's segmentation and the results state HSV color space based K-means clustering achieved the highest accuracy 96.71%.

Archana and Arun (2018) have presented an article on the Automatic segmentation of the infected rice leaf using image processing techniques. The study opted for the Brown Spot and bacterial leaf blight as the study object. In this study, the images acquired are processed and transformed from RGB color space to HSV (Hue, Saturation, and value) lab color space and the Hue component is used for further analysis. K-means clustering is used to segment the images to calculate the mean value from each cluster and compute the distance of each cluster to the corresponding mean.

4.4 *Oryza sativa* diseases classified by Deep Learning algorithms

Chowdhury et al. (2019) have presented an article on the recognition of rice diseases and pests that used a convolutional neural network. The study opted for six different diseases: Brown spot, neck blast, Bacterial leaf Blight, Sheath rot, False smut, and Sheath blight, as well as the following three pest varieties: Brown Plant Hopper, Stem Borer, and Hispa as the study object. In this study, they have taken nine classes (five classes for diseases, three classes for pests, and one class for a healthy plant). They have collected images from

Bangladesh Rice Research Institute, Bangladesh. The study proposed identifying whether the leaf is healthy or not, if not healthy what kind of disease it is infected with. The main best aspect of the work is that it can detect disease at any part of plant like leaf, stem, and grain. The study compared the performance of five different CNN architectures, they are VGG16, ResNet50, Inception V3, InceptionResNet V2, and Xception by using fine-tuning, transfer learning, and training from the scratch and the performance of the VGG16 was overwhelming out of the rest architectures in terms of accuracy. Due to the large size and a huge number of parameters, the above-mentioned models are not suitable for mobile devices. The study proposed a new CNN architecture and named it as" Stacked CNN". Stacked CNN proposed can exploit training in two phases to reduce the size of the model and it can also maintain high classification accuracy. The stacked CNN accuracy performance was 95% in classifying the diseases and pests.

Lu et al. (2017) have acquired images from Heilongjiang Academy of land reclamation sciences, China. The study opted for the Brown Spot, Rice blast, false smut, Bakanae, sheath blight, sheath rot, bacterial sheath, seedling blight, bacterial wilt and Bacterial leaf blight diseases as the study object. In this study, the author collected 500 images of 10 various rice diseases from various sources. To reduce the running time, the images are compressed to 512*512 and then applied the ZCA whitening method Eq. (3) to remove the correlation between data. Further, the images are labeled with 10 kinds of rice diseases. By applying the Preprocessing procedure and selecting randomly 10,000 12*12 patches from the 500 images, an image patch is achieved along with the corresponding feature map. In this paper, the hierarchical architecture of CNN is proposed which contains three convolutional layers to extract low-level features and high-level features and Stochastic pooling layer is used to reduce the variance and Softmax regression is applied in multiclass classification problem. To train the network they have used a supervised learning algorithm. The study results state that stochastic pooling has achieved 95.48% better recognition performance when compared with the mean pooling, max pooling, and stochastic pooling using the 10-fold cross-validation strategy to train and test. The study compared the recognition performance of different convolutional filters of sizes 5*5, 9*9, 16*16, and 32*32, the results state that 16*16 filter better performed with 93.29%. The study compared the proposed Convolutional neural network (CNN) results with BP Method, support vector machine (SVM), and Particle swarm optimization (PSO) and achieved 95% accuracy.

$$X_{ZCA,Whitening} = U \cdot \frac{X_{rot}}{S + \varepsilon} \cdot U^T \tag{3}$$

where $U = SVD(S)$, $X_{rot} = U^T.X$, and $0.01 < \varepsilon < 0.1$

4.5 *Oryza sativa* diseases classified by both artificial neural network algorithms and instance based algorithms

Bhagyashri and Gajanan (2017) have acquired rice disease images from Rice Research center Karjat, Maharastra, India to detect rice diseases using texture and color descriptors Eq. (4). The study opted for the Brown Spot, and leaf blast as the study object. In this study, 40 images of each disease have been collected. The images acquired have been preprocessed by converting them from RGB color space to L*a*b color space. To provide better accuracy, K means clustering algorithm is used to do image segmentation and for feature extraction, three types of features were extracted like area, grey level co-occurrence matrix (GLCM), and color moment. The genetic algorithm was opted to reduce computational

complexity and to have low dimensionality while selecting the extracted features. The study compared Support vector machine (SVM) and Artificial neural network (ANN) classifiers and the results show that using the SVM classifier resulted in 92.5% accuracy when compared with ANN classifier accuracy, which was only 87.5%.

$$A_{affed} = \frac{P_{affed}}{P_{Total}} \tag{4}$$

where A_{affed} = Total Percentage of affected area, P_{affed} = Affected Pixels in the image, P_{Total} = Total number of pixels in the image.

Narmadha and Arulvadivu (2017) have presented a study taking image sources from internet on the paddy Disease symptoms identification using image processing. The study opted for the Narrow brown spot, Brown Spot, and Blast as the study object. In this study, the images acquired are processed to 109*310 size and as a step of preprocessing the images are converted from RGB color space to gray color space. K means clustering algorithm is used to segment the images and shape and color features like lesion shape and lesion color are extracted. Four characteristic features of the leaf diseases like lesion type, boundary color, spot color, and broken paddy leaf color were used to establish a classification system. This study reported that the Support vector machine (SVM) classification is performing well in identifying whether the leaf is diseased or not whereas Artificial Neural Network (ANN) and Fuzzy classification are good at identifying the disease of the rice plant (Table 7).

Gayathridevi and Neelamegam (2018) have presented a study acquiring images from Tamil Nadu Rice research Institute, Aduthurai, Thanjavur, India. The study was confined to rice diseases in Thanjur, India and has opted for the Leaf streak, Leaf blast, False smut, Brown Spot and Bacterial leaf blight as the study object. In this study, the collected 500 images of both healthy and infected rice plants were collected. As the initial step, the images were pre- processed by applying the median filter and they were resized to reduce the image file size. To classify the images into clusters, K-means clustering was applied to the images. Employing hybrid methods like Scale-invariant feature transform (SIFT), Discrete wavelet transform (DWT), and Gray level co- occurrence matrix (GLCM) various features are extracted from the segmented images to train various classifiers. The study compared four classifiers namely K-nearest neighbor's (KNN), Artificial Neural Network (ANN), Bayesian classifier and Multiclass Support Vector Machine and the results tabulated show that Multiclass support vector machine (SVM) classifier has attained highest accuracy i.e. 98.63% when compared with other classifiers.

Maohua et al. (2018) have worked on Rice blast recognition collecting the images from Nanjing Agricultural University. Diseases were analyzed using principal component analysis and neural network. The study chose four different types of Rice blast lesions as the research object. In this report, the images obtained both from healthy and infected rice plants and have been transformed from RGB color space to HSI and YCbCr color space to reduce the noise and the impact of inconsistent lighting. There is always variation in the color, morphology, and texture between four forms of rice blast lesions. Because of this discrepancy, the research considered the Color feature extraction, Morphological feature extraction, and texture feature analysis to extract the characteristics. In this paper, they have applied Principal component analysis (PCA) to reduce the dimension of the feature vector. The dimensionality reduced feature vectors are fed as input to the Backpropagation neural network (BPNN). The study has published results comparing the recognition rate of PCA-BPNN and PCA-SVM where PCA- BPNN performed well by achieving 95.83% accuracy.

Table 7 Summarization table of the complete review

Reference article	*Oryza sativa* Diseases	Total number of images	Image Pre-processing	Image segmentation	Feature extraction	Classifier	Accuracy (%)
Phadikar and Sil (2008)	Rice Blast and Brown spot	300	Contrast and Brightness enhanced, RGB Color space to HSI color space	Entropy-based bi-level thresholding	Boundary detection algorithm and Spot detection algorithm	Self-Organizing Map (SOM)	92
Sanyal and Patel (2008)	Rice leaf blast and Brown spot	400	RGB		Color texture extracted	Multilayer perceptron classier	89.26
Libo and Guomin (2009)	Rice Brown Spot	400	RGB Color space to Lab color space		RGL components of images	Back Propagation Neural Network	90
Yao et al. (2009)	Rice leaf blast, Sheath Blight, Bacterial Leaf Blight	216	Median filter	Otsu's Method	Shape and color features extracted with GLCM	Support vector machine	97.20
Orillo et al. (2014)	Brown Spot, Rice blast, Bacterial leaf blight	134	RGB Color space to HSV color space, Histogram equalization	Otsu's Method, Blob Extraction	Blob Extraction	Back Propagation Neural Network	100
Lu et al. (2017)	Brown Spot, Rice blast, false smut, Bakanae, sheath blight, sheath rot, bacterial sheath, Seedling Blight, Bacterial wilt, Bacterial leaf blight	500	ZCA whitening method	PCA		Deep Convolutional neural network	95.48
Prajapati et al. (2017)	Leaf smut, Brown Spot and Bacterial Leaf Blight	120	RGB color space to HSV color space, Background removal	K-means Clustering Method	Min–Max Normalization, GLCM	SVM	96.71

Table 7 (continued)

Reference article	Oryza sativa Diseases	Total number of images	Image Pre-processing	Image segmentation	Feature extraction	Classifier	Accuracy (%)
Bhagyashri and Gajanan (2017)	Brown Spot, and Leaf Blast	80	RGB Color space to Lab color space	K-means Clustering Method	GLCM	SVM & ANN	92.5 & 87.5
Narmadha and Arulvadivu (2017)	Narrow brown spot, Brown Spot, and Blast		RGB Color space to gray scale color space	K-means Clustering Method	Shape, Color,& Correlation by GLCM	Artificial Neural Network (ANN) and FUZZY classification	
Archana and Arun (2018)	Brown Spot and Bacterial Leaf Blight		RGB Color space to HSV color space	K-means Clustering Method			
Gayathridevi and Neelamegam (2018)	Leaf streak, Leaf blast, False smut, Brown Spot, Bacterial Leaf Blight	500	Median Filter	K-means Clustering Method	SIFT, GLCM, DWT	Multi Class SVM	98.63
Maohua et al. (2018)	Rice blast lesions	387	RGB Color space to HSV and YCbCr color space	Otsu's Method	Principal Component Analysis	Back Propagation Neural Network	95.83
Takuya and Yutaka (2018)	Rice blast	167	RGB	K-means Clustering Method	Principal Component Analysis	Support Vector Machine	95
Ramesh and Vydeki (2018)	Rice blast	300	RGB Color space to HSV color space	K-means Clustering Method	Mean Value, Standard Deviation, GLCM	Artificial Neural Network	99
Taohodul et al. (2018)	Rice blast, Rice bacterial blight, and Rice brown spot	60	RGB	Masking Green Pixels		Gaussian Naive Bayes	90
Abu Bakar et al. (2018)	Rice leaf blast		Histogram Equalization, RGB to HSV	Thresholding	Edge Detection	Greens theorem	

Table 7 (continued)

Reference article	*Oryza sativa* Diseases	Total number of images	Image Pre-processing	Image segmentation	Feature extraction	Classifier	Accuracy (%)
Komal et al. (2019)	Bacterial Leaf Blight, False Smut, Brown spot	400	RGB Color space to gray scale color space	K-means Clustering Method	SIFT algorithm, Bag of Words, Brute Force matcher	Support Vector Machine	94.16
Wan-Jie et al. (2019)	Rice Blast	5808	RGB		Haar-Wavelet LBPH	CNN (With Softmax) & CNN (With SVM)	95.83% & 95.82
Chowdhury et al. (2019)	Brown spot, Neck blast, Bacterial leaf Blight, Sheath Rot, False Smut, Sheath blight, Brown Plant hopper, Stem borer, and Hispa	1426	RGB			Stacked CNN	95
Neha et al. (2019)	Rice blast, Sheath Blight, Brown Spot	Not mentioned		Otsu's Method	Texture analysis, Blob Extraction	Support Vector Machine	

Wan-Jie et al. (2019) have presented a research on the Rice blast disease recognition by gathering images from the Institute of Plant protection Institute, Jiangsu, China & Agricultural Sciences Academy, Nanjing, China. The study opted on Rice Blast as the study object. In this study, the acquired images are preprocessed and extracted to 128*128 pixels from the original images by moving stride of 96 pixels and rice blast lesions found by professional in the sphere. On the whole, they have identified 2906 positive and 2906 negative samples. In this research, a correlating study based on the data set was conducted and they have analyzed the experimental results with the t-SNE qualitative assessment. The results of the assessment indicated that CNN is more discriminative and representative when compared with the LBPH (Local binary pattern histogram) and Haar-WT (Haar wavelet transform). Further, a qualitative analysis of the CNN (Convolutional neural network) with Softmax, CNN with SVM (Support Vector Machine), LBP (Local binary pattern) with SVM and Haar-WT with SVM were compared in terms of classification of rice blast, where the results prove that the CNN with SVM and CNN with Softmax has achieved 95.82% and 95.83% accuracy, which was better than the other two classifiers.

4.6 *Oryza sativa* diseases classified by both Bayesian algorithms and artificial neural network algorithms

Taohodul et al. (2018) have presented an article by acquiring images from International Rice research Institute. The study opted on Rice blast, Rice bacterial blight, and rice brown spot as the study object. In this study, the backdrop and an unimportant part of the acquired image is removed to minimize processing time. Among the RGB pixels, the green pixels are masked with blue pixels so that the green part (unaffected) of the leaf can be removed. The leftover image now contains only the real RGB form, which was further used to extract the RGB percentage inside the disease area. The pixels are classified into four categories namely Color A, Color B, Color C, and Color D, as well as the portion of each color class is computed depending on the number of pixels falling under the infected area. By applying the Gaussian Naive Bayes classifier using the neural network on the RGB percentage attained as in Eq. (5), the study has achieved 90% accuracy in classifying Bacterial blight and Brown spot disease whereas 89% accuracy in classifying the rice blast.

$$P\left(x_i|y\right) = \frac{1}{\sqrt{2\pi\sigma_y^2}} exp\left(-\frac{\left(x_i - \mu_y\right)^2}{2\sigma_y^2}\right) \qquad (5)$$

4.7 *Oryza sativa* diseases classified by Image Thresholding

Abu Bakar et al. (2018) have presented a study using multi-level color image thresholding for rice disease detection. The study opted on a Rice leaf blast as the study object and collected the images from Malaysian Agricultural Research and Development Institute, Malaysia. In this study, the acquired images are resized to 640*480 pixels and noise restoration is performed. Histogram Equalization (HE) method is used to enhance image contract and afterwards the RGB color space images are transformed to HSV color space. Only the V component is enhanced and merged to color space. As part of image segmentation, the background is subtracted and then the features are extracted by applying the Edge

detection. Green's theorem Eq. (6), is used to calculate the infected area and classified the severity of the disease in three stages.

$$\int_C Pdx + Qdy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \tag{6}$$

5 Summarization of the survey

All the literature study conducted in this article has been tabulated (see the Table 7: Summarization table of the complete review) to better understand the various research work undertaken in identifying and classifying plant diseases of *Oryza sativa*.

Many important consequences can be taken into account from the summary table, such as understanding the prominent role that image acquisition and image preprocessing play. Since all images captured are in RGB color space, some preprocessing techniques are always required to achieve better results and from our study, we understood that median filter applications on RGB color space images will have a big impact on the segmentation process. The most widely used segmentation algorithm is also the K-means clustering. Different classifiers play a vital role in the classification of *Oryza sativa* diseases, but we can point out from our summary table that backpropagation neural network has reached 100% (Orillo et al. 2014), but here the author has taken only 134 photos. As discussed in (Table 6: Summary of various Image Classifiers) BPNN Image Classifiers works well and gives good accuracy with large size data set. Convolutional neural network with Softmax outperformed the rest of the classifiers by reaching 95.83% (Wan-Jie et al. 2019) and this can be taken as a benchmark because the authors created their own image dataset of 5808 images in this research work, which was the highest among the rest of the works presented and even in the work they compared CNN with Softmax and CNN with SVM, showing that CNN with Softmax provides good classification accuracy for *Oryza sativa* diseases.

6 Challenges

Detection of *Oryza sativa* disease is a difficult field for the researcher working on the domain of image processing and artificial intelligence because there are many challenges that come up during the six-stage method. In this section, we have also addressed some of the key challenges that researcher faces in this section and still remain after so much progress.

The major challenge in this area is image acquisition phase, as further phases depend on this phase. The challenges are (a) the noisy background due to varying climatic conditions; (b) the differences in contrast between the leaves and panicles; (c) variations in climate causes variations in lighting which might reflect in low intensity values of an input image. The challenges directly affect the results of segmentation phase so do the feature extraction phase.

In the segmentation phase of *Oryza sativa* disease identification, the major task would be separating the disease area and the background. The challenges that exist during the segmentation process are (a) Varying sizes of the disease on the leaf; (b) Huge processing time is needed during region growing segmentation method; (c) varying disease color create

large variability in colors and make the color based segmentation process difficult; (d) varying changes in the scale and origin of the diseased part and the background. These above challenges directly have impact on the accuracy and also degrade the overall performance.

In the feature extraction phase, the color, texture and share features of the *Oryza sativa* plant are extracted. The challenges are (a) High dimension; (b) Redundancy which lead to high computation time (c) features extracted are sometime irrelevant.

7 Future scope

The research community working on *Oryza sativa* disease identification and classification have brought in some very fine deep learning architecture's, but still there is a major challenge to be addressed.

The challenge is that the whole community is working on the images which are been captured in the *Oryza sativa* fields. If the real scenario is considered collecting images of infected leaves on one to one basis would be a very difficult situation.

So here comes the need and scope for any new deep learning architecture which can overcome the above scenario by processing an entire sequence of data i.e. video. As recent advancements in deep learning addressed Long Short Term Memory (LSTM) (Alex Sherstinsky 2020) been playing well in classification of objects in videos. The researchers have good scope in developing some new deep learning architecture with LSTM to identify and classify *Oryza sativa* Diseases.

8 Conclusion

In this study, many forms of image processing methodologies and artificial intelligence involved in identifying and classifying the plant at *Oryza sativa* were discussed. The survey covers six key steps such as Image acquisition, Image Preprocessing, Image Enhancement, Image Segmentation, Feature Extraction, and Image Classification. From the literature review, it was understood that all the research addressed was mostly confined to Rice blast or Brown spot, and precision was achieved only by using the auto weight allocation within the context of deep learning architecture. We conclude that the collection of images and preprocessing of images play an important role in the overall process, as they impact on the accuracy of segmentation. We also conclude that K-means clustering was the most used technique of segmentation to segment the infected area from the background. Color features are also the most prominent features which can be feed forwarded to different classifiers to identify and classify *Oryza sativa* diseases. In this manner, their endeavors are expected to execute a powerful, quick, precise and automatic framework that is used for disease identification on the *Oryza sativa* unaffected leaves.

References

Abu Bakar MN, Abdullah AH, Abdul Rahim N, Yazid H, Misman SN, Masnan MJ (2018) Rice leaf blast disease detection using multi-level colour image thresholding. J Telecommun Electron Comput Eng (JTEC) 10:1–15

Adam S, Castilla N, Vera Cruz C (2013) Rice Knowledge Bank. International Rice Research Institute, [Online]. Available: http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases/item/blast-leaf-collar

Alex Sherstinsky (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena 401

Archana KS, Arun S (2018) Automatic rice leaf disease segmentation using image processing techniques. Int J Eng Technol (UAE)

Bhagyashri SG, Gajanan KB (2017) Computer vision based approach to detect rice leaf diseases using texture and color descriptors. In: International conference on inventive computing and informatics (ICICI)

Capms O, Penn S, Ponce J, Fortsyth D (2019) George Mansion University, 2019. [Online]. Available: https://cs.gmu.edu/~kosecka/cs682/lect-segmentation-part1.pdf

Chowdhury RR, Preetom SA, Eunus Ali M, Mohammad AIK, Sajid HA, Farzana N, Abu W (2019) Identification and recognition of rice diseases and pests using deep convolutional neural networks

Dipalee G, Siddhartha C (2015) Discrete wavelet transform for image processing. Int J Emerg Technol Adv Eng 4(3)

Divya S, Rajesh W, Manasi G (2015) A review: color feature extraction methods for content based image retrieval. IJCEM Int J Comput Eng Manag 18(3):9–14

Dorothy R, Joany R, Joseph RR, Santhana SP, Rajendran S (2015) Image enhancement by Histogram equalization. Int Nano Corrosion Sci Eng 2(4):21–30

Duncan AV, Bao-Rong L, Norihiko T (2008) The evolving story of rice evolution. Plant Sci 174(4):394–408

Fatemeh H, Mohammad KS (2017) A survey on classification techniques for opinion mining and sentiment analysis. Artif Intell Rev 52:1495–1545

Food and Agriculture Organization of the United Nations. [Online]. Available: http://www.fao.org/statistics/en/

Food and Agriculture Organization of the United Nations. [Online]. Available: http://www.fao.org/faostat/en/#home

Gayathridevi T, Neelamegam P (2018) Image processing based rice plant leaves diseases in Thanjavur, Tamilnadu. Cluster Comput 1–14

Ghassan Mahmoud HA, Ahmed Mohamed A (2015) Edge detection methods. In: 2nd World symposium on Web applications and networking (WSWAN), Sousse

Gonzalez Fernandez I, Iglesias-otero MA, Esteki M, Moldes OA, Mejuto JC, Simal Gandara J (2019) A critical review on the use of artificial neural networks in olive oil production, characterization and authentication. Crit Rev Food Sci Nutr 59(12):104

Gramene. [Online]. Available: https://archive.gramene.org/species/oryza/rice_intro.html

Humeau H, Anne (2019) Texture feature extraction methods: a survey. IEEE Access 7:8975–9001

Huy MD (2019) "kaggle," [Online]. Available: https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset

International Rice Research Institute. [Online]. Available: https://www.irri.org/our-solutions/irri-education

Jiang J (1999) Image compression with neural networks—A survey. Sig Process Image Commun 14(9):737–760

Jurgen S (2014) Deep learning in neural networks: an overview. Neural Networks 61:85–117

Kamlesh L, Murarka PD, Narendra SC (2015) Color space transformation for visual enhancement of noisy color image. Int J ICT Manag 3(2)

Kennedy G, Burlingame B, Nguyen (2002) Nutrient impact assessment of rice in major rice-consuming countries. [Online]. Available: http://www.fao.org/3/y6159t/y6159t04.htm

Komal B, Mehwish B, Mariam R (2019) Detection and classification of rice diseases: an automated approach using textural features. Mehran Univ Res J Eng Technol 38:239–250

Konstantinides K, Bhaskaran V, Beretta G (1999) Image sharpening in the JPEG domain. IEEE Trans Image Process 8(6):874–878

Libo L, Guomin Z (2009) Extraction of the rice leaf disease image based on BP neural network. In: International conference on computational intelligence and software engineering

Lu Y, Yi S, Zeng N, Liu Y, Zhang Y (2017) Identification of rice diseases using deep convolutional neural networks. Neurocomputing 267

Maohua X, You M, Zhixiang F, Ziang D, Shishuang H, Lei S (2018) Rice blast recognition based on principal component analysis and neural network. Comput Electron Agric 154:482–490

Mehwish R, Muhammad S, Mudassar R (2014) Image compression: a survey. Res J Appl Sci Eng Technol 7(4):656–672

Naresh K (2019) The professionals point. [Online]. Available: http://theprofessionalspoint.blogspot.com/2019/03/advantages-and-disadvantages-of_4.html

Narmadha RP, Arulvadivu G (2017) Disease symptoms identification in paddy leaf using image processing. Adv Comput Sci Technol 10:2213–2223

Naveen T, Shruti K, Manoj Kumar P (2016) Comparison of digital image segmentation techniques- a research review. Int J Comput Sci Mobile Comput 5(5):215–220

Neha M, Priyanka BR, Soumya GH, Pooja R (2019) Paddy leaf disease detection using image processing and machine learning. Int J Innov Res Electr Electron Instrum Control Eng 7(2)

Ning C, Wenjing L, Ruizhen B, An C (2017) Application of computational intelligence technologies in emergency management: a literature review. Artif Intell Rev 52(3):2131–2168

Orillo JW, Dela Cruz J, Agapito L, Satimbre PJ, Valenzuela I (2014) Identification of diseases in rice plant (*Oryza sativa*) using back propagation artificial neural network. In: IEEE international conference humanoid, nanotechnology, information technology communication and control, environment and management (HNICEM), Palawan, Philippines

Pankaj Kumar (2014) Image enhancement using histogram equalization and histogram specification on different color spaces, Rourkela

Phadikar S, Sil J (2008) Rice disease identification using pattern recognition techniques. In: International conference on computer and information technology

Prajapati HB, Shah J, Dabhi V (2017) Detection and classification of rice plant diseases. Intell Decis Technol 11:357–373

Ramesh S, Vydeki D (2018) Rice blast disease detection and classification using machine learning algorithm. In: International conference on micro-electronics and telecommunication engineering

Rana A, Rawat AS, Bijalwan A, Bahuguna H (2018) Application of multi layer (perceptron) artificial neural network in the diagnosis system: a systematic review. In: International conference on research in intelligent and computing in engineering (RICE), San Salvador

Rice Knowledge Bank. International Rice Research Institute, [Online]. Available: http://www.knowledgebank.irri.org/training/fact-sheets/pest-management/diseases

Sano Y, Sano R, Morishima H (1984) Neighbour effects between two co-occurring rice species, *Oryza sativa* and *O. glaberrima*. J Appl Ecol 21(1):245–254

Sanyal P, Patel SC (2008) Pattern recognition method to detect two diseases in rice plants. Imag Sci J 56:319–325

Shah JP (2019) UCI Machine Learning Repository, 04 2019. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases

Shakeel A, Abdul H, Khushal K, Isfahan T, Mehboob A, Faisal S (2019) Evaluation of synergistic effect of nanoparticles with antibiotics against enteric pathogens. Appl Nanosci

Shakeel A, Isfahan T, Kashif SH, Khushal K, Muhammad S, Mehboob A, Faisal S (2019b) Synthesis of silver nanoparticles using leaves of *Catharanthus roseus* and their antimicrobial activity. Appl Nanosci, DEC

Sparks A. "irri.org," [Online]. Available: http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases

Suganya Devi K, Vaishnave MP, Srinivasan P, Arulperumjothi G (2019) Detection and classification of groundnut leaf diseases using KNN classifier. In: 2019 IEEE international conference on system, computation, automation and networking (ICSCAN)

Sujung K, Wook-joong K, Seong-Dae K (2008) Automatic white balance based on adaptive feature selection with standard illuminants. In: 5th IEEE international conference on image processing, San Diego

Takuya K, Yutaka H (2018) Development of classification system of rice disease using artificial intelligence. In: IEEE international conference on systems, man, and cybernetics (SMC)

Taohodul I, Manish S, Sudipto B, Rudra R (2018) A faster technique on rice disease detectionusing image processing of affected area in agro-field. In: International conference on inventive communication and computational technologies (ICICCT 2018)

Thao N, EUn Ae P, Jiho H, Dong Chul P, Soo Young M (2013) Object detection using scale invarinat feature transform. In: Genetic and evolutionary computing, Prague, Springer, pp 65–72

Vaibhav (2015). [Online]. Available: http://computervisionwithvaibhav.blogspot.com/2015/10/otsu-thresholding.html

Wan-Jie L, Hong Z, Gu-feng Z, Hong-XIn C (2019) Rice blast disease recognition using a deep convolutional neural network. Sci Rep 9

Wu J, Zhiming C, Victor SS, Pengpeng Z, Dongliang S, Hengrong GS (2013) A comparative study of SIFT and its variants. Measure Sci Rev 13(3):122–132

Xiao L, Ying-lan M, Li-zhuang M, RUi-Ling Z (2014) A survey for image resizing. J Zhejiang Univ Sci C 697–716

Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B (2009) Application of support vector machine for detecting rice diseases using shape and color texture features. In: International conference on engineering computation

Zahid I, Muhammad AK, Muhammad S, Jamal HS, Muhammad HU, Kashif J (2018) An automated detection and classification of citrus plant diseases using image processing techniques: a review. Comput Electron Agric 12–32

ZIegler RS, Leong AS, Teng PS (1994) Rice Blast Disease, CAB International

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.