Comparative study on Functional Machine learning and Statistical Methods in Disease detection and Weed Removal for Enhanced Agricultural Yield

Sudha D. *

*Corresponding author, Associate Professor, School of Information Technology & Engineering [SITE], VIT, Vellore, Tamil Nadu, India. E.mail: msudha@vit.ac.in

Menaga D.

Research Scholar, School of Information Technology & Engineering [SITE], VIT, Vellore, Tamil Nadu, India. E-mail: menaga.india@gmail.com

Abstract

Agriculture is one of the essential sources of occupation and revenue in India. Conferring to existing statistics, most agriculturalists are facing severe losses due to poor farming yield. Farming activities are challenged by various environmental factors that affect agricultural productivity to a greater extent. The present farming situation is above the average of the process involves more biochemical bases for managing the diseases and other destructing facts. The foremost problems they are facing in day-to-day farming tasks are crop or plant diseases affecting productivity. Also, the growth of weeds along with field crops has been another challenge. The technology has developed to rectify the problems using some machine learning algorithms like Random Forest algorithms, Decision trees, Naïve Bayes, KNN, K-Means clustering, Support vector machines. The result has been evaluated and observed through the performance evaluation metrics using confusion matrix, accuracy, precision, Sensitivity, specificity with the observations, research, and studies. The statistics have expressed the overall accuracy of 98% by achieving the detection of diseases in plants and by removing the weeds that ruin the growth of plants.

Keywords: Machine learning; Statistical Techniques Hyperspectral Data; Image classification and accuracy.

Introduction

Plants are the most important living organisms to the farmers in their life. Plants are giving food and giving life to the farmers. The diseases affecting plants are the major issue in the agriculture field. The growth of plants and irrigation fields may face more challenges due to this problem [1]. The authors have selected some particular plants and applied some research work to them. Identifying the diseases in plants is very useful to obtain some information regarding the research. The pathogens are the main cause of the diseases of plants. Most of the problems can be identified and rectified using technology like Image Processing techniques with the help of hyperspectral imaging, and object detection. Image segmentation, image classification, image pre-processing, and image acquisition are the main process used in image processing techniques [2]. Images of plants are used by the techniques like thermography, chlorophyll fluorescence, and canopy reflectance. KNN, RGB imaging, Texture analysis filled by the PCR amplification R-DNA sequences, and pathogenicity tests. Most of the algorithms are mainly applied to the diagnosis of plant diseases and the removal of unnecessary weeds in the plants. They are K-Nearest Neighbours, Support Vector Machines, Random forest, reinforcement learning, Decision trees, and K-Means clustering algorithms (Cintora-Martínez et al.,2017). The datasets are the collection of data or images with the availability of tools and images. Images can be RGB images, hyperspectral images, captured images. Datasets can be Plant Village datasets, GenBank data, and Google data. Data restoration may be completely managed by some of the morphological characteristics in the technological process. The decision support systems provide the authorized information with the help of the system, software, and some operations (Raghunath, K. M. Karthick et al., 2022). This will process the report to decide with the support system. The statistical approaches validate the materials and methods which has been used for the results with some performance evaluation metrics. The accuracy is still increasing to improve the exactness of the research work.

Literature Review

Hlaing & Zaw (2017) the authors proposed the plant disease and the pathogens depends on the plant diseases using the statistical features. Th healthy plants and unhealthy plants are differentiated. The test data contains the 93.5% of accuracy and the training data contains 33% of accuracy.

Karthik et al., (2020) the authors proposed earlier and accurate detection of plants. The ANN and image processing techniques have proposed the classification of plant diseases and identify the diseases depends on the classifier.

Dhakate & Ingole (2015), have proposed the pomegranate fruits are affected by various diseases with the pathogens of fungus and bacteria. The diseases are identified by the climatic conditions. The diseases are bacterial blight, fruit spot, fruit rot, leaf spot. The decision

support systems have been used for testing and training applications. The GLCM method is used for feature extraction and ANN method used for classification method. The overall accuracy of the paper is 90%.

Ashourloo et al (2016), have proposed developing the spectral disease index (SDI). This system is the capable of identifying the different stages of wheat rust diseases.

Methodologies of Plant Disease Detection

Plant Pathology

Plant pathology is the detailed review of plant diseases and diagnosing the diseases using various techniques or methods(Praveen Sundar, P.V et al., 2021). It contains a list of common diseases affecting the plant and how to control those diseases using effective techniques with the help of machine learning techniques. Pathogens are caused by environmental factors and infectious organisms. Diseases causing damage to the plants can substantially diminish costs in the fields. Diseases are mainly inferred by viruses, bacteria and fungi (Mahesh, T. R et al., 2022).

Plant disease identification and classification

The most cost-effective management of diseases is planting a resistant variety. Nearly identified diseases are a blight, blast, false smut, leaf scald, tungro, and yellow mottal virus. Pre-planting task concerns prefer the correct diversity, emerging a cropping calendar, and arranging the place for transplanting (Islam et al., 2017). Essential components should be examined at the time of maturation of the plant. After gardening, the rice receives postharvest procedures like drying, storage, and milling to ensure good eating quality and marketability (Karthik et al., 2020). The following figure shows various plant disease classification and its pathogen.

Factors influencing plant disease diagnosis

Due to improper precautions, the diseases diffuse in the middle of fields and spread depending on the climate. Infectious plants will be prevented from the other plants by cleaning the instruments of the harvesting method after the process is completed. Fertilizers of the plants should be reduced because of more growth. More growth may ruin the field and also other plants. Other ways to encourage natural pest enemies are to allow plants on the bunds and between fields to flower. Store grain at moisture content below 13-14%, preferably available in an airtight container (Luvisi et al., 2016). Grains should be cleaned before storing them in the storehouse to avoid dust. Storing the new grains with the old grain may affect the new grains through the infection. The figure 2 shows the step by step classification affected region with the help its image processing.
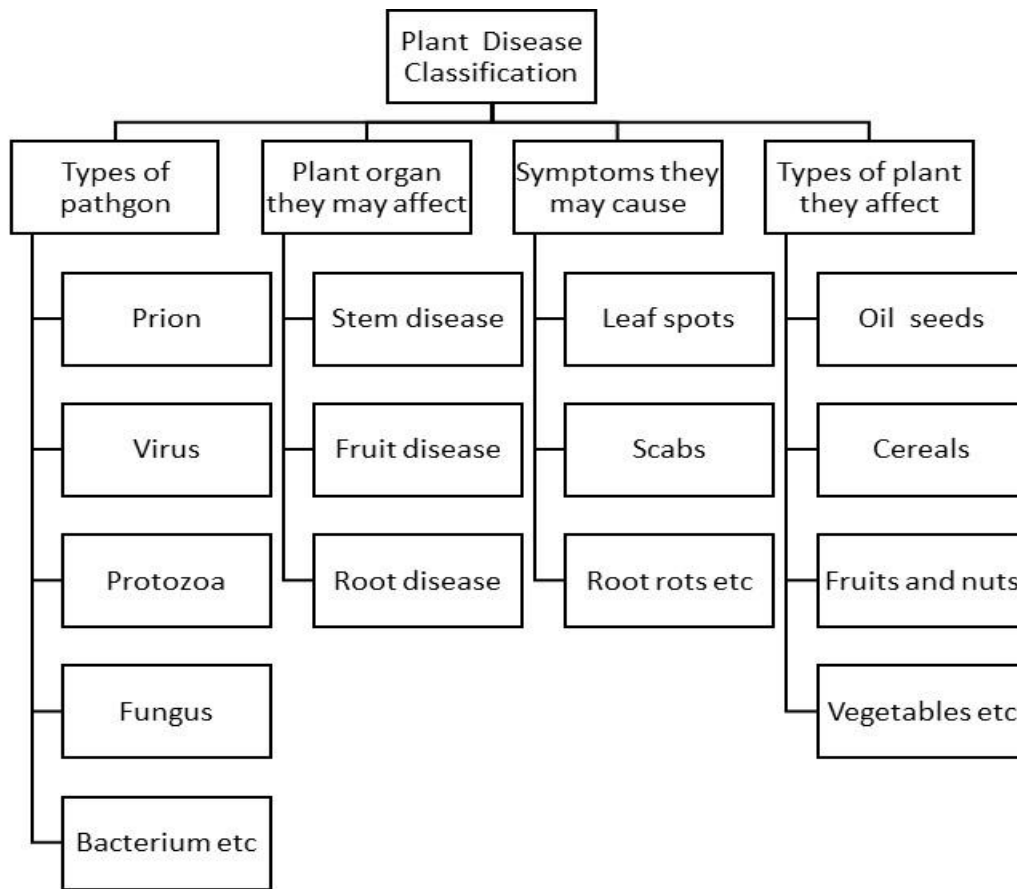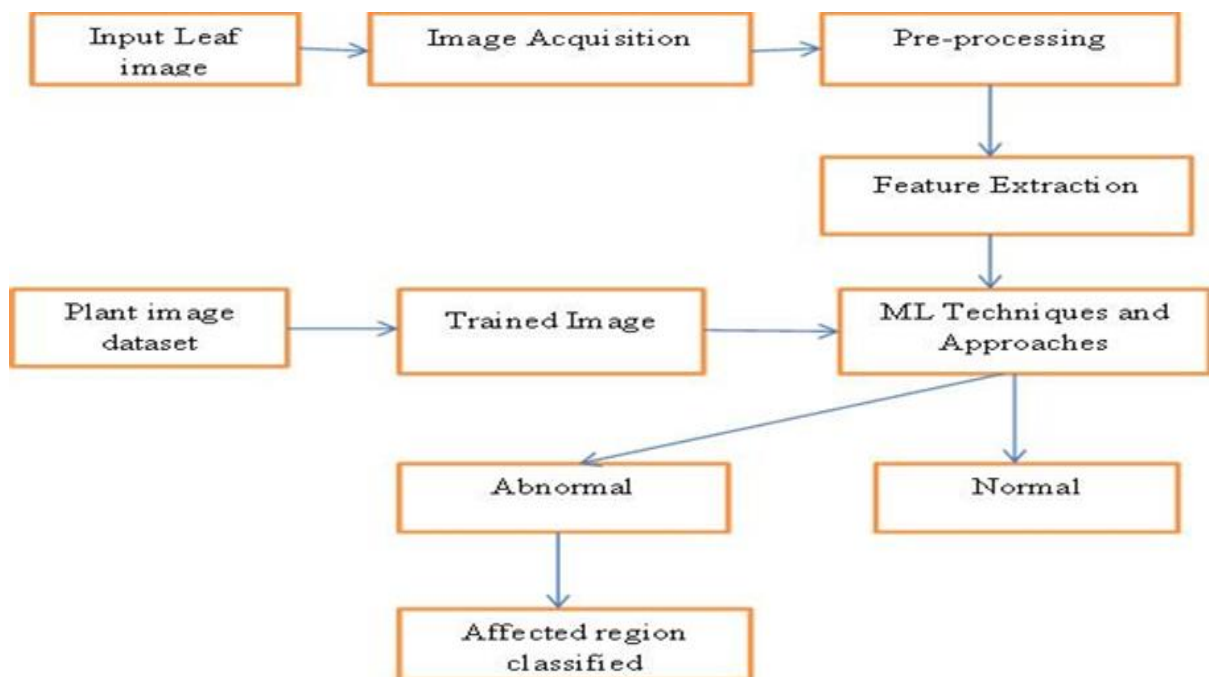
Figure 1. Plant disease classification

Figure 2 . Classification of affected region with the help of image processing

Factors influencing Weed control systems

Weed control is significant to intercept disappearances in yield and manufacturing prices and to conserve good cereal aspects. Exclusively, weeds reduce the value by straight opponent for sunlight, nutrients, and water. Weeds rise the manufacturing values e.g., higher labour or input costs. It reduces grain value and cost (Badage, 2018). For example, weed seeds in grain can cause the buyer price to be reduced. Weed management should be practised during specific stages of rice production like during land preparation, in the nursery, and during early crop growth. Controlling methods of weeds growth are stale seedbed technique, herbicides, dirty dozen, Manual and mechanical weeding (Rangarajan et al.,2018).

Leaf Colour Chart

Basically, pick any of the plants without diseases or the place where the number of people using the same variety of crop with the same number of the crop. Plants should be very younger than the other diseased plants in the hill station. The topmost plant should be more in the level of the height. Keep the mid-portion of the plant on the LCC and compare its components like colour, size, and nature with the corresponding panels. The room temperature should be maintained or under the body should be maintained for reading the plant with its panel. Because sunlight may affect the reading or disturb the reading from the panels. Each and every time the same member can read the value of the leaf colour chart for preventing unnecessary confusion (Kumar, V. V et al., 2021).

Plant Disease Detection and Weed Control System Using Various Techniques

Image Processing Technique

The main algorithms used in the image processing techniques are SVM classifiers, Complete Local Binary Pattern, K-means clustering method. The image processing techniques acquire 97% of the overall process (Kranth., 2018). Very few computational efforts are done through image processing techniques. Handcrafted or custom features based employed with statistical classifiers. Precision, Recall, Accuracy, and F1score are the studies approved in the statistical approaches. Plant Village dataset with 3000 tomato images is collected. All image processing operations can be grouped into some techniques with the analysis.

Satellite Technology

Satellites are foremost to all the people in many ways that we consume at all times. They distribute their work with electrical waves like radio and cable television. They connect us with the satellite power through cellular phone calls from long distances talks. Everyday benefits of space observation or investigation are Improving health care, protecting our earth and our surrounding, creating scientific and technical jobs, improving our day-to-day lives, enhancing safety on Earth, making scientific discoveries, sparking youth's interest in science, cooperating with countries around the world. Satellites such as NASA's TERRA satellite,

RESAT-1, PSLV-C16, and PSLV-C36 are the dataset used in the identification of wheat and cotton plant diseases. The diseases are Leaf Rust, Powdery Mildew Stem Rust, and Yellow Rust, Alternaria, Cercospora. These diseases are extracted from pests, insects, bacteria, and fungus. Histogram analysis and edge detection, Canny Edge Detection Algorithm, and the decision support system are the techniques used to identify the diseases of wheat and cotton plant using these satellites (Ramesh & Vydeki, 2018). Government stores, price lists for pesticides, and nearby open markets are some statistical approaches maintained to get higher accuracy.

Scrapping Technique

Web scraping is also known as web data extraction is data scraping used for extracting information from websites. The web scraping software may directly access the World Wide Web using the Hypertext Transfer Protocol or a web browser. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot or web crawler (Thandapani et al., 2018). It is a form of copying in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis (Vipinadas & Thamizharasi, 2016).

The scraping technique is aligned with the user interface model to find the diseases of wheat and pest. The diseases that occurred in wheat and pests are false smut, blast, and brown spot. The improper agricultural sector is the main reason for these diseases (Engelhardt et al., 2018). Python libraries such as Beautiful Soup, Scrapy, and XML, are the best example of the dataset of web scraping. Dispersion, c-value, and TF-IDF is the statistical approach to get higher accuracy (Nema & Dixit, 2018).

CFI Innovative technique

The description of research methods lacked detail and did not address innovativeness, making it difficult for the committee to assess the feasibility of the potential for breakthroughs. The research program was diverse and unfocused with no cohesive objectives and a low degree of synergy among the different projects or themes (Kouser, R.R et al., 2018). The proposal required more detail on the current state of the field and the international research context. The proposal did not clearly outline the motivation, key questions, objectives, and hypotheses. The committee questioned the feasibility of the proposed research design (Abdu et al.,2019). The proposal did not address potential research challenges and did not include a contingency plan. The Canada Foundation for Innovation's [CFI] Innovation Fund [IF] provides continued investments in infrastructure, across the full spectrum of research, from the most fundamental to applied through to technology development (Devaraj et al.,2019). The objectives of the program are to (Arzanlou et al.,2017) enable global leadership by supporting world-class research or technology development (Hlaing & Zaw, 2017) enhance and optimize the capacity of institutions and research communities to conduct the proposed research or technology development program[s] and (Cintora-Martínez et al., 2017) lead to social, health,

environmental and/or economic benefits for Canadians. Applicants need to balance readability with a level of detail that permits reviewers to assess whether the standard for each criterion has been met. This applies to all sections of the application—Research or Technology Development, Team, Research Capacity, Infrastructure, Sustainability, and Benefits. Committees have noted that "generic" responses and a lack of details/clarity prevent an effective determination of the extent to which applicants meet the stated category criteria (Zhan et al.,2019). The committees were not merely requesting longer responses but asked that applicants provide the relevant and concrete details that would allow them to determine the degree to which the relevant criteria are satisfied. All of the elements of the project should be linked to one another.

This technique was analyzed to detect the disease of wheat. The diseases of wheat are Head Blight [Fusarium]. The reason for this disease is the FHB disease complex. The algorithm used for this kind of technique is SVM [Support Vector Machine]. Thermal, Fluorescence, and Hyperspectral Imaging, IRT, CFI, and HIS are the dataset that helped to explore the images in the SVM. Superior Performing Software System SPSS 24 was the statistical approach to detect the diseases using the CFI-Innovative technique. The maximum accuracy reached is 89 % (Panchal et al., 2019).

Noise Filtering Technique

Noise is always present in digital images during image acquisition, coding, transmission, and processing steps. It is very difficult to remove noise from digital images without prior knowledge of filtering techniques. This article is a brief overview of various noise filtering techniques. These filters can be selected by analysis of the noise behaviour. In this way, a complete and quantitative analysis of noise and their best-suited filters will be presented over here (Shelake et al., 2019).

Filtering image data is a standard process used in almost every image processing system. Filters are used for this purpose. They remove noise from images by preserving the details of the same. The choice of filter depends on the filter behaviour and type of data (Singh et al.,2020). Noise is the random variation in picture element values in an image. So when it comes to the filtering of images, the first intuition that comes is to replace the value of each pixel with an average of pixel around it. This process smooths the image. For this, we consider two assumptions (Gao et al., 2019).

This technique is applied to the apple plant. The diseases obtained in the apple plant are Early, middle, and end stages disease, leaf-like target spots, downy mildew, powdery mildew, and anthracnose. Hyperspectral images, Plant Village open-access database with 50,000 datasets are used in the noise filtering with the Convolutional neural network, Data analysis, and deep learning algorithms. Noise filtering techniques obtained 99% of accuracy and also handle disease identification in real field climatic conditions (Chen et al.,2019).

Machine Learning Techniques

Plant diseases can be identified and diagnosed the diseases to achieve by many learning techniques. Using machine learning techniques, diseases can be predicted with the symptoms and easily diagnose the disease with the help of some algorithms like supervised and unsupervised techniques. Weeds can be predicted when the starting stage is in between the plants using machine learning techniques (Ma et al.,2019). It is always good to have a practical insight into any technology that you are working on in our problem statement. These machine learning ideas will comfort the researchers in learning all the solutions for the problem that needs to succeed in plant disease detection and weed management system in the industry (Barreto et al., 2020).

Supervised Algorithm

The reason for applying the supervised algorithm in our problem statement is too large data consumption in disease detection. Each plant has many diseases affected by different pathogens or different environments. The supervised algorithm is the classification process that which the plant diseases can be classified into many parts and identified according to the classification using image processing techniques. Weeds are classified depending upon the plants and removed or controlled using ML techniques. Some of the supervised algorithms that researchers have used in our problem are naïve Bayes, decision trees, support vector machines, random forests, and neural networks (Mallhi et al., 2020).

Naive Bayes

Naïve Bayes can be used for multiple classes of classification processes by using the Bayes theorem. Bayes theorem can be applied to all sorts of plants and their diseases. Due to an unfavourable environment, the diseases like Wilt, spot, powdery mildew, galls, and dryness are affected to all the plants and some of the common leaves. Web images are the source of this algorithm to detect diseases. With the help of some datasets, the weeds can be removed through this algorithm. Unusual care of plants and other issues can be repaired. Leaf blight and brown spot diseases can be resolved. Mostly 97% of accuracy was gained in this algorithm (Iqbal & Talukder 2020).

Decision Tree

The decision tree allows reviewers to predict whether the plant is affected by disease or not and weeds are controlled or cannot be controlled. Answering a question from the farmers and filling out the survey helps to fix our problem statement. Some of the information will be missing by the farmers, but the decision trees support us to deliver accurate results for this problem. Some of the categorical attributes are delivered to this problem. Corn plant diseases are abnormal. Plants with wither, turn into stunts that change their colour, for example, leaves that turn yellow or dry are identified and rectified by the generic feature extraction. Scale-invariant feature transform speeded up robust features and Oriented FAST and rotated, and

object detector such as histogram of oriented gradients are the datasets applied to the weed control system results from higher accuracy (Ennadifi et al., 2020).

Random Forest Algorithm

Random forest algorithm assists in the plant disease identification and weed removal system. The random forest supplies more prediction detection of plant diseases with the help of more information or data collected from the dataset and database. This algorithm is mainly used in genetic feature extraction to detect pathogens like a virus, bacteria. Abnormal plants that wither, turn into stunted and some that change colour, for example, leaves turn yellow or dry. These are the symptoms that can occur in the plant identified by using a random forest algorithm. RGB images and dimensional features are the studies declared in the random forest algorithm. The skills of random forest algorithms are used to evaluate the diseases in the identification of plant diseases. Plant Village dataset/ Image training libraries and RGB images are the main datasets classified with the help of image processing. Rice leaf contains spots, Rust, Rots, wilt, cankers, and dwarfing diseases. 140 images of tomato diseases can be identified with the help of a color histogram, Hue moments, and feature extraction (Nagaraju & Chawla,2020) [58]. The following figure shows working of Random Forest algorithm
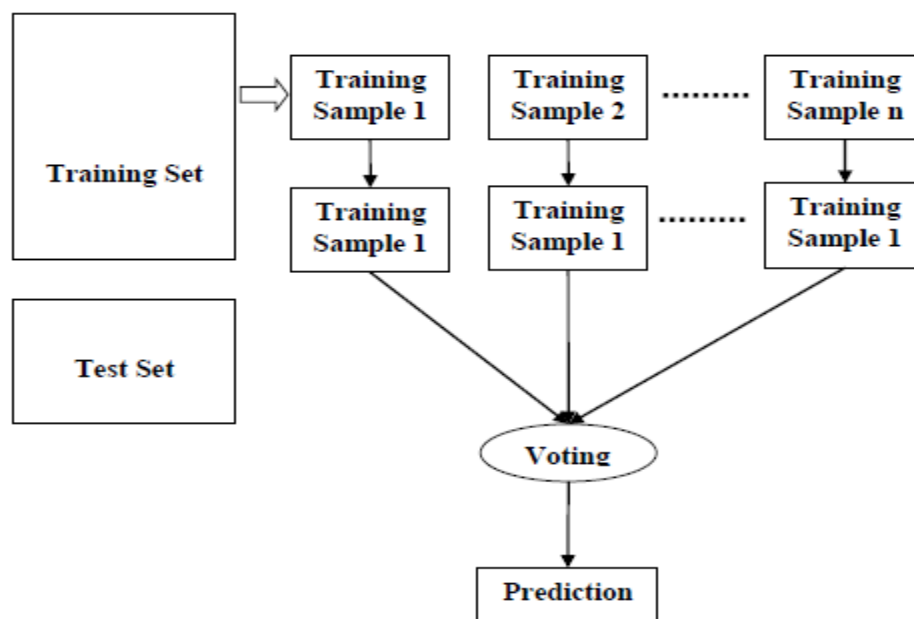

Figure 3. Working of Random Forest algorithm

Convolutional Neural Networks

Researchers have made neural networks for developing new ideas, classification of problems, and consolidating the more complex structures. Our work is defined and classified mainly using this neural network. Because more diseases are identified in the major plants. The neural networks are used to classify those plants with diseases and the solution to overcome the problems. Apple fruit has count ripe and half-ripe disease in the species of Phytophthora.

The diseases are collected using ImageNet/Real and synthetic images in the multiscale studies and provided 91% accuracy using the deep convolutional neural networks. Neural network algorithms are mostly applied in vegetables, cotton, cereals, fruits, leaves, and groundnuts. The mean, variance, SD, specificity, sensitivity are the studies to produce the maximum accuracy of the result (Vishnoi et al.,2021). Threshold values handle disease identification in real field climatic conditions.

Table 1. Comparison of CNN models with Accuracy

CNN Models	Train Accuracy	Test Accuracy
CNN	98.21%	96.62%
MCNN	98.51%	94.08%
ModCNN	98.52%	97.69%
Faster RCNN	66%	60%

Unsupervised Algorithm

Unsupervised learning is like learning without a teacher. One basic and important process that you might want to perform with data is to visualize the data and also to identify the problem. But it is very difficult to visualize things in more than two (or three) dimensions and most data is in hundreds of dimensions. Dimensionality reduction is the problem of performing the high dimensional data and embedding it in a lower dimension space. The main performance of unsupervised machine learning has automatically derived a partitioning of the data into clusters. Some of the unsupervised algorithms used in the disease detection of plants are K-Means clustering, Dimensionality reduction, K-Nearest Neighbours (KNN), and Principal Component Analysis (PCA) (Guo et al., 2020).

K-Means Clustering:

The method of proving the convergence is to specify a clustering quality objective function, and then to show that the K-Means algorithm converges to a [local] optimum of that objective function is optimizing is the sum of squared distances from any data point to its assigned k-centre (V., Muthukumaran et al.,2021). This is a natural generalization of the definition of a mean: the mean of a set of points is the single point that minimizes the sum of squared distances from the mean to every point in the data. Rice leaf diseases are detected by Bacterial, fungal, and viral pathogens using Image processing in the unsupervised algorithm. Clustering is the main progress in this algorithm. It is preferable as it is easy to get unlabelled data in comparison to labelled data. It is used for more complex tasks as compared to other algorithms. Those data points which are near to the particular k-centre, create a cluster (Sun et al.,2020).

K-Nearest Neighbour (KNN) algorithm

K-Nearest Neighbour is the main concept of a supervised learning algorithm that can work with both classification and regression problems. This algorithm works by assuming the similarities between the new data point and available data points. Based on these similarities,

the new data points are put in the most similar categories. It is also known as the lazy learner algorithm as it stores all the available datasets and classifies each new case with the help of K-neighbours. The new case is assigned to the nearest class with the most similarities, and any distance function measures the distance between the data points (Zou et al.,2019).

Principal Component Analysis (PCA)

Principle Component Analysis (PCA) is an unsupervised learning technique, which is used for dimensionality reduction. It helps in reducing the dimensionality of the dataset that contains many features correlated with each other. It is a statistical process that converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation. It is one of the popular tools that is used for exploratory data analysis and predictive modelling. The diseases identified in the cassava are Cassava brown streak, Cassava Mosaic disease. The techniques are Spectrometry, GMLVQ used in plant disease identification. The 760 numbers spectral images continue to identify the diseases.

Table 2. Some reported disease detection and its algorithm with accuracy

Plants	Disease	Algorithm	Accuracy	Reference
Tomato	Spot,Blight, Virus	SVM Classifier	92%	Hlaing & Zaw -2017
Apple fruit	Count ripe and half-ripe	CNN	95.7%	Raghunath, K. M. Karthick et al -2022
Potato	Late blight, early blight	SVM	95%	Islam et al -2017
Perennial	Anthracnose,Bacterial blight	SVM, KNN, ANN,	89%	Bajwa et al-2017
Wheat, Cotton	Mildew, Rust, Alternaria, Cerespora	Canny Detection	93%	Badage- 2018
Apple	Gemini virus Proteins,	RNA	88%	Zhan et-2018
Pea plant	Rust	SVM	89.60%	Singh et al-2019
Jute,Grape, Paddy, okra	Brown and yellow spots	K-means, ANN	95%	Halder et al-2019
Septoria	Leaf spot	SVM, KNN, RF, NB	60%	Jumat et al-2018
Rice	Rice blast	ANN	99%	Ramesh & Vydeki-2018
Tomato	Early and Late blight lesions	SVM and CNN	96.9% and 94.17%	Abdu et al -2019
Chilli plant	Affected spots, Background spots	K-Means	79.6%	Wahab et al -2019
Wheat	Head Blight Fusarium	SVM	89%	Mahlein et al -2019
Citrus fruit	Anthracnose, canker, scab, melanoses, black spot	K-Means,ANN SVM	92%	Dhiman, G et al -2021
Cucumis melo	Cryptic virus, Melon Necrotic Spot	Neighbour-joining algorithm	57.7%	Zhan et al -2019
Pepperell Tomato	Early Blight, Late Blight, and Bacterial spot	K-means clustering and HSV dependent classification	98%	Panchal et al -2019
Sugar beet	Root and crown rot	RFE	99%	Barreto et al -2020
Potato	Early and Late Blight	CNN and RF	96%	Iqbal & Talukder 2020

Performance Evaluating Metrics

Evaluating performance in the machine learning algorithm lies under accuracy, confusion matrix, precision, specificity, and recall or sensitivity. Good evaluation is compromised with a good questionnaire. The evaluation consists of the impact, production, outcome, and development in agronomy.

Confusion matrix

This metric is used for classification algorithm, finding a better outcome. Solving the problem of whether the citrus frosting the disease of citrus canker or not? Consider the trained data set and target variable associated with citrus fruit and disease citrus cancer. When the fruit has disease 0: When the fruit N do having the disease. Now we should differentiate the dimensions of the actual solution and predicted solution. Based on the performance, the confusion matrix is calculated using TP [True Positive], TN [True Negative], FP [False Positive], and FN [False Negative].

Accuracy

Accuracy can be calculated by the total number of true positive and true negative by the total predictions. Accuracy is the possible outcome of a problem with specific algorithms. It is the measuring of good outcomes with the target variables and predictions (Dandawate & Kokare, 2015). For example, consider the 100 apple trees are planted in a field. Out of 100apple trees, 5 apple trees are affected by the scab disease which made the tree lose leaf and trees. Therefore, 95% accuracy is predicted, and also the actual outcome was 95% as calculated by the formula:

$$Accuracy = TP+TN/TP+TN+FP+FN \tag{1}$$

Precision

Precision is a measure of which proportion of plants that we diagnosed as having cancer, actually had cancer. The predicted positives [people predicted as cancerous are TP and FP] and the people having cancer are TP.

Sensitivity

Sensitivity is a measure that tells us what proportion of plants had cancer was diagnosed as having cancer by the algorithm. The actual positives [plants having cancer are TP and FN] and the plants diagnosed by the model having cancer are TP. [Note: FN is included because the person had cancer even though the model predicted otherwise].

Specificity

Specificity is a measure that tells us what proportion of patients that did NOT have cancer, were predicted by the model as non-cancerous. The actual negatives [People NOT having cancer are FP and TN] and the people diagnosed by us not having cancer are TN.

Applications of Disease detection systems and weed removal systems

Plant disease detection and weed removal system have many applications in the field of machine learning and image processing. The variety of plants like wheat, barley, fruits, and vegetables are improved in the part of agriculture. Crop improvement, soil management, and water management are powerful applications of disease diagnosis. Biosensors and image sensors are the applications of controlling weeds in crop management. The soil will be purified in the process of leaf disease diagnosis (Cambra Baseca et al., 2019).

Detecting the plant disease helps the farmer to monitor the large work in wide area of farms and detect the symptoms of the diseases so that easily can identify the solution in the earlier time. Develops the health status of the leaves that programs statistical analysis.

As mentioned before, farmers will get relief from their problems. Due to technological development, researchers can easily overcome agricultural problems. Enrichment of food provides a healthier supplement to all the people. Weed removal systems mainly avoid the disease of plants.

Research gaps

There are many advantages in detecting plant disease and removing the weeds using ML approaches that enhance the healthy plants to the farmers. But issues are never destroyed in any techniques. The factor influencing the existing problem approaches are discussed below.

- Problem.1. Plant disease and the growth of weeds are the main issues in agriculture in India. The existing methodologies of plant disease diagnosis are spray methods, capturing the images by using a mobile camera, an objective lens for identifying the disease, microscope for identifying the portion of the affected area in the plant.

- Problem.2. Soaking onto seeds provides lesser accuracy. Only a small portion is detected in the plants. Poor water-soluble is attained by existing pesticides. Sedimentation cause heritage problem in existing technologies.

- Problem.3. The existing datasets were approached in the previous publication. The existing datasets enlarge the accomplished risk due to climate change. The total numbers of image datasets are 79265. These datasets decapitate the diseases and issue uncertain methods. It provided dummy information about the technologies, algorithms, and models.

- Problem.4. The existing equipment cannot control the diseases due to laboratory imaging or environmental risks. The equipment cannot measure clear accuracy in the problem. Cadmium dot and zinc sulphide dot will grow in all plants due to insufficient equipment.

- Problem.5. Human interaction of identifying the plant diseases may consume less energy and produce more obstacles in the development of research process. The process is very slow. Economic growth of the farmers is less developed. The growth of the plants without any disease may show the growth of the farmers.

- Problem.6. Reduction of production and poor quality of production are the main challenges in the field of agriculture. Adequate instruments are needed to diagnose the diseases and cost is very high for the particular instruments.

- Problem.7. The other challenge in the plant disease detection is incapable of capturing all the images through the datasets. The values are very less. The missing information may affect the procedure of identifying the diseases.

- Problem.8. Proper utilization of techniques can provide the great solution. But the data sets are very hard to find out from the websites and also lack of images also affects the procedure.

Current challenges of plant disease detection:

1. Multiple number of infections are identified in single or multiple leaves.

2. Complex backgrounds associated with different real time conditions.

3. Computational time complexity is implemented.

4. Multiple fruit or multiple leaves detection are detected in same branch.

5. Infection level may over-increase or over-decrease in time variant.

6. Overlapped stems are surrounded and their shades on fruits are identified.

Future trends

In the future pest management is involved to destroy the efficacy of current problems. The release of carbon-di-oxide can be controlled in the future. Potential vulnerabilities and biotic loss can be managed. CRISPR tools should be evolved the laboratory research to land research. Nutrient levels and stress resistance levels can be increased in the future. Treatment of disease detection and weed removal system should be demonstrated earlier. Successful implementation can be done using new technology. The arable lands will appear in the next generation period. In the future, researchers planning to use multispectral images, and field images for better accuracy.

The machine algorithms associated with agro-technology will be used for integrated tools and applicable tools. The investigation will grow at different stages. The proposed work will add the field test in the ML algo8rithms. Variables and Virulent factors have been established to reduce the risk in the future. More pathogens and their components are discussed and survey online or field will be investigated. Building the technologies, the disease will be recruited. The plant according to their production in India must be addressed. Taking the future prospectus to induce the farmer's wealth for growth of crops. From the review of plant disease detection and weed removal systems, we can decide that Agriculture will have an extraordinary future without any obstacles to farmers.

Conclusion

In this paper, a systematic review of plant disease detection and weed management is absolutely discussed. A variety of plants according to the diseases are disclosed with the pathogens. The pathogens are identified according to the plant diseases and in the survey. The components of ML techniques are applied to this system. The algorithms used in this problem are mainly highlighted and explained to identify the disease and identify the weed in the middle of plants. Feature extraction played a vital role and image classification and disease classification are revealed. Types of ML algorithms made the review easier to study. An effective way of detecting the disease provides better accuracy and problems of the weed removal system. Technologies and systems are analyzed in the advanced activities of plants. Computational models are used to analyse the problem very clearly. Many datasets are included and databases are acquired through the samples. The problem statement and related works are clearly explained. Existing system access the problems from previous works. Research gaps are illustrated and depicted to overcome complex diseases.

Acknowledgement

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

References

Abdu, A. M, Mokji, M. M., and Sheikh, U. U.(2019), A Pattern Analysis-based Segmentation to Localize Early and Late Blight Disease Lesions in Digital Images of Plant Leaves. In 2019 IEEE International Conference on Signal and Image Processing Applications [ICSIPA], pp. 116-121.

Abdu, A. M., Mokji, M. M., Sheikh, U. U., and Khalil, K.(2019), Automatic disease symptoms segmentation optimized for dissimilarity feature extraction in digital photographs of plant leaves. In 2019 IEEE 15th International Colloquium on Signal Processing & Its Applications [CSPA], Vol. 8, pp. 60-64

Aruraj, A., Alex, A., Subathra, M.S., Sairamya, N.J., George, S.T., and Ewards, S.V.(2019) Detection and classification of diseases of the banana plant using local binary pattern and support vector machine. In2019 2nd International Conference on Signal Processing and Communication [ICSPC], Vol. 29, pp. 231-235

Arzanlou, M., Torbati, M., & Narmani A. (2017), Podosphaera clandestine causes powdery mildew on sour cherry in Iran. Australasian Plant Disease Notes. Vol. 1;No. 12.

Aslam, W., Noor, R.S., Hussain, F., Ameen, M., Ullah, S., and Chen, H.(2020), Evaluating morphological growth, yield, and postharvest fruit quality of cucumber [Cucumis sativus L.] grafted on cucurbitaceous rootstocks. Agriculture. Vol. 10,No.4.

Astonkar, S. R., & Shandilya, V. K. (2018), Detection and Analysis of Plant Diseases Using Image Processing.Vol .5, No. 4, pp. 3190-3193

Badage, A (2018), Crop disease detection using machine learning: Indian agriculture. Int. Res. J. Eng. Technol.[IRJET].Vol.5, No.9,pp. 866-869.

Bajwa, S,G., Rupe, J,C., and Mason, J.(2017), Soybean disease monitoring with leaf reflectance. Remote Sensing. Vol. 9, N0.[2].

Balducci, F., Impedovo, D., and Pirlo, G. (2018), Machine learning applications on agricultural datasets for smart farm enhancement. Machines. Vol. 6, No. 3.

Barreto, A., Paulus, S., Varrelmann, M., and Mahlein, A.K.(2020), Hyperspectral imaging of symptoms induced by Rhizoctonia solani in sugar beet: Comparison of input data and different machine learning algorithms. Journal of Plant Diseases and Protection. Vol. 127,No.4.

Behmann, J., Bohnenkamp, D., Paulus, S., and Mahlein, A. K. (2018), Spatial referencing of hyperspectral images for tracing of plant disease symptoms. Journal of Imaging. Vol. 4, No.12.

Cambra Baseca, C., Sendra, S., Lloret, J., and Tomas.,J.(2019), A smart decision system for digital farming. Agronomy. Vol. 9,No.5.

Chaware, R., Karpe, R., Pakhale, P., and Desai, S. (2017), Detection and recognition of leaf disease using image processing. International Journal of Engineering Science and Computing.

Chen, Y., Mi,Y., Sun, X., Zhang, J., Li, Q., Ji, N., and Guo, Z.(2019), Novel inulin derivatives modified with Schiff bases: Synthesis, characterization, and antifungal activity. Polymers. Vol. 11,No.6.

Chhillar, A., and Thakur S. (2021), Plant Disease Detection Using Image Classification. In Proceedings of International Conference on Big Data, Machine Learning and their Applications , pp. 267-281.

Cintora-Martínez, E.A., Leyva-Mir, S.G., Ayala-Escobar, V., Ávila-Quezada, G.D., Camacho-Tapia, M., & Tovar-Pedraza, J.M.(2017) Pomegranate fruit rot caused by Pilidiella granati in Mexico. Australasian Plant Disease Notes. Vol. 1;No.12.

Dai, T., Wang, A., Yang, X., Yu, X., Tian, W., Xu, Y., and Hu, T.(2020), PHYCI_587572: an RxLR effector gene and new biomarker in a recombinase polymerase amplification assay for rapid detection of Phytophthora cinnamomi. Forests. Vol. 11,No.3.

Dandawate, Y., & Kokare, R.(2015), An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective. In2015 International conference on advances in computing, communications and informatics [ICACCI],Vol. 10.

Dasgupta, I., Saha, J., Venkatasubbu, P., and Ramasubramanian, P.(2020), AI Crop Predictor and Weed Detector Using Wireless Technologies: A Smart Application for Farmers. Arabian Journal for Science and Engineering.Vol. 45,No.12.

Debeljak, M., Trajanov, A., Kuzmanovski, V., Schröder, J., Sandén, T., Spiegel, H., Wall, D.P., Van, de Broek M., Rutgers, M., Bampa, F., and Creamer, R.E. (2019), A field-scale decision support system for assessment and management of soil functions. Frontiers in Environmental Science. Vol. 5, No. 7.

Devaraj, A., Rathan, K., Jaahnavi, S., and Indira, K. (2019), Identification of plant disease using image processing technique. In2019 International Conference on Communication and Signal Processing [ICCSP], Vol. 4, pp. 0749-0753

Dhiman, G., Vinoth Kumar, V., Amandeep Kaur & Ashutosh Sharma., (2021), DON: Deep Learning and Optimization-Based Framework for Detection of Novel Coronavirus Disease Using X-ray Images. Interdiscip Sci Comput Life Sci 13, 260–272

Doh, B. , Zhang, D., Shen, Y., Hussain, F., Doh, R.F., Ayepah, K.(2019), Automatic citrus fruit disease detection by phenotyping using machine learning. In 2019 25th International Conference on Automation and Computing [ICAC], Vol .5, pp. 1-5.

Engelhardt, S., Stam, R., and Hückelhoven, R. (2018) Good riddance? Breaking disease susceptibility in the era of new breeding technologies. Agronomy. Vol. 8, No. 7.

Ennadifi, E., Laraba, S., Vincke, D., Mercatoris, B., and Gosselin, B.(2020),Wheat Diseases Classification and Localization Using Convolutional Neural Networks and GradCAM Visualization. In2020 International Conference on Intelligent Systems and Computer Vision [ISCV] Vol. 9.

Gao, Y., Lu, Y., Lin, W., Tian, J., and Cai, K.(2019), Biochar suppresses bacterial wilt of tomato by improving soil chemical properties and shifting soil microbial community. Microorganisms. Vol. 7,No.12.

Govardhan, M., & Veena, M. B.(2019), Diagnosis of Tomato Plant Diseases using Random Forest. In2019 Global Conference for Advancement in Technology [GCAT] pp. 1-5.

Guo, Y., Zhang, J., Yin, C., Hu, X., Zou, Y., Xue, Z., and Wang, W.(2020), Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society.

Halder, M., Sarkar, A., and Bahar, H.(2019), Plant Disease Detection by Image Pro-Cessing: A Literature Review. image. Vol.1, No.3.

 Hlaing, C.S., & Zaw, S.M.(2017), Plant diseases recognition for smart farming using model-based statistical features. In 2017 IEEE 6th global conference on consumer electronics [GCCE]. Vol. 24,[pp. 1-4].

Iqbal, M.A., & Talukder, K.H.(2020),Detection of potato disease using image segmentation and machine learning. In2020 International Conference on Wireless Communications Signal Processing and Networking [WiSPNET] Vol. 4.

Islam, M., Dinh, A., Wahid, K., & Bhowmik, P.(2017), Detection of potato diseases using image segmentation and multiclass support vector machine. In2017 IEEE 30th Canadianonference on electrical and computer engineering [CCECE].

Jumat, M.H., Nazmudeen, M.S., and Wan, A. T.(2018), Smart farm prototype for plant disease detection, diagnosis & treatment using IoT device in a greenhouse. 7th Brunei International Conference on Engineering and Technology 2018 [BICET 2018].

Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., and Menaka, R.(2020), Attention embedded residual CNN for disease detection in tomato leaves. Applied Soft Computing.Vol. 1;No.86.

Koshy, S.S., Sunnam, V.S., Rajgarhia, P., Chinnusamy, K., Ravulapalli, D.P., and Chunduri, S.(2018), Application of the internet of things [IoT] for smart farming: a case study on groundnut and castor pest and disease forewarning. CSI Transactions on ICT. Vol. 6,No.3.

Kouser, R.R., Manikandan, T., Kumar, V.V (2018), "Heart disease prediction system using artificial neural network, radial basis function and case based reasoning" Journal of Computational and Theoretical Nanoscience, 15, pp. 2810-2817

Kranth, G. P., Lalitha, H.M., Basava, L., and Mathur, A.(2018), Plant disease prediction using machine learning algorithms. Int. J. Comput. Appl. Vol. 182, No.25, pp.1-7.

Kumar, V. V., Raghunath, K. M. K., Rajesh, N., Venkatesan, M., Joseph, R. B., & Thillaiarasu, N. (2021). Paddy Plant Disease Recognition, Risk Analysis, and Classification Using Deep Convolution Neuro-Fuzzy Network. Journal of Mobile Multimedia. doi:10.13052/jmm1550-4646.1829

Kusumo, B. S., Heryana, A., Mahendra, O., and Pardede, H. F.(2018) Machine learning-based for automatic detection of corn-plant diseases using image processing. In2018 International Conference on Computer, Control, Informatics and its Applications [IC3INA] Vol.1, pp. 93-97.

Luvisi, A., Ampatzidis,Y,G., and De Bellis, L.(2016), Plant pathology and information technology: Opportunity for management of disease outbreak and applications in regulation frameworks. Sustainability. Vol. 8, No,8. https://doi.org/10.3390/su8080831

Ma, Y., Zhang, J., Xiao, Y., Yang, Y., Liu, C., Peng, R., Yang, Y., Bravo, A., Soberón, M., and Liu, K.(2019), The cadherin Cry1Ac binding-region is necessary for the cooperative effect with ABCC2 transporter enhancing insecticidal activity of Bacillus thuringiensis Cry1Ac toxin. Toxins. Vol. 11,No.9.

Mahesh, T. R., V. Dhilip Kumar, V. Vinoth Kumar, Junaid Asghar, Oana Geman, G. Arulkumaran, and N. Arun. (2022) "AdaBoost Ensemble Methods Using K-Fold Cross Validation for Survivability with the Early Detection of Heart Disease." Computational Intelligence and Neuroscience

Mahlein, A. K., Alisaac, E., Al Masri A., Behmann, J., Dehne, H.W., Oerke, E.C. (2019), Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring fusarium head blight of wheat on spikelet scale. Sensors. Vol .19, No. 10.

Mallhi, A.I., Chatha, S.A., Hussain, A.I., Rizwan, M., Bukhar, S.A., Hussain, A., Mallhi, Z.I., Ali, S., Hashem, A., Abd_Allah, E.F.,and Alyemeni, M.N.(2020), Citric acid assisted phytoremediation of chromium through sunflower plants irrigated with tannery wastewater. Plants. Vol. 9,No.3.

Nagaraju, M., & Chawla, P.(2020), Systematic review of deep learning techniques in plant disease detection. International Journal of System Assurance Engineering and Management. Vol. 11,No.3.

Nema, S., & Dixit, A.(2018), Wheat leaf detection and prevention using support vector machine. In2018 International Conference on Circuits and Systems in Digital Enterprise Technology [ICCSDET], pp.1-5.

Neupane, S., Purintun, J.M., Mathew, F.M., Varenhorst, A.J., and Nepal, M.P.(2019) Molecular basis of soybean resistance to soybean aphids and soybean cyst nematodes. Plants. Vol. 8,No.10.

Owomugisha, G., Melchert. F., Mwebaze, E., Quinn, J.A., and Biehl, M.(2018), Machine learning for diagnosis of disease in plants using spectral data. InInternational Conference on Artificial Intelligence [ICAI], pp. 9-15.

Panchal, P., Raman, V.C., and Mantri, S.(2019), Plant diseases detection and classification using machine learning models. In2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution [CSITSS], Vol. 4, pp. 1-6

Praveen Sundar, P.V., Ranjith, D., Vinoth Kumar, V. (2020). Low power area efficient adaptive FIR filter for hearing aids using distributed arithmetic architecture. *Int J Speech Technol* 23, 287–296 (2020). https://doi.org/10.1007/s10772-020-09686-y

Raghunath, K. M. Karthick, V. Vinoth Kumar, Muthukumaran Venkatesan, Krishna Kant Singh, T. R. Mahesh, and Akansha Singh (2022). "XGBoost Regression Classifier (XRC) Model for Cyber Attack Detection and Classification Using Inception V4." *Journal of Web Engineering.* doi:10.13052/jwe1540-9589.21413

Ramesh, S., & Vydeki, D.(2018), Rice blast disease detection and classification using machine learning algorithm. In 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering [ICMETE] Vol. 20, pp. 255-259.

Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., and Vinod, P.V.(2018), Plant disease detection using machine learning. In2018 International conference on design innovations for 3Cs compute communicate control [ICDI3C].

Rangarajan, A. K., Purushothaman, R., and Ramesh, A.(2018) Tomato crop disease classification using pre-trained deep learning algorithm. Procedia computer science. Vol.133, pp.1040-1047.

Shelake, R. M., Pramanik, D., and Kim, J.Y.(2019), Exploration of plant-microbe interactions for sustainable agriculture in CRISPR era. Microorganisms. Vol. 7, No. 8.

Singh, S., Sharma, M. P., Alqarawi, A. A., Hashem, A., Abd_Allah, E. F, Ahmad, A.(2020), Real-time optical detection of isoleucine in living cells through a genetically-encoded nanosensor. Sensors. Vol. 20, No.1.

Singh, V., & Misra, A,K. (2017), Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture. Vol. 1;No.4.

Singh. K., Kumar, S., and Kaur, P. (2019), Support Vector Machine classifier based detection of fungal rust disease in Pea Plant [Pisam Sativa]. International Journal of Information Technology. Vol.11, No.3, pp. 485-492.

Sun, G., Jia, X., and Geng, T. (2018), Plant diseases recognition based on image processing technology. Journal of Electrical and Computer Engineering. Vol.1.

Sun, Y., Wang, M., Mur, L.A., Shen, Q., and Guo, S.(2020), Unravelling the roles of nitrogen nutrition in plant disease defences. International journal of molecular sciences. Vol. 21,No.2.

Thakur, T.B., & Mittal, A.K.(2020), Real Time IoT Application for Classification of Crop Diseases using Machine Learning in Cloud Environment. International Journal of Innovative Science and Modern Engineering [IJISME]. Vol. ;6,No.4.

Thandapani, S. P., Senthilkumar, S., and Priya, S.S.(2018) Decision Support System for Plant Disease Identification. In International Conference on Advanced Informatics for Computing Research,Vol. 14, pp. 217-229.

Tlhobogang, B., & Wannous, M. (2018) Design of plant disease detection system: a transfer learning approach work in progress. In 2018 IEEE International conference on applied system invention [ICASI], Vol.13, pp. 158-161.

Tripathi,A., Singh, A.K., Singh, K.N., and Singh, K.K., Choudhary, P., Vashist, P.C.(2021) 2 Food security and farming through IoT and machine learning. In Internet of Things and Machine Learning in Agriculture, pp. 21-40

V., Muthukumaran, Satheesh Kumar S., Rose Bindu Joseph, Vinoth Kumar V., and Akshay K. Uday (2021). "Intelligent Medical Data Analytics Using Classifiers and Clusters in Machine Learning." Advances in Computational Intelligence and Robotics : 321–335. doi:10.4018/978-1-7998-6870-5.ch022.

Vipinadas, M. J., & Thamizharasi, A.(2016), Detection and Grading of diseases in Banana leaves using Machine Learning. International Journal of Scientific & Engineering Research. Vol. 7, No. 7, pp. 916-924.

Vishnoi, V.K., Kumar, K., and Kumar, B.(2021), Plant disease detection using computational intelligence and image processing. Journal of Plant Diseases and Protection. Vol. 128,No.1.

Wahab, A. H., Zahari, R., and Lim, T. H.(2019), Detecting diseases in chilli plants using K-means segmented support vector machine. In 2019 3rd International Conference on Imaging, Signal Processing and Communication [ICISPC] , pp. 57-61

YP., Abdul, W., Zhan, J., and Yang, L.N.(2020), Temperature-mediated plasticity regulates the adaptation of Phytophthora infestans to azoxystrobin fungicide. Sustainability. Vol. 12,No.3.

Zhan, B., Cao, M., Wang, K., Wang, X., Zhou, X. (2019), Detection and characterization of cucumis melo cryptic virus, cucumis melo amalgavirus 1, and melon necrotic spot virus in Cucumis melo. Viruses. Vol.11, No.1.

Zhan, B., Zhao, W., Li, S., Yang, X., and Zhou, X. (2018) Functional scanning of apple geminivirus proteins as symptom determinants and suppressors of posttranscriptional gene silencing. Viruses, Vol.10, No.9.

Ziska, L. H., Bradley, B. A., Wallace, R. D., Bargeron, C .T., LaForest, J. H, Choudhury, R.A., Garrett, K.A., and Vega, F. E,(2018), Climate change, carbon dioxide, and pest biology, managing the future: coffee as a case study. Agronomy. Vol. 8, No.8.

Zou, Z., Liu, F., Chen, C., and Fernando, W.G.(2019), Effect of Elevated CO2 Concentration on the Disease Severity of Compatible and Incompatible Interactions of Brassica napus–Leptosphaeria maculans Pathosystem. Plants.Vol. 8,No.11.