Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IDENTIFICATION AND CLASSIFICATION OF DISEASES IN BASIL AND MINT PLANTS USING PSORBFNN

V. SATHIYA[1], DR. M.S. JOSEPHINE[2], DR.V.JEYABALARAJA[3]

[1]Research scholar, Department of Computer Science and Engineering,
Dr. M.G.R. Educational and research institute (deemed to be university), Chennai, Tamilnadu, India

[2]Professor, Dr. M.G.R. Educational and Research institute (deemed to be university), Chennai, Tamilnadu, India

[3]Professor, Velammal Engineering College, Chennai, Tamilnadu, India

[1]deviviji2000@yahoo.co.in,[2]josejbr.07@gmail.com,[3]jeyabalaraja@gmail.com

ABSTRACT

The world is a place for many living things apart from human beings, like plants, trees, animals, bacteria, insects, mammals, reptiles, etc., in which plants and trees are the vital source for oxygen for us humans. Every organism has its own life-time where it could live prolong for several hundred/ thousand/ billion years like a 'Bristlecone pine-tree' or algae like 'Proterocladusantiquus'. During their lifespan each plants and trees also experience sickness or illness due to diseases which results in withering away, fruits and leaves drop. Studying these factors affecting the plants, plant's lifespan, solutions and classifications of plants and diseases are known as "Plant Pathology". Though the Plant Pathology could be achieved through machine languages and automated machine approaches, the involvement of humans for classifying and categorizing diseases have been the only approach till-date. It is costlier, time consuming and labour intensive. Hence the proposed research aims at developing an algorithm that could automatically identify, classify and categorize the plant inputs through RBFNN (radial basis function neural network) with image segmentation through weighing function, where the optimization is done through PSO aiming at efficient and higher accuracy rate based rapid results. The developed RGM (region growing algorithm) increases network efficiency for speed and clustering the common attribute-based seeds towards extraction process of plant's feature. By focusing on fungal diseases classification with factors like, leaf spots, leaf curl, late blight, common rust, early blight and cedar-apple rust the study was carried-out. The developed algorithm and outcome through test and train method shows efficiency, accuracy in classifying and categorizing the plant diseases.

Keywords: *K-Means, Particle Swarm Optimization, Radial Basis Function, Neural Networks, Region Growing Algorithm*

1. INTRODUCTION

The role of computer and computing in the current technology-based era have not only impacted the communication-based industries but also many industries like banking, healthcare, food and tourism etc. Development of applications and computing software for identification and classification of information had been booming lately in the medical and science-based field. Information in science and researches could be an image, a text, an audio file, etc to evaluate and examine the root-cause of a solution seeking valid truth, facts and reliable outcomes. Image segmentation as a process of computing has been adopted majorly by researchers in the scientific research towards analysing the images into varied several parts to focus upon a particular region of the input seeking more clear results. The process of image partitioning is generally known as recognizing/ identifying a particular area of the selected sample for better categorization and classification. The process of image partition is done as segmentations by grouping similar color, shape, texture and pattern, etc of the object under computing. This is a pre-processing phase in a research, prior processing the image as a whole via *traditional* or *soft-computing* techniques.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In the traditional methods of image segmentation region-based, thresholding, edge-based and clustering-based techniques are utilized and whereas in the soft-computing (SC), neural networks (NN), genetic algorithm (GA) and fuzzy logics (FL) are the mostly adopted and popular techniques. Nowadays in the soft-computing the 'uncertainties' of the sample/ objects utilized could be dealt with easily and thus for image segmentation and other uncertainty-based studies, many researchers tend to adopt 'soft-computing' techniques since it could be modified according to the researcher's necessity and demands [1]. In SC, two methods are combined to form a unified technique to perform or evaluate the object according to the algorithm developed. Generally, the FL, GA and NN are combined with either BFO or PSO (Particle Swarm Optimization) based swarm intelligence (SI) methods in plant pathologies. Human interventions are deduced by implementing the automatic approach upon the samples with written algorithm in SC.

Similar to human beings the plants and trees also gets affected by diseases and die due to lack of disease classification for curing the root cause. Hence researchers tend to detect plant diseases at their early stages under 'Plant pathology', a study upon plant's growth, diseases, monitoring plant diseases and its growth, identifying and classifying the diseases towards seeking solutions, manually. This study aims at similar SC concept where the methods could be adopted and implemented in the plant pathology to deducing human intervention. Though the traditional plant pathology processes include human intervention and their naked eye-based evaluations, it has been witnessed and argued [2] that its time consuming, costlier and also labour intensive that delays the work far from deadlines/ due dates resulting in over-budget. By adopting SC humans could implement their thinking processes as algorithms into the machine language where the system would be able to identify and classify the ailment of plants at early stages. Through this technique, the time could be minimized through focusing upon "image segmentation" where the automated approach could widely focus on a sample rapidly and move onto the next sample which reduces the time consumption and human intervention.

The research developed focuses upon a RBFNN based SC approach where the identifying the ailments in plants and classifying the leaves accordingly is done through test-and-train method. Through PSO algorithm, K-means clustering and RGM identification techniques the study is trained and tested for optimal outcome. The PSO algorithm in RBFNN is assigned with optimal weights, towards identifying optimal regions in the ailed plant-leaves. RBFNN in this research is a linear-function that could cluster the ailed plant-leaves based upon the segmented images. The RBFNN is additionally enhanced towards RGM approach where seed points of the samples would be clustered together by focusing on the factors like features, attributes and patterns. The PSO as an optimization process in the research field had been gaining prominence towards its ease-of-application under "unsupervised environment" where it is referred and identified as a meta-heuristic paradigm of optimization [3].

Simultaneously the PSO algorithm as the optimization technique is very efficient towards weight optimization in this study. The outcomes of the traditional technique along with the developed technique would be compared for precision, accuracy and recall rate, by utilizing the statistical and mathematical calculations.

1.2. Problem statement

Herbal plants (also known as medicinal plants) are identified and found majorly in India. More-than 8000 species are found with rich history of healing power to heal several diseases in a traditional way. Among the most commonly used medicinal/herbal plants in the world, Chamomile (MatricariaChamomilla), Calendula (Calendula officinalis), Plantain (Plantago major), Peppermint (Metha x piperita), Holy Basil (Ocimumtenuiflorum) (a.k.aTulsi), Basil (Ocimumbascilicum) are the top-most herbs (Khan et al., 2015). Mint and Basil are the most identified and widely used herbal plants in major "Ayurvedic medicines" (Marimuthu et al., 2018). Hence preventing these herbal plants from diseases has been a primary concern of agriculturalists. Several techniques are being adapted in order to monitor and maintain the health of plants both by quantitative and qualitative approaches. However, plant diseases among the herbal plants for a long time have been manually identified, studied, classified and tested. The aim of the research is to eradicate the manual intervention, time and labour involved in identifying and classifying the herbal plant diseases, especially in the Basil and Mint leaves through an automated approach. The research aims at developing a machine learning model to identify fungal diseases namely, leaf curl, common rust, early blight, leaf spots, late blight and cedar-apple rust in basil and mint leaves.

1.3. Research Question

1. How can the process of identification and classification of diseases in mint and basil leaves be automated through deep learning?

2. LITERATURE REVIEW

The literature upon classifying and categorizing the plant diseases has been not attempted by many researchers successfully, since its complicated, time consuming, costlier, human labour intensive and lack of information towards classifying plant diseases [4]. Howsoever the existing researches and studies upon the successive plant pathology by adopting the SC along with traditional techniques that gained recognition had been achieved in the recent years (refer with:Table 1). The SC methods to identify/ classify the plant diseases had been adopted by handful researchers for varied purposes. In this research the focus of plant-leaves is upon the basil and mint and identifying the diseases through automated machine language instead of human intervention, especially through the PSO algorithm through RGM method in RBFNN.

2.1 Particle swarm optimization

PSO as a highly efficient technique has been identified and adopted by the researcher in the early 20s. The studies by Rini[5] and Esmin and Matwin[6] had insisted that PSO algorithm in the genetic classification could be utilized to identify diseases in plants and also the level of mutation through advanced hybrid optimization algorithm. Though the PSO is the recently advised and adopted algorithm in plant-pathology it is not a new trend and thus could be identified as a practice of modified algorithm by researchers to identify varied classes of plant and diseases. Similarly, the authors Zhang[7],El-Shorbagy and Hassanien,[8], Aje and Josephat [9], had analysed the application and uses of PSO in varied fields and the theorization of the PSO. According to the studies it could be found that, PSO as hybrid technique could be effective and efficient in attaining the goals rather than single-technique. Authors Xu [10] and Yan [11] studied how the PSO is utilized for complex and large problems in research where it uses the search-space of a binary or real variable. Application of PSO could be majorly identified in constrained optimizations, Minimum/Maximum problems and Multi-objective optimizations, especially in researches that seek complex procedures and outcomes.

2.2 Machine language and soft computing in plant pathology

In 2017, authors Zhang [12], Lu [13], Akram[14], Singh[15], and Yuan [16] adopted the Soft Computing techniques, majorly the K-means for identifying plant diseases. Similarly, in 2016 the authors Padol[17], Barbedo [18], Pujari[19], Kaur [20]adopted the SC techniques where SVM was most popular towards plant pathology-based researches.

In the year 2015 the authors Manikar[21] and in 2014 authors Patil [22],Waidyarathne [23] and Zhou [24] along with Phadikar [25] also utilized the SC techniques especially the SVM based and image processing techniques to identify plant-diseases. Though earlier the studies utilizing SC techniques were lesser, there were recognisable studies in 2010 authors Rumpf [26] and Liu [27] made use of SVM and neural network methods to evaluate and identify the plant diseases. Though there are no studies towards automated plant diseases identification and thus the proposed study would offer huge contribution.

Authors De-Luna [28] utilized the ANN model to identify the diseases in the tomato plants through inspection V3 model through size classification. Similarly, the authors Maktedar and Keskar[29] studied about eight varied herbal plant-based leaves by adopting the SONN and DNN techniques with image-edge detection processes. Authors Mangayarkarasi and Venkataraman [30] utilized PNN model through traditional human intervention analysis on herbal plant-based leaves in identifying medicinal values with PNN technique unlike plant diseases identification. Though the approaches and the techniques vary in each study the aim of the research was to examine the plants for classification and identification using machine learning. However, the researches considered ML as an assisting tool instead on independent tool that could be utilized to examine, identify and classify on its own through structured algorithm.

2.3. Diseases in Basil and Mint

Authors McGrath [31], Kalamartzis [32],Topolovec-Pintaric and Martinko[33] studied about the diseases in Basil and the authors Johnson and Santo [34], El-Mougy[35], Dung [36], studied about the diseases in the Mint. According to them.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Literature Review on Mint and Basil Plant Disease Identification

S. No.	Author	Year	Technique	Implementation
1.	El-Mougy et al.,	2001	Phaseolus Vulgaris L. (root rot and wilt in Mint, Rose, Lemon and Geranium)	Researchers implemented tests on Mint, lemon, rose and Geranium through MSTAT-C and Duncan's Multiple-Range towards diseases identification
2.	Johnson and Santo	2007	Verticillium dahliaeand Pratylenchus penetrans in Mint (wilt development)	Authors conducted study upon mint with two isolations: VCG 2B and VCG 4A
	McGrath	2019	Downy Mildew presented in Basil	Study was conducted upon Basil plants in Florida, USA through manual observation
3.	Kalamartzis et al.,	2020	Cultivator of Basil and Wilt	Authors focused Basil plants in Mediterranean areas through LSD technique
4.	Topolovec-Pintaric and Martinko	2020	Basil - Downy Mildew	The authors analysed the *Peronospora belbahrii*in basil plants in Europe through optimization process and PCR technique
5.	Dung	2020	Verticillium dahlia in Mint plants in US (wilt development)	Author studied the mint plants through double-root technique towards diseases identification

the diseases through fungal and bacterial infections on plants, i.e., infected leaves (wilting, curling, spotting, etc), infected stems and roots could be observed and classified under human intervention. The leaves show symptoms like, leaf spots (brown spots of septoria), rusting (stripe-rusts), Chlorosis (yellowish leaves), curling, powdery mildew, tar spots, blisters, etc.

Once the symptom is identified and the leaf is classified it would be easier for the researcher to study the plant diseases and introduce/ develop a treatment procedure to cure the diseases. To carry these processes of identification and classification, researchers have been indulging in manpower and human labour that costs time, budget and patience which ultimately extends each work and researcher to indulge with single task for longer duration. Hence to minimize the gap of non-existence of automation in plant diseases identification and classification the current research aims to adopt the PSO algorithm with RGM method to identify and classify the plant diseases in mint and basil where human intervention is unnecessary.

Through reviews and inference of Table 1, it is very clear that none of the authors adopted hybrid algorithm-based studies upon plant diseases identification, especially towards mint and basil. Thus, RBFNN with PSO algorithm, the research would be the first of its nature and outcomes would be compared for better comparison against the traditional methods.

2.4. Literature contribution

From Table 1 it can be understood that, the existing studies use either human intervention or analyse the plant diseases through basic machine learning techniques by focusing on only one herbal plant. To fill-in the gap in the current research, the researcher has developed the model to identify and classify the disease in two or more herbal plants through deep learning approach. The research has made the study unique by employing combining the optimization algorithm (PSO) and RBFNN model in order to identify and classify mint and basil leaves and detect the diseases that occur in those leaves. The research adapts a two stage approach and in the first stage the leaves are classified through labelling as Mint and Basil respectively. In the second stage, three individual classification for basil leaves has been made as healthy, mildew and wilted and four individual classification for Mint leaves has been made as healthy, rusty, powdery and wilted. Thus, this study has proposed a solution to classify diseases of two herbal plants through an automated approach.

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. PROPOSED METHOD

The research is focused upon two plant species Ocimumbasilicum (Basil) and Menthaceae of Lamiaceae family (Mint). The proposed technique basically focuses on the 'image' of the plant leaves as inputs and it segments the picture for closer look to classify the diseases and categorise each leaves of similar diseases as one group for analysing the plant diseases. By utilizing the PSO in this research the RBFNN is constructed, trained and tested (refer with: Fig. 1) to attain reliable and valid outcomes towards attaining higher accuracy, precision and recall rate along with f-1 score.

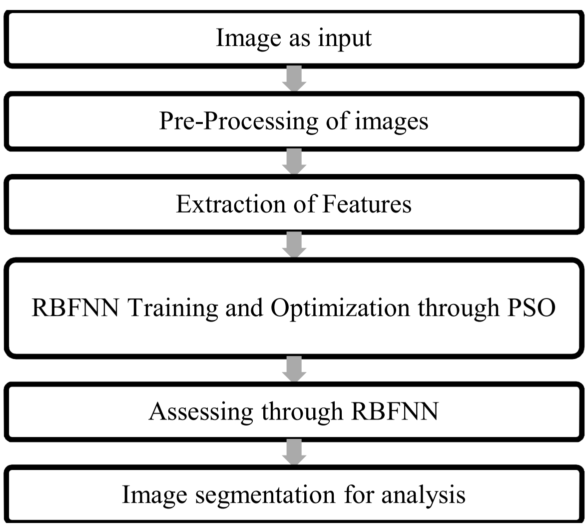


Figure. 1Proposed Work – Methodology

The Mint and Basil leaves are utilized here as samples for image segmentation and plant diseases' classification. The plant diseases are classified as three classes in Basil, namely: healthy, wilted and mild-dew basil leaves and the Mint leaves are classified as four classes namely: healthy, wilted, rusty and powdery. For the segmentation in image pre-process a combined algorithm-based evaluation technique is developed to classify the leaves and categorize them for analysis.

3.1 Acquisition of image and processing methods

The image acquired is added with a 'white background' for analysis since the image segmentation will be accurate with visible areas of the images of leaves that has variations in color ranges from green to red variations. The datasets from popularity-based domains have been acquired with .jpg, .jpeg, .gif, .bmp formats for the processing and similarly the datasets for real-time analysis had been captured with camera and inserted as input directly.

In this research the focus is upon the Basil leaves and Mint leaves, totalling 1628 original leaves (refer with: Table 2) has been used

Table 2. Leaf-Disease dataset for Image Classification

Class	No. of Original Images
1. Healthy Basil	550
2. Unhealthy Basil (Mildew)	245
3. Unhealthy Basil (Wilted)	100
4. Healthy Mint Leaf	148
5. Fusarium Wilt Mint Leaf	220
6. Mint Leaf Rust	115
7. Powdery Mildew Mint Leaf	250
Total	1628

The samples gathered and pre-processed for the analysis in the study would be utilized for image segmentation and also for the classification of the leaf classes of the plant under the disease categorization.

3.2 Proposed Algorithm and System Flow

Steps of Proposed Method: The algorithm for the "plant-leaf disease identification" based computation in this research is done through the following six steps

1. The class of every image is evenly categorized and distributed into folders prior processing
2. The gained pre-processed images are converted into Grey-Scale images
3. To construct the pre-processed images, the inputs (images) are passed via Region Growing technique (RGM)
4. The obtained Grey-Scale images will be flattened and then passed through RBF Neural-Network (RBFNN)
5. An optimization algorithm named 'PSO' is utilized to train the RBFNN
6. Finally, once the network is trained, Accuracy, Precision and Recall calculations are computed upon test-datasets.

Based on the algorithm developed the work is carried-out through test-and-train method where basically around 3-5 sample datasets are initially trained with neural network algorithm and the outcome is tested out for reliability with obtained outcome of previous research for same dataset. Once the outcomes are valid and reliable and the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

NN is trained with datasets, the complete dataset from databank would be tested for outcomes and the results would be compared for accuracy and precision/ reliability. To explain the processes a flowchart of the proposed research is necessary.

System Flow Diagram**:** The flow of the developed research is depicted and developed through flowchart representation through which the training and testing phases are pre-explained prior detailed methods and examination upon each process is done in the research. The system flow (refer with : Fig 2) is as follows.

3.3 Adopted Methods and Data Evaluation Techniques

The data analysis techniques in this research mainly focus upon four methods: RGM, K-means clustering, RBFNN and PSO. The developed research focuses on classifying and identifying the plant diseases by utilizing the intelligence of the computational approaches. In this approach the RBFNN (*Radial Basis Function Neural Network*) technique is adopted which is eventually trained through PSO (*Particle Swarm Optimization*) towards identifying the affected or damaged area/ region of plant leaves (mint and basil). Basically, RBFNN works as a linear function that is of unique competence approach which either decreases or increases monotonically, distancing from centre-point that handles complexity of the damaged/ affected region.

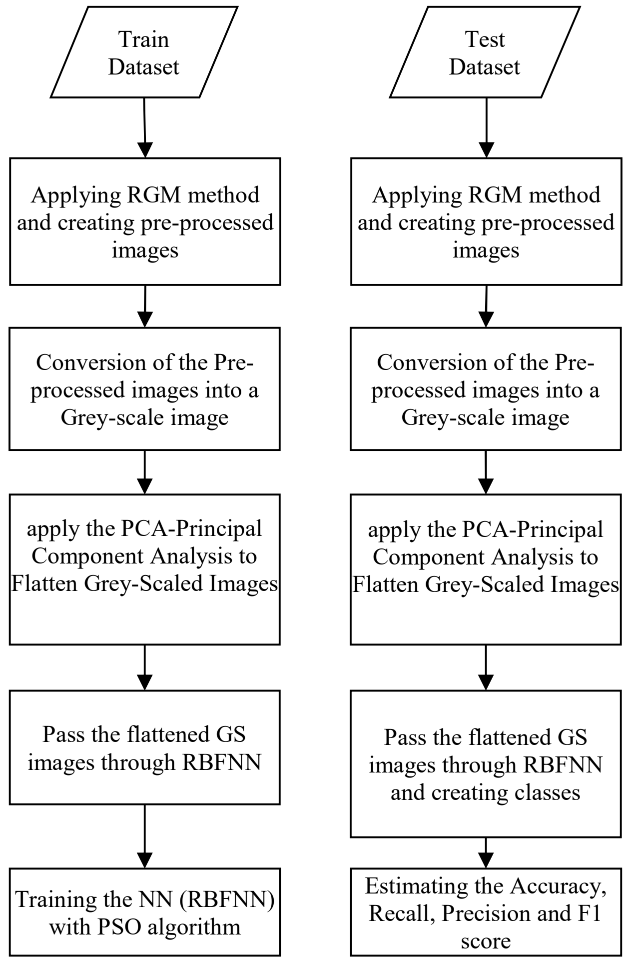


Figure. 2 System flow-diagram

PSO is a meta-heuristic optimization process adopted globally for its adaption and ease-of-application. The PSO algorithm basically occupies particles as swarms that attempts to cross a multidimensional search-space towards gaining optima, where every particle is considered as a valuable solution and it's prejudiced by its neighbours' experience and also its' own pathway. Henceforth adopting the PSO and RBFNN in this research would provide the estimated outcome. The classification of the plant disease in this study follows the approaches below

a) **RGM (REGION GROWING METHOD):**The RGM is an image segmenting process (aka pixel-based segmenting process) is used to segment the regions of an image for better visual investigation, given that it entails the initial-seed point selection criterion. Here the neighbouring initial-seed pixels are examined through image segmentation, towards determining whether neighbour pixels are to be appended along with the focused region. Images from pre-process pipeline is shown in Fig.3.

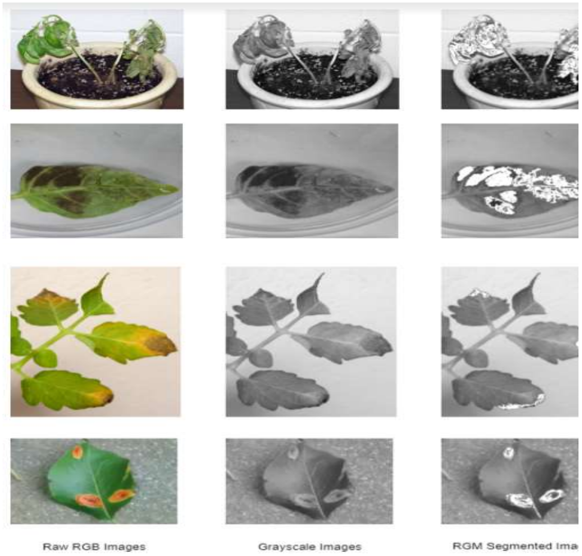


Figure. 3 Sample images from image to pre-processing pipeline

This technique is similar with the data clustering processes where the iterations are carried out through the following formulae: (a)

- ***a)*** $\bigcup_{i=1}^{n} A_i = A$
- ***b)*** A_i represents connected region where $i = 1, 2, ..., n$
- ***c)*** $A_i \cap A_j = \emptyset, i \neq j$
- ***d)*** $P(A_i) = TRUE$ for $i = 1, 2, \ldots, n$
- ***e)*** $P(A_i \cup A_j) = FALSE$ for random adjacent – region of set A_i and A_j

  *-where **P(Aᵢ)** represents the logical predicates the points of the set A_i and **Ø** defines the null-set*

Fig.4. shows the dimensions were used in PCA for visualization purposes it is not the same as the number of dimensions used in the actual pipeline.

Figure. 4PCA Visualization of a sample of Basil Leaves data in two dimensions

b) **K-MEANS CLUSTERING:** The clustering through K-means is normally defined as an iterative process since it partitions the acquired dataset as non-overlapping subgroups of *Kpre-defined distinct* (clusters) in which every single data-point belongs to individual groups and only clustering with one single-group. Through clustering the K-means generally clusters the similar intra-cluster data-points and contrarily differentiates each cluster with distancing from one-another. The process of minimization in the intra-clusters is done through two varied metrics of distances, *Euclidian* and *Manhatten*. The research applies and adopts the Euclidian distance method as follows:

$$E = \sum_{i=1}^{a} \sum_{b=1}^{B} x_{ib} \parallel y^i - \mu b \parallel^2 \quad (1)$$

c) **PSO (Particle Swarm Optimization)**: Though in general plant disease-based identification and classification, the researchers adopt optimization algorithms it differs as per utilization and necessity. In this research a variant of PSO approach is adopted where it works with *swarm* as population of *particles* known as candidate solutions. These particles work with a pre-defined formula where they freely move around within search-space through guidance of their own position (i.e., best-known positions) along with other swarms in total, until a new best-position is found/ discovered through improvisation of positions. This process of identifying the best position and guiding the swarms is repeated until the satisfying outcome is gained, however this technique has limits and disadvantage where it is not guaranteed to provide the researchers with expected outcomes every-time. The algorithm for the PSO is as follows:

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

GENERAL ASSUMPTIONS AND REPRESENTATION OF THE VARIABLES

Assume that **N** represents the number of particles that are present in a particular swarm. The position of each particle is represented by $\mathbf{p_t}$ in the search space $\mathbf{X^{m.}}$

In the same the search space $\mathbf{X^m}$, the velocity of each particle is represented by $\mathbf{Vel_t}$

Assume that the particle is denoted by **t,** with$\mathbf{b_t}$ representing the particle's best possible position and $\mathbf{s_t}$ representing the entire swam's best possible position.

D) RBFNN (RADIAL-BASIS FUNCTION NEURAL-NETWORK): The RBFNN is a unique machine language-based algorithm which is considered as an effective, intuitive and extremely rapid approach. In the RBFNN based approaches the algorithms are developed based upon the layers utilized for classification and identification processes. To solve issues and problems that are critical to examine, the RBFNN adopts 3-layered network towards both classifying and examining the outcomes, especially with regression-based problems. Clusters in RBFNN are formed with K-means approach of clustering algorithm where different approaches could also be utilized based upon the necessity and complexity of the problems and processes. Generally, the RBFNN models out the process in a "data plane" form b adopting the circular shapes for representation, in 2D.

Accordingly, by considering the radii and centroids of clusters, the data point of the cluster could be assumed. However, it is also possible that none of the data-points of clusters could be identified if the centroid radii are far from each other. Apparently, this could only lead a researcher towards ambiguity of data-points of clusters and their class. Henceforth to identify and argue this theoretical gap RBF had been introduced in the NN which is the significant element in the RBFNN. Implementing the RBF would enable a researcher to be attentive towards the closeness rate in-between the cluster data-points and centroids irrespective of the distances found. RBF generally utilizes circular shapes of smooth transitioning models instead of circles that are of sharp cut-off in shapes.

Similarly, the RBF also provides information on prediction's confidence rate that cannot be offered by the K-means clustering.

– **RBF Function:** formula utilized for calculating RBF

$$e^{-\beta * R^2} \quad (2)$$

a) BETA THROUGH SD OF CLUSTER:

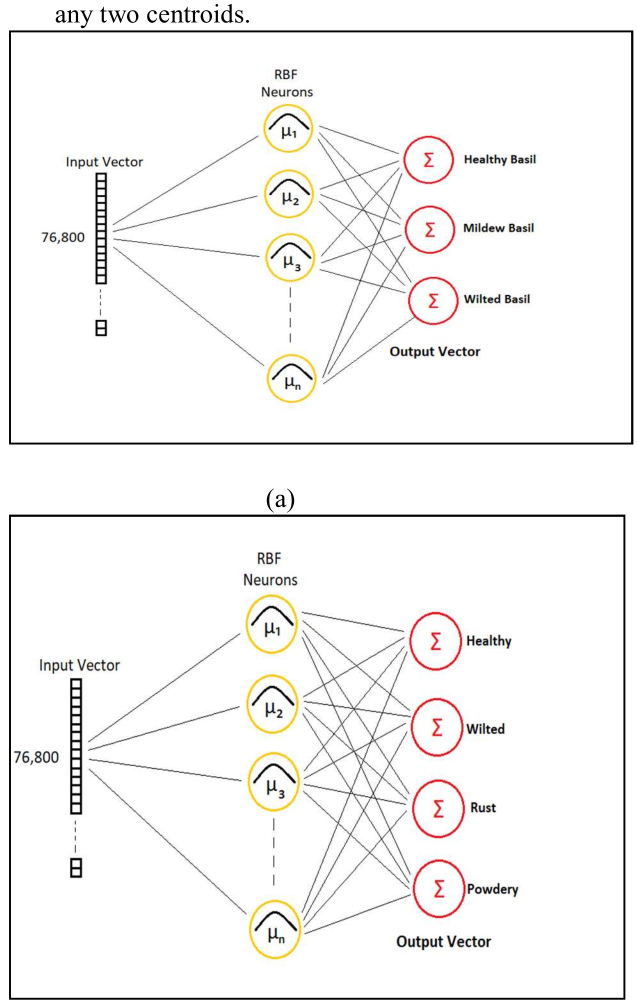
$$\beta = \frac{1}{std^2} \quad (3)$$

-where,the ***R*** defines the *distances* between the centroid and data-point of cluster;

b) BETA THROUGH CENTROID ESTIMATION:

$$\beta = \frac{\sqrt{2*c}}{Rmax} \quad (4)$$

-where,the ***c*** defines the centroids in total and ***Rmax***defines the *maximum-distance* between any two centroids.



(b)

Figure. 5 Neural Network Diagram: (a) Basil and (b) Mint

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The neural network diagram (refer with: Fig. 5.) for proposed research is upon two plants, ***(a) Basil*** (with three leaf classes such as: Healthy Basil, Mildew: Unhealthy Basil and Wilted: Unhealthy Basil) and ***(b) Mint*** (with four leaf classes as: Healthy Mint, Fusarium: Wilted Mint, Rusty Mint and Mildew: Powdery Mint) (refer with: Fig. 5a and 5b). Based on the developed neural network, the data is tested and trained along with the evaluation methods and algorithm.

4. RESULTS

The results of the research-based plants: Basil and Mint as inputs where the Basil is classified as, healthy, mildew and wilted; Mint is classified as healthy, mildew, rusty and powdery. The following are the images passed through the machine learning as inputs along with the grey-tone and diseases identified outputs for Basil and Mint leaves:

Basil: The first image is the original input where the latter two images of the plant classifications are the pre-processed and identified damaged/ diseases affected areas in plants.

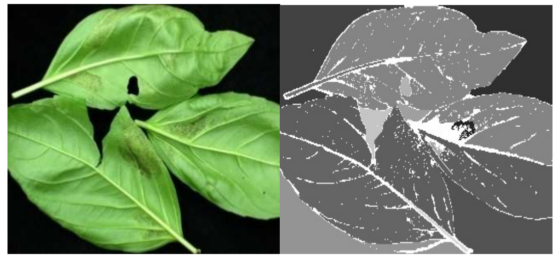
- HEALTHY BASIL

(a) (b)

(c)

Figure. 6 Healthy Basil leaves: (a) test (b) test pre-processed (c) test pre-processed intermediate

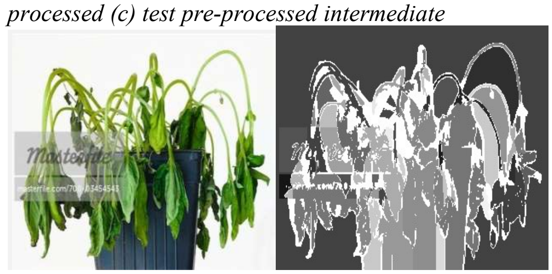
- UNHEALTHY BASIL



(a) (b)

(c)

Figure. 7Mildew Basil leaves:(a) test (b) test pre-processed (c) test pre-processed intermediate



(a) (b)

(c)

Figure.8Wilted Basil leaves: (a) test (b) test pre-processed and (c) test pre-processed intermediate

The Fig 6(a) is the healthy Basil for plant diseases identification as input; the Fig.6(b) is the grey-tone image of the input de-noised and smoothened for pre-processing intermediate test phase (refer Fig 6c). Similarly, the Fig. 7(a) and 8(a) are the Unhealthy (Mildew & Wilted) Basil for plant diseases inputs in test phase and the Fig. 7(b) and 8(b) are the grey-tone image of the inputs

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

as pre-processed. Similarly, Fig. 7(c) and 8(c) are the test pre-processing intermediate images retrieved after de-noised and smoothened.

Mint: The first images of each mint classification below represent the original image of mint as an input, where, the latter two images of the classifications are the pre-processed and processed image of diseases affected areas in plants.

- HEALTHY MINT

(a) (b)

(c)

Figure.9 Healthy Mint leaves: (a) test (b) test pre-processed and (c) test pre-processed intermediate

The Fig. 9(a) is the healthy Mint leaves for plant diseases identification as input; the Fig. 9(b) is the grey-tone image and the Fig. 9(c) is the pre-processed intermediate image that has been de-noised and smoothened for final test phase.

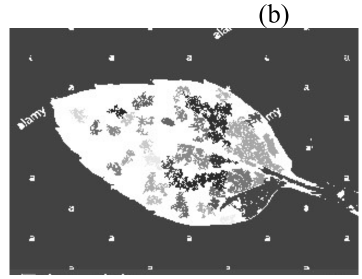
- UNHEALTHY MINT

(a) (b)

(c)

Figure.10 Wilted Mint leaves: (a) test (b) test pre-processed and (c) test pre-processed intermediate

(a) (b)



(c)

Figure11. Rusty Mint leaves:(a) test (b) test pre-processed and (c) test pre-processed intermediate

The Fig. 10(a) and 11(a) are the Unhealthy (Wilted, Rusty) Mint leaf images are the inputs of plant diseases identification process. The Fig.10(b) and 11(b) are the respective grey-tone images and the Fig. 10(c) and 11(c) are the test pre-processing intermediate images retrieved after smoothening and de-noising.

INTERPRETATION: The Figures above represent the RBFNN based outcomes based upon the rates of accuracy, precision, recall and F-1 score. Through RGM for seeking seed-points and with RBFNN network the algorithm had been applied for processing in the evaluation technique for image processing. The first image of each class represents the original image of Mint and Basil chosen for the research towards image processing, the second image portrays the grey-scale conversion of the original image introduced. Once the images are converted into greyscales, masking is done upon the converted images for identifying the diseases and the final outcome achieved is the "plant diseases identified" region of the leaves Basil and Mint plants. Through the findings the outcomes are obtained and the higher accuracy rate, precision rate, recall rate along with f-1 score have been gained.

5. PERFORMANCE EVALUATION

From the tested and trained sets of data in this research, the plants (Basil and Mint) with the diseases are classified under 'classes' according to their patterns and diseases identified. The

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

outcomes of the datasets prior applying the developed algorithm is tested against the gained results to authenticate the developed algorithm and its efficiency under “validation technique” (performance evaluation by Confusion matrix) and compared for its performance to validate the reliability, robustness and validity.

The Table 3 represents the accuracy, recall, precision and F-1 scores of the tested-and-trained sample images of Basil and Mint under confusion matrix method.

From the evaluation metrics table, it is clearly inferred that the 7th K value (metric) has high scores in all four estimations than other 7 metrics.

To evaluate these metrics, the study makes utilization of the following key metric estimations:

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \qquad (5)$$

$$Recall = \frac{TP}{TP+FN} \qquad (6)$$

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \qquad (7)$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \qquad (8)$$

*where **TP** is True Positive, **TN** is True Negative, **FN** is False Negative and **FP** is False Positive.*

Table 3. Evaluation Metrics Table

K Values / Metrics	2	3	4	5	6	7	8	9
Accuracy	0.45	0.51	0.57	0.54	0.54	**0.59**	0.59	0.54
Precision	0.30	0.57	0.58	0.55	0.54	**0.63**	0.58	0.59
Recall	0.43	0.49	0.56	0.53	0.54	**0.58**	0.58	0.52
F1 Score	0.39	0.49	0.55	0.51	0.53	**0.58**	0.57	0.51

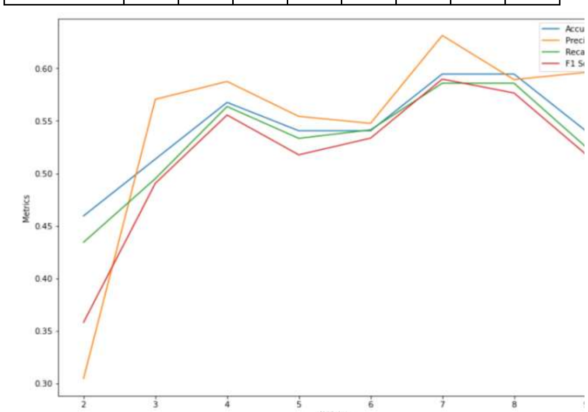


Figure. 12 Evaluation metrics: plot-1

The Fig.12 represents the evaluation metrics upon: Accuracy, Recall, F-1 and Precision scores where the K-value of 7 is higher in all 4-evaluation metrics reaching the peak point denoting that the optimal clusters are at K-7.

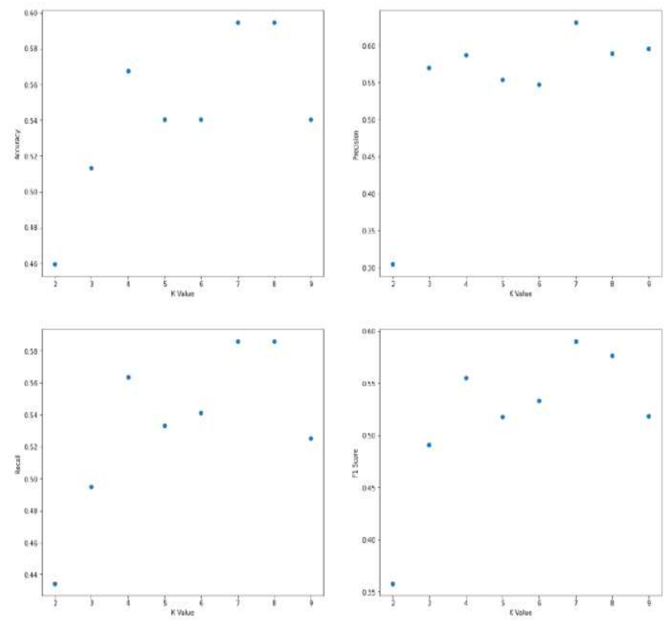


Figure. 13 Evaluation metrics: plot-2

The above Fig.13 represents the plot-2 evaluation metrics’ individual plotting where the clusters are accumulated at K7 denoting that the metric value is applicable and acceptable in this research for successive outcomes. The Figure shown below (refer with: Fig.14) represents the visualization of true labels versus predictions made by RBF for basil leaves.

Journal of Theoretical and Applied Information Technology
15th November 2022. Vol.100. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

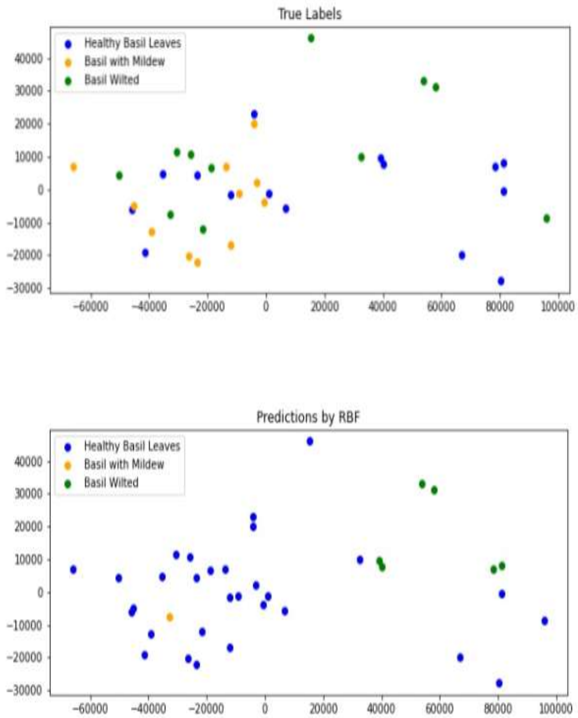


Figure. 14 Visualisation of Predictions vs True Labels for a sample of Basil Leaves Images

The existing plant diseases based RBFNN classifiers in the plant pathology either utilizes the human intervention for classification [37] and low rate in accuracy, precision, F1 and recall scores. Hence in RBFNN model with PSO as primary estimation optimizer was adopted and the algorithm was found to be effective with average scores, ***accuracy (0.59), precision (0.63), recall (0.58) and F1 (0.58)*** respectively. The algorithm was trained and tested with datasets until the expected outcome for Basil attains and Mint attains stability and reliability with higher k-value.

Finally, through this research, the aim of attaining higher accuracy, precision and recall rate for Mint and Basil along with f-1 score has been achieved with the ***7th Cluster*** through testing**,** (refer with: Table 3)

Thus, the developed RBFNN technique in automated plant diseases identification along with K-means, PSO and RGM approaches is a success, with limitations of two plant categorizations.

6. CONCLUSION AND FUTURE WORK

The study mainly focused on soft computations in plant pathology that adopts the hybrid networking functions (RBFNN) and swarming algorithm (PSO) towards identifying the plant diseases in Mint and Basil. The classification of plants is done through K-means clustering and RGM image segmentation process where, pre-processing, grey-toning, pre-processing intermediate and processing the data acquired and datasets accessed through popular databank sites. Mint and Basil are the focused plants and hence the algorithm was initially trained with small samples and later the whole datasets are tested for compatibility with existing results of the datasets in databank. The healthy images are initially loaded as inputs and the test is carried out through clipping, smoothening and de-noising the images for quality-based outcomes.

The results obtained were accurate where the similarity rate of Basil and Mint was attained at the evaluation and outcome analysis of k-values, where the 7thmetric k-value shows a good fit with: ***accuracy (0.59), precision (0.63), recall (0.58) and F1 (0.58)***. Thus, the developed algorithm is effective, rapid, reduces load time and diminishes human intervention in identification and classification of plant leaves diseases.

The study thus contributes as the first attempt in basil and mint towards automated machine language based RGM identification process where the researchers would be able to do classification and identification of diseases rapidly without "manual" labour and observation throughout. In suture the same research could be conducted upon varied plants and also to achieve higher accuracy, precision, recall and f1 score with different optimization techniques. Thus, the research will contribute the future investigators and the developers with base values for reference and comparison.

6.1. Contribution

The current research has developed a machine learning model with RBFNN layers and algorithms (RGM, PSO and K-means clustering) in order to identify and classify diseases in two types of herbal plants namely Basil and Mint.

The study thus contributes as the first attempt in basil and mint towards automated machine language based RGM identification process where the researchers would be able to do classification and identification of diseases rapidly without "manual" labour and observation throughout. In suture the same research could be conducted upon varied plants and also to achieve higher accuracy, precision, recall and f1 score with different optimization techniques. Thus, the research will contribute the future investigators and the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

developers with base values for reference and comparison

6.2. Limitations

The research is limited to herbal plants and thus other crops and plants have not been taken into consideration. The number of classifications made on mint leaves is restricted to three and basil leaves is restricted to four in this research. In future the researcher aims at improving the model to achieve higher accuracy and precision

REFERENCES:

[1] Akhtar. A, Khanum. A, Khan. S.A and Shaukat. A, "Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques", in :*11th International Conference on Frontiers of Information Technology*, Islamabad, pp. 60-65, 2013.

[2] Mohanty. S.P, Hughes. D.P and Salathe. M, "Using Deep Learning for Image-Based Plant Disease Detection", *Front. Plant Science,7(1419).* Retrieved on 26thDecember 2020 from https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full, 2016.

[3] Sengupta.S, Basak. S and PetersII. R.A., "Particle Swarm Optimization: A Survey of Historical and Recent Developments with Hybridization Perspectives", *Mach. Learn. Knowl. Extr,**2019(1),*** pp: 157-191,2019.

[4] Saleem, M. H., Potgieter, J., & Mahmood Arif, K. "Plant Disease Detection and Classification by Deep Learning", *Plants (Basel, Switzerland)*, ***8(11),*** p. 468,2019.

[5] Rini. D.P, Shamsuddin. S.M and Yuhaniz. S.S, "Particle Swarm Optimization: Technique, System and Challenges", *International Journal of Computer Applications (0975 – 8887),* ***14(1),***pp: 19-27,2011.

[6] Esmin. A.A.A and Matwin. S, "HPSOM: A Hybrid Particle Swarm Optimization Algorithm with Genetic Mutation", *International Journal of Innovative Computing, Information and Control,* ***9(5),*** pp: 1919-1934, 2013.

[7] Zhang. Y, Wang. S and Ji. G, "A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications: Review Article", *Hindawi Publishing Corporation Mathematical Problems in Engineering Journal,* ***2015,***pp: 1-38,2015.

[8] El-Shorbagy. M.A and Hassanien. A.E, "Particle Swarm Optimization from Theory to Applications", *International Journal of Rough Sets and Data Analysis,* ***5(2),*** pp: 1-24,2018.

[9] Aje. O.F. and Josephat. A.A., "The particle swarm optimization (PSO) algorithm application – A review", Global Journal of Engineering and Technology Advances, 03(03), pp: 001-006, 2020.

[10] Xu. X, Rong. H, Trovati. M, Liptrott. M and Bessis. N, "*CS-PSO: Chaotic Particle Swarm Optimization Algorithm for Solving Combinatorial Optimization Problems".* Retrieved on 5th Feb. 2021 from https://core.ac.uk/download/pdf/227102534.pdf,2016.

[11] Yan. D, Lu. Y, Zhou. M, Chen. S, et al., "Recent Advances in Particle Swarm Optimization for Large Scale Problems", *Journal of Autonomous Intelligence,* ***1(1),***pp: 1-14,2018.

[12] Zhang. S, Wu. X, You. Z and Zhang. L, "Leaf image based cucumber disease recognition using sparse representation classification", *Comput. Electron. Agricult.*, ***134***, pp: 135-141,2017.

[13] Lu. Y, Yi. S, Zeng. S, Liu. Y and Zhang. Y, "Identification of rice diseases using deep convolutional neural networks", *Neuro-computing*, ***267***, pp: 378-384,2017.

[14] Akram. T, Naqvi. S.R, Haider. S.A and Kamran.M,"Towards real time crops surveillance for disease classification: Exploiting parallelism in computer vision", *Comput. Electr. Eng.*, ***59***, pp: 15-26, 2017.

[15] Singh.V and Misra.A.K, "Detection of plant leaf diseases using image segmentation and soft computing techniques", *Inf. Process. Agricult.*, ***4***, pp: 41-49,2017.

[16] Yuan. L, Bao. Z, Zhang. H, Zhang. Y and Liang. X, "Habitat monitoring to evaluate crop disease and pest distributions based on multi-source satellite remote sensing imagery", *Optik_Int. J. Light Electron Opt.*, ***145***, pp: 66-73,2017.

[17] Padol. P.B and Yadav. A.A, "SVM classifier based grape leaf disease detection," In *Proc. Conf. Adv. Signal Process. (CASP)*, pp: 175-179,2016.

[18] Barbedo J. G. A., Koenigkan L. V. and Santos. T.T, "Identifying multiple plant diseases using digital image processing", *Biosyst. Eng.*, ***147***, pp: 104-116, 2016.

[19] Pujari. D.J, Yakkundimath. R and Byadgi. A.S, "SVM and ANN based classification of plant diseases using feature reduction technique", *Int. J. Interact. Multimedia Artif. Intell.*, ***3(7),*** pp: 6-14,2016.
[20] Kaur. I, Aggarwal. G and Verma. A, "Detection and classification of disease affected region of plant leaves using image processing technique", *Indian J. Sci. Technol.*, ***9(48),*** pp: 1-13,2016.
[21] Mainkar. P.M, Ghorpade. S and Adawadkar. M, "Plant leaf disease detection and classification using image processing techniques", *Int. J. Innov. Emerg. Res. Eng.*, ***2(4),*** pp: 139-144,2015.
[22] Patil. S.S and Suhas. K.C, "Identification and classification of cotton leaf spot diseases using SVM classifier," *Int. J. Eng. Res. Technol.*, ***3(4),*** pp: 1511-1544,2014.
[23] Waidyarathne. K. P and Samarasinghe. S, "Artificial neural networks to identify naturally existing disease severity status", *Neural Comput. Appl.*, ***25(5),*** pp: 1031-1041,2014.
[24] Zhou. R, Kaneko. S, Tanaka. F, Kayamori. M and Shimizu. M, "Disease detection of Cercospora Leaf Spot in sugar beet by robust template matching", *Comput. Electron. Agricult.*, ***108***, pp: 58-70,2014.
[25] Phadikar. S, Sil. J and Das. A.K, "Rice diseases classification using feature selection and rule generation techniques", *Comput. Electron. Agricult.*, ***90,*** pp: 76-85,2013.
[26] Rumpf. T, Mahlein A.K, Steiner. U, Oerke. E.C, Dehne. H.W and Plümer. L, "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance", *Comput. Electron. Agricult.*, ***74(1),*** pp: 91-99,2010.
[27] Liu. Z.Y, Wu. H.F and Huang. J.F, "Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral relectance and principal components analysis", *Comput. Electron. Agricult.*, ***72(2)***, pp: 99-106,2010
[28] De-Luna, R. G., Baldovino, R. G., Cotoco, E. A., De-Ocampo, A. L. P., Valenzuela, I. C., Culaba, A. B., & Gokongwei, E. P. D., "Identification of Philippine herbal medicine plant leaf using artificial neural network", In: *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM),* pp. 1-8, 2018.
[29] Keskar M and Maktedar D, "Enhancing Classifier Accuracy in Ayurvedic Medicinal Plants Using WO-DNN", *International Journal of Engineering and Advanced Technology,**9(1),*** pp: 6705-6713,2019.
[30] Venkataraman D and Mangayarkarasi N "Computer Vision Based Feature Extraction of Leaves for Identification of Medicinal Values of Plants", In : IEEE International Conference on Computational Intelligence and Computing Research,2016.
[31] McGrath. M.T, "Expect And Prepare For Downy Mildew In Basil", pp: 1-7. Retrieved on 6th February 2021 from http://vegetablemdonline.ppath.cornell.edu/NewsArticles/Basil%20Downy%20Mildew-VegMD-McGrath-2019.pdf,2019.
[32] Kalamartzis. I, Dordas. C, Georgiou. P and Menexes. G, "The Use of Appropriate Cultivar of Basil (Ocimumbasilicum) Can Increase Water Use Efficiency under Water Stress", *AgronomyReview Journal,* ***10(70),***pp: 1-16,2020.
[33] Topolovec-Pintaric. S and Martinko. K, "*Downy Mildew of Basil: A new destructive disease worldwide*", Chapter: Plant Diseases-Current Threats and Management Trends, IntechOpen Publications, pp: 1-15,2020
[34] Johnson, D. A., and Santo, G. S. "Development of wilt in mint in response to infection by two pathotypes of *Verticillium dahliae*and co-infection by *Pratylenchuspenetrans",Plant Diseases*, ***85***, pp: 1189-1192,2001.
[35] El-Mougy. N.S, El-Gamal. N and Abdel-Kader. M.M, "Control Of Wilt And Root Rot Incidence In *Phaseolus Vulgaris* L. By Some Plant Volatile Compounds", *Journal Of Plant Protection Research,* ***47(3),***pp: sek1:255-sek1:265, 2007.
[36] Dung. J.K.S, "Verticillium Wilt of Mint in the United States of America", *Plants-A Review Journal,* ***9(1602),***pp: 1-17, 2020.
[37] Chouhan. S.S, Kaul. A, Singh. U.P and Jain. S, "Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology", *Digital Object Identifier, IEEE,* ***6,*** pp: 8852-8863,2018