CAHIERS MAGELLANES-NS
Volume 06 Issue 1
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

SUGARCANE DISEASE IDENTIFICATION USING CONVOLUTION NEURAL NETWORK

[1]Hardik Singh, [2]Harsh Bhatnagar, [3]Achint, [4]Yash Kansal, [5]Mohini Preetam Singh

Dept. Of Computer Science And Engineering
Meerut Institute Of Engineering And Technology, Meerut

[1]hardik.singh.cse.2020@miet.ac.in , [2]harsh.bhatnagar.cse.2020@miet.ac.in,
[3]achint.a.cse.2020@miet.ac.in, [4]yash.kansal.cse.2020@miet.ac.in , [5]mohini.singh.cse@miet.ac.in

Abstract:
In the contemporary era, the upsurge of plant ailments stands out as a significant factor impacting diminished agricultural yield and escalating losses among farmers. Thus, it becomes paramount to utilize a method capable of furnishing prompt and accurate outcomes. The recent proliferation of deep learning has emerged as pivotal in effectively addressing both traditional and unconventional hurdles. The Convolutional Neural Network (CNN) has surfaced as an advanced strategy for cutting-edge identification and detection. In confronting the challenge of plant diseases, we have curated an exhaustive dataset encompassing 37 distinct plant disease varieties corresponding to 5 seedling varieties (238, 223, 268, 218, 214) for the training and validation of our model. Our approach entails the utilization of Resnet (Residual Neural Network), a specialized CNN architecture. We capture images of afflicted plant foliage and leverage CNN-driven classification for disease identification. Our model showcases superior precision in contrast to various previously employed techniques for disease detection.

Keywords: Convolution Neural Network, Resnet, EfficientNet, Neural architecturesearch, Focal loss function.

1. Introduction:

In India, agriculture stands as the primary economic sector, sustaining roughly half of the nation's population. A wide array of crops, vegetables, fruits, and pulses are cultivated, collectively adding about 17% to 18% to the GDP. Despite its pivotal role, a considerable portion of crops succumbs to destruction annually due to diseases and pests. Plant diseases are inevitable, often discerned by farmers only after they have ravaged around 10% of the plant. Moreover, current methods for disease detection often fall short, potentially generating inaccurate outcomes basedon observed symptoms. The ramifications of crop failure profoundly affect the financial stability of farmers, exacerbating unfavorable conditions. Notably, there has been a marked uptick in the suicide rate among farmers in recent times. Hence, timely disease detection holds significantsway in the agricultural domain to mitigate substantial losses. Achieving this entails meticulous diagnosis and surveillance. Leaves, constituting a vital element of plants, furnish valuable clues regarding disease presence, rendering their thorough examination indispensable.

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

1.1. Literature review:

The sugarcane diseases mentioned by Pujari, Jagadeesh are the most important fungal diseases present in the plants and effects the productivity of the crop. Crop diseases have a very bad impact on the crop production and growth of the plant. By using this CNN model farmers can take live image of the leaf of the plant taken with the smartphone and upload it on the system and analyses the disease present in their crop. Using a deep learning trained model farmers can find the diseases present in their crop by uploading real time images. S.Militante analyzed the diseases in the sugarcane which causes a very bad impact on the growth of the plant and their productivity. Deep Machine Learning models can detect these diseases early and it can minimize the loss and productivity can be increased. To prevent the spread of fungal diseases can be controlled by the early diagnosis of the crop by the farmers using well trained Deep Learning Models. CNN can detect more diseases and have a high level of accuracy. CNN can identify a high level of illness and diseases across various crops and detect the diseases.
The Machine Learning algorithm can detect several diseases present in the sugarcane crop.

1.2. Deep Learning:

Internationally, deep learning frameworks, as delineated in citation [1], are extensively utilized for image recognition and disease detection in both human and plant domains. Challenges such as image identification and categorization have been notably streamlined through the integration of deep learning methodologies, encompassing Artificial Neural Networks (ANN) [2], Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN), among others. Considerable research endeavors have already been undertaken employing these advancements.

1.3. Convolution Neural Network:

- CNN, at its core, is a neural network crafted for discerning shapes, hues, and patterns in images with streamlined processing, rendering it especially adept for a myriad of identification and detection undertakings. The computer perceives each image as a pixel matrix. By convolving it with a filter matrix, a fresh matrix embodying a distinct functionis derived, facilitating operations such as edge detection.

- This principle enables the pinpointing of particular patterns or spots within an image. Khatri employed deep learning, particularly linear vector quantization ANN, for the identification of handwritten digits, attaining a peak accuracy of 94.9% [9]. Our methodology centers on diagnosing diseases in brinjal leaves, utilizing image processingalongside artificial neural network methodologies.

- The sugarcane plant leaf data undergoes processing utilizing the K-means clustering algorithmfor segmentation, followed by the application of an artificial neural network for image classification [1]. Pujari introduced image processing techniques for identifying

CAHIERS MAGELLANES-NS
Volume 06 Issue 1
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

and categorizing fungal diseases in diverse crops, juxtaposing SVM and ANN classifiers, with SVM exhibiting superior accuracy [2].

- Prakash devised a system for leaf disease detection and classification, employing K-meansclustering for segmentation, GLCM for feature extraction, and SVM for classification, attaining a commendable accuracy of 90%.

- In our study, we frame plant disease detection as a classification challenge and utilize CNNto tackle it. The devised CNN model exhibits enhanced performance in contrast to alternative methods utilized for disease identification and detection. Subsequent sections delineate our CNN model, analyze results, and draw conclusions.

2. Method Used:

The flowchart depicted in Figure 1 furnishes an overview of the procedures entailed in constructing the model and assessing its performance. The initial step involves acquiring the image dataset, followed by preprocessing tasks encompass activities such as shaping, resizing, and converting images into a fixed-size array format. Comparable preprocessing steps are employed during image testing. The processed image is subsequently augmented and incorporated into the CNN model constructor.

Training of the CNN model is conducted using the training dataset, facilitating its ability todiscern test images and the associated diseases they depict. Evaluation of results occurs post-successful training and preprocessing stages. Subsequent to training, the model is subjected to comparison with the test image, and the accuracy of the model's assessment is documented and validated. This thorough process ensures the model's proficiency in accurately identifying diseases in test images.

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

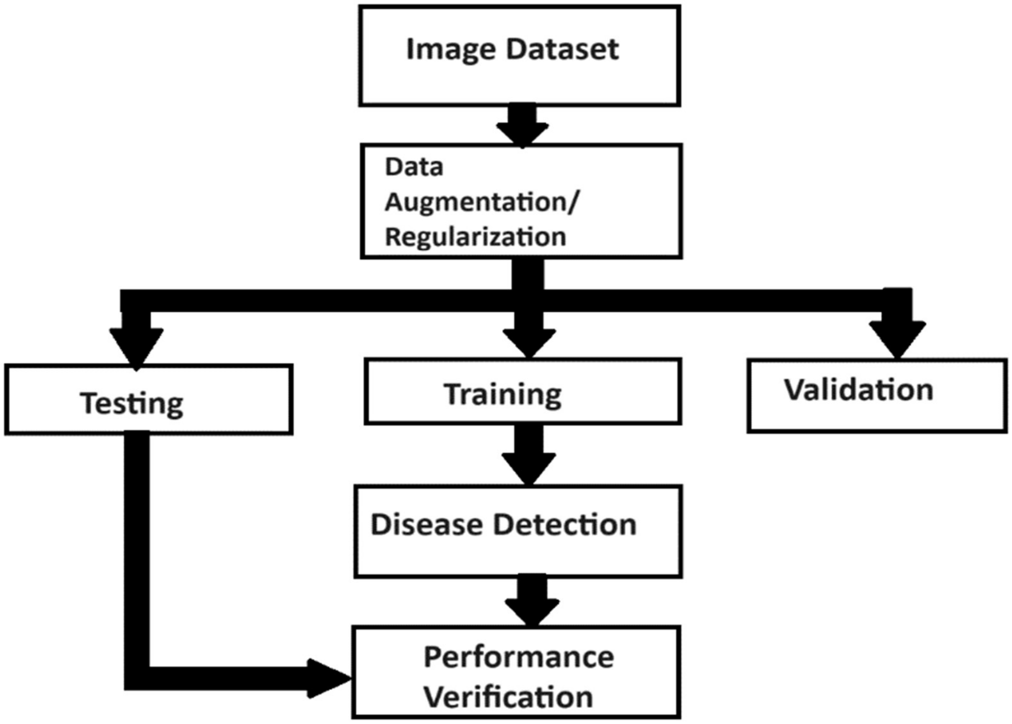

Figure 1– Flowchart of the steps involved inbuilding the model used.

Table 1- Accuracies for different networks

Model	**Testing Accuracy (Training FC layers only, for 5-10 epochs)**	**Testing Accuracy (Training all layers, 10-15 epochs)**	**Testing Accuracy (Training all layers, testing with augmentation)**
ResNet-34	**95.66%**	**98.52%**	**98.76%**
ResNet-50	**98.13%**	**99.77%**	**99.67%**
EfficientNet	**96.14%**	**98.89%**	**99.72%**

2.1. CNN classification model:

CAHIERS MAGELLANES-NS
Volume 06 Issue 1
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

To summarize, the model architecture outlined utilizes the ResNet architecture implemented in PyTorch, comprising four primary layers: convolution, pooling, dropout, and fully connected. These layers are subdivided into six convolution layers, four pooling layers, two dropout layers, and four fully connected layers. The integration of bottleneck layers is geared towards optimizing loss. Key particulars encompass:

- Convolution layers employ strides of (1,1) and (2,2).
- Padding is applied in convolution layers to avert information loss and retain the originalimage size as output.
- Test and train data are resized to dimensions of 224 pixels in height and width and loadedin batches of 128 images.
- Max Pooling is utilized, and a dropout layer is incorporated to mitigate overfitting andenhance the model's generalization capabilities
- Instance normalization is implemented to enhance performance, normalizing each datum individually.
- The ReLU activation function is employed for all layers except the final one.
- The softmax activation function is utilized on the final layer to derive probabilities ranging between 0 and 1.

Regrettably, the comprehensive structure of convolution and pooling layers, along with the overarching model architecture, is not delineated in the summary text but can be referencedin a figure

CAHIERS MAGELLANES-NS
Volume 06 Issue 1
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

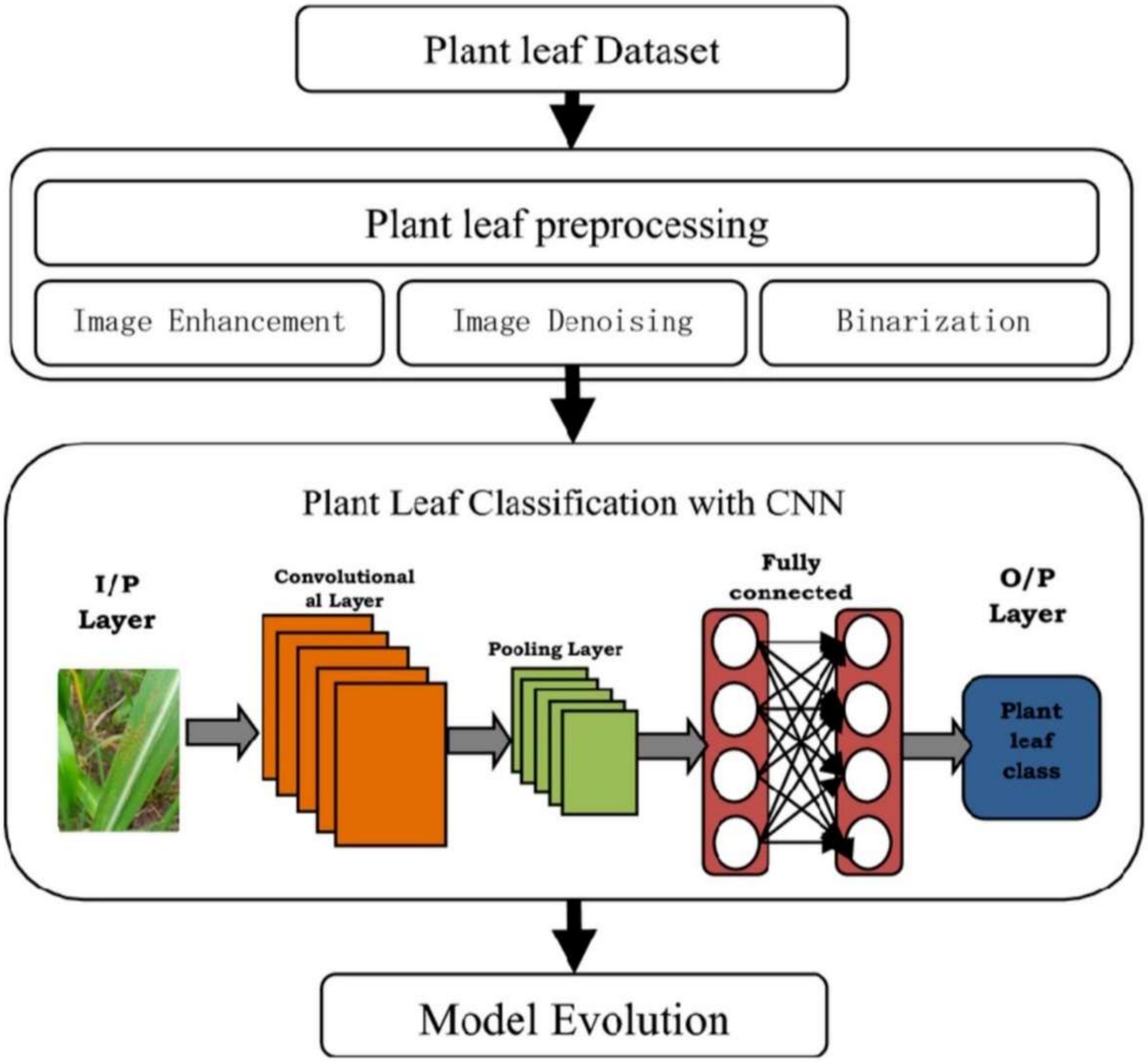

Figure 2- Detailed diagram of CNN model

In brief, the described neural network model comprises various layers and processes:

- Conv2D Layer: This layer convolves the input image into multiple images using aspecified activation function.
- MaxPooling2D Layer: It employs max-pooling to extract the maximum values fromthe dataset images, facilitating feature map calculation.

Flatten Layer: This layer flattens the dimensions of the images into a single column subsequent toconvolution.

- Dense Layer: Establishes a fully connected layer with a specified number of nodes.
- Dropout Layer: Mitigates overfitting by randomly discarding a fraction of

CAHIERS MAGELLANES-NS
Volume 06 Issue 1
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

nodes duringtraining.

- Image Data Generator: Augments the dataset by implementing transformations such asresizing, shearing, zooming, and horizontal flipping to enhance model resilience.
- Training Process: Entails loading images from the designated directory, utilizing the Flow from directory function, and fitting the data into the CNN model using the fit generator. Parameters like steps per epoch signify the number of batches within a singleepoch.
- Epochs: Signifies the number of cycles during the training of the dataset.
- Validation Process: Integrates validation data to evaluate model performance, with thevalidation dataset superseding the training dataset during this phase.

2.2. Training & testing model:

It seems like your message got cut off. Could you please provide more details or complete your question about the dataset preprocessing and the 37 different diseases in plant leaves? I'm here to provide any information or assistance you require, whether it's about data or classification.

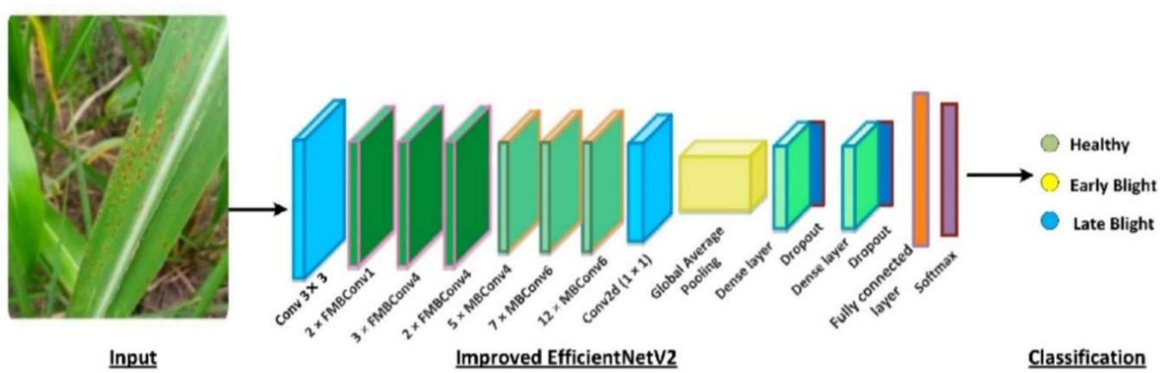

Figure 3- Training Model

Table 2- Dataset Table

S. No	Class	Count
1.	Bud Proliferation	**878**
2.	Root-Knot	**769**
3.	Spiral	**984**
4.	Dwarf	**634**
5.	Yellow Leaf	**798**

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

6.	Mosaic	**864**
7.	Grassy Shot	**977**
8.	Early Bird Sprouting	**880**
9.	Leaf Scald	**694**
10.	Top Root	**932**
11.	Black Rot	**900**
12.	Black stripe	**789**
13.	Brown spot	**577**
14.	Brown stripe	**872**
15.	Downy mildew	**955**
16.	Eye spot	**666**
17.	Fusarium sett and stem rot	**756**
18.	Iliau	**934**
19.	Leaf blast	**843**
20.	Leaf blight	**990**
21.	Leaf scorch	**809**
22.	Marasmius sheath and shoot blight	**912**
23.	Myriogenospora leaf binding	**737**
24.	Phyllosticta leaf spot	**866**
25.	Phytophthora rot of cuttings	**596**
26.	Pineapple disease	**842**
27.	Pokkah boeng	**751**
28.	Red leaf spot	**675**
29.	Red rot	**987**
30.	Red rot of leaf sheath and sprout rot	**509**
31.	Red spot of leaf sheath	**831**
32.	Rhizoctonia sheath and shoot rot	**786**
33.	Rind disease	**432**
34.	Ring spot	**654**
35.	Rust	**987**
36.	Schizophyllum rot	**968**
37.	Sclerophthora disease	**965**

The dataset has undergone preprocessing, including tasks like resizing, shaping, and conversion into an array format of fixed size. It encompasses 38 distinct diseases found in plant leaves. Any image from this dataset can be chosen and utilized as a test image within the software.

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

The model is trained using a dataset containing various layers within a Convolutional Neural Network (CNN), including Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D layers. Following successful training and preprocessing, the model can analyze and detect diseases in a test image based on its learning from the training dataset. Prediction is made by comparing the features of the test data with the patterns learned during training.

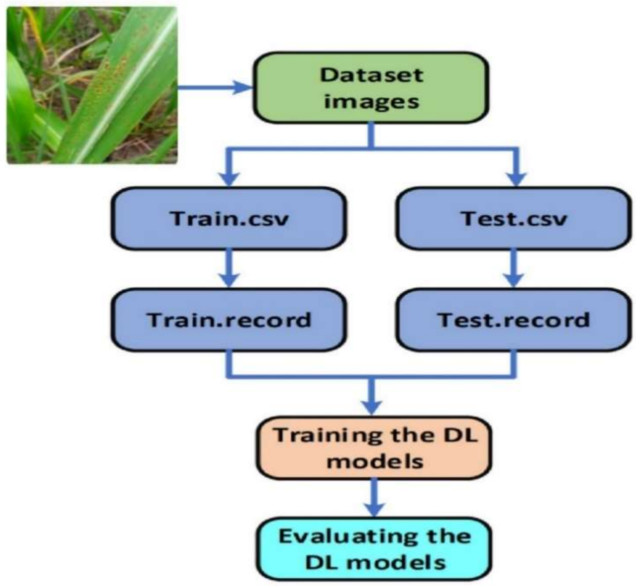

Figure 4 -Testing Model

Table 3- Mean metrics (precision, recall, F1) and overall accuracies across various experimental configurations on sugarcane dataset.

Model	Mean Precision	Mean Recall	Mean F1 Score	Overall Testing Accuracy (Training all layers, testing with augmentation)	Testing Accuracy (Training all layers, testing without augmentation)	Testing Accuracy (Training only FC layers)
ResNet-34	**0.9134**	**0.9056**	**0.9034**	**0.92345**	**0.9123**	**0.8326**
ResNet-50	**0.9197**	**0.9213**	**0.9065**	**0.9256**	**0.9254**	**0.8431**
EfficientNet	**0.9354**	**0.9234**	**0.9311**	**0.9372**	**0.9278**	**0.8611**

3. Experimental Results:

The outcomes stem from training a Convolutional Neural Network (CNN) on a dataset comprising both original and augmented images. CNNs prove effective in learning features

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

forvisual imagery, particularly when trained on extensive datasets, resulting in enhanced performance. The dataset utilized for this purpose comprises 37 plant diseases, encompassing twenty nine thousand colored leaf images.

The model is trained on the disease dataset, and the Cross Entropy function is computed to assessits performance. Optimization of the model is achieved using the Stochastic Gradient Descent technique, with a learning rate set at 0.005. The network is trained for 100 epochs, utilizing a dataset of 37 plant diseases, totaling twenty nine thousand leaf images, all in color.

Figure 5 depicts a graph illustrating the cost function with respect to the number of epochs. It'sevident from the figure that as the number of epochs increases, the cost function decreases andeventually stabilizes towards the end of training.

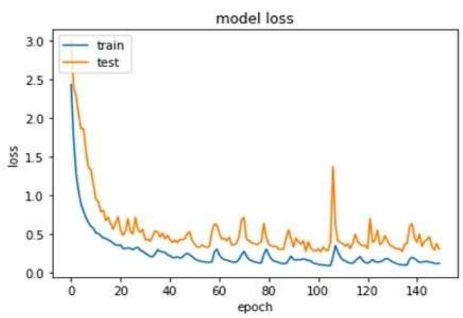

Figure-5 Depicts a graph showing the cost function's variation with the number of epochs during training.

The figure unmistakably illustrates that with the increasing number of epochs, the cost function exhibits an initial decrease followed by stabilization or convergence towards the end of training.

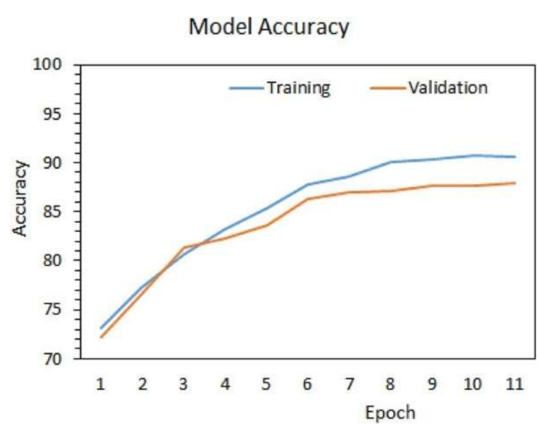

Figure-6 Illustrates a graph displaying the test accuracy of the model.

$$F(x) = H(x)-x$$

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

We attained a remarkable accuracy with our model. Upon comparing it with various other models and presenting the findings in a table, our model emerged as the top performer, showcasing superior accuracy.

4. Conclusion:

In current time it is impossible for the growers and farmers to keep eye on every plant in their fields. Automated plant disease detection systems harness the power of convolutional neural networks to effectively identify and classify plant diseases, amalgamating the knowledge of plant pathology experts with the capability to extract symptomatic traits. Our findings signify substantial advancements in predicting plant diseases using CNNs. The shift towards innovating new CNN architectures for plant disease recognition is apparent, showcasing enhanced accuracy in comparison to conventional methods. The CNN model can detect diseases in the plant with 98.69% accuracy. The model can differentiate between a healthy and unhealthy plant. With the help of the model the growers would be able to identify their plants health. The web application based on CNN developed is targeted to help the farmers identifying diseases in the sugarcane crop and find their cure and stop the spread of the diseases in their other plants.

5. References:

1. Anand R, V. S. J., "An Application of image processing techniques for Detection of Diseases on Brinjal Leaves using K-Means Clustering Algorithm," Fifth International Conference On Recent Trends In Information Technology, Coimbatore, 2016.

2. Pujari, Jagadeesh D., et al., "Identification and classification of fungal disease affected on agriculture/horticulture crops using image processing techniques," 2014 IEEE InternationalConference on Computational Intelligence and Computing Research.

3. Prakash, R. Meena, et al., "Detection of leaf diseases and classification using digital image processing," 2017 international conference on innovations in information, embedded and communication systems (ICIIECS).

4. Phadikar, Santanu, and Jaya Sil, "Rice disease identification using pattern recognition techniques," 2008 11th International Conference on Computer and Information Technology. Narmadha, R. P., and G. Arulvadivu, "Detection and measurement of paddy leaf disease symptoms using image processing," 2017 International Conference on ComputerCommunication and Informatics (ICCCI).

5. Orillo, John William, et al., "Identification of diseases in rice plant (oryza sativa) using back propagation Artificial Neural Network," 2014 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM).

ISSN:1624-1940
DOI 10.6084/m9.figshare.26090680
http://magellanes.com/

6. Şengür, Abdulkadir, et al., "A retinal vessel detection approach using convolution neural network," 2017 International Artificial Intelligence and Data Processing Symposium (IDAP).

7. Pooja, V., et al., "Identification of plant leaf diseases using image processing techniques," 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR).Khatri, Sunil Kumar, et al., "Recognizing images of handwritten digits using learning vector quantization artificial neural network," 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO).

8. Financial Express (2018) " India economic survey 2018: Farmers gain as agriculture mechanisation speeds up, but more R&D needed." [Online] Accessed on 8 March, 2019https://www.financialexpress.com/budget/india-economic-survey-2018-for-farmers- agriculture-gdp-msp/1034266/.
9. Bengio, Yoshua, et al., "Deep learning." Nature 521.7553 (2015): 436-444.Zurada, JackM.,"Introduction to artificial neural systems," St. Paul: West publishing company, 1992.

10. Nagasubramanian K., Jones S., Singh A. K., Sarkar S., Singh, A., and Ganapathysubramanian B., "Plant disease identification using explainable 3D deep learning on hyperspectral images." Plant Methods, pp 1-10, 2019. https://doi.org/10.1186/s13007-019-0479-8.

11. Cruz A.C., Luvisi A., De Bellis, L., and Ampatzidis, Y. "Vision-based plant disease detection system using transfer and deep learning." Proceedings of the ASABE Annual International Meeting, Spokane, USA, pp. 1-19, 2017. https://doi.org/10.13031/aim.201700241

12. U, Shruti, V. Nagaveni, and B.K. Raghvendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection", 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore India, pp. 281-284, 15-16 March 2019. https://doi.org/10.1109/ICACCS.2019.8728415.

13. S. Militante, and B. Gerardo, "Detecting Sugarcane Diseases through Adaptive Deep Learning Models of Convolutional Neural Network", 6th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, pp. 1-5, 20-21 December, 2019.https://doi.org/10.1109/ICETAS48360.2019.9117