

CROP DISEASE IDENTIFICATION THROUGH CROP LEAVES USING MACHINE LEARNING

Ashish Bishnoi¹, Shambhu Bharadwaj², Surjeet Dalal³

¹Research Scholar, Teerthanker Mahaveer University, Moradabad, India
ashishbishnoi05@gmail.com

²Associate Professor, Teerthanker Mahaveer University, Moradabad, India
shambhu.bharadwaj@gmail.com

³Professor, Amity University, Gurugram, India
profsurjeetdalal@gmail.com

Abstract

Crops in India suffer from various kinds of diseases that impact the overall yield production. These diseases not only cause less production per hectare but also affect the quality of yield. Modern techniques available in Information technology specifically Machine Learning and Deep Learning were used to improve the yield both in terms of quality and quantity could be done. The Machine Learning model developed for disease identification for one crop could also be used for disease identification in similar crops without requiring it to build the model from scratch.

This approach could be implemented on paddy or rice as their demand is increased every year and their production have to be increased per hectare as lots of fertile land is lost due to urbanization and their production is affected due to floods and cyclones. Also the paddy or rice suffers from various diseases caused by various microbes which negatively affect their production. In this paper we discuss about the various factors affecting the paddy production and how modern day techniques like machine learning could be used to resolve them.

Keywords

Artificial Intelligence, Machine Learning, Deep Learning, Crop Diseases

1. INTRODUCTION

India is commonly known as land of agriculture where most of the people survival is based on it either these people are directly or indirectly involved with the process of agriculture. There are two main agriculture seasons which are rabi and kharif where rabi crops includes mustard, cumin, wheat etc. and there season begins from November to April where as Kharif crop includes rice or paddy, bajra, groundnut etc. as they requires a lot of water for their growth due to which they begins from July to November. Government of India aims to double the income of farmers by the year 2025 for which it is required that the latest computing technologies like Machine learning, Deep learning and Transfer learning must be utilized in the domain of agriculture. When the researcher's from both domains join their hands only then it is possible to achieve the goals.

The Engineering Science that makes the computing machines intelligent is proposed by the researcher John Mc Carthy in the year 1956. This provides the ability to machines specifically computers to think and take decision like a human expert. [1] This capability of machines used exhaustively by Humans in different domains like education, health care, logistics, robotics, agriculture etc. To improve the efficiency, capability and decision making power of machines is further enhanced by Machine Learning. We use it in the domain of agriculture where per hectare production of crop should be increased by analyzing the various diseases at their initial stage and providing cure of them to farmers. Models produced by machine learning behave in the following way:

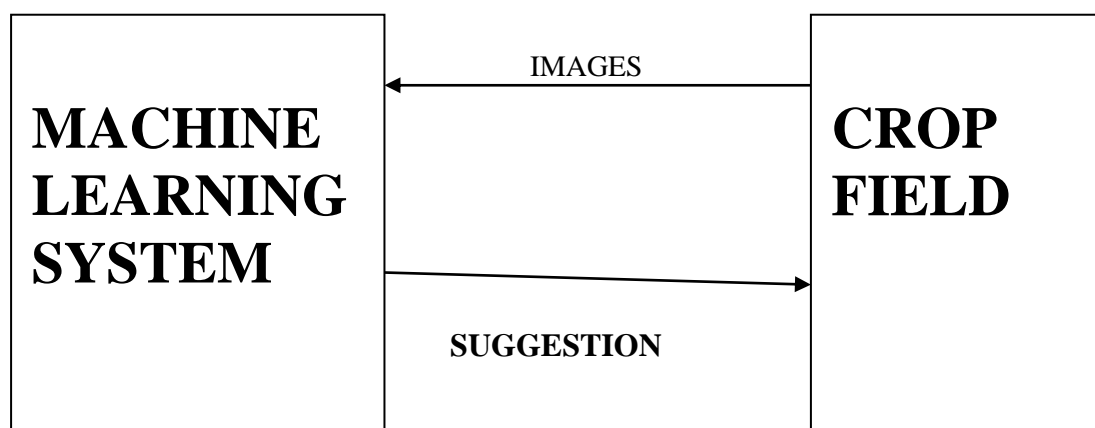


Figure 1. Machine Learning used by farmer

This research presents a hybrid model of Machine Learning that is able to detect common rice disease at early stage so that production of rice is further increased. These diseases like brown spot, narrow brown spot, sheath rot, sheath blight, false smut etc. which were spotted on leaf of crop plants and mainly caused by fungus. [2]



Figure 2. Various Diseases in Rice

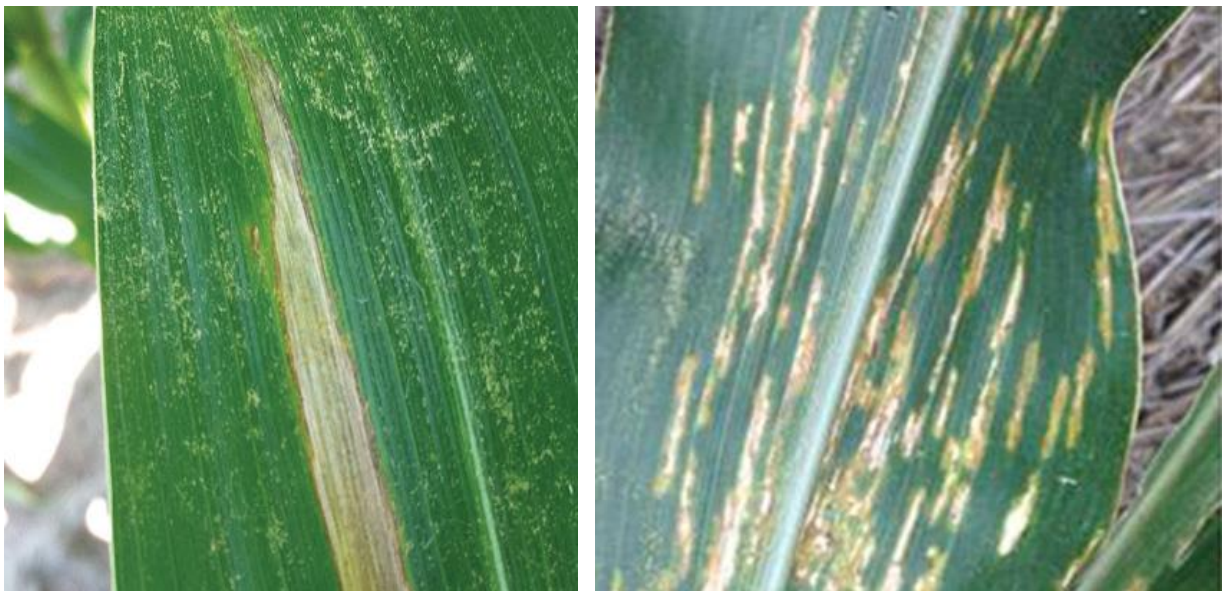
These disease effects the production of rice which was surplus for previous year but ever increasing population of India and their dependency on rice creates huge pressure on farmers to produce more. Out of total production of rice is 5% about basmati rice while 95% belongs to the non basmati category. [3]

This gap is filled by the Millet crops like Maize, Jowar etc which are cultivated in large amount due to their multiple usage not only for human population but also used for feeding the cattle. A lot of work is done on the Maize but Jowar is lacking behind while Jowar is more nutritious than other millets. Maize and the Jowar crops possess lots of similarity both in terms of diseases, water requirement and cultivation period. [4] These diseases includes Rust, Downy Mildew and Charcoal rot etc.



Figure 3. Healthy Maize Leaf

Above picture from the dataset collected is an example of healthy maize leaf detecting there is no disease present in the crop maize plant. With the help of Machine Learning models it is possible to identify the disease at its beginning stage for which we have to create dataset of infected plants with various diseases. Here are some leaves of maize plant infected with common diseases.



(i)

(ii)

Figure 4. Maize Leaves infected with blight



Figure 5. Maize Leaves infected with blight

Sometimes these diseases in crops are due to the soil category in which they were cultivated. [5] The diseases in maize and jowar crops not only affect the quality but also their yield. For identifying the disease in their beginning stage various models based on different algorithms of machine learning are used which enables the farmers to take appropriate measures to protect their crop. In present scenario, farmer uses the pesticides and fertilizers in vast amount to improve the productivity of maize per hectare that in turn is emerging as a major factor for various diseases in human beings or animals which is disturbing the balance created by mother nature or disturbing the food cycles. Selection of the algorithm in machine learning depends on various factors like the amount of data set available, environmental factors like temperature, rains during harvesting season etc, accuracy produced by particular algorithm.

Above all the factors is the amount of data set and accuracy produced by particular algorithms. With different models of identifying disease in Maize used by the farmers. The accuracy of these models is increased with every passing season of maize crop. In this research paper, we consider the work done on maize crops by the researchers and various factors considered by them. We analyze the various works and identify the factors that is still not considered by them and important for Indian farmers. We also discuss the various maize disease and their impact on Indian context. In today's world, the use of technology like Artificial Intelligence not only increases the yield of crop but also reduces the cultivation cost for farmers. In this research paper, we will devise the model that could be applicable on multiple crops provided that the crops should have similar characteristics and crop disease so that learning made on one crop could be applicable on another crop. It not only reduces the development cost for the model but also reduces the effort of the researchers for developing model for every crop and disease. It make possible for researchers to learn and train the model for disease identification for the crops where large data set is available but using that learning for least cultivated crop or for which data-set is not available. It reduces the time and cost for developing the new model that has limitation of similarity or similar factors.

2. LITERATURE SURVEY

Machine Learning is a subset of Artificial Intelligence which was developed by Arther Samuel in the year 1959. It is used to build models by the machine itself for identifying the objects without any human intervention. The data set used in these models specifies both the input set and the corresponding output. The power of Machine Learning is not restricted to computer science but it could be implemented in multiple domains like health care, education, agriculture, robotics etc. The main approaches of machine learning were supervised learning, unsupervised learning and reinforcement learning. But in identifying the corn disease on the basis of image of corn leaves and the common diseases are: Rust, Downy Mildew, Black & Charcoal Rot etc. We use the algorithms primarily from super-vised learning if the data received is labeled; we hereby compare the different approaches. Presence of Nitrogen in land used for cultivation of corn is mandatory which increases the production of crop and gives more profit to the farmer. According to agriculture scientists the value of Nitrogen in

soil is between 0.95-1.05. [6] The method of computing nitrogen in soil is machine learning models and sensors.

In most of the cases manual procedures were used to determine the disease in maize through its leaves, for which skilled and experienced farmers required. It's a time consuming process and sometimes farmer's are not able to identify the disease in maize crop using their skills precisely. To replace this several unique models for disease identification in maize are developed that uses new generation computer technologies like machine learning and deep learning. These diseases not only affect the growth of plant but also quantity and quality of the produce. [7] Identification as well as finding of maize disease in its beginning [8] is very helpful for the farmers not only to reduce their cultivation cost but also to increase their yields both in terms of quantity and quality per hectare. In machine learning widely uses KNN (k-Nearest Neighbor) algorithm to identify the disease in maize plant in four different phases. These phases may include filtering of images, segmentation of leaf images, feature extraction of healthy leaves and classification of leaves with different kinds of diseases that occurs in maize crop.

Another approach for finding diseases in maize crop through its leaves is CNN (Convolution Neural Network) deep learning algorithm is also used. Here features are extracted the machine itself for the set of images presented to the system and does not require the features extracted by human beings for training of the model. It is achieved by the computational power given to computers by NVIDIA [9]. Different deep learning models that were used to classify and detect corn diseases in maize or corn plants are compared for accuracy in different segments.

Saeed et al. in his model to find the diseases present in the crop of potato, tomato and maize used PLS (Partial Least Square) regression are used for feature extraction in CNN model after classification is able to achieve accuracy of 90.1% on an average. To identify and classify the most common disease occurring in Guava automatically using AI (Artificial Intelligence) framework by Almadhor et al. [10]. To identify the infected area the model uses chromatic image segmentation instead of individual channels for feature extraction in order for the best results. The technique proposed by Oyewola et al. can expand the disease data set and the images were balanced by unique block processing. [11] This technique uses gamma correction and de-correlation stretching improving the separations of color from the images. In order to evaluate the model researcher selects eight performance metrics.

Bayes Theorem is used for detecting the diseases commonly occurring and affecting the yield of maize crop per hectare. It is a method used to calculate data uncertainty into definite data. This uncertainty is obtained by comparing between yes data and no data. Bayes probability is one way to overcome the uncertainty of data by using the stated bayes formula:

$$P(H | E) = P(E | H).P(H) / P(E)$$

Where :

P (H | E) = probability hypothesis H if given evidence E

P (E | H) = probability of the emergence of the evidence E if hypothesis H is known

P (H) = H probability without containing any evidence

P (E) = probability evidence E

The accuracy for disease identification in maize plants is 90% so that farmers could easily identify the diseases and took appropriate measures to control it. [12]. In order to identify the blight disease in maize plants through the images of their leaves cannot be alone possible through colors captured by the device but we had to consider the RGB value for each pixel present in the image that should be look on computers. CIElab is develop to perform the required transformation.



Figure 6. Acquisition of Image

CIElab is composed of binary feature that helps to perform separation where “0” means background and “1” means foreground. It performs clear separation so that one could identify the disease pattern. As to know that objects are part of the foreground and that is proved by examining 8-connectivity pixel.

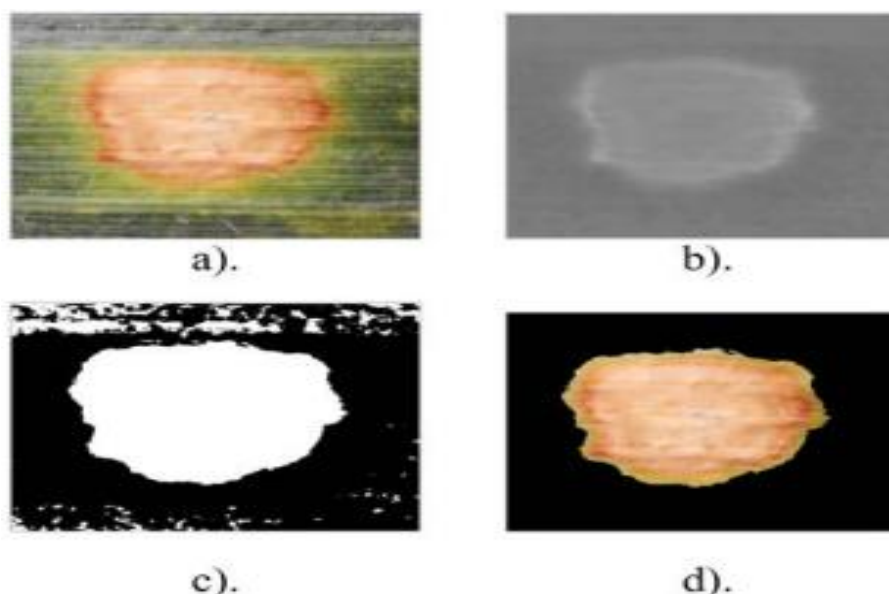


Figure 7. Disease Segmentation Result

In today’s World, various algorithms, techniques and models of machine learning were used not only to predict but also to classify the maize plant diseases, so that farmers could take appropriate measures with in time. Due to various diseases occurring in humans due to food and also extensive use of chemicals by farmers disturbs the food chain. As a result some farmers are moving towards organic farming which requires the early detection of crop diseases using Artificial Intelligence techniques like Machine learning and Deep learning models. These models regularly updated due to change in data or increase in data that increases their accuracy and efficiency.

3. MODELS

We apply the various models on the dataset collected from kaggle [13] that includes 2188 images which were categorized in four categories- Healthy, Blight, Common Rust, Grey Leaf Spot that represents different classification accuracy.

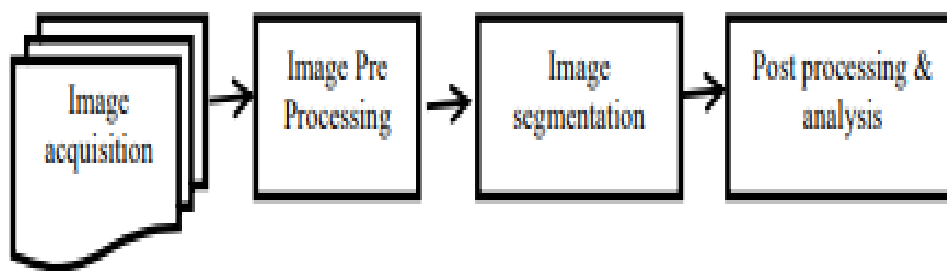


Figure 8. DPI Stages

Images used by the algorithms increased to the total of 4000 by performing augmentation on them and each category have equal images of equal size i.e. 256 x 256 pixels where each pixel had three channels that were divided by 255 in order to get values between 0 & 1.

Table1.Comparison of Models [15]

Methods	Classification Accuracy
SEG-KNN	80.13
GA-SVM	82.69
SIMPLE-CNN	87.58
RESNET50	95.48
GOOLENET	94.06
VGGNET	63.99
DMS-Robust Alex Net	98.63

After analyzing the performance of various predesigned models for detecting disease in corn crop that there is a requirement of the new model based on the existing one to get better classification accuracy while considering the different weather parameters such as temperature, humidity, wind speed etc. These parameters play the vital role in occurrence of the specific disease in corn crop plants.

Model designed by the researchers are used to predict the occurrence of various diseases in corn plants. To perform the statistical assessment of performance of various models confusion matrix is commonly used. It is used to predict a particular category for every given input instance. The confusion matrix is applicable for both binary and multi class classification.

4. CONCLUSION

Farmers cultivating the various crops not only for them but for the entire human population and if they suffer losses due to various diseases caused by various micro-organisms and lose to the entire human society. Due to which agriculture is the backbone of economy of various Nations. In this paper, we discuss primarily about corn or maize crop diseases that could be identified by using crop leaf agronomy and new computer aided techniques like machine learning. This research paper provides the approaches or methodologies to design and develop autonomous models for predictions or detections of diseases. These impact both the quantity/yield per hectare of cultivation land and quality of the yield. Models discussed and proposed in the paper had the 97% their average accuracy.

5. FUTURE WORK

The models that had to be designed in future can have more performance and average accuracy while considering more parameters and diversity in the dataset. These models could combine the existing and future techniques to use large number of parameters for improving the performance of the models over vast amount of dataset showing the greater accuracy in the result. In future work, feature extraction and classification accuracy will be improved in proposed models.

REFERENCES

- [1] Joseph P. Bigus & Jennifer Bigus, Constructing Intelligent Agents using java, 2nd edition, wiley
- [2] <http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases>
- [3] <https://www.cotecna.com/en/media/articles/india-rice-market-trends-and-prospects> & <https://www.mordorintelligence.com/industry-reports/india-rice-market>
- [4] Bishnoi A., Bhardwaj S., Dalal S., “*Disease Detection in Crop’s using Transfer Learning: A Review*”, International Conference on Communication, Security and Artificial Intelligence, IEEE Conference ID:59793, 23-25 NOV. 2023 @Galgotias University, Greater Noida
- [5] Dwivedi Kr. R., Kumari N., Bishnoi A., Pandey Prasad R., “*Soil Identification and Classification using Machine Learning A Review*”, 11th International SMART conference proceedings, published in IEEE explore (2022).
- [6] Dan Li, Yuxin Miao, “Corn Nitrogen Nutrition Index Prediction Improved by Integrating Genetic, Environmental, and Management Factors with Active Canopy Sensing Using Machine Learning”, Remote Sens. 2022, 14, 394. <https://doi.org/10.3390/rs14020394> or <https://www.mdpi.com/journal/remotesensing>
- [7] M. Minervini, H. Scharr, and S. A. Tsafaris, “Image analysis: The new bottleneck in plant phenotyping,” IEEE Signal Processing Magazine, vol. 32, no. 4, pp. 126–131, 2015.
- [8] G. Xu, F. Zhang, S. G. Shah, Y. Ye, and H. Mao, “Use of leaf color images to identify nitrogen and potassium deficient tomatoes,” Pattern Recognition Letters, vol. 32, no. 11, pp. 1584–1590, 2011
- [9] A. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol, and A. Srikaew, “Grape leaf disease detection from color imagery using hybrid intelligent system,” in 2008 5th international conference on electrical engineering/electronics, computer, telecommunicat
- [10] F. Saeed, M. A. Khan, M. Sharif, M. Mittal, L. M. Goyal, and S. Roy, “Deep neural network features fusion and selection based on PLS regression with an application for crops diseases classification,” Appl. Soft Comput., vol. 103, May 2021, Art. no. 107164.
- [11] D. O. Oyewola, E. G. Dada, S. Misra, and R. Damaševičius, “Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing,” PeerJ Comput. Sci., vol. 7, p. e352, Mar. 2021.
- [12] Frans Ikorasaki, Muhammad Barkah Akbar, “Detecting Corn Plant Disease with Expert System Using Bayes Theorem Method”, The 6th International Conference on Cyber and IT Service Management (CITSM 2018) Inna Parapat Hotel – Medan, August 7-9, 2018
- [13] <https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset>
- [14] Ranajana Sharma, Priyanka Suyal, Sarthika Dutt, Shambhu Bhardwaj, “Detection of Wheat Crop Quality using Deep Convolution Neural Network” SMART-11, IEEE Conference ID: 55829, 16-17 DEC. 2022 @CCSIT, Teerthanker Mahaveer University, Moradabad
- [15] Daneshwari A. Noola, Dayanand R B, “Computer Aided Corn Leaf Identification System”, Proceedings of the Second International Conference on Smart Electronics and Communication (ICOSEC). IEEE Xplore Part Number: CFP21V90-ART; ISBN: 978-1-6654-3368-6
- [16] Dionis A. Padilla, Ramon Alfredo I. Pajes III, Jerome T. De Guzman, “Detection of Corn Leaf Diseases using Convolutional Neural Network with OpenMP Implementation”, 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) | 978-0-7381-1177-3/20/\$31.00 ©2020 IEEE
- [17] Sharma, R., Suyal, P., Dutt, S., Bharadwaj, S., “Detection of Wheat Crop Quality using Deep Convolution Neural Network”, Proceedings of the 2022 11th International Conference on System Modeling and Advancement in Research Trends, SMART 2022, 2022, pp. 1266–1269