Literature Survey

Artificial intelligence (AI) is a rapidly evolving area that offers unparalleled opportunities of progress and applications in many Agriculture fields. This proposal is regarding automatic detection of diseases and pathological part present within the leaf pictures of plants and even within the agriculture Crop production it is through with advancement of technology that helps in farming to extend the production. Primarily there is downside of detection accuracy and in neural network approach support vector machine (SVM) is exist already.

Indian economy greatly depends on agriculture. Nowadays detection of crop disease is very important topic for analysis. It is one of the issues that cause reduction in quality and quantity of crop. So detection and classification of crop disease is necessary task to increase crop productivity and economic process.

The proposed research work is to analyze various machine learning and image processing techniques applied to detect crop disease. In this paper we will review, different machine learning techniques, such as Random Forest, Decision tree, Multilayer Regressor, Regression algorithms, image processing techniques, Extreme Machine Learning to get better accuracy for system. Here, crop leaf images are taken as input and after processing that image, it will detect whether there is any disease or not. If disease is detected, then it will tell what type of disease it is and will provide solutions such as pesticides or chemicals to cure that disease. It will increase the productivity and economic process.

With increase in population the need for food is on rise, in such circumstances, plant diseases prove to be a major threat to agricultural produce and result in disastrous consequences for farmers. Early detection of plant disease can help in ensuring food security and controlling financial losses. The images of diseased plants can be used to identify the diseases. Classification abilities of Convolutional Neural Networks are used to obtain reliable output.

In the paper **—Deep learning for Image-Based Plant detection" [1]** the authors Prasanna Mohanty et al**.,** has proposed an approach to detect disease in plants by training a convolutional neural network. The CNN model is trained to identify healthy and diseased plants of 14 species. The model achieved an accuracy of 99.35% on test set data.

Malvika Ranjan et al. in the paper **—Detection and Classification of leaf disease using Artificial Neural Network"** proposed an approach to detect diseases in plant utilizing the captured image of the diseased leaf. Artificial Neural Network (ANN) is trained by properly choosing feature values to distinguish diseased plants and healthy samples. The ANN model achieves an accuracy of 80%.

According to paper **—Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features" [3]** by S. Arivazhagan, disease identification process includes four main steps as follows: first, a color transformation structure is taken for the input RGB image, and then by means of a specific threshold value, the green pixels are detected and uninvolved, which is followed by segmentation process, and for obtaining beneficial segments the texture statistics are computed. At last, classifier is used for the features that are extracted to classify the disease.

Kulkarni et al. in the paper **—Applying image processing technique to detect plant diseases" [4],** a methodology for early and accurately plant diseases detection, using artificial neural network (ANN) and diverse image processing techniques. As the proposed approach is based on ANN classifier for classification and Gabor filter for feature extraction, it gives better results with a recognition rate of up to 91%.

Azadbakht et al. [2019] analyzed the performance for identifying leaf rust of wheat at canopy scale and at a dierent level of leaf area index (LAI) - high, medium and low. Four methods had been used for analyzing the performance in identifying the rust detection namely Gaussian process regression, random forest regression, v-support vector regression and boosted regression tree. The severity level of disease is also predicted in this. For analyzing the performance the experiment was conducted at North West of Iran in 7000 hectares of wheat cultivation. Hyperspectral reectance data recorded at dierent environmental condition is used for it. Based on the results from the experiment hyperspectral images can be used for identifying the wheat rust. Comparison of spectral vegetation indices (SVIs) and ML shows that ML performs better than SVIs.

Ip et al. [2018] presented a review about crop protection using big data and how it helps in controlling the weed. Discussed mainly about invasive species detection, predicting and modeling of resistance in herbicide, support system used for the purpose of crop protection and robotic weed control system. The machine learning approaches used for the above-mentioned problem are discussed clearly.

Markov random _eld uses spatial component for analyzing is explored in it. Probabilistic graphical model, traditional neural networks and support vector machine are the generative approach used in it. For the purpose of modeling and prediction of resistance in herbicide rule-based approach and CART is used. For detecting species and weeds random forest classi_er is used. RIM is used to control the ryegrass in the _eld. AgBot is robotics used as autonomous weed control. Markov Random Field helps in content-based fresh fruit grading and it has the ability to handle spatial data with irregular spacing. For modeling, the crop disease in corn Bayesian network is used and SVM helps in the semi-autonomous estimation of vegetables.

Chlingaryan et al. [2018] nitrogen estimation is an important factor in precision agriculture. Remote sensing system there is a problem in processing huge amount of data from a di_erent platform. The machine learning techniques help in handling the non-linear task. The mainly hybrid system of integrating machine learning techniques and remote sensing technique is used in nitrogen estimation. For estimating the chemical concentration in dry leaf Continual Removal (CR) method is used. For identifying more informative feature (BPNNs) is used. For calculating the vegetation indices back propagation neural networks are used. Feature extraction can be done by combining a convolution neural network and Gaussian process. The Gaussian process also helps in identifying di_erent plant leaves characteristic. For the purpose multi-class crop prediction M5 prime regression tree is used. And for nitrogen estimation, least square support vector machine is used. Fuzzy Cognitive Map is used in the _eld of crop management.

Lu et al. [2017] presented a system based on the deep learning framework named in _eld automatic wheat diagnosis system (DNIL- WDDS). In wild condition, the image-level annotation is done for the wheat disease. VGG-FCN-S and VGG- FCN-VD16 are the two architecture used in it. The proposed system achieves the accuracy of 97.95% and 95.12% when compared to CNN architecture achieving the accuracy of 93.27% and 73.00 %. It has an added advantage of integrating the system with the mobile application. Nearly four types of deep learning method are used for analyzing the disease prediction accuracy in wheat. The DMIL-WDDS model shows improvement in accuracy when compared to other models.

Tavakoli and Gebbers [2019] presented an analysis of winter wheat nitrogen and assessment of water in the _eld by using a camera. This experiment was conducted during a period of three years (2012, 2013, and 2014). Nitrogen fertilization and dierent level of water are applied in the _eld for the purpose of the experiment. Two machine learning algorithm was developed for the purpose of analysis namely Random Forest (RF) and Partial Least Square Regression (PLSR). Specter radiometer was used for radial measurement. Separately Vegetation Index (VI) is also calculated. For analyzing the nitrogen content R2 (RMSE) model is used and it is calculated separately for both data type. Random forest algorithm performs better in combined-date data. Nitrogen estimation calculation performs better while using the digital camera. It can also be integrated with the smartphone. It has a limitation of accessing only three spectral bands so that the analysis of plant status is also limited.

dos Santos Ferreira et al. [2017] proposed a method to identify unwanted weeds in the soybean _eld. Unwanted weed includes unwanted grasses and broadleaf. Convolution neural network technique is applied in the process of identifying the weeds in the soybean _eld. For the purpose of capturing the image, drones were used in it. The database used for analyzing purpose includes _fteen thousand pictures weeds, soil, soybean, grass weed, and broadleaf. SafeNet architecture is used for training the neural network. The cafe software includes Alex Net in it. Pynovisao algorithm is used to build a robust image database. The results are compared with Support Vector Machine, Ada Boost, and Random Forest. The accuracy of 99% is achieved using the convolution neural network.

Barbedo [2018] the problems faced in the machine learning technique has been overcoming by the deep learning concepts such as Convolution Neural Networks (CNN). Large data sets are needed for processing this technique. This paper mainly focuses on how the size of data and its variety a_ects the performance of the deep learning concepts. 12 plant species with di_erent samples, di_erent disease, and di_erent character are taken into consideration. This analysis describes the di_erent CNN network used for disease classication along with where this large amount of data are collected for classi_cation. Accuracy is also calculated for each deep learning concepts. The number of correctly classi_ed sample divided by the total number of samples provides the accuracy value. List of di_erent plant species and its disease are listed in it. Removing background from image improves the prediction accuracy. This analysis was performed mainly using dataset obtained from di_erent sources.

Conclusion

The key objective of this work is to analyze di_erent machine learning techniques widely used in the prediction of plant diseases and how advancement can be made in the future in this technique to achieve higher accuracy, robustness, cost-e_cient disease prediction system. The steps involved in image processing techniques like pre-processing, segmentation, extracting feature and classi_cation based on symptoms in the plant are discussed in this survey. Machine learning techniques play a key role in the machine vision system. In the future, deep learning framework can be used for disease prediction system. Integrating image processing techniques and deep learning techniques proved to be more potential in disease prediction system. Still, more investigations have to be made in these techniques for achieving better prediction system. If disease is detected, then it will tell what type of disease it is and will provide solutions such as pesticides or chemicals to cure that disease.