International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 40-44
http://www.publishingindia.com/ijkbcs/

Mango Leaf Diseases Detection using Deep Learning

Amisha Sharma[1*], Rajneet Kaur Bijral[2], Jatinder Manhas[3] and Vinod Sharma[4]
[1]M.Tech. Student, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: amishasharma26111997@gmail.com
[2]PhD Research Scholar, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: rajbijral@gmail.com
[3]Sr. Assistant Professor, Department of Computer Science and IT, Bhaderwah Campus, University of Jammu, Jammu and Kashmir, India. Email: manhas.jatinder@gmail.com
[4]Professor, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: vnodshrma@gmail.com
*Corresponding Author

Abstract: Diseases and pests cause great economic loss to the mango industry every year. The detection of various mango diseases is challenging for the farmers as the symptoms produced by different diseases may be very similar, and may be present simultaneously. This research paper is an attempt to provide the timely and accurate detection and identification of mango leaf diseases. Convolutional Neural Networks are end-to-end learning algorithms which perform automatic feature extraction and learn complex features directly from raw images, making them suitable for a wide variety of tasks like image classification, object detection, segmentation etc. In the proposed study, we develop a Convolutional Neural Networks based model for detection and classification of mango leaf diseases at the initial stages. Data augmentation is performed on a collected dataset. We applied data augmentation techniques like rotation, translation, reflection and scaling. Convolutional Neural Networks model has been trained on the augmented data for detection and classification of mango leaf diseases. The proposed CNN based model attains 90.36% of accuracy. The results validate that the proposed method is effective in detecting various types of mango leaf diseases and can be used as a practical tool by farmers and agriculture scientists.

Keywords: Convolution Neural Network (CNN), Crop, Deep learning, Image classification, Mango.

I. Introduction

About 40% of mangoes production is produced by India and stands first in the various mango growing countries of the world. Amongst the fruit crops in India vital place gets occupied by mangoes and plays an essential role in the economy of the country. About 30% to 40% of the crop yield got infected by various diseases and due to unaided eye perception the mango leaf disease went undetected. The different diseases affecting mango leaves cannot get acknowledged by the farmers which cause less production of mango fruit. Different diseases [1] cause different effects on mango crops. Some cause white patches and some cause black and all these patches seem over the surface of the leaf or early grown fruits as well while some other diseases cause white fungal powder on leaves and some affect the young leaves and shoots also. All these different types of illnesses need to be discerned in the initial stages and should be managed before it grows more and causes a severe loss to the plant life. To get this detection done in the initial stage, farmers and agricultural scientists need to keep an eye continuously on the plant parts which is a sluggish process. For the advance detection of disease in the plants some technique is needed as the prior acclaim of disease is the first step in the detection and expansion of mango diseases. Conventional ways to identify diseases is time consuming and expensive as it needs the expertise, knowledge and continuous monitoring. Still it lacks correct recognition of disease because of the complex structure and pattern of the leaf. With the advent of computational methods in the field of image recognition [2] and classification [3] this problem can be solved with greater accuracy. By using technology one can detect diseases on a large scale. In the case of mango leaves; there are various types of diseases [4] present like powdery mildew, anthracnose, red rust etc. In the present work, a deep learning (DL) based model has been proposed for the classification of various mango leaf diseases (powdery mildew, anthracnose, red rust) at the initial stages. Accuracy, Recall, Precision and F-Score have been used to evaluate the model.

II. Literature Review

Early Disease Detection and discussion are important for better yield and quality of crops. Disease Plants can lead to the huge Economic Losses to the Individual farmers. From the extensive literature review it has been found that various Machine Learning (ML) and Deep Learning (DL) based techniques have been used for recognition and classification of mango leaf disease to prevent the loss of harvest [5] a novel segmentation approach is proposed in this study to segment the diseased part by considering the vein pattern of the leaf. Afterward, features were extracted and fused using canonical correlation analysis (CCA)-based fusion. As a final identification step, a cubic support vector machine (SVM) is implemented to validate the results. The highest accuracy achieved by this proposed model is 95.5%. In 2013 a deep Convolutional Neural Network [6] was presented to identify 14 crop species and 26 diseases (or absence thereof) by using a public dataset of 54,306 images of diseased and healthy plant leaves collected under controlled conditions. The trained model achieves an accuracy of 99.35%. [7] Proposed the DL based approach for image recognition and examined the three main architecture of the Neural Network: Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully CNN (R-CNN) and Single shot Multibook Detector (SSD). System Proposed can detect the different types of disease efficiently and have the ability to deal with complex scenarios. The accuracy of 94.6% was attained. [8] Built a model for plant diseases and pests detection, and put forward a comparison with traditional plant diseases and pests detection methods. [9] Proposed a work that includes finding a solution to the problem of 38 different classes of plant diseases detection using the simplest approach while making use of minimal computing resources to achieve better results compared to the traditional models. VGG16 training model is deployed for detection and classification of plant diseases. Neural network models employ automatic feature extraction to aid in the classification of the input image into respective disease classes. This proposed system has achieved an average accuracy of 94.8% indicating the feasibility of the neural network approach even under unfavorable conditions. [10] Used a pre-trained CNN architecture called AlexNet that is modeled for automatic feature extraction and classification. The system is developed with MATLAB and achieves an accuracy rate of the detection of 99% and 89% for Grape leaves and Mango leaves respectively.

III. Methodology

A deep learning based model is designed for the proposed framework and for this study; three commonly found diseases are considered viz. Anthracnose, Red rust and Powdery mildew. Data is collected from the Sher-e-Kashmir University of Agriculture Sciences and Technology, Jammu (SKUAST-J) where they grow different varieties of fruits for educational and research purposes. There are around 980 images of 4 classes i.e. Normal, Anthracnose, Red rust and Powdery mildew. Fig. 1 shows some images of the collected dataset. The brief flowchart of the proposed model is shown in Fig. 2.

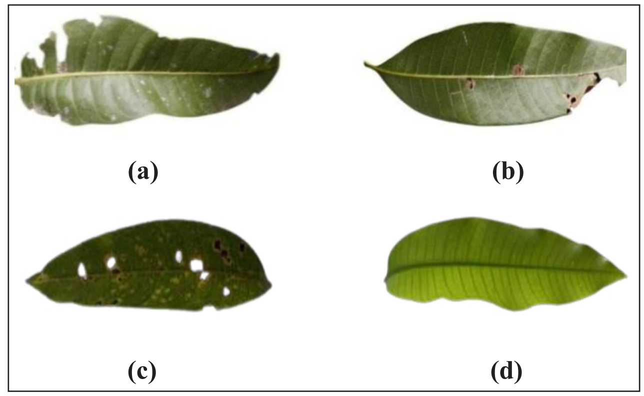

Fig. 1: Showing the Samples of Images of Mango Leaves (a) Anthracnose (b) Powdery Mildew (c) Red Rust and (d) Normal in Collected Dataset

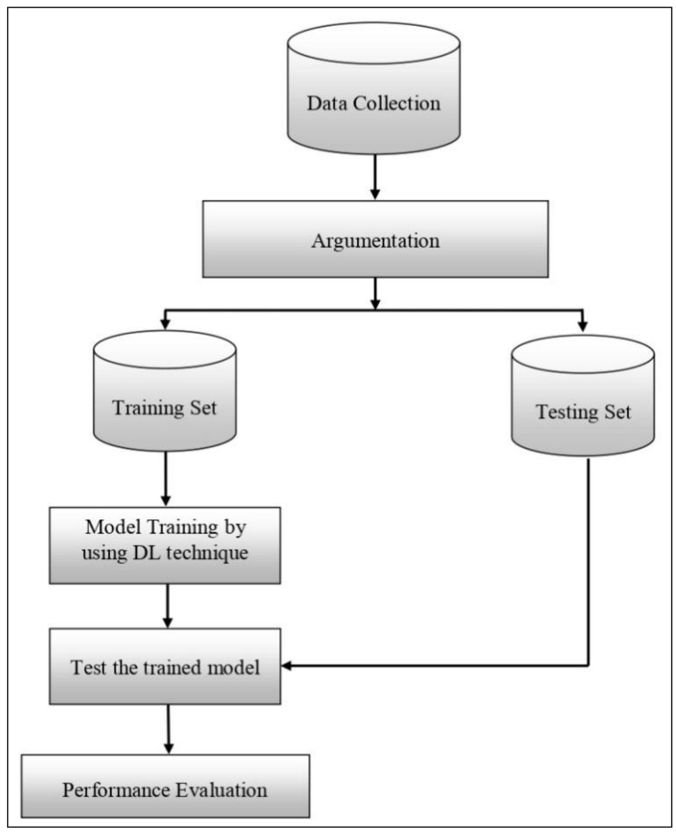

Fig. 2: Research Methodology for the Mango Leaf Diseases Detection using CNNs

Deep Learning

DL is the subset of machine learning (ML) [22] that emulates the functioning of the human brain in coursing data and producing patterns for use in decision making. It is a broader and more advanced part of ML methods. It is also called deep neural learning or deep neural networks. DL technique uses various layers consisting of nonlinear units. Every layer utilizes the output of the previous layer and considers it as its input [11].

Convolutional Neural Networks

Convolutional Neural Network (CNNs) [23] is a specific type of artificial neural networks (ANNs) that uses perceptrons, a ML unit algorithm, for supervised learning, to analyze data. CNNs are capable of performing complex jobs with images, sounds, texts, videos, etc, and most commonly used for analyzing visual imagery. CNNs use a wide variety of multilayer perceptrons designed that require minimal preprocessing. They are also known as shift invariant, based on their shared-weights architecture and translation invariant characteristics. CNNs are best for predictions and most widely used than other algorithms. It consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of CNNs typically consist of a series of Convolutional layers that convolve with a dot product. The activation function is commonly a RELU layer (rectified linear unit), and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers because their inputs and outputs are masked by the activation function and final convolution layer [12]. The basic architecture of the CNNs is shown below in Fig. 3.

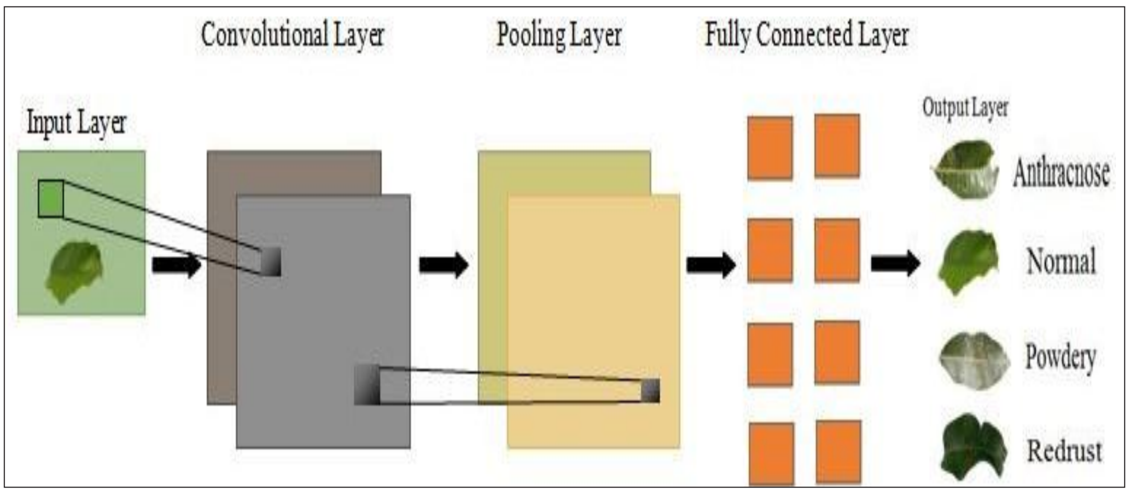

Fig. 3: Basic Architecture of Convolutional Neural Networks (CNNs)

Convolution Layer: Convolutional layer performs action on the input layer and convolves it and submits its result to the following layer. Here the series of images i.e. different mango leaves images acts as Input. This process is similar to the response of a neuron present in a human body. Sometimes a “kernel” is passed over the image, viewing a few pixels at a time (for example, 3X3 or 5X5). Convolution operation is there as well which is a dot product of the original pixel values with weights defined in the filter. The results obtained are then summarized into one number that represents all the pixels the filter observed [13].

Activation Layer: Matrix generated by the convolution layer is much smaller in size than the original image. This matrix passes through an activation layer that introduces non-linearity to allow the network to train itself via back propagation. The activation function generally used is ReLU [14].

Pooling Layer: Pooling layer is another building block of CNN. In this layer there is further down sampling and reduction of the size of the matrix. Over the results of the previous layer, a filter is passed and selects one number out of each group of values (typically the maximum value, this is called max pooling). This allows the network to train much faster and efficiently by focusing on the major information in each feature of the image [15].

Fully Connected Layer: A one-dimensional vector which is representing the output of the previous layers acts as input for this layer. Its output is a list of probabilities for different possible labels attached to the image (e.g. Anthracnose, Red rust, Powdery mildew and Probability [16]. The label that receives the highest probability is the classification or detection decision.

Output Layer: Output layer in CNN is that layer in which the input from the other layers is flattened and sent so as to transform the output into the number of classes as desired by the problem undertaken [17].

There may be multiple activation and pooling layers, depending on the CNN architecture.

Performance Evaluation

The performance of the model will be analyzed on the basis various parameters derived from its confusion matrix represented as under in Table I.

Table I

		Actual Result	
		LOW	HIGH
Predicted Result	LOW	True Positive	False Positive
	HIGH	False Negative	True Negative

True Positive (TP): These are the correctly predicted positive values which mean that the value of the actual class is yes and the value of predicted class is also yes.

True Negative (TN): These are the correctly predicted negative values which means that the value of actual class is no and the value of predicted class is also no.

False Positive (FP): When the actual class is no but the predicted class is yes.

False Negative (FN): When the actual class is yes but the predicted class is no.

Accuracy - Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations. [18]

Precision - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations [19].

Recall - Recall is the ratio of correctly predicted positive observations to the all observations in actual class – yes [20].

F1 Score - F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives in observations [21].

These performance measures can be calculated as shown in the Table II below.

Table II

Parameter	Formula
Accuracy	(TP + TN) / (Total Cases)
Recall	TP / (TP + FN)
Precision	TP / (TP + FP)
F1 Score	(2 * Recall * Precision) /(Recall + Precision)

IV. Results and Discussion

In the proposed study, a dataset of 980 images is collected from SKUAST-J. Dataset consists of 4 classes i.e. Normal, Anthracnose, Red rust and Powdery mildew. Argumentation is performed on the collected dataset to increase the number of images for model training and testing. Then the splitting of the dataset has been performed to train and test the dataset. 80% of the data has been used to train the model and the rest 20% of images have been used to test the model. A CNN based model has been developed to classify the images. The model has been trained with epochs equal to 40 and the batch size of 35. The test accuracy of 90.36% is attained. The training and Validation accuracy graph attained by the proposed model is shown below in Fig. 4. The Training and Validation loss graph is shown in Fig. 5. A generalized model is built in the proposed study as the validation loss is very less and the Validation accuracy is more than the training accuracy as shown in Fig. 5. The confusion matrix thus generated by the model shown in Table III. The various parameters are used to measure the performance of the model like precision, recall and F1 score and the values of these for the proposed model are 0.9, 0.90 and 0.9 respectively.

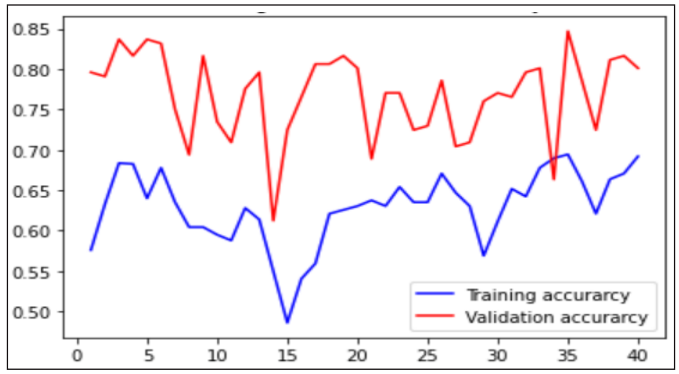

Fig. 4: Shows the Training and Validation Accuracy of the Proposed Model

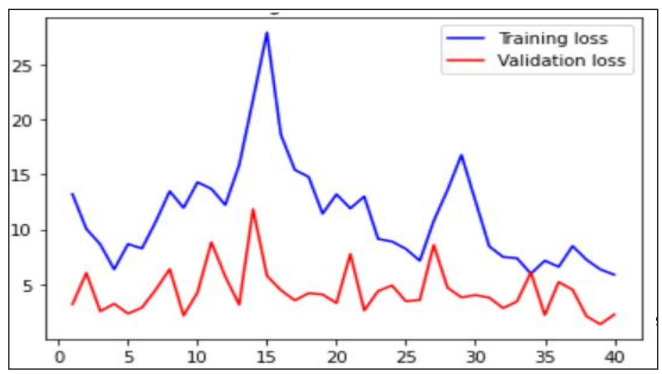

Fig 5. Shows the Training and Validation Loss for the Proposed Model

Table III: Shows the Confusion Matrix of the Proposed CNNs based Framework

Predicted class / True class	Anthracnose	Normal	Powdery	Red rust
Red rust	3	3	2	50
Powdery	7	0	40	0
Normal	0	47	0	0
Anthracnose	40	0	4	0

V. Conclusion

Mango is one of the most cultivated fruit crops in India. So it is important to protect it and detect the various diseases in the initial stages. For that, a model based on deep learning approach called CNN is proposed for the identification of 3 different plant leaf diseases, detection and recognition systems. This approach utilized a minimum set of layers to identify the diseases of four classes. The CNN is trained with SKUAST-J dataset. The Model works with accuracy of 90.36%. In order to increase accuracy and to make the proposed model more efficient in future, the number of the dataset can be increased, the concept of transfer learning can also be used. Moreover, the model can be made more efficient to detect images in the natural background as well.

References

[1] S. Gulvani, and R. Patil, "Deep learning for image based mango leaf disease detection," *International Journal of Recent Technology and Engineering (IJRTE*), vol. 8. no. 3S3, pp. 54-56. 2019.

[2] M. M. Krishna, M. Neelima, M. Harshali, and M. V. G. Rao, "Image classification using deep learning," *International Journal of Engineering & Technology*, vol. 7, no. 2.7, pp. 614-617, 2018.

[3] G. Geetha, S. Samundeswari, G. Saranya, K. Meenakshi, and M. Nithya, "Plant leaf disease classification and detection system using machine learning," *Journal of Physics: Conference Series,* vol. 1712, 2020.

[4] O. Parkash, and A. K. Mishra, "Important diseases of mango and their effect on production," *Biological Memoirs*, vol. 18, no. 1&2, pp. 39-55, 1992.

[5] R. Saleem, J. H. Shah, M. Sharif, M. Yasmin, H.-S. Yong, and J. Cha, "Mango leaf disease recognition and classification using novel segmentation and vein pattern technique," *Appl. Sci.*, vol. 11, no. 24, 2021, Art. no. 11901.

[6] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, 2016.

[7] S. Kumar, V. Chaudhary, and S. Chandra, "Plant disease detection using CNN," *Turkish Journal of Computer and Mathematics Education,* vol. 12, no. 12, 2021.

[8] J. Liu, and X. Wang, "Plant diseases and pests detection based on deep learning," *Plant Methods*, vol. 17, 2021, Art. no. 22.

[9] R. Rinu, and S. H. Manjula, "Plant disease detection and classification using CNN," *International Journal of Recent Technology and Engineering (IJRTE),* vol. 10, no. 3, pp. 152-156, 2021.

[10] U. S. Rao, R. Swathi, V. Sanjanaa, L. Arpitha, K. Chandrasekhar, Chinmayi, and P. K. Naikb, "Deep learning precision farming: Grapes and mango leaf disease detection by transfer learning," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 535-544, 2021.

[11] H. Muresan, and M. Oltean, "Fruit recognition from images using deep learning," *Acta Univ. Sapientiae Inform.*, vol. 10, pp. 26-42, 2018.

[12] S. Indolia, A. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network," *Procedia Computer Science*, vol. 132, pp. 679-688, 2018.

[13] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural network," 2017.

[14] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural network: An overview and application in radiology," *Insights into Imaging*, vol. 9, pp. 611-629, 2018.

[15] R. Nirthika, S. Manivannan, A. Ramanan, and R. Whang, "Pooling in convolutional neural network for medical image analysis: A survey and an empirical study," *Neural Comput Appl.*, vol. 34, no. 7, pp. 5321-5347, 2022.

[16] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, ... and L. Farhan, "Review of deep learning: Concepts, CNN, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, 2021.

[17] E. Irmark, "Implementation of convolutional network approach for COVID-19 disease detection," *Physiological Genomics*, vol. 52, no. 12, pp. 590-601, 2020.

[18] S. Sorooshain, N. F. Aziz, A. Ahmad, S. N. Jubidin, and N. M. Mustapha, "Review on performance measurement systems," *Mediterranean Journal of Social Sciences*, vol. 7, no. 1, 2016.

[19] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation," *International Journal of Machine Learning Technology*, vol. 2, no. 1, pp. 37-63, 2011.

[20] C. Goutte, and E. Gaussier, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," 2014.

[21] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation," 2014.

[22] S. Rajesh, P. Amudhavalli, and Dhanapriya A. P., "Prediction of air pollution using supervised machine learning," *Journal of Applied Information Science*, vol. 10, no. 1, pp. 10-16, 2022.

[23] Chinmaya H. S., J. M. Balaji, G. N. Sharma, and N. Divakar, "Dragonfly-net: Dragonfly classification using convolution neural network," *Journal of Applied Information Science*, vol. 10, no. 1, pp. 60-66, 2022.