



Automatic Classification Of Plant Leaf Disease Using A Novel Hybrid Approach Of LSTM-CNN

Kumud Pant Department of Biotechnology, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 kumud.pant@geu.ac.in

Bhasker Pant Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 pantbhasker2@gmail.com

Dibyahash Bordoloi Head of the Department, Department of Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand India, 248002 dibyahashbordoloi@geu.ac.in

Abstract: When a plant is growing, it is vulnerable to many diseases. One of the trickiest issues in agriculture is the early detection of plant diseases. The diseases must be detected early otherwise, they may negatively impact the overall production, lowering the farmers' profits. Many researchers have presented various cutting-edge frameworks relying on Neural Networks to address this issue. However, the majorities of these processes either have limited prediction performance or use massive amounts of input variables. In this system, a hybrid approach of Long Short Term memory (LSTM) and Convolution Neural Network (CNN) is used to predict and classify different types of plant leaf diseases automatically. With just 9,825 training parameters, the suggested system achieves 98.65% testing accuracy and 99.23% training accuracy. When comparison to other methods described in the literature, the proposed hybrid model calls for fewer training parameters. As a result, the time required for training this model for detecting the plant diseases automatically and the time needed to diagnose the plant disease using the trained mode both can decrease significantly. It helps to farmer to diagnose various diseases easily at an initial stage, so that they can go for different treatments and some preventive measures.

Keywords: Neural Networks, Long Short Term memory, Convolution Neural Network, LSTM, CNN.

Introduction

The backbone of our Indian economy is agriculture. Agriculture accounts for 10% of total GDP, with India accounting for 16%. In India, the agricultural sector either directly or indirectly supports nearly 70% of the population. Therefore, the production of high-quality, disease-free crops are crucial for the development of the nation's economy. Similar to

humans, different ripening stages or growth stages of plants are identified to different diseases. As a result, the overall production of crops and consequently the farmer's net profit are negatively impacted. The initial diagnosis of plant viruses is required to address this problem. Plant disease is manually detected by either farmers or agricultural scientists. But doing this requires a lot of effort and time.

In order to address this issue, several researchers in the world presented variety of cutting-edge processes for instantaneous disease detection in plants using various Learning techniques [1]. These cutting-edge systems employ a wide range of training variables. As a result, the training and prognostication times of these processes are extremely long, or they necessitate the use of a machine with greater computational power. Figure-1 shows the images of various plant leaf diseases. This study attempts to decrease the number of features used for prognostication using the LSTM network while maintaining classification performance of detecting plant diseases. This, in turn, substantially reduces the number of input variables, resulting in a reduction in training and detection accuracy.

The neuron architecture in the human brain serves as inspiration for deep learning techniques. These methods use Artificial Neural Network and their discrepancies, i.e. Convolutional neural network and Recurrent Neural Networks, to find hidden frameworks in the data. Compared to machine learning techniques, deep-learning methods have two major advantages [2]. To begin with, they do so automatically, doing away with the requirement for a completely separate feature extraction module. Second, processing large datasets with numerous dimensions takes less time when using Deep Learning techniques. Deep Learning methods are consequently suggested. Due to their efficiency with image data, DL techniques like CNN and LSTM are frequently used in applications involving computer vision.



Figure-1: Various types of plant leaf diseases

These two methods use convolution operations for extracting diverse temporal and features from various image data. CNN used to assign input-images to the appropriate class, whereas LSTM reduce dimensionality of images efficiently [4]. Here suggests a novel-hybrid approach for fully automated plant disease prediction based on LSTM and CNN with very few training images, in contrast to other cutting edge systems that really are currently available in the field. Although there are many methods for automatically detecting plant diseases in the literary works, as far as we know, no scientific work has still proposed a combination type of LSTM with CNN [10]. The number of features can be minimized using LSTM in the proposed works; there is a decrease in feature methods training [11]. This paper is structured in this way as follows, section-2 discuss about literature review. The proposed methods and working procedures are explained in section-3. Implementation and results-discussion in section-5. Conclusion is at the last session of this paper.

2. Related works

In 2016, He et al., created a mobile application that allows farmers to upload leaf images to quickly identify diseases in banana plants, where pictures are taken with the help their smartphones [3]. With 99% confidence, InceptionV3 algorithm used in the mobile application to predict diseases. According to the ResNet paper, their model performed best among other is ResNet152, which utilise 60 million training parameters to artificially increase the dataset for their research, they likewise utilized data augmentation [12].

According to their paper, the model of VGG-19 surpassed the InceptionV3 where 97% accuracy in training and 94% accuracy in testing. In 2015, according to Simonyan and Zisserman et al., discussed their good working method, VGG-19, used 150 millions of training parameters. It achieves greater accuracy comparatively [2]. In 2018, Ferentinos et al, used different modern CNN-architectures, including AlexNet, AlexNetOwTbn, GoogLeNet [8]. Plant diseases were identified using the Plant Village dataset and VGG [6].

In 2016, based on the VGG-Net article, he found that VGG significantly outperform other Cnn models with a precision of 99% utilising 140 millions of trainable parameters were discussed by Himani [9]. Zaki et al. investigated the efficacy of the AlexNet and GoogLeNet CNN frameworks for identifying diseases in plants using the Plant-Village dataset [5]. They carried out 58 different experiments using 85 different working methodologies. The transfer learning technique i.e. GoogLeNet significantly outperform the competitive performance with an accuracy of 99.3%. GoogLeNet used ~ 8 millions of training parameters, according to Feng et al., in 2015 [7].

3. Proposed Approach

The proposed work depicts, a detection of diseases in plant automatically is made possible with a novel hybrid method of LSTM and CNN. This method can identify the plants disease

with very less time consumption. This technique notifies the area where the disease is caused whether it is in the bottom of leaf or at the top of leaf etc., by analyzing and processing the leaf image. The proposed system is divided into three modules, which are classified as shown in the Figure-2.

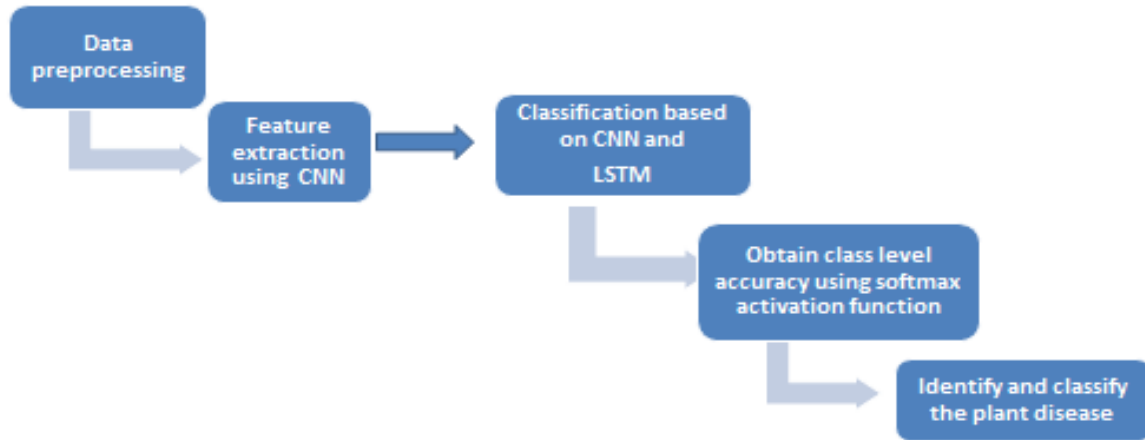


Figure-2: Workflow of proposed model

3.1 Data preprocessing

The input images are pre-processed to improve the contrast of the image for segmentation in better way, so that the accuracy classification will get increase. Resizing the image into 256*256 is the first process. Then the RGB image i.e. color images are converted into gray-scale image using the RGB to grey conversion method. The histogram is then normalized using the contrast enhancement method to improve image contrast. The intensities would be better distributed using this technique. The benefit of this technique is that it is simple. A Histogram function are defined as,

A discrete function of histogram is defined in the equation (1),

$$H(a = k) = mk \quad (1)$$

Histogram normalization is done in the equation (2)

$$p(a = k) = \frac{mk}{m} \quad (2)$$

Where,

$p(a=k)$ is histogram normalization function.

mk is intensity (k) in each pixels

n is a number of an image pixel in total

3.2 Feature extraction using CNN

The main process of feature extraction is dimensionality reduction. Initially the pre-processed input images are taken, CNN is used to extract the features from the given image, where a raw data is converted and reduced to more manageable parts, in order to do the process easily.

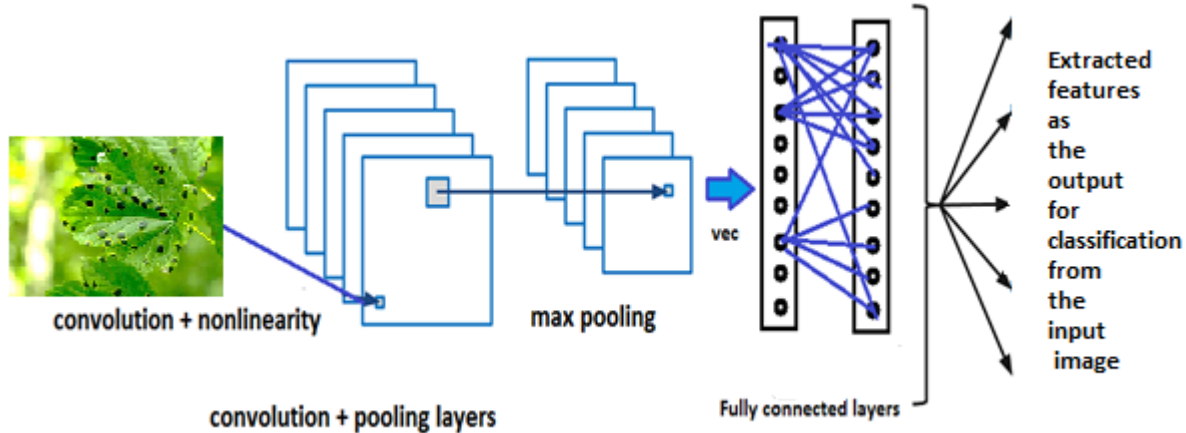


Figure-3: Feature extraction process using CNN

The convolution layer and pooling layer helps to extract the features at first. Then it send to the max pooling layer, Finally we get an output as the extracted feature of input image as shown in Figure-3. These layers send the images to fully connected layer as a vector values which contains lots of hidden layers to process and extracts the features. Finally we got an output of extracted features for classification from the given input images of plant leaves. From this output we can easily classify the plant leaf images whether it is affected by diseases or not.

3.3 Classification using LSTM-CNN

A good architecture can provide more accurate classification results rapidly. Recently, network types such as recurrent neural network, convolutional neural network, Generative adversarial networks etc. are being used to obtain good accuracy. Here in this proposed structure, we create a merging of CNN and LSTM networks, where CNN extract the details from the images and LSTM serves as a separator to obtain higher accuracy of plant leaf diseases prediction. Here the Figure-4 shows the good normal leaf and bacterial affected diseased leaf.



Figure-4: Classification of normal and diseases leaf

CNN is the most widely used network for feature extraction for detection of plant leaf diseases. LSTM-CNN is used to produce good accuracy for classifying the diseases. The 20 layers of the network are divided into 13 flexible layers, 6 integration layers, 2 FC layer, 2 LSTM layer, and 2 output layers with a Softmax function. Each convolution block consists of 2 or 3 CNN's, one layer of integration, and a drop-out layer. Extract features using softmax function, a flexible layer with 3*3 character size is used. By using a max-pooling layer with 2x2, the input image size is reduced as shown in Figure-3. Finally using the concept of LSTM-CNN, we get a better accuracy of 99.35% when compared to all other techniques.

4. Results Analysis

The proposed hybrid model of CNN-LSTM used to classify whether the leaf is diseased leaf or healthy leaf efficiently. The suggested model outperform with other existing models such as CNN, DC-GAN, C-GAN, ResNet50, UNet,etc., The below table shows the comparison of accuracy obtained from various learning algorithms. By this comparison, CNN-LSTM model achieves greater accuracy and precisions.

Table-1: Comparison of various models

Models Used	Accuracy	Precision
CGAN	96.5%	97.2%
CNN	98.3%	97.7%
ResNet-50	98.5%	98.1%
DC-GAN	97.3%	97.5%
UNet	97.3%	96.9%

LSTM-CNN	99.3%	98.9%
----------	-------	-------

Here Table-1 shows the comparison of various model based on the classification accuracy. From these models we conclude that LSTM-CNN achieves better. The performance chart for this comparison is shown in Figure-5. It shows a clear diagrammatic view of the comparison with various models proposed.

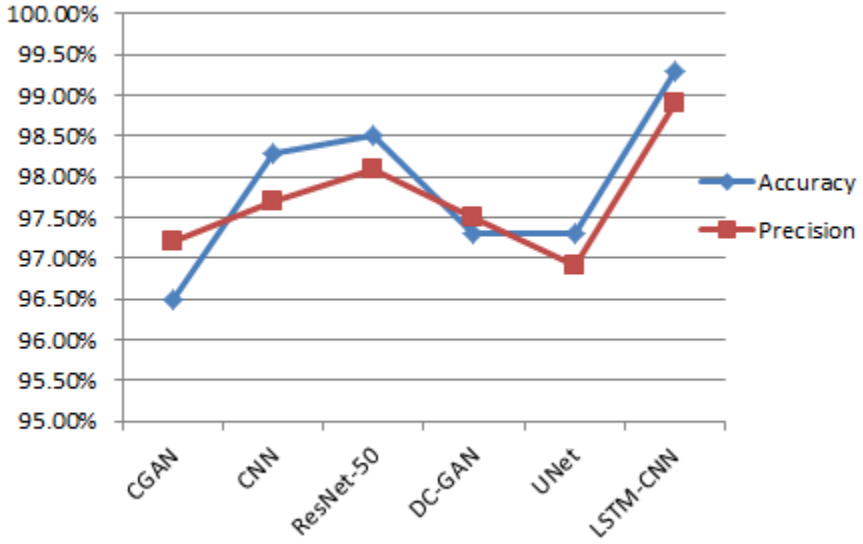


Figure-5: Performance graph of various models.

With just 9,825 training parameters, the suggested system achieves 98.65% testing accuracy and 99.23% training accuracy. In comparison to other methods described in the literature, the proposed hybrid model calls for fewer training parameters achieves greater accuracy.

5. Conclusion

Plant leaf disease detection at the early stage is more challenging task. A novel hybrid model of LSTM-CNN classifies the plant-leaf images as healthy leaf and unhealthy leaf effectively. Initially CNN extract the features and LSTM-CNN combined to classify the image. In order to get the output rapidly we used a softmax activation function. As a result, the time required for training this model for detecting the plant diseases automatically and the time needed to diagnose the plants diseases using trained mode both can decrease significantly. It helps to farmer to diagnose various diseases easily at an initial stage, so that they can go for different treatments and some preventive measures. By using this method we achieve an accuracy of 99.3% and precision of 98.9%. In future, the plant leaf images are saved in the cloud in order to reduce the latency as well as the privacy of the farmers.

References

4226 | Kumud Pant Automatic Classification Of Plant Leaf Disease Using A Novel Hybrid Approach Of LSTM-CNN

- [1] Ahmed, K., Shahidi, T.R., Irfanul Alam, S.M., Momen, S., 2019. Rice leaf disease detection using machine learning techniques. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE, Dhaka, Bangladesh, pp. 1–5.
- [2] Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. arXiv 1–14 preprint arXiv. Singh, V., 2019. Sunflower leaf diseases detection using image segmentation based on particle swarm optimization. *Artific. Intellig. Agric.* 3, 62–68. <https://doi.org/10.1016/j.aiia.2019.09.002>.
- [3] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Las Vegas, NV, USA, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- [4] Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S., Bhardwaj, S., 2020. Potato leaf diseases detection using deep learning. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, Madurai, India, India, pp. 461–466. <https://doi.org/10.1109/ICICCS48265.2020.9121067>.
- [5] Zaki, M.J., Meira Jr., Wagner, 2014. *Data Mining and Analysis: Fundamental Concepts and Algorithms*. 2nd ed. Cambridge University Press, Cambridge, United Kingdom. P. Bedi and P. Gole *Artificial Intelligence in Agriculture* 5 (2021) 90–101 101
- [6] Szegedy, C., Liu, Wei, Jia, Yangqing, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Boston, MA, USA, pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>.
- [7] Feng, D., Feng, M., Ozer, E., Fukuda, Y., 2015. A vision-based sensor for noncontact structural displacement measurement. *Sensors* 15 (7), 16557–16575. <https://doi.org/10.3390/s150716557>.
- [8] Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145, 311–318. <https://doi.org/10.1016/j.compag.2018.01.009>.
- [9] Himani, 2014. An analysis of agriculture sector in Indian economy. *IOSR J. Human. Soc. Sci.* 19 (1), 47–54. <https://doi.org/10.9790/0837-191104754>.

- [10] Hughes, David P., Salathe, M., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv 1–13 preprint arXiv.
- [11] Islam, M., Dinh, Anh, Wahid, K., Bhowmik, P., 2017. Detection of potato diseases using image segmentation and multiclass support vector machine. 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE). IEEE, Windsor, ON, Canada, pp. 1–4. <https://doi.org/10.1109/CCECE.2017.7946594>.
- [12] Jearanaiwongkul, W., Anutariya, C., Andres, F., 2018. An ontology-based approach to plant disease identification system. Proceedings of the 10th International Conference on Advances in Information Technology - IAIT 2018. ACM Press, New York, New York, USA, pp. 1–8. <https://doi.org/10.1145/3291280.3291786>.