Leaf Disease Detection and Classification

Conference Paper · April 2020

2 authors, including:

Samhitha Suda
University of Texas at Dallas

Leaf Disease Detection and Classification

Samhitha Suda
Department of Electronics and Communication
Sreenidhi Institute of Science and Technology
Hyderabad, India
sudasamhitha@gmail.com

Dr. C N Sujatha
Department of Electronics and Communication
Sreenidhi Institute of Science and Technology
Hyderabad, India
cnsujatha@sreenidhi.edu.in

Abstract— Pests damage plants and crops, this impacts the country's agricultural production. Farmers or specialists typically notice the plants with their bare eyes for disease detection and recognition. But this approach can be time-consuming, costly and unreliable. Automatic detection using various image processing techniques provides fast and accurate performance. This paper is concerned with a new methodology to plant disease recognition model by the use of convolutional neural networks, centered on the classification of the leaf pictures. New ways of training and the methodologies (Tensorflow, Keras) used enable a quick and easy model employment in practice. All the critical measures required to incorporate this model of disease identification are thoroughly outlined in the paper, beginning with collecting photos to build a database, a deep learning system for performing deep CNN training. To test and train the model, we use updated VGGNet architecture and transfer learning process. We also implemented the collection of validations. Where the only research is to improve accuracy and find the best model, then we deploy it to work. From the extracted features in the neural network we get the approx. 98 percent accuracy for corn crops. The project presents diagnosis of leaf disease using image processing techniques for the automated vision system that is used in the agricultural sector. The proposed decision-making method incorporates characterization of image content and with convolutional neural network.

Keywords— Deep Learning, Plant disease, CNN, VGGNet, Tensorflow, Python, Keras, Accuracy.

I. INTRODUCTION

The origins of agriculture in India dates to the civilization in the Indus Valley and to some areas of South India before that. India is the second largest agricultural producer worldwide. Agriculture employed more than 50% of India's population in 2018 and accounted for 17–18% of its national GDP. Agriculture is the centerpiece of the economy of India. The largest producer of pulses, rice, wheat, spices and spice products in the world is located in India. Economic growth for farmers depends on the quality of the products they make, which are dependent on plant growth and yield. The identification of diseases in plants therefore plays a crucial role in agriculture. Plants are particularly susceptible to diseases impacting plant growth and in effect impacting the farmers' ecology. Plant disease signs are visible in different parts of the plant, including leaves, etc. Manual plant disease detection with leaf pictures is an intricate task. Also there are no publicly accessible databases of plant diseases from the local area. There is also no real time algorithm that detects the crop disease accurately. Aim of this project is to reduce the crop loss faced due to the ignorance or lack of knowledge of the diseases that attacked the crop, automatic disease detection methods to use in order to identify a plant disease at an early stage, consequently computational methods must be developed which automatically detect and classify the disease using leaf pictures, apply the public data collection to the Regional plant-disease list and to provide a real-time plant disease detection algorithm. Thanks to the booming field of Artificial Intelligence, the most significant reasons why object detection has gained more popularity in recent years. In the deep learning domain itself, the approach to object detection can be obtained a high precision rate in the last few years. Various models of the principle of deep learning are used to train and test networks for Object Detection Method. AI has several research areas, but deep learning detection of objects is a trend to develop robust object detection systems for different applications. In this project, we considered that, depending on the application or need, each architecture should be capable of being merged with any feature extractor. Some trade or cash crops, cereal crops, and vegetable crops and fruit plants such as corn, apple, cherry, grape, tomato, eggplant and groundnut can be considered. This project is mainly concerned with the diseases in the corn crops and their detection with highest accuracies that can be achieved. These leaves images are chosen for our purpose. Early identification of plant leaf diseases can be a helpful source of knowledge for effective disease prevention, plant growth techniques, and disease management measures to deter disease production and dissemination by the use of the appropriate pesticides. Since it is an automatic disease detection algorithm rather than a manual model, it processes large number of inputs and gives output regarding the diseases in an inconsiderable time.

II. LITERATURE SURVEY

Jiang Lu, et al. [1] identify their project as an automated in-field diagnosis program of wheat disease, based on a weekly supervised framework of deep learning, which integrates detection for wheat diseases and localizes disease zones with only annotations of image level to wildlife-related training images. In addition, the Wheat Disease Database 2017 (WDD2017) new infield image data collection for wheat diseases is being collected to verify the system's performance. Under two architectures, i.e. The mean accuracies of VGG-FCNVD16 and VGG-FCN-S, which have been surpassed by 93.28 percent and 73.01 percent by two traditional CNN sets, are 97.94 percent and 95.11 percent respectively for five folds over cross-validation on the WDD2017 VGG-CNN-VD16 and CNN-S-VGG The experimental results show that, while retaining accurate monitoring for the corresponding disease areas, the device proposed exceeds the traditional CNN architectures for recognition accuracy in a comparable amount. In addition, the program proposed for the treatment of agricultural diseases has been developed into a smartphone device for the first time. A 40 research efforts using profound learning methods applied to various challenges in farming and food production were addressed and surveyed in paper [2]. Examine the specific agricultural issues under examination, the different modeling and framing of the sources used, the design and pre-processing of the data used and the overall performance of the sample. Comparisons are also being studied regarding the classification and regression performance of deep learning and other common techniques. Findings show that deep learning offers high precision, surpassing current widely used strategies of image processing. In paper [3] models of the neural convolutional network for the identification and diagnosis of plant diseases with simple photographs of healthy and diseased plants through methods of deep learning. Model training took place with an open database of eighty seven thousand eight forty nine photos of twenty five different plants in a collection of fifty seven different categories, including good plants. Several models have been educated, with the best performance achieving a success rate of ninety eight percent in identifying the corresponding combination of [plant, disease] (or healthy). The high success rate of this model makes this a valuable early warning or consulting tool and a way to support an integrated framework for the identification of plant diseases in working farm circumstances. A methodology for the detection of early and precise plant diseases, using artificial networks and various image processing techniques is defined in the paper [4]. As the proposed method is based on the Artificial neural network classifying and the Gabor feature removal filter, better results are obtained with a recognition rate of up to ninety one percent. A classifier based on Artificial neural network is used to classify various plant diseases and to combine them with textures, color and morphology. The paper [5] introduced an important approach such as K mean clustering, texture and color analysis for the identification of diseases in Malus domestica. This uses the texture, color characteristics that normally occur in natural and affected zones to distinguish and identify specific agriculture. In paper [6] the efficiency of traditional multiple analysis, artificial neural network (BPNN) and supporting vector machine (SVM) were examined by the authors. A statistic approach based on the SVM was concluded that the relationship between environmental conditions and the severity of diseases could be better defined in order to manage diseases. Authors submit an examination in paper [7] and recognize the need to develop a swift, economical and trustworthy health monitoring sensor to facilitate agricultural progress. Researchers discussed techniques currently being used to build a ground sensor network for monitoring the health and disease of plants in field conditions, including spectroscopic, photographic and dynamic profiling-based methods for detecting plant diseases. Upon reviewing the study and evaluating the findings of their work on papers [8–11], it has been determined that image processing approaches such as double-stranded ribo-nucleic acid (RNA) analytics, nucleic acid samples and microscopy should also be applied to known diseases among other approaches widely used for plant disease diagnostics. Many methods for the identification of plant diseases by computer vision are currently in use. One is the identification of diseases by removing color as described in paper [12]. This paper used coloring models for YcbCr, HSI and CIELB, and thus the noise from different sources, for example camera flash, was observed in disease spots and not affected. Furthermore, the method of extracting shape features may achieve plant disease detection. In sugarcane blades where threshold segmentation for determining the area of the leaf and the triangulate threshold for the lesion region was used by Patil and Bodhe for the detection of the disease, they obtained an average 98.59% accuracy in the final experiments [13]. The texture extraction feature can also be used for the identification of plant diseases. Patil and Kumar suggested a model for the detection of plant disease by measureing the gray level co-occurrence in a picture using texture characteristics such as inertia, homogeneity and correlation [14]. They experimented with the identification of diseases on maize leaves in conjunction with color extraction. The integration of all of these features offers an extensive set of features to enhance and classify images. A study of well-known conventional methods for extracting characteristics was presented in paper [15]. The research in this article focuses on the application of these methodologies and techniques, while artificial intelligence sciences (AI) are rapidly advancing.

III. REQUIRED TECHNOLOGIES & SOFTWARE

A. *Deep Learning*

Deep learning is a particular subset of Machine Learning, which is an Artificial Intelligence sub-set. Artificial Intelligence is the broad domain of computer production that can think intelligently. Machine learning focuses on the development of computer programs that can access and use the data to learn for themselves. Deep learning is a set of algorithms used in machine learning utilizing multiple nonlinear transformations for top-level abstractions in data through the use of model architectures.

B. *Convolutional Neural Networks*

A Convolutional Neural Network is one of a Deep Learning algorithms capable of taking an input image, assigning value to various aspects / objects in the image, and being able to differentiate between them. Such are also known as shift invariant or space invariant artificial neural networks (SIANN), based on the characteristics of the design and invariance translation of their shared weights.

C. *Object detection*

Detection of objects has a very wide field of deep learning. Object recognition using the convolution neural network used in many fields of real time. To detect the object in profound learning, it uses the deep learning neural network. In deep learning, identification of objects simply means recognizing objects from pictures.

D. *Tensorflow*

Tensorflow is an open source software library for numerical computation with high accuracy. Initially developed by Google Brain team researchers and engineers in the Google AI organization, it provides strong support for machine learning and in-depth learning and the scalable digital computing foundation is used in many other science fields.

E. *Keras*

Keras is a high-level neural network API which can run on Tensorflow, Theano, and CNTK and is written in Python language. It was designed with a focus on allowing for easy experimentation. Keras takes into account quick and simple prototyping (through ease of use, measured quality, extensibility). Recurrent and Convolution systems both are bolstered. Runs on both GPU and CPU.

F. Anaconda

Anaconda is an open source and free R and Python programming language distribution that is designed to simplify package management and implementation. Package versions are managed through a package management system.

G. Python

Python is a language of programming which is multi paradigm, general purpose, and interpreted, high level. Python helps programmers to build basic or complicated programs utilizing various programming types, get better results and compose code just as if communicating in a human language.

H. VGGNet

VGG networks are developed by Simonyan et al with the idea of deeper CNNs. There are two main factors which distinguish VGG networks from AlexNet: network depth and filter size. It incorporates multiple network depth settings. It is a deep convolutional network used for object recognition. VGGNet parameters range in between 133M–144M.

IV. Proposed System

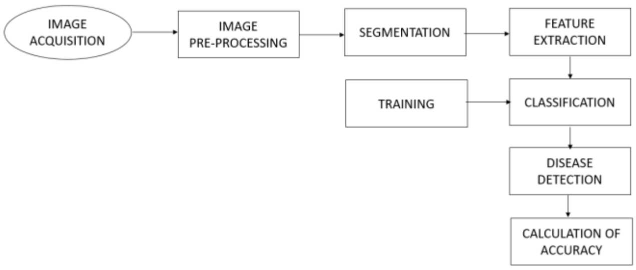

Fig. 1. Block diagram for leaf disease detection and classification

A. Image Acquisition

This is the advancement of a digitally encoded portrayal of the visual characteristics of the object, such as the physical scene or the internal structure of the object. Often the word means or involves encoding, compressing, storing, printing, and displaying these images. The potential for digitally continuous copying and copying of copies without any reduction in image quality is a major benefit of a digital image over an analog image such as a film photograph.

B. Image Pre-Processing

Pre-Processing is a term often used for operating images at lowest level of extraction, here both the input and output images are intensity images. Intensity image is a matrix containing data which represents intensities inside a range. Preprocessing aims to enhance image data that removes undesirable distortions or improves other image features that are necessary for further processing.

C. Segmentation

The segmentation of images is a method of splitting a digital image into multiple sections (pixel sets, also known as image objects). The objective of segmentation is to simplify and/or alter the portrayal of an image that is much more vivid and easier to interpret.

D. Feature extraction

It is a process in which the initial set of raw data is decreased to more manageable groups for processing by reducing the size. The basic feature of such massive datasets is a large number of variables that require a great deal of computing properties.

E. Training

Deep learning models of the neural network learn to map inputs to outputs, given a set of example collection of training data. The training method includes finding in the network a set of weights that appears to be sufficient or sufficient enough to solve the particular problem.

V. Design And Implementation

An image of a leaf with corn disease is uploaded to the model system. Image acquisition is carried out on it there by producing a digitally encoded representation of the image. This image is then pre-processed i.e. converted into an intensity image. Intensity image is a data matrix which represents intensities within a range. Objective of pre-processing is to remove any kind of distortions and improve other intensity feature at the output end. Segmentation is carried out on this image where the digital image is separated into pixels. This is done to shorten the representation of image which can be easily analyzed further. In feature extraction we reduce the dimensions of image, where the initial set of data can be grouped which are manageable while processing. Classification of these groups takes place by subjecting them to training. Training in deep learning is a process where inputs are mapped to the outputs on the basis of a given set of training data. After the mapping is done, the disease of the leaf is detected and therefore accuracy is calculated by using the existing formulae. If the input image is not of a corn leaf then arises the false positive and true negative conditions which are the type I and type II errors, can be studied further deeply in the machine learning.

VI. Results And Discussion

Fig. 2. Image before and after segmentation

Figure 2 shows the input image ‘a’ and the image after segmentation process i.e. ‘a_marked.’ During the segmentation process, the digital input image is separated into pixels, also the background of the image is blurred and noise is removed.

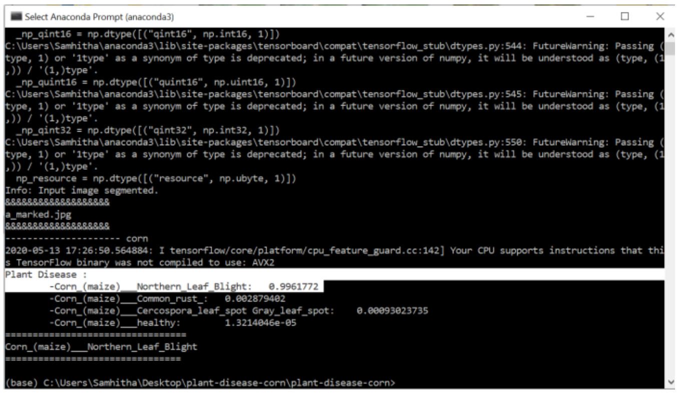

Fig. 3. The disease is detected

Figure 3 shows that input leaf image is a corn leaf suffering from the northern leaf blight disease and also the accuracy with which the model detected is given as 99.617 %. This means that of all the images that the system has been trained with, it can detect that the input image is 99.617% precise to the northern leaf blight disease and since it has the highest percentage of all the other diseases, the system declares it to be the attacking disease.

VII. Conclusion

Detection of objects with convolution neural network is commonly used in the developing generation. These have various uses in medicine and in agriculture. The accuracy can be improved with the increase or decrease in the number of convolution layers. CNN validation is a kind of production environment where we can assess accuracy and enhance accuracy. Product implementation stage will be after the accuracy measurements in test model and final stage. The proposed system has been developed taking into account the benefits of farmers and the agricultural sector. The developed system can detect plant diseases and can also provide remedies that can be used against the disease. In order to improve the health of the plant, proper knowledge of the disease and the cure can be used. The proposed system is based on python and provides an accuracy of around 98 percent for every disease detected. Accuracy and speed can be increased by using Google's GPU for processing. The system can be installed on Drones so that aerial monitoring of crop fields can be carried out. Our main aim of the project is to reduce the burden on farmers by acknowledging them with the diseases that attacked their crop in early stages before spraying the wrong pesticide to their crops resulting in further harm to the crops. By using our algorithm we can further include new plant disease pairs to extend the vastness of the system and can be used for various crops. Also the proposed system is a cheaper way to avoid loss of crops and enlighten the farmers. Since India is a country that hugely depends on agriculture i.e. our exports in the agriculture sector form a huge difference in the GDP of the country, we intend to protect the crops by proposing this model owing to the present day agricultural scenario in India.

References

[1] Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, Changshui Zhang, “An in-field automatic wheat disease diagnosis system”, Computers and Electronics in Agriculture, Vol. 142, Pages 369–379, 2017.

[2] Andreas Kamilaris, Francesc X. Prenafeta-Boldu, “Deep learning in agriculture: A survey”, Computers and Electronics in Agriculture, Vol. 147, Pages 70–90, 2018.

[3] Konstantinos P. Ferentinos, “Deep learning models for plant disease detection and diagnosis”, Computers and Electronics in Agriculture, Vol. 145 Pages 311–318, 2018.

[4] Kulkarni Anand H, Ashwin Patil RK, “Applying image processing technique to detect plant diseases”, Int J Mod Eng Res, Vol. 2, Issue 5, Pages 3661–3664, 2012.

[5] Bashir Sabah, Sharma Navdeep, “Remote area plant disease detection using image processing”, IOSR J Electron Commun Eng ISSN, Vol 2 Issue 6, Pages 2278-2834, 2012.

[6] Rakesh Kaundal, Amar S Kapoor and Gajendra PS Raghava “Machine learning technique in disease forecasting: a case study on rice blast prediction,” BMC Bioinformatics, 2006.

[7] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, “A review of advanced techniques for detecting plant diseases,” Computers and Electronics in Agriculture, vol. 72, Issue 1, Pages1–13, 2010.

[8] P. R. Reddy, S. N. Divya, and R. Vijayalakshmi, “Plant disease detection techniquetool—a theoretical approach,” International Journal of Innovative Technology and Research, Pages 91–93, 2015.

[9] A.-K. Mahlein, T. Rumpf, P. Welke et al., “Development of spectral indices for detecting and identifying plant diseases,” Remote Sensing of Environment, vol. 128, Pages 21–30, 2013.

[10] W. Xiuqing, W. Haiyan, and Y. Shifeng, “Plant disease detection based on near-field acoustic holography,” Transactions of the Chinese Society for Agricultural Machinery, vol. 2, article 43, 2014.

[11] A.-K. Mahlein, E.-C. Oerke, U. Steiner, and H.-W. Dehne, “Recent advances in sensing plant diseases for precision crop protection,” European Journal of Plant Pathology, vol. 133, Issue 1, Pages 197–209, 2012.

[12] P. Chaudhary, A. K. Chaudhari, A. N. Cheeran, and S. Godara, “Color transform based approach for disease spot detection on plant leaf,” International Journal of Computer Science and Telecommunications, vol. 3, Issue 6, Pages 65–69, 2012.

[13] S. B. Patil and S. K. Bodhe, “Leaf disease severity measurement using image processing,” International Journal of Engineering and Technology, vol. 3, Issue 5, Pages 297–301, 2011.

[14] J. K. Patil and R. Kumar, “Feature extraction of diseased leaf images,” Journal of Signal & Image Processing, vol. 3, Issue 1, Pages 60, 2012.

[15] T. R. Reed and J. M. H. Dubuf, “A review of recent texture segmentation and feature extraction techniques,” CVGIP: Image Understanding, vol. 57, Issue 3, Pages 359–372, 1993