E-AGRI KIT AGRICULTURAL AID USING DEEP LEARNING

[1]Pokkili Sampath Kumar, [2]Muthe Akhil Kumar,[3] Reddy Rudhramsh Reddy,
[4]Guntoja Kalpana Devi

[1,2,3]B.Tech Student, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad, Telangana, India. sampathpokkili@gmail.com*,akhilvarmamuthe@gmail.com ,reddyrudhramsh@gmail.com*

[4]Assistant Professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad,Telangana, India.gkalpana15@gmail.coms.

Abstract- *This project presents an agricultural aid application, developed and designed, to help farmers by utilizing Image Processing, Machine Learning and Deep Learning concepts. Our application provides features such as early detection of plant disease, implemented using various approaches. After evaluation, results showed that Convolutional Neural Network was per forming better for plant disease detection with an high accuracy. It further helps the farmer to forecast the weather to decide the right time for agricultural activities like harvesting and plucking. To avoid reoccurrence of disease due to loss in soil minerals, a crop-specific fertilizer calculator is incorporated which can calculate the amount of urea, diammonium phosphate and muriate of potash required for a given area. This project showcases an agricultural aid app that was built and designed to assist farmers by employing Image Processing, Machine Learning, and Deep Learning. Features like early detection of plant disease are available in our application and are implemented in a number of ways. It was determined that Convolutional Neural Network was superior for detecting plant diseases with a high degree of accuracy. The farmer can use the weather forecast to plan out agricultural tasks like harvesting and plucking at the optimal time. A crop-specific fertilizer calculator is in the works to determine how much urea, diammonium phosphate, and muriate of potash should be applied to a given area to prevent the recurrence of disease caused by depleted soil minerals.*

I. INTRODUCTION

According to a study by the Associated Chambers of Commerce and Industry of India, annual crops losses due to pests and diseases amount to Rs.50,000 crore ($500 billion), which is tantamount to a country where at least 200 million go to bed hungry every night. Agriculture being a vital sector has a majority of the rural population in developing countries relying on it. The sector is faced by major challenges like unprecedented pest attack and unforeseen weather conditions affecting their produce leading to major loss of food and effort. Technology plays a vital role in uplifting the livelihoods of the rural populace which can be done by using a simple agro-android application system.Plant diseases can affect vast produce of crops posing a major menace to food security as well as leading to major losses to farmers. An extensive review of existing research was conducted by us on this domain and in an effort to help farmers overcome this problem, we have designed an android application, Agricultural Aid which utilizes machine learning to provide plant disease detection. This detection is combined with an android application which provides features like weather forecast of up to 7 days, fertilizer calculator and language translation in up to 4 languages which has been implemented and integrated using Android Studio and its APIs. For disease classification, we followed two approaches: Image Processing with Machine Learning and Deep Learning models.

The first approach i.e. Image Processing approach usually includes multi-step preprocessing techniques such as: Filtering, color space conversion, thresholding and finally, contouring to mark out the infected region. These methods can be used with Machine Learning concepts to provide classification of infected regions. However, the accuracy for such methods isn’t very high. As an alternative to these steps, “GrabCut” Algorithm can also be used which is an optimized method of foreground extraction to eliminate background noises using minimal user interaction [4]. It has better accuracy in terms of background elimination and can be used for better classification however for the time being this method wasn’t used in the application but can be incorporated in the future to improve accuracy. For the second approach i.e. Deep Learning approach, a deep neural architecture is used to train and test on leaf image databases to classify the disease. The paper provides a comparison of results obtained after applying Deep Learning Models such as CNN, ResNet-152 and Inception v3. In our agriculture aid,CNN Model is used to train and form an automated plant disease system based on images of leaves of both healthy and diseased plants.

II. EXISTING SYSTEM

Image Processing is a popular first step in the process of plant disease detection and often includes multi-step processes to achieve processed, ROI centric input images for further classification. In this approach, various Image Processing techniques were used on the input image to get a final output image which would mark the infected area and also calculate the percentage of area infected in the leaf. The major advantage observed in this approach was that this approach eliminated the need to extract the leaf and place it on a black background during image capture for the algorithm to work. In real time scenario, the infected leaf would be present amongst a cluster of mixed crops and this approach was able to segregate the infected leaf from the healthy ones in an input image.

III. PROPOSED SYSTEM

In the Deep Learning Approach, we decided to take a subset of the Plant Village dataset along with the cotton dataset to train and test the CNN model. The input image is fed into this model, which initially took a portion of the Plant Village and Cotton Dataset, with a training - validation split of 70-30, therefore getting 4200 images for training and 1800 images for validation. A CNN (Convolutional Neural Network) is a deep learning model that takes inputs which are assigned weights depending on various features. CNN is a widely used neural network for image-based datasets.

Our CNN model consists of 4 main convolutional layers with 32, 64, 128, 128 filters consecutively, each followed by a ReLU activation function, max pooling and dropout layer. This set of convolutional layers is followed by a flatten and then a dense layer which is finally followed by the soft max activation function that tells us which class has the maximum probability.

IV. FEASIBILITY STUDY

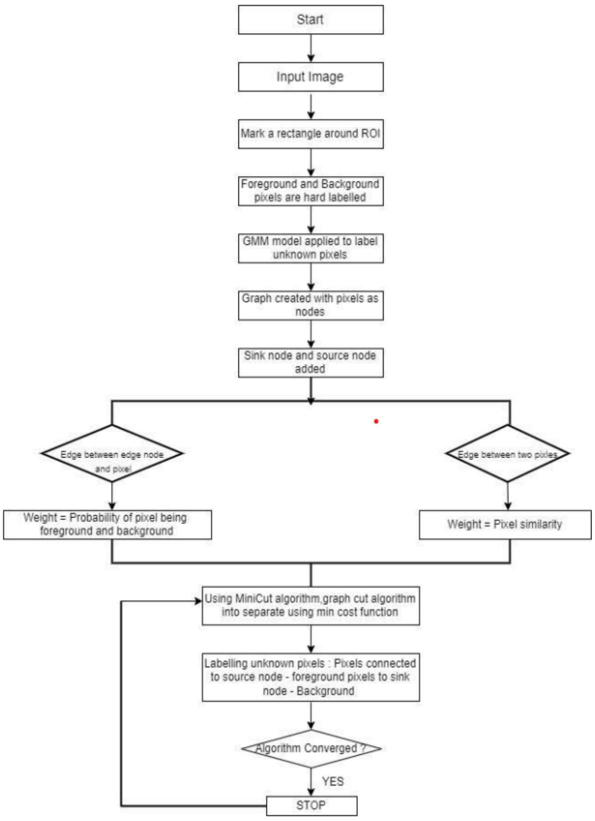
The feasibility of the project is analysed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system isnot a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are

*** ECONOMICAL FEASIBILITY :** This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

*** TECHNICAL FEASIBILTY :** This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

*** SOCIAL FEASIBILITY :** The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by thesystem, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods thatare employed to educate the user about the system and to make him familiar with it. His level of confidence must beraised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

V. ARCHITECTURE


Start
Input Image
Mark a rectangle around ROI
Foreground and Background pixels are hard labelled
GMM model applied to label unknown pixels
Graph created with pixels as nodes
Sink node and source node added
Edge between edge node and pixel
Edge between two pixles
Weight = Probability of pixel being foreground and background
Weight = Pixel similarity
Using MiniCut algorithm,graph cut algorithm into separate using min cost function
Labelling unknown pixels : Pixels connected to source node - foreground pixels to sink node - Background
Algorithm Converged ?
YES
STOP

After selecting the image for extraction, the user is required to mark out a rectangle selecting the region/object of interest (ROI) where everything outside the rectangle is taken as background. The computer labels the pixels as “foreground” and “background” pixels which is also known as hard labelling. After labelling, a Gaussian Mixture Model (GMM) is applied on the foreground and background. GMM learns and tries to label unknown pixels inside the rectangle as either probably foreground or background according to hard-labelled pixels and color statistic which is similar to clustering. Using this pixel distribution, a graph is created where pixels are denoted as nodes. Additionally, 2 nodes are added, namely: Source node and Sink node such that every foreground pixel is connected to source node and every background pixel is connected to sink node. The weight of the edges connecting pixels to source node is calculated as the probability of the pixel being a part of the foreground or the background. The weight of the edges connecting two pixels is defined by pixel similarity such that if there is a small difference in pixel color the weight will be high. Using a min-cut algorithm on the graph, the graph is cut into separate source node and sink node using minimum cost function where the cost function is sum of all the edges that are cut. After the cut, nodes (pixels) connected to source node are labelled as foreground and nodes(pixels) connected to sink node becomes background.

VI. MODULES

1.Tensorflow

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.Tensor Flow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

2.Numpy

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.It is the fundamental package for scientific computing with Python. It contains various
features including these important ones:

➢ A powerful N-dimensional array object

➢ Sophisticated (broadcasting) functions

➢ Tools for integrating C/C++ and Fortran code

➢ Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using Numpy which allows Numpy to seamlessly and speedily integrate with a wide variety of databases.

3.Pandas

Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem.

4.Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter Notebook, web application servers, and four graphical user interface toolkits. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatter plots, etc.,with just a few lines of code. For examples, see the sample plots and thumbnail gallery.For simple plotting the pyplot module provides a MATLAB-like interface, particularly when combined with IPython. For the power user, you have full control of line styles, font properties, axes properties, etc, via an object oriented interface or via a set of functions familiar to MATLAB users.

5.Scikit – learn

Scikit-learn provide a standard Python interface for a variety of supervised and unsupervised learning techniques. It is licensed under a BSD-style license that allows for both academic and commercial use, and is available on a variety of Linux versions.

VII. CONCLUSION

During our analysis, we have understood the need for efficient plant disease identification & classification algorithms and prevention methods. Due to a large number of crops and diseases available, it is crucial that the detection system should be able to adapt to the changing variables and trends. Hence, Machine learning and Deep learning approaches were employed for this project which ensures that the code trains itself against as many possible numbers of different crops and diseases as possible The paper consists of an android application covering plant disease detection and other functionalities such as language translation, weather forecasting and fertilizer calculator. With this application we aim to provide aid in the unprecedented agricultural activities and ensure a healthy plant. We have also tested our application on cotton dataset and performed real time analysis on a diseased tomato crop to ensure our model does not overfit and performs well in a live environment. In future, we aim to expand our dataset to include more varied types of crops and disease so that the algorithm can adapt better to real time conditions and provide wide coverage.

ACKNOWLEDGMENT

We thank CMR Technical Campus for supporting this paper titled “E-AGRI KIT AGRICULTURAL AID USING DEEP LEARNING”, which provided good facilities and support to accomplish our work. Sincerely thank our Chairman, Director, Deans, Head Of the Department, Department Of Computer Science and Engineering, Guide and Teaching and Non- Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work.

REFRENCES

1. CropLife International (May 2015). India's farmers fighting pests. Retrieved from:https://croplife.org/news/keeping-indias-pests-in-line/
2. Economic Times (Sept 2018). India sets record farm output target for 2018-19. Retrieved from: https://economictimes.indiatimes.com/news/economy/agriculture/indiasets-record-farm-output-target-for-2018-19/articleshow/65858058.cms
3. Sharada P. Mohanty David P. Hughes and Marcel Salathé."Using Deep Learning for Image-Based Plant Disease Detection."Front. Plant Sci., 22 September 2016.
4. Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. 2004. "GrabCut": interactive foreground extraction using iterated graph cuts. In ACM SIGGRAPH 2004 Papers (SIGGRAPH '04). Association for Computing Machinery, New York, NY, USA, 309–314.
5. R. Chapaneri, M. Desai, A. Goyal, S. Ghose and S. Das, "Plant Disease Detection: A Comprehensive Survey," 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA),Mumbai, India, 2020, pp. 220-225, doi: 10.1109/CSCITA47329.2020.9137779.
6. Raghavendra, B. K. (2019, March). Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), (pp. 313-316). IEEE.