Software Tool for Plant Pathometry & Estimation of Nutrient Content

Submitted in partial fulfilment of the requirements

of the degree of

Bachelor of Technology

by

Parv Bhatt (17115063)
Harshit Khetan (17115038)

Supervisor

Prof. Indra Gupta

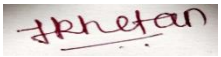
Department of Electrical Engineering

Indian Institute of Technology Roorkee

2021

Declaration

We hereby declare that the work which is presented here, entitled **Software Tool for Plant Pathometry & Estimation of Nutrient Content**, submitted in partial fulfilment of the requirements for the award of the Degree of **Bachelor of Technology** in the Department of Electrical Engineering, Indian Institute of Technology Roorkee. We also declare that we have been doing our work from August 2020 under the supervision and guidance of **Dr. Indra Gupta, Electrical Engineering Department, Indian Institute of Technology Roorkee**. The matter presented in this dissertation report has not been submitted by us for award of any other degree of institute or any other institutes.

Date: May 30, 2021 Signature:  Signature: 

Name: Parv Bhatt Name: Harshit Khetan

Enrolment No.: 17115063 Enrolment No.: 17115038

Certificate

This is to certify that the above statement made by the candidates is true to best of my knowledge and belief.

(Signature)

Dr. Indra Gupta

Associate Professor

Department of Electrical Engineering

Indian Institute of Technology Roorkee

Acknowledgement

We would like to thank Department of Electrical Engineering, Indian Institute of Technology Roorkee for having the course EEN-400B which gave us the opportunity to learn and explore our own idea. We would also like to show our gratitude towards our supervisor Dr. Indra Gupta for sharing her valuable opinions and guiding us throughout the project.

Table of Contents

List of Figures and Tables

ABSTRACT

Agriculture plays a critical role in providing food supply to the growing population of the world. It is estimated that, on average, annual global food supply loss due to plant diseases stands at 40%. In developing countries, smallholder farmers generate more than 80% of the agricultural production. For them, the loss of crops has devastating consequences. Sometimes, farmers can lose almost 100% of their crop due to plant diseases. This makes crop diseases a significant threat to food security around the world. Reliable, accurate assessments of nutrient content & disease intensity are critical for farmers & also for many research areas in plant pathology. We aim to employ deep learning tools to estimate the nutrient content & identify the intensity & type of disease in a plant by scanning its leaves & helping our fellow farmers by providing treatment options for the plant.

1. INTRODUCTION

We tried to understand what types of diseases occur in plants, mainly crops, and tried to learn how to determine the plants' nutrient content by examining the leaf of that particular plant. We reviewed some research papers and found out that for precise crop management, nitrogen (N) content management which determines the optimal amount of nitrogen for a specific location based on the yield potential, is the most frequently practiced operation. Its accurate assessment in plants is key to nutrient management. We used the Plant Village Dataset collected by Penn State University's research and development team. The dataset contains more than 54,000 images of 14 different crop species for example Apple, Tomato, Orange, Blueberry Corn, etc. The images include plants affected by 21 different kinds of diseases with 25% healthy leaves images. The data was pre-processed. The steps included converting images to grayscale, generating histograms of oriented gradients, creating the feature matrix, standardizing the matrix and applying PCA for feature reduction (15,876 fs to 970 fs).

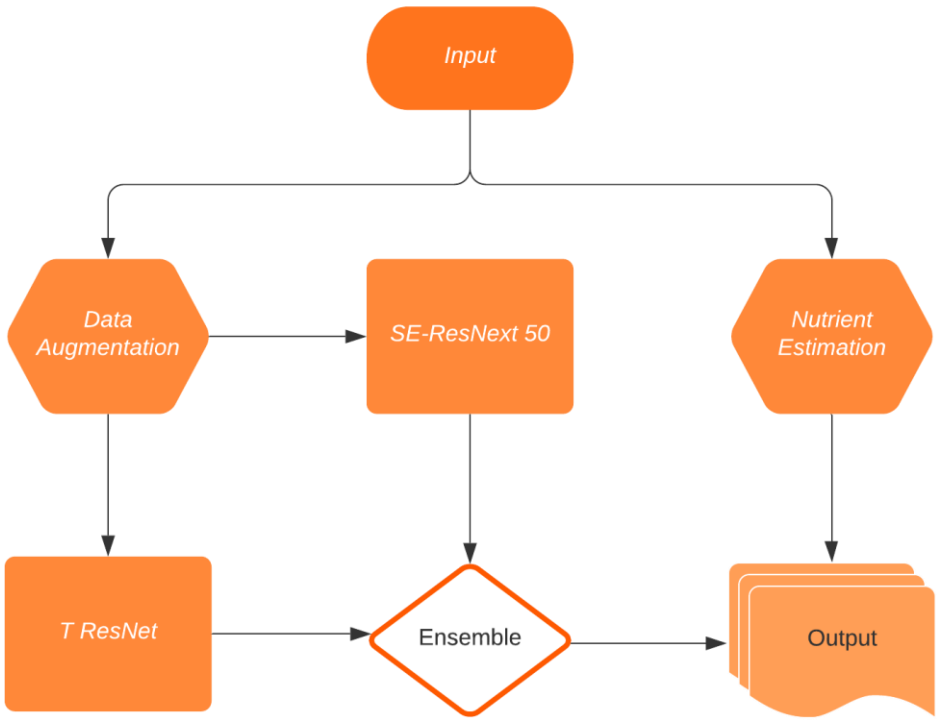


Fig. 1: *Flowchart of the Proposed Method*

2. PROPOSED METHOD

The flowchart of the proposed approach is given in Fig. 1. We use two base models. These base models' outputs are then combined through an ensemble model. In the remainder of this section, we describe these steps in more detail.

2.1 Base Models

We use two model frameworks: SE-ResNext 50 & T ResNet as base models. In the SE-ResNext 50 model, the prefix se is the process of squeeze and excitation. The principle of this process is to enhance the important features and weaken the unimportant features by controlling the size of the scale. It is the same as the attention principle. It is to make the extracted features more directional, so as to better recognize the fine features in the FGVC (Fine-Grained Visual Categorization) task.

TResNet provides a very good speed - accuracy - batch size trade-off for GPUs. We were able to train on input resolution 600x600 with batch size of 64 and training rate of 90 image/sec, which enabled us to experiment fast and efficient.

2.2 Ensemble of Models

The ensemble of models method is used here to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone. The ensemble model used was 0.6*Model1 + 0.4*Model2. This combination provided a very good speed – accuracy trade-off.

2.3 Nutrient Estimation

Digital image processing, as a tool to analyse the colour plant image and extract the important features of the image, was used in the current study. Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. It is an improvement of pictorial information for human interpretation and processing of image data for storage, transmission, and representation for autonomous machine perception. Initially, pre-processing of all images was done to enhance their visual quality, further to achieve the various features of colour images and to transform colour (RGB) images into normalized r, g, and b chromaticity

coordinates. The composite colour images were decomposed into red spectrum image (R), green spectrum image (G), and blue spectrum image (B) components. Subsequently, the images were also converted into hue, saturation, and intensity (HIS) coordinates to extract the intensity component. In order to extract the colour information and to obtain different features, the entire image features were segmented from its background using an automatic segmentation technique based on a modification of Otsu's algorithm. The automatic threshold technique selects a threshold to segment the background from the object. Thus segmented images were used to calculate the mean values of images. Four image features (mean, variance, average energy, and entropy) from each normalized 'r' and 'g' segmented image histogram of leaves were calculated. Fig. 2 shows the flow chart for extracting the colour image features.

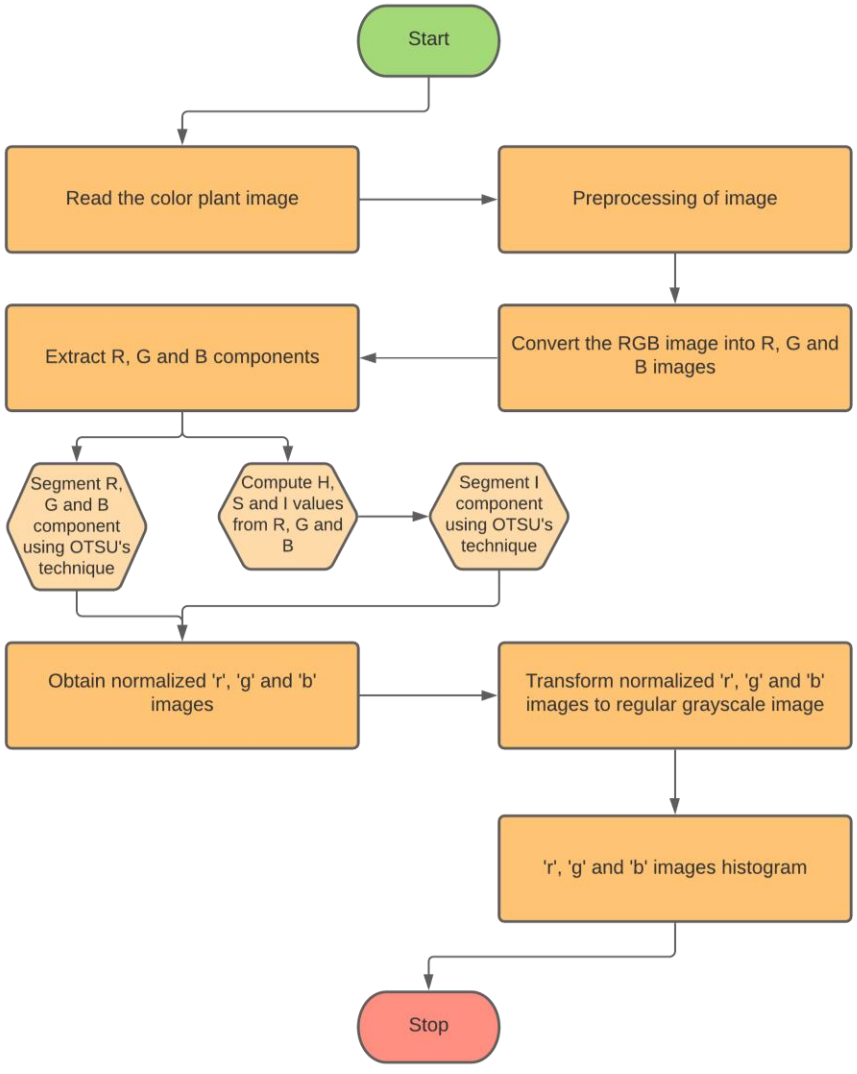


Fig. 2: *Flowchart of the Embedded Image Processing*

3. RESULTS & DISCUSSIONS

Method	F - Score
SE-ResNext 50	93.53
T ResNet	87.12
Ensemble	**94.13**

Table 1: *Experimental Results for Disease Detection*

Object detectors are generic & they aren't developed considering the domain-specific challenges. These networks also have many internal parameters & these parameters need to be tuned for the particular application. Hence, it isn't sufficient to use more advanced models and a comprehensive understanding of the characteristics of the data is of the essence.

We performed the Kjeldahl Method to determine the nutrient content of the plant. The nitrogen level obtained from the Kjeldahl method was then fed in the regression model along with the values of R, G, B, r, g and b to train the model. Total 140 samples were used to develop the prediction model and 80 samples were used for validation.

$$\text{Nitrogen Percent} = 8*10^{-5}*(r^2) - 0.0162*(r) + 0.8778$$

This trained model was used to estimate the nutrient content of other plants with an average error of 15.54%.

Fig. 3: *Heat Map*

4. CONCLUSIONS

In this study, we have trained two different object detectors for disease detection. We have used ensemble techniques to utilize both the individual networks. In this work, we have focused on using lighter networks & taken the ensemble of weak classifiers approach. The use of lighter networks made the hyper-parameter tuning possible in feasible periods & allowed us to experiment with various network parameters.

We have also developed an algorithm to extract important features from a digital image to estimate the nutrient content of a plant.

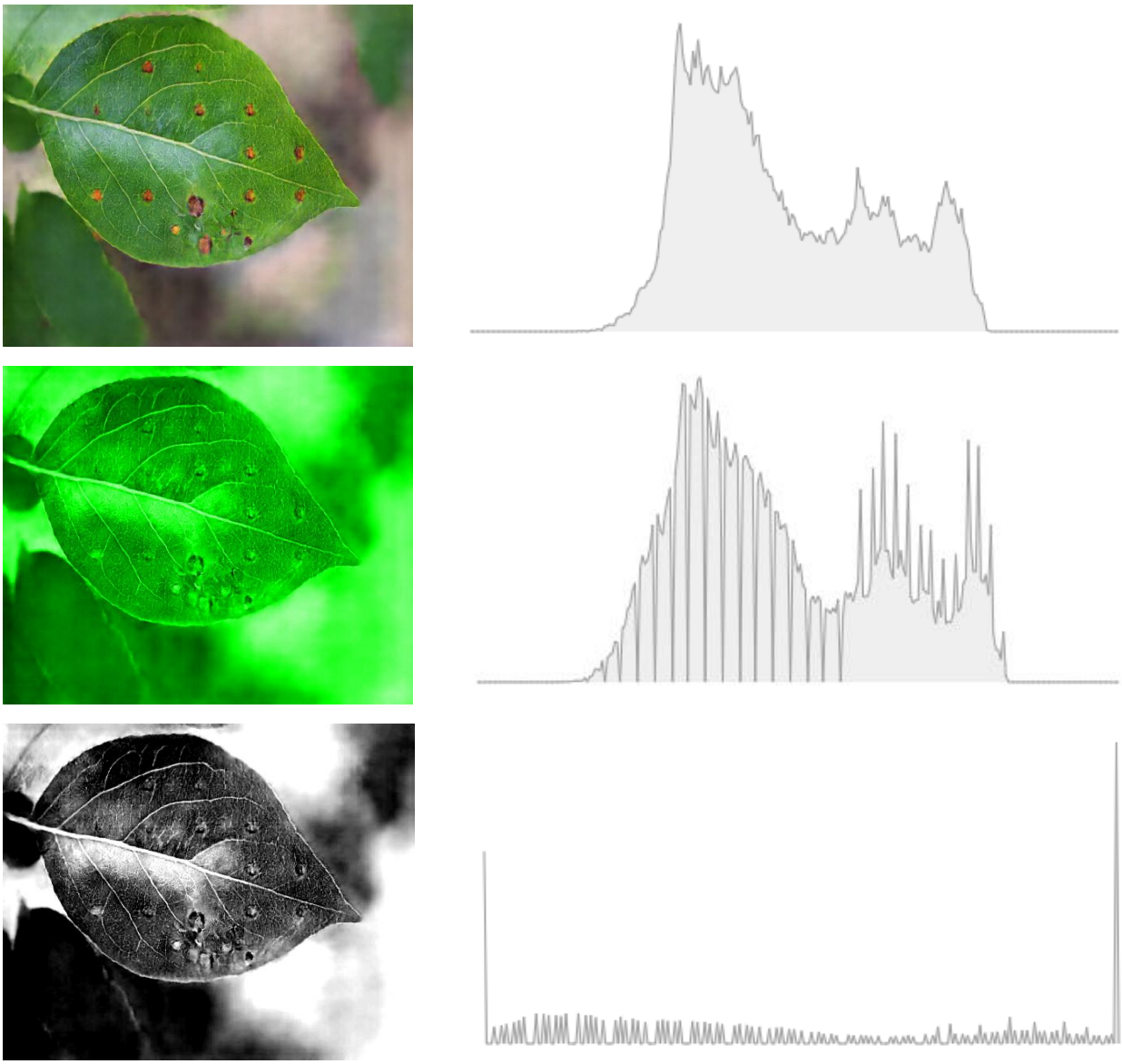


Fig. 4: *Original Image; G Image; normalized 'g' image*

5. WEB APPLICATION

We also created a web application for farmers to use. Django as python backend framework is used for creating its APIs and React as the Javascript frontend framework. We also used Semantic as its CSS library. In this application, the farmers need to upload the photo of the leaf and click on upload. Then the APIs take the image and after processing it, returns back with the results. The name of the leaf, the health status of the plant and its nitrogen content is displayed.

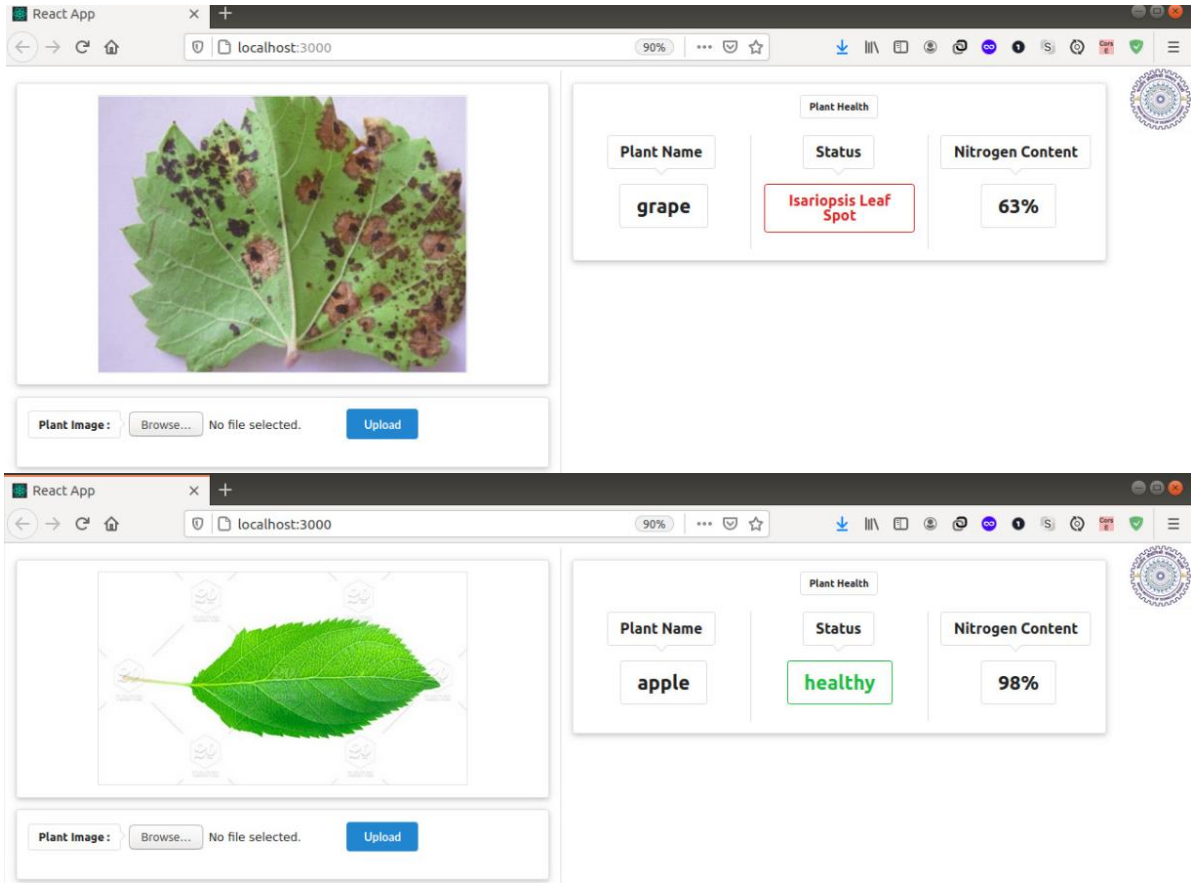


Fig. 5: Web Application Preview

6. FUTURE APPLICATIONS

Similar application can be created for mobile for better use. Even a wider and larger scale industrial initiative that can be taken is that the above given facility is installed on a drone that can provide the automated plant pathometry services.

7. APPENDIX I

The Kjeldahl method is used for quantitative determination of nitrogen in chemical substances developed by Johan Kjeldahl. Initially a leaf sample of 0.2 g is oven dried for 72 hrs. and then properly crushed. The sample is then mixed with 5 mL H_2SO_4 in the presence of K_2SO_4 and $CuSO_4$ and then heated in the digestion flask on the heater for 4 hrs. Heating the substance with sulphuric acid decomposes the organic nitrogen to ammonium sulphate. In this step potassium sulphate is added in order to increase the boiling point of the medium (from 337℃ to 373℃). Chemical decomposition of the sample is supposed to be completed after the medium (initially very dark) become clear and colourless. The solution is distilled with sodium hydroxide (added in small quantities approximately 10 ml) to convert the ammonium salt into ammonia. The amount of ammonia present (hence the amount of nitrogen present in the sample) is determined by back titration. The end of the condenser is dipped into a solution of hydrochloric acid or sulphuric acid of precisely known concentration (generally 0.2 to 0.4 N). The ammonia reacts with the acid and the remainder of the acid is then titrated with a sodium carbonate solution with a methyl orange pH indicator. Percentage Nitrogen is calculated using the following formula as in equation below.

$$\% \text{ Nitrogen} = \frac{(0.014 * \text{Volume of } H_2SO_4 \text{ required} * \text{Normality of the } H_2SO_4)}{\text{Sample weight of the collected leaves}}$$

8. REFERENCES

1. **Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn:** Delving into high-quality object detection. In Proceedings of the IEEE conference on computer vision & pattern recognition, pages 6154–6162, 2018.

2. **Tsung-Yi Lin, Piotr Doll' ar, Ross Girshick, Kaiming He, Bharath Hariharan, & Serge Belongie.** Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision & pattern recognition, pages 2117–2125, 2017.

3. **Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, & Piotr Doll' ar.** Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988, 2017.

4. **Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun.** Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision & pattern recognition, pages 770–778, 2016.

5. **Otsu, N.** 1979. A threshold selection method from gray-level histogram. IEEE Trans. Systems Man Cybern. 9 (1): 62-66.

6. **Chapman, H. D., and P. F. Pratt.** 1961. *Method for Analysis of Soils, Plants and Waters.* Univ. of California, USA.

7. **Sandino, J.; Pegg, G.; Gonzalez, F.; Smith, G.** Aerial mapping of forests a_ected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. Sensors **2018**, 18, 944.

8. **Siqueira, R.; Longchamps, L.; Dahal, S.; Khosla, R.** Use of fluorescence sensing to detect nitrogen and potassium variability in maize. Remote Sens. **2020**, 12, 1752.

9. **Alvar-Beltrán, J.; Fabbri, C.; Verdi, L.; Truschi, S.; Dalla Marta, A.; Orlandini, S.** Testing proximal optical sensors on Quinoa growth and development. Remote Sensing **2020**, 12, 1958.

10. **Aranguren, M.; Castellón, A.; Aizpurua, A.** Crop sensor based non-destructive estimation of nitrogen nutritional status, yield, and grain protein content in wheat. Agriculture **2020**, 10, 148.