Nutritional Deficit Detection in Crops Using Machine Learning

Dileep P [1] , Nagaraju I [2], Revathy P [3]
[1,2] Associate Professor, Department of Computer Science and Engineering
[3] Assistant Professor, Department of Computer Science and Engineering
[1,2] Malla Reddy College of Engineering and Technology,Kompally,Hyderabad,India.
[3] Narsimha Reddy Engineering College,Kompally,Hyderabad,India

Abstract— IP and ML are used to analyze images of crops for signs of nutrient deficiency. Vitamins and minerals are essential to a plant's healthy development and growth. Nitrogen, calcium, phosphorus, potash, sulphur, and magnesium (mg), to name a few, are essential for consistent and vigorous crop growth. Reduced crop output is the direct outcome of nutritional inadequacies, which make it harder to carry out routine agricultural tasks. Therefore, it is essential to have a quick evaluation of food consumption. Many crop leaflets exhibit glaring shortages, with customized layouts for each component. Our planned work is to provide a self-sufficient, trustworthy, low-cost alternative for identifying nutritional deficiency. Datasets for both unhealthy and full-functioning branches are built using IP methods including RGB color feature extractor, real-time texture recognition, edge identification, and so on. The resulting database will serve as training data for supervised ML, which will then be used to spot signs of nutrient deficiency and choose the strongest seedlings for further cultivation.
keywords : Plants, nutrient deficit, nutrients, Feature Extraction, and healthy leaves.

I. INTRODUCTION

India has the most irrigated land and the 12th-highest agricultural GDP in the world. It also accounts for 7.68 percent of global agrarian output. A large percentage of Indians rely only on farming for their livelihood, yet advances in technology have done nothing to break through the traditional divide between agriculture and the scientific community. The goal of the research we propose is to devise a simple, trustworthy, and precise method of tracking crop development, an area where micronutrients play a significant role. Micronutrient analysis in crops and the use of several techniques to maximize yield are focal points. Because pH has such a significant effect on nutrients' availability and digestibility, it may be possible to alter the fluid's pH as a strategy. In addition to reducing pesticide use (weakened crops are more susceptible to insects), identifying nutritional shortage may help offer correct fertilisers. Identifying the nutritional contents will help in supplying the proper number and kind of vitamin in the most current agricultural strategy, where food is produced without land, where nutritional supplements are provided to seedlings in the form of liquid. In addition to strengthening the economics of producers throughout the country, the recommended technique will be crucial in optimizing output to feed the world's ever-growing population. Therefore, we describe a system that can analyze a photograph of a plant and provide a diagnosis of its nutritional status. First, the image quality will increase when distortion is corrected [2]. This refined image is then segmented and features gathered for use in database creation and classification, all with the help of data analytics. The optical illusion created when an object is lighted by a radioactive generator is known as an image. The three pillars of each picture are its source, its topic, and its creation process. IP is the technique used to probe those vital properties, analyze them, and provide the result or notification required for the specific task at hand.

II. STANDARD IP AND ML APPROACHES

A. Supervised ML

Supervised ML is the more popular method because it produces reliable results. Some input variable (Y) and its related output (X) are supplied in the training database. The translation function may be calculated with the help of eq. (1).

$$Y=f(X) \quad (1)$$

Training datasets may also be used to find the translation equation. In order to accurately predict the output (X) from a new input (Y) outside of the database, the translation equation must be determined. It gets its supervised moniker from its capacity to learn from datasets used for training and to consistently predict output based on input. When the benchmark is met, learning stops and the information may be put to lawful use. The approach requires a proper training database; the more records there are, the more reliable the result; and the training and maintenance steps are simple. Extrapolation and classification problems are two of supervised learning's subsets. Categorization is used when the output may be altered in some way (defects, color, etc.). Valid outcomes, such as True/False, heating rate, etc., are modeled using regression.

B. Segmentation

The image is broken down into homogeneous chunks based on predetermined criteria. This may be done using a variety of techniques, including K-means clustering and the conversion of RGB to HIS. Using edge recognition techniques, the broken part will be isolated. The image may be efficiently subdivided into its HIS version and its RGB version by dividing the color pattern.

C. Thresholds

Using thresholding, the leaves will be separated into groups according to the characteristics and/or RBG levels of the samples used. This is a method for eliminating the foreground elements. In simple thresholding, if a pixel's brightness is below a predefined cutoff, it is replaced with a black pixel, and if it is above the limit, it is replaced with a white pixel.

III. RELATED WORKS

In the literature, coffee plant intellectual property is underrepresented. Subsequently, in [4], it is proposed that IP of coffee plants be used for autonomous diagnosis of nutritional deficiencies. Botanical classification is discussed at length in several books, journals, and websites [7, 16, 9]. Information about species relationships is encoded in the veins of leaves. Using Support Vector Machine [5], Penalized Discriminant Analysis, and Random Forests, [7] provides a technique for identifying three legume varieties using just structural factors calculated on fragmented venation. Both [16] and [9] propose new methods for vein retrieval. Other studies [6, 15, 14, 8, 13] use ML methods to classify crops based on their multilateral properties. In [6], we obtain geometry, color, and texture information from 1800 leaflets, and in [15], we retrieve 12 structural characteristics. In both articles, 32 plant species are classified using a stochastic neural net. Using knowledge-related visual qualities gleaned from structure, indentation, and vascular data, [14] provides a different method for classifying crops using an artificial neural network. This study presents a Random Decision Forest (RDF) technique for dynamically diagnosing nutrient shortages in crops [12], which combines universal and regional information gleaned from leaflets. Local features are recovered using SIFT descriptors from a bag of components, and thirteen universal characteristics are recovered using geometry and color data. According to the results, universal elements performed better than local elements in the classification task. Another leaf categorization method is proposed in [8], whereby features are obtained based on location, and classification is determined by heterogeneity measurements between the query image and the prior information. Using the SIFT method to recover localized features and the Geometry Relevance method to recover universal characteristics, as well as the balanced K-NN approach for classification [10], [13] combines locally and globally characteristics to improve crop categorization.

IV. PROPOSED METHODOLOGY

The results of each shortfall are different for different crops. In Table I, we see the symptoms for each deficiency. The primary method is to identify the missing qualities and recover them. Red, blue, and green color indices and mean values, edges, texture, and so forth are among the recovered features. The deficiencies of several nutrients in the leaves are shown in Figures 1 through 6. The intended layout is shown in Figures 7 and 8.

TABLE I. NUTRIENT INADEQUACY INDICATORS

Vitamins/ Minerals	**Indicators Of Deficiency**
Nitrogen	Extremely pale tint, upright leaves with light green/yellow coloration, burnt in severe lack
Calcium	Delicate leaves are pale and brittle from the apex and have a dark green tint.
Magnesium	Pale margins, cup-shaped folds, and, in severe cases, leat dies
Sulphur	Light green leaves, pale green veins
Phosphorus	Underneath a leaf, the leaf goes brown/black and bronze.
Potassium	Pale leaves with little rusty dots.

Fig. 1. Nitrogen insufficiency leaf image

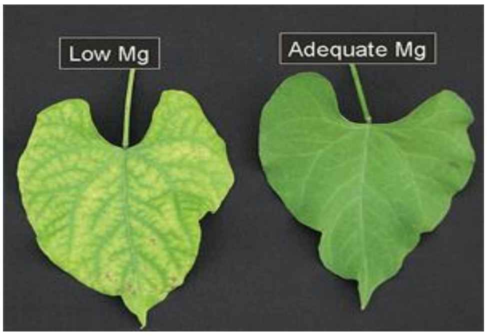


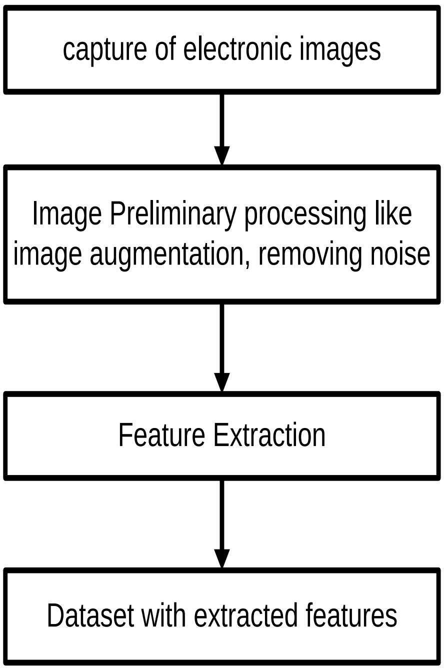
Fig. 2. Magnesium insufficiency leaf image

Fig. 3. Potassium insufficiency leaf image

Fig. 4. Sulphur insufficiency leaf image

Fig. 5. Calcium insufficiency leaf image

Fig. 6. Phosphorus insufficiency leaf image



Methodology employed in this work

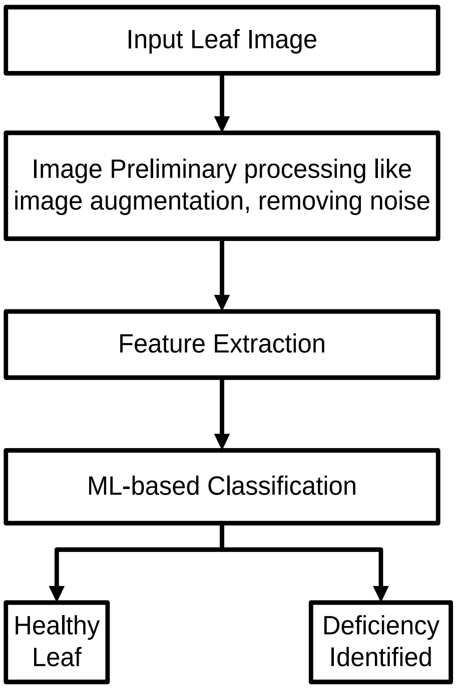


Fig. 7. Classification Process

A. Image Archive in Digital Form

Initiate Supervised ML database creation. About 1700 photographs are required for typical plants, plus another 100 for each of the six nutrient deficiencies. Every plant, both bad and healthy, is photographed against a white background under standard lighting conditions using an automated system.

B. Primitive Image Processing

The resulting image may include distortion or extraneous data. The region of interest may be achieved by removing the background. If problems are discovered, they are removed, and the problematic area, in this case the leaves, is expanded so that more data can be retrieved and analyzed for signs of inadequacy. In order to get rid of the distortion and get smoother photos, the mean filtering is used. By replacing each pixel value with an estimate of the neighboring pixel values, the mean filter smooths out fluctuations in the picture. The core serves as a symbol for the size and shape of the area under examination. Histogram equalization is used to spread severity for the purpose of picture enhancement.

C. Extraction of Features

Features are extracted from the first processed image. Here we have the red (R), green (G), and blue (B) components, as well as the band ratios of green to red and blue to green. The primary color is green because of the standard leaf used for comparison. The average frequencies of colors with values from 0 to 255 are also calculated for R, G, and B. Gavg/Bavg and Gavg/Ravg are used to determine the ratio of their means.

D. Detecting Edges

A dietary deficiency may exist if the input image provided does not have a predominance of Green. In this case, the image will be processed using edge recognition to spot smudges and an inadequate defect detection zone. Several methods have been developed to identify edges, such as the Laplacian of Gaussian (LOG), Roberts, Prewitt, Sobel, Zero Crossing, and Canny. First-degree derivatives may be interpolated using Canny, LOG, or Zero Crossing, while second-degree derivatives can be extrapolated using Roberts, Prewitt, or Sobel.

Gradients in images are used to put a numeric value on subtle shifts in pixel brightness. The gradient is an integral of the first degree. The rate of change in the image is determined by the magnitude of the gradients, while the direction of change is indicated by the orientation. The borders of first order variations are more open and susceptible to noise. Second-degree derivatives are used to build the best edges when there is a drastic change in the feature space. A higher score indicates a crisper image.

Construction of Datasets

The photographs have been sorted based on the flaws identified by experts in the field. Improved ML support has been added to the database properties that may be accessed. All characteristics, including edges, blotches, suggested deficiency, color indices, the mean indicator, and band proportions, have been recalculated for each shortcoming. Possible applications for this database include using the ML algorithm to troubleshoot supervised learning. At this time, 70% of the information is being utilized for instructional purposes. The remaining 30% is put into testing the efficiency of the system and building a database. Comparing the predicted actual and simulated outcomes will establish the system's accuracy.

Grouping (F)

Classification is a supervised learning method in ML, meaning that the source is already known and the output is controlled by the input database. In this instance, the tree diagram was used to spot problems. A feature extractor technique will be used to preprocess the photos. Now, a tree structure will be used to assess these metrics using the input database; if the maximum characteristics match the data source amount for a certain shortfall, the outcome will be that shortfall. If the comparison fails, the characteristics from the input image are subtracted from the attributes in the database, and the least dissimilar point is returned as the result.

V. RESULTS AND DISCUSSION

A total of 335 images of coffee plants were used for the functional evaluation: 255 for the training data set and 80 for the test dataset (10 images per category). The photographs included in the review were provided by Cenfrocafe in Peru. The photographs were taken against a white background, at an angle with the peak at the top. Table II displays the number of images included in the learning dataset for each class. We utilized RDF to set up four different scenarios: 1) Using terms that are specific to the area. 2) Relying on generalizable features. Third, taking into account both local and global factors. 4) Combining the global and local features of SIFT. A different number of trees, denoted by T, are used to construct each system. The results of each scenario's classification efforts are shown in Table III.Table III shows that training RDF with T = 20, 80, and 100 trees using universal parameters improves classification accuracy. The SIFT classifiers' accuracy will be too low in this case. Siscafe was used to classify the leaflets in the test dataset so that the proposed classifier could be tested for its universal properties. Table IV displays the results for each critical shortage. In conclusion, the proposed approach was able to pinpoint nitrogen, calcium, and phosphorus deficiencies in the diet. Similar visual issues are brought on by magnesium and potassium deficiencies. In order to more accurately classify such micronutrient deficiencies, better features for depicting veins are necessary. In order to improve the diagnostic accuracy of potassium deficiency, it is necessary to enhance the features that identify apoptosis on the top leaves. In conclusion, the proposed model achieves higher classification accuracy than Siscafe. However, the efficacy of the classifier may be affected by the number of deficiencies, since coffee plants often show two or more nutrient deficits simultaneously. A growing lack of essential nutrients might reduce agricultural yields. Morphological changes are more apparent when a species experiences a significant lack of nutrients. In addition, the data included includes images taken at varying stages of nutritional deficit, which may affect the classification scheme.

TABLE II. COUNT OF IMAGES PER CLASS

Deficiency	**No. of images**
Nitrogen	40
Calcium	43
Magnesium	21
Sulphur	24
Phosphorus	44
Potassium	20

TABLE III. CLASSIFICATION ACCURACY

Model	**Accuracy(%)**				
	T=20	**T=40**	**T=60**	**T=80**	**T=100**
1	39, 1	32, 9	39, 2	42, 5	42, 3
2	68, 4	63, 6	68, 3	68, 4	68, 5

3	57, 2	58, 0	54, 9	59, 5	59, 6
4	49, 8	52, 4	56, 0	54, 6	54, 0

TABLE IV. ACCURATELY PREDICTED CLASSES BY PROPOSED MODEL

Deficiency	**Correct classes**	
	Proposed model	**siscafe**
Nitrogen	10	2
Calcium	5	3
Magnesium	6	4
Sulphur	3	0
Phosphorus	9	5
Potassium	8	4

VI. CONCLUSION AND FUTURE SCOPE

Focusing on maximum output is crucial in order to satisfy the demands of a growing society. If crops are given adequate nutrients to flourish, this is always possible. The nutritional makeup of crops is seldom considered, despite its importance. This article focuses on the use of IP and ML methods for determining nutritional content. This will help farmers determine the safety of their crops and what measures may be taken to ensure their continued success. This will be helpful in hydroponics, hydroponics, and hydroponics, as well as vertical gardens. Recognizing leaf age, identifying pathogens, recommending fertilizer, and recognizing natural leaf age would all complement the proposed endeavor.

REFERENCES

[1] D. Monsalve, M. Trujillo and D. Chaves, "Automatic classification of nutritional deficiencies in coffee plants," 6th Latin-American Conference on Networked and Electronic Media (LACNEM 2015), 2015, pp. 1-6, doi: 10.1049/ic.2015.0317.

[2] R. Guerrero, B. Renteros, R. Castañeda, A. Villanueva and I. Belupú, "Detection of nutrient deficiencies in banana plants using deep learning," 2021 IEEE International Conference on Automation/XXIV Congress of the Chilean Association of Automatic Control (ICA-ACCA), 2021, pp. 1-7, doi: 10.1109/ICAACCA51523.2021.9465311.

[3] Luis Alberto Vives Garnique, Hever Ivan Mejıa Cabrera, Kevin Leandro Vilcherrez Chavarry, and Marcelo Jes´us Vassallo Barco. Vision artificial: Aplication defiltros segmentation in imagenes de hojas decafe. Ingenierıa: Ciencia, Tecnologıa e Innovacion, 1(2), 2014.

[4] Kristen Grauman and Bastian Leibe. Visual object recognition. Number 11. Morgan & Claypool Publishers, 2010.

[5] Jacques Hadamard. Sur les probl`emes aux d´eriv´ees partielles et leur signification physique. Princeton university bulletin, 13(49-52):28, 1902.

[6] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto, and Paulus Insap Santosa. Leaf classification using shape, color, and texture features. arXiv preprint arXiv:1401.4447, 2013.

[7] Monica G Larese, Rafael Nam´ıas, Roque M Craviotto, Miriam R Arango, Carina Gallo, and Pablo M Granitto. Automatic classification of legumes using leaf vein image features. Pattern Recognition, 47(1):158–168, 2014.

[8] Chia-Ling Lee and Shu-Yuan Chen. Classification of leaf images. International Journal of Imaging Systems and Technology, 16(1):15–23, 2006.

[9] Yan Li, Zheru Chi, and David D Feng. Leaf vein extraction using independent component analysis. In Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on, volume 5, pages 3890–3894. IEEE, 2006.

[10] David G Lowe. Object recognition from local scaleinvariant features. In Computer vision, 1999. The proceedings of the seventh IEEE international conference on, volume 2, pages 1150–1157. Ieee, 1999.

[11] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27(10):1615–1630, 2005.

[12] Nobuyuki Otsu. A threshold selection method from graylevel histograms. Automatica, 11(285-296):23–27, 1975.

[13] ZhiyongWang, Bin Lu, Zheru Chi, and Dagan Feng. Leaf image classification with shape context and sift descriptors. In Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on, pages 650–654. IEEE, 2011.

[14] Qingfeng Wu, Changle Zhou, and Chaonan Wang. Feature extraction and automatic recognition of plant leaf using artificial neural network. Advances in Artificial Intelligence, 3, 2006.

[15] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu- Xuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. A leaf recognition algorithm for plant classification using probabilistic neural network. In Signal Processing and Information Technology, 2007 IEEE International Symposium on, pages 11–16. IEEE, 2007.

[16] Xiaodong Zheng and Xiaojie Wang. Leaf vein extraction using a combined operation of mathematical morphology. In Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference on, pages 1–4. IEEE, 2010.