Spectral and Numerical Analysis of Hyper spectral Data using Vegetation Indices

Nikhil M. Sapate, Ratnadeep R. Deshmukh

Abstract: Indian economy is majorly influenced by Agriculture and its allied sectors. More than 50% of the population of India is dependent on agriculture and its allied sectors for their survival. According to the Report of Indian Council of Agricultural Research, 30-35% of the crop yield gets wasted due to disease. Using modern day remote sensing techniques plant health can be monitored and it can be specified whether a plant is diseased or healthy. Hyperspectral Remote Sensing is the technique by which fine and minute information of vegetation can be obtained with the help of narrow wavebands. Data of 80 leaf samples of Tomato crop collected in spectra form and text form using ASD FieldSpec4 spectroradiometer and ViewSpec PRO. This information of plant leaves was used to identify vegetation attributes and its status. Vegetation Indices are calculated using mathematical formulae published in the previous study. Random forest classification used to discriminate among Healthy and Diseased plants. Algorithm works with an accuracy of 93.75% with misclassification rate 0.0625. Along with Green wavelength range and Red edge of the spectrum, specific disrupted behavior was observed in Shortwave Infrared Region of the spectra. The research paper focuses on Spectral and Numerical study and analysis of Tomato Leaf disease with the help of ASD FieldSpec4.

Keywords: ASD FieldSpec4 Spectroradiometer, Hyperspectral Remote Sensing, Vegetation Indices, VNIR-SWIR Spectroscopy, 

I. INTRODUCTION

Hyperspectral Remote Sensing is the technique nowadays majorly considered to study vegetation. Babadoost M. et al. studied about it that encourages non-destructive methods of identifying plant behavior and its properties [1]. Crop yield faces severe issues due to the challenging environment. Pests cause crop loss in India and it is almost 10-30% of the total crop production. Sreekala G. Bajwa et al. stated that In order to manage diseases effectively, Disease Detection is the multidisciplinary approach that integrates with cultural, biological, physical and chemical strategies [2]. In order to analyze and to detect changes in plant physiology and chemistry, Spectral Vegetation Indices and its use play a vital role. Indices are based on sensitive waveband information and their significance directly relates to the presence of Water Content, Leaf Area, and Pigmentation status. A.K. Mahlein et al. observed and concluded that Vegetation indices play a crucial role to detect whether the plant is infected with disease or not infected [3]. Application of Spectroscopy consists of emission of electromagnetic radiation which interacts with pigment content, leaf structure and water content. Photosynthetic pigments, mainly chlorophyll and carotenoids absorb the electromagnetic radiation in visible region causing less reflectance. Alfadhl Yahya Khaled et al discussed that in short-wave near-infrared region also lack of reflectance due to structural discontinuities occurred in leaf [4]. Benjamin Dechant et al presented a study which stated Chlorophyll which is useful pigment for a plant to perform photosynthesis occurs in red wavelength [5]. Another important factor is Carotenoids and it absorbs energy and protects chlorophyll from photo-damage. One of the key carotenoids is xanthophyll. It also protects the plant from excessive sunlight energy and prevents plants from further damage. A third and equally important factor in plants health is Canopy. It is an uppermost or topmost layer of trees or branches. Carotenoids and Chlorophyll incorporate with each other to minimize the impact of variable canopy structure which automatically results in good health of plants. T. Rumpf et al. stated in a study that these factors are affected during pathogenesis and it can be recorded and can be studied using Hyperspectral remote sensing and Vegetation Indices are found to be highly correlated to these physiological parameters [6]. A previous study established the use of such vegetation attributes such as water pigments, nitrogen-rich compound, structural materials and Vegetation Indices that can predict about the health status of plants. The bands in green, red and NIR of the spectrum are considered to be useful for it. So-Ra Kim, Lammert Kooistra, R. Devadas et al. stated that changes in the pigment can be observed with the spectral characteristics and it emphasizes on health-related aspect of plants [7, 8, 9]. The objectives of this study was to Collect Hyperspectral Non- Imaging data of crop Tomato, to perform spectral and numerical analysis of this data, to study Vegetation Indices and to study its relation with pigmentation and health of plant, to discriminate healthy and diseased plant using spectral and numerical analysis.

II. EXPERIMENT

A. Design of Experiment

The study has shown that disease detection in plants can be performed promisingly with the help of Spectroscopy. Sindhuja Sankaran et al presented Biotic and Abiotic stresses in plants can be identified using Spectroscopy [10]. To perform these components in vegetation, ASD FieldSpec4 is a spectroradiometer used for laboratory tests. Spectral Range of this device is highly resolving and covers wavelength of 350-2500 nm. It has a
Spectral and Numerical Analysis of Hyper spectral Data using Vegetation Indices

sampling rate of 0.2 seconds per spectrum. Recording of spectra is indulged by three detectors:

1. Silicon-Photodiode array composed of 512 elements for the VNIR (350-1000 nm)
2. Thermoelectrically cooled and equipped InGaAs-photodiodes (Indium, Gallium, Arsenid) for SWIR1 (1000-1800 nm).
3. Thermoelectrically cooled InGaAs-photodiodes (Indium, Gallium, Arsenid) for SWIR2 (1800-2500 nm).

B. Data Acquisition

Leaf samples of crop Tomato are collected from a farm on Paithan Road, Aurangabad, Maharashtra, India on 12 October 2018 and 30 October 2018. GPS coordinates of the study area are mentioned in the figure below. Samples consist of healthy as well as diseased leaves of Tomato plant. 2 sets of 40 samples each are formed to study, test, categorise and further classify the data. In order to avoid air contact and to preserve its moisture, these leaves were taken Laboratory and spectral measurements were taken with ASD FieldSpec4 device and packaged software RS3.

C. Leaf Spectra Reading

Initially, all the samples were closely observed and sorted into categories like healthy, unhealthy, and dry. Leaves size vary throughout the samples. So, the leaves having long and spreading structure were considered for its lab test purpose. In a dark room inside the lab spectral measurements were taken with the device ASD FieldSpec4 (ASD Co. USA). Before taking a sample the spectrum of white reference panel (99% reflectance provided by ASD Co. USA) was established and plotted. Leaves consisting of 20 samples of healthy and diseased leaves undergone spectral reading and spectra of each and every leaf were recorded. This device has a range of 350 nm – 2500 nm.

D. Spectra Analysis by ViewSpec PRO

During the experiment at one end, the spectra can be viewed and recorded using RS3 software. To process each file of each sample the file is stored to the .asd file format. This file is encrypted file and can be viewed in the form of graphs using ViewSpec Pro. As and when the Spectroradiometer was used software packages followed equal importance for the examination and analysis of Hyperspectral Data.

Fig (2). Spectral Graph of Healthy Leaves

Recorded readings can be graphically visualized using ViewSpec PRO. Samples show standard vegetation spectra. It can be concluded that Spectral Graphs match with standard vegetation Spectra available from previous study.

Fig (3). Spectral Graph of Diseased Leaves

Also, this software package is able to perform pre-processing such as first derivative, second derivative, log and statistical analysis. To numerically analyse the required .asd file it needs to be stored in text file format. Spectral Graphs of Healthy and Diseased plants are shown in fig. (2) and fig. (3).

E. Vegetation Indices

In the past few years, research has shown that there is a strong correlation between vegetation properties and photosynthetic pigment tissues such as chlorophyll and carotenoids. Driss Haboudane John R. Miller, Minghua Zhang et al provides an approach that at different wavelengths (e.g. 700-900 nm) these pigments behave differently with the help of which healthy and diseased leaves spectra can be distinguished and discrimination can be performed [11, 12]. Sujan Sarkar et al studied that the differentiation between healthy and diseased can be roughly recorded by recording visual test of leaf and specific behavior of its spectra [13]. Hui Wang et al presented an approach of Vegetation indices that exhibit the increase in their own values as the disease severity gets increased [14]. Abel Chemura et al studied that in some cases, relation between wavebands and vegetation
indices could not be affirmed because they share complex, non-linear relationship with each other [15]. Xin Yang studied different factors such as local and global values of indices help to understand the changing nature of various pigments and phenotypes [16]. Raymond F. Kokaly et al studied that canopy levels signify the status of absorption activity in fresh, dry and diseased leaves [17]. Sushma D Guthe et al stated that Phosphorous is important in plants for it contributes to normal growth and maturity of plant while potassium content is a nutrient to plants growth. These two contents can be analysed with the help of Vegetation indices [18]. Nikhil M. Sapate et al presented a study that significant level of nutrient prediction can be indicated with the indices like NDVI, GNDVI, SAVI, and OSAVI [19]. Wenjiang Huang et al stated that in addition to traditional indices, HI (Health Index), PMI (Powdery Mildew index), YRI (Yellow Rust Index) were the newly introduced indices [20]. Anne-Katrin Mahlein et al studied use of more two or more than two indices helps to detect plant disease in better manner [21]. After studying the previous methodologies and after deciding the crucial factors for affecting plants health we have used various Vegetation Indices.

Every Vegetation Index focuses on specific behaviour and functionality. Wenjiang Huang et al studied that Carotenoid Reflectance index (CRI1) measures the sensitivity of carotenoid pigment in plant foliage. Due to the overlap between chlorophyll and carotenoid absorption peaks and due to higher concentration of chlorophyll than carotenoid in most leaves, it becomes difficult to estimate carotenoid with the help of reflectance. Carotenoid concentration relative to chlorophyll can be concluded by higher value of CRI [22]. Photochemical Reflectance Index (PRI) reacts when changes in carotenoids take place in live foliage, especially with xanthophyll. This index is considered to be most important for measuring health of plants. StructureInsensitive Pigment Index (SIPI) is useful for measuring high variability in canopy and leaf area index. Water Band Index (WBI) is sensitive to changes in canopy water content. Physiological Reflectance Index (PhRI) provides ample amount of physiological information and follows diurnal changes especially with xanthophyll pigments and photosynthetic rates of controlled and nitrogen limited leaves. Prasad S. Thenkabail stated that Normalized Pigment Chlorophyll Ratio Index (NPCI) is nitrogen sensitive index and it helps to measure the nitrogen content in plant foliage [23]. Here are the results for Disease and Healthy plants obtained after performing calculations on numerical data. Priyanka U. Randive et al studied that Structural, Physiological, photochemical changes can be observed by using sensitive bands and applying indices formula on it [24]. Andre Grobe-Stoltenberg et al studied that there are numerous studies which indicate that selecting different indices makes it easier for us to classify healthy and unhealthy plant with increasing accuracy [25]. If we observe the fig.(3) It can be easily understood that chlorophyll absorption in the red region is moderately disturbed. Akshay V. Kshirsagar et al stated that Red edge can be used as an indicator to study chlorophyll absorption [26]. After 485 nm, the reflectance fluctuates and does not show sustained growth and does not reach a standard peak of NIR region. Rumiana Kancheva et al studied about top bands that show different behavior in spectra are 730 nm and 733 nm which belong to red- Near Infrared Region (R-NIR region) and irregularly show reflection [27]. Wavelengths like 690, 695,775 and 780 nm are helpful for understanding the health of plant and are detector of abnormalities up to more extent. Also from the figures (2) and (3) significant difference is observed in the range of 1000-1400 nm bands of healthy and diseased plants and these bands are correlated with the combined water in the leaf. The wavelength range of 900-1500 nm is usually a critical range which is used in previous studies for discrimination. For the discrimination of plants as healthy or diseased, selecting the right indices for numerical examination is very important. Hence all the indices are studied, tried and tested before use on various datasets available. Classification method is considered to be the preferred method for discriminating two classes. Indices of samples if shows high value then it have more chances of infestation or occurrence of disease. Also with the help of indices estimation of damage to a plant can also be worked out. Hence our analyses also include that few plants that have damaging on its leaf, show different spectra compared to healthy. Plants having scars and spots show low reflectance at 2000 nm- 2200 nm (Shortwave Infrared 2 Region) shows abnormal spectral behavior.

### III. RESULTS AND DISCUSSION

After applying formulae to numerical data obtained from spectra using ViewSpec Pro, results were calculated using Vegetation Indices Formulation. And the results were ready for analysis. Records of vegetation indices were categorized as Results for Diseased Plants and Results for Healthy Plants. We have used vegetation Indices as an indicator for deciding the whether plant is healthy or diseased.

**Table 1. Results of Diseased Plants.**

| Sr. No. | CR | PRI | SIPI | WBI | PhRI | NPCI |
|--------|----|-----|------|-----|------|------|
| 1      | .05| -.21| .96  | 2.47| .21  | .80  |
| 2      | 3.8| .01 | .75  | 1.02| .04  | .07  |
| 3      | 6.6| -.03| .94  | 1.02| .11  | .14  |
| 4      | 3.3| 0   | .75  | 1.03| .05  | .06  |
| 5      | 6.0| .00 | .81  | 1.02| .07  | .03  |
| 6      | 7.6| .006| .83  | 1.02| .08  | .05  |
| 7      | 7.7| .01 | .80  | 1.02| .07  | .02  |
| 8      | 6.4| .003| .80  | 1.02| .07  | .03  |
| 9      | 6.1| 0   | .77  | 1.03| .06  | .14  |
| 10     | 6  | .003| .78  | 1.03| .06  | .08  |
| 11     | 6.5| -.01| .77  | 1.0  | .06  | .06  |
| 12     | 5.2| -.02| .76  | 1.03| .06  | .09  |
| 13     | 4.2| .005| .73  | 1.03| .05  | .04  |
Spectral and Numerical Analysis of Hyper spectral Data using Vegetation Indices

Table 2. Results of Healthy Plants.

| Sr. No | CR I1 | PRI | SIP | WBI  | PhRI | NPCI |
|--------|-------|-----|-----|------|------|------|
| 1      | 7.6   | .02 | 0.86 | 1.03 | 0.084| 0.03 |
| 2      | 6.1   | 0   | 0.78 | 1.02 | 0.076| 0.007|
| 3      | 5.1   | 0   | 0.78 | 1.03 | 0.078| 0.01 |
| 4      | 5.1   | .01 | 0.78 | 1.03 | 0.069| 0.02 |
| 5      | 7.5   | 0   | 0.78 | 1.02 | 0.072| 0.02 |
| 6      | 7.6   | 0   | 0.80 | 1.04 | 0.070| 0.02 |
| 7      | 6.3   | .02 | 0.80 | 1.02 | 0.061| 0.03 |
| 8      | 6.1   | 0   | 0.78 | 1.02 | 0.076| 0.007|
| 9      | 5.1   | 0   | 0.78 | 1.03 | 0.07 | 0.01 |
| 10     | 5.1   | 0   | 0.78 | 1.03 | 0.07 | 0.01 |
| 11     | 4.5   | 0   | 0.72 | 1.04 | 0.064| -0.07|
| 12     | 3.7   | .01 | 0.71 | 1.03 | 0.05 | -0.04|
| 13     | 5.0   | 0   | 0.78 | 1.03 | 0.07 | 0.05 |
| 14     | 4.8   | .01 | 0.71 | 1.03 | 0.055| -0.06|
| 15     | 3.8   | 0   | 0.71 | 1.04 | 0.057| -0.03|
| 16     | 4.2   | 0   | 0.74 | 1.04 | 0.05 | 0.02 |
| 17     | 4.3   | 0   | 0.7 | 1.04 | 0.05 | -0.04|
| 18     | 4.1   | .01 | 0.75 | 1.04 | 0.05 | -0.034|
| 19     | 5.2   | .01 | 0.77 | 1.02 | 0.05 | 0.06 |
| 20     | 5.3   | .01 | 0.77 | 1.02 | 0.05 | 0.01 |

Following is the Graphical Representation and analysis of Vegetation Indices Values:

Fig. (4) Graph of CRI1 values for Healthy (CRI1_H) and Diseased (CRI1_D) Plant

From Table 1 and Table 2 and above graph we can analyse that Carotenoid pigmentation is high in diseased plant and it represents abnormality in functioning of pigments. It is majorly distributed in 495-570 nm.

Fig. (5) Graph of PRI values for Healthy (PRI_H) and Diseased (PRI_D) Plant

Photochemical Reflectance is high in healthy plants and signifies good working of xanthophyll and confirms good plant health. PRI mainly formulates reflectance in green region. From the analysis of above graph we can conclude that Canopy and leaf area index varies by more margins in diseased plants whereas in healthy plants it does not vary by more margins.

Fig. (6) Graph of WBI values for Healthy (WBI_H) and Diseased Plant (WBI_D)

From the above graph we can conclude that water content in the Healthy plant is observed to be high and in the diseased plant is observed to be low. If Water availability in plant is low, then it can cause starvation of invading cell of water which results in malfunctioning of photosynthesis and directly reduces the immunity of a plant. Because of such actions and reactions on a plant there are high possibilities that plant may have already been infected and may have a disease. Also it may cause stunted growth due to lack of water content.
SIPI hence can be used as a conclusive parameter. From the above graph analysis can be done that Canopy water content fluctuates among diseased plants while they show negligible fluctuation in healthy plants.

As we observe the graphical representation, it is well understood that indices of diseased plant fall towards lower range of values in a graph. This indicates the lack of pigmentation, abnormality, and suspicious photochemistry.

Physiological Reflectance is represented by PhRI represents the physiological diurnal changes specially with xanthophyll and from the above graph we can conclude that it has high values in diseased plants as compared to healthy plants.

From the above graph we can conclude that Nitrogen content represented by NPCI index was recorded to be more in diseased plant and it directly shows poor and stunted plant growth.

Supervised Classification was performed using Random forest algorithm which is well known for its high accuracy. It is an ensemble method having multiple decision trees. Correlation of vegetation indices and health status of plant resulted in accuracy to 93.75%.

Following are the results obtained after Random forest classification:
Prediction: [1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1]
Test Results: [1 1 1 0 1 1 0 1 1 1 1 1 0 1 1 1]
Confusion Matrix for Dataset is: [[2 1] [0 13]]

Performance of model designed by us can be measured using metrics like precision, recall, F1-Score, Support.

|   | Precision | Recall | F1Score | Support |
|---|-----------|--------|---------|---------|
| 0 | 1.00      | 0.67   | 0.80    | 3       |
| 1 | 0.93      | 1.00   | 0.96    | 13      |
| Acc |          |        |         | 16      |
| Mac. Avg | 0.96 | 0.83   | 0.88    | 16      |
| Wei. Avg | 0.94 | 0.94   | 0.93    | 16      |

Only accuracy is not enough for understanding the efficiency of model. Along with accuracy, when F1-Score is analysed to conclude that model has learned and is predicting with fine accuracy. There are some terminologies that must be studied to understand this numeric representation is as follows:

| N = 16 | Predicted : NO | Predicted : YES | Total |
|-------|----------------|-----------------|-------|
| Actual No | TN : 2 | FP: 1 | 3    |
| Actual Yes | FN: 0 | TP: 13 | 13    |
| Total | 2 | 14 | |

Accuracy = (TP + TN) / Total
= (13 + 2) / 16
= 0.9375

Accuracy of algorithms is 93.75%.

Misclassification Rate = (FP + FN)
= 1 + 0 / 16
=0.0625
IV. CONCLUSION

Pigmentation plays vital role in nutritional status of plant, its immune system and growth. Chlorophyll is very important among such pigments. Previous studies have shown that specific sensitivity is observed at narrow band of spectra having a sensitive leaf to chlorophyll concentration and canopy scales. Stunted growth, malnutrition of plant, irregularities in pigmentation may occur in diseased plants. Photosynthetic capacity or productivity of a canopy can be measured with the help of formulae of Vegetation Indices and their range indicates certain features. If assessment of plants health is to be performed then it becomes important to obtain information about leaf chlorophyll concentration and physiological status of plant and level of nutritional stress etc. In green range of wavelength spectrum (495-570 nm) absorption was very less in diseased plant. For diseased leaves, spectra differed in red edge (680-750 nm). After 1000 nm, diseased sample showed radical difference in the spectra. Some samples having scars, spots showed more reflection in 1450-1820 nm. In shortwave infrared region of reflectance (SWIR2) spectra graph showed specific changes that were indication of infestation. Vegetation indices focusing on each different important vegetation attribute were chosen. CR1, PRI, SIRI, WBI, PhRI, NPCI indices showed numerically less valued results for diseased plant. Some abnormal growth was also noticed in graphs of indices of diseased plants. That abnormalities in spectra and values falling towards lower bound of vegetation index range concluded infestation, stunted growth, malnutrition, irregularities in pigmentation of diseased plant.

ACKNOWLEDGEMENT:

This work is supported by Dept. of Computer Science and Information Technology under the funds for Infrastructure under science and Technology (DST-FIST) with sanction no. SR/FST/ETI-340/2013 to Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India. The authors would like to thank Department and University Authorities for providing the infrastructure and necessary support for carrying out the research.
19. Nikhil M. Sapate , Dr. R.R. Deshmukh, Pooja V. Janse, “Spectral Discrimination of Healthy and Diseased Plants Using Non Imaging Hyper spectral Data – A Review”, International Journal for Research in Engineering Application & Management (IJREAM), Special Issue - NCRICE , 2019 , pp. 85-87.

20. Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Dongyan Zhang, “New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases”, IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 7, No. 6, June 2014, pp. 2516-2522.

21. Anne-Katrin Mahlein , Erich-Christian Oerke , Ulrike Steiner , Heinz-Wilhelm Denne, “Recent advances in sensing plant diseases for precision crop protection”, European Journal of Plant Pathology, Volume 133, 2012, pp.197–209.

22. Wenjiang Huang, Junjing Lu, Huichun Ye, Weiping Kong,” Quantitative identification of crop disease and nitrogen-water stress in winter wheat using continuous wavelet analysis”, International Journal of Agricultural and Biological Engineering, Vol. 11 No.2 , March 2018, pp. 145-151.

23. Prasad S. Thenkabail, John G. Lyon, Alfredo Huete, “Hyperspectral Remote Sensing of Vegetation”, Taylor and Francis Group, 2012, pp. 309-324.

24. Priyanka U. Randive , Ratnadeep R. Deshmukh, Pooja Vinod Janse , Jaypalsing N. Kayte, “Study of detecting Plant diseases using Non-Destructive Methods: A Review”, International Journal of Emerging Trends & Technology in Computer Science, Volume 7, Issue 1, January – February 2018, pp. 66-71.

25. Andre Grobe-Stoitenberg, Christine Hellmann, Christiane Werner , Jens Oldeland, Jan Thiele, “Evaluation of Continuous VNIR-SWIR Spectra versus Narrowband Hyperspectral Indices to Discriminate the Invasive Acacia longifolia within a Mediterranean Dune Ecosystem”, MDPI, Remote Sensing, Volume No. 8, Article Number 334, 2016, pp. 1-18.

26. Akshay V. Kelarsagar, Ratnadeep R. Deshmukh, Pooja V. Janse, Rohit Gupta, Jaypalsing N. Kayte, “Detection of disease from Chilly Plant Using Vegetation Indices”, International Journal of Computer Sciences and Engineering, Vol.-7, Issue-1, January 2019, pp. 333-336.

27. Rumiana Kancheva, Denitsa Borisova, “ Plant Spectral Signatures as Growth Stress Indicators”, 2006 2nd International Conference on Information & Communication Technologies, IEEE, Vol. 2, 24-28 April 2006, pp. 355-359.

AUTHORS PROFILE

1* Corresponding Author: Nikhil M. Sapate Pursued Bachelor Of Technology In Computer Science And Engineering In Maharashtra Institute Of Technology, Aurangabad (Ms)431001 India In 2014, And Currently Pursuing Master Of Technology Under The Faculty Of Engineering And Technology In Computer Science And Engineering In Department Of Computer Science And IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS) India. He Received the International Award as the Best Researcher from World Research Council for his research on Non-Imaging Hyperspectral Data.

2: Dr. Ratnadeep R. Deshmukh, has completed Ph.D. from Dr. B.A.M. University in 2002. He is working as a Professor in Computer Science and Information Technology (CSIT) Department, at Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS) INDIA. He is a Sectional President of Information and Communication Science & Technology (including Computer Sciences) section, Indian Science Congress. He is a fellow and Chairman of IETE, Aurangabad Chapter and life member of various professional societies as ISCA, CSI, ISTE, IEEE, IAEEng, CSTA, IDES, Etc. He has published more than 160 research papers in various National and International Journals and Conferences.