International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-10 Issue-9, July 2021

Principle Component Analysis for Crop Discrimination using Hyperspectral Remote Sensing Data

Pooja Vinod Janse, Ratnadeep R. Deshmukh

Abstract*: Crop discrimination is still very challenging issue for researcher because of spectral reflectance similarity captured in non-imaging data. The objective of this research work is to focus on crop discrimination challenge. We have used ASD FieldSpec4 Spectroradiometer for collection of leaf samples of four crops Wheat, Jowar, Bajara and Maize. We used vegetation indices and some spectral reflectance band for featuring our dataset. We applied Principle Component Analysis (PCA) for discrimination and it has been observed that when we use first and second principle component, it will give poor result but if third principle component is used then we get accurate and fine results.

***Keywords*: Crop Discrimination, ASD FieldSpec 4 Spectroradiometer, Principle Component Analysis (PCA), vegetation indices.**

I. INTRODUCTION

Monitoring crop condition applications like Land Use Land Cover (LULC), biophysical characteristics, crop yield prediction, crop growth monitoring, etc. requires high dimensional and detailed data. Hyper spectral data proven to be satisfactory data for providing additional information and improvements in results obtained by many researchers.

But crop type classification found to be challenging task for many researchers till date. It happens due to similarity of reflectance pattern in spectral signature. Thenkabail *et al*. [1] performed arduous hyperspectral data analysis for classification of crops based on many technique consisting of principal components analysis (PCA), lambda-lambda models, stepwise Discriminant Analysis (SDA) and vegetation indices. These hyperspectral approaches increases accuracy for crop classification from 9% to 43%.

Miglani et.al. [2] evaluated hyperspectral remote sensing satellite data of Hyperion classifying different winter crops such as mustard, sorghum, wheat, sugarcane and potato using PCA and band-to-band correlation analysis for the feature selection step. In the crop spectra, near 690 nm shows chlorophyll absorption, a steep slope in the red edge region (700–750 nm) and near 940 and 1104 nm leaf water absorption were extraordinarily marked.

Carlos et.al. [3] performed hyperspectral data analysis and discriminated soybean plant's spectral behavior to detect genetic seperability of soybean crop with remote sensor and he used vegetation indices and multivariate statistics to discriminate soyabean varieties. He used multivariate analysis and vegetation indices like EVI, NDVI, GNDVI, SAVI, TVI and OSAVI. It is observed that simulated Discriminant and discriminant analysis shows satisfactory results, with average global hit rates of 99.28 and 98.77%, respectively.

Jan Rudolf Karl Lehmann et.al. [4], have performed on Acacia longifolia (Native Shrub) to spectrally discriminate from other non-native and native species. He used jump correction followed by a first-derivative Savitzky-Golay smoothing with a second polynomial order and a filter width of nine points, PLS regression, PCA-LDA and found that better estimation of A. longifolia was achieved by using regions of wavelength between 1360–1450 nm and 1630–1740 nm, and got the accuracy of 98.9%.

Shreedevi Moharanaa et.al. [5], performed rice crop classification using hyperspectral remote sensing data. She used hierarchical clustering technique, waveform classification approach significant Wavebands. From this hierarchical clustering, library of spectral signature of rice crop were identified which will benefit to create classification maps of rice crop and critical wave bands like 519nm, 559 nm, 649 nm, 729 nm, 779nm, and 819 nm were marked subtle to nitrogen which will additionally helps in mapping of nitrogen from paddy agriculture.

Jeffrey H. Wilson et.al. [6], discriminated five cash crops Soybean, Canola, wheat, barley and oat using ViewSpec Pro for extracting text data, Stepwise Discriminant Analysis (SDA). He found hyperspectral bands in the visual and near infrared (NIR) regions (400–900 nm) can be used to excellently differentiate between five crop species under investigation.

II. MATERIALS AND METHODS

A. Study area

Leaf samples of Bajara, Cotton, Jowar, Maize and Wheat from Paithan Road, Chikalthana and Harsool road Area of Aurangabad region were collected. The latitude value of Aurangabad, Maharashtra is 19.8762° N and longitude value is 75.3433° E. The mentioned study area is consist of both man made material and agricultural fields.

Manuscript received on June 28, 2021.
Revised Manuscript received on July 18, 2021.
Manuscript published on July 30, 2021.
* Correspondence Author

Miss. Pooja V. Janse*, Department of Computer Science and Engineering and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (Maharashtra), India.

Dr. R. R. Deshmukh, Professor, Department of Computer Science and Engineering and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (Maharashtra), India.

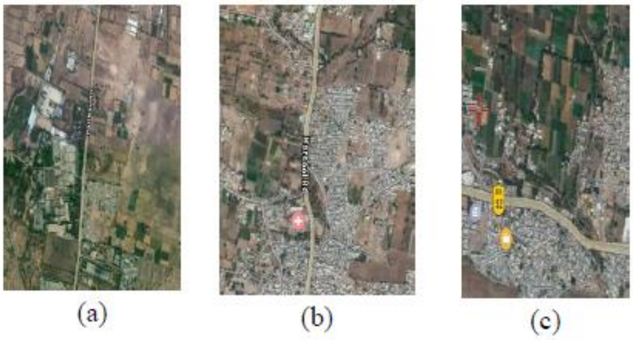

Fig. 1: Satellite image of the study area (a) Paithan road, (b) Harsool road, and (c) Chikalthana

B. Leaf Sample Preparation and Laboratory Setup

Leaf samples of Bajara, Cotton, Jowar, Maize and Wheat were cut from plant and immediately kept in air tight plastic bag so that there will be no or very less loss of its biological properties. These samples then brought to laboratory in Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra for spectral signature generation.

We have collected spectral signature of leaf samples by using RS3 software and ASD FieldSpec4 spectroradiometer. We collected database in laboratory in controlled condition, it was a dark room specially created for spectral data collection because other colors and light sources will affect the spectral signature. The spectrometer setup is shown in figure 3.

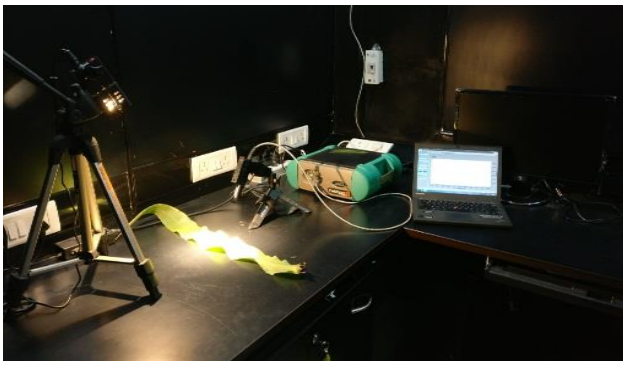

Fig. 3: Spectrometer Setup in data collection process

C. Instrumentation and Software

Analytical Spectral Device (ASD) FieldSpec 4 Spectroradiometer is a prime instrument used for generating hyperspectral signature data. It is general-purpose instrument which has demonstrated effectiveness of application in several areas which require measurement of radiance, reflectance, irradiance, or transmittance. FieldSpec 4 Spectroradiometer is a portable and field transferrable device of incredible precision, with provides a range of spectra from 350 nm to 2500 nm.

RS3, ViewSpec Pro, Microsoft Excel are essential software's which are used in this study. RS3 software circumstances to the version of the Analytical Spectral Devices application. Its main purpose is to store and receive the signature of material data transmitted from ASD Spectroradiometer. The software ViewSpec pro is used for converting .asd file into ASCII file of .txt format. ENVI is useful for analysis, visualization and presentation of data.

We have warmed up ASD Fieldspec4 Spectroradiometer for 30 minutes before data collection. Distance between spectral gun and light source was kept 50 cm. Distance between spectral gun and sample was set to 8 cm. and 8^0 Field of View (FOV) was used. Then we used RS3 Software for capturing spectral signature. The splash screen of RS3 is shown in figure 4.

Then Spectroradiometer have been optimized first to set the appropriate light source settings. Then white reference panel reading have been measured. Then spectral signature of leaf samples have been collected. Leaf spectral measurement provides reflectance values from 350nm-2500nm.

III. METHODOLOGY

A. Spectral Data Processing for analysis

The spectral data which is captured using ASD Spectroradiometer were acknowledged and saved using RS3 software in file format of .asd. Viewing of the spectral data can be possible with ViewSpec Pro after required transformation from .asd format to ASCII in .txt format. Then for the data conversion process ViewSpec Pro was used. All data saved in .asd format were transformed to ASCII and later saved as .txt format. Furthermore, the .csv data preparation process was done by using MS-Excel to average all the spectra data obtained from ASD Spectroradiometer device.

The reflectance band showing absorption characteristics of particular molecule are taken into consideration for statistical analysis. Table I shows spectral band and its absorption properties.

Table-I: Vegetative characteristics and their centered spectral band

Reflectance Band	Vegetative Characteristics
370	Phototropism
420	a - Carotene
425	b - Carotene
430	Chlorophyll a
440	a - Carotene
445	Xanthophyll and synthesis of chlorophyll
450	b - Carotene
453	Chlorophyll b
470	a - Carotene
475	Xanthophyll
480	b - Carotene
650	Synthesis of Chlorophyll
960	Chlorophyll absorption
1100	Chlorophyll absorption
1400	Water absorption
1930	Water absorption
2200	Peak Al – OH, Mg – OH, CO_3

B. Vegetation indices

Vegetation indices are used as potential variables for crop type discrimination. Some of those are used in this research work for getting more clear result of classification. Table II shows vegetation indices used in crop discrimination process.

Table-II: Vegetation indices calculated from hyperspectral data

Sr. No.	Vegetation Indices	Equation	Use
1	Lipidium Index (LI)	R_{630}/R_{586}	Sensitive to Lipidium in visible range which shows bright reflectance display
2	Normalized Difference Vegetation Index (NDVI)	$(R_{864}-R_{671}) / (R_{864} + R_{671})$	React to variation in amount of green biomass and more competently in vegetation
3	Simple Ratio (SR)	R_{864}/R_{671}	Same as NDVI
4	Pigment Specific Normalized Difference (PSND)	$(R_{800} - R_{470}) / (R_{800} + R_{470})$	Estimates LAI and Carotenoids
5	(RVI)	R_{1088}/ R_{1148}	Calculate water content and LAI at canopy Level
6	Water Index (WI)	R_{900} / R_{970}	Calculate leaf level relative water content
7	SGI (Sum Green Index)	$(R_{508}+R_{518} +R_{528}+R_{538}+R_{549} +R_{559}+R_{569}+R_{579} +R_{590}+R_{600}) / 10$	Closely related to greenness and leaf pigments like chlorophyll and carotenoids
8	Red Edge NDVI (RENDVI)	$(R_{752} - R_{701}) / (R_{752} + R_{701})$	Shows spectral variations associated with red edge wavelength position which may be affected by changes in chlorophyll concentration or water stress
9	Vogelmann Red Edge Index- I (VOG-I)	R_{742} / R_{722}	Same as RENDVI

By using vegetation indices and selecting wavelengths which shows particular molecule absorption properties we have prepared feature dataset and it is further used for classification.

We have applied Principle Component Analysis (PCA) and Random Forest classifier on created dataset.

IV. RESULT AND DISCUSSION

Principle Component Analysis (PCA) is dimensionality reduction tool used in this research and then Random Forest Classifier is applied over dataset. When we selected principle component equal to one then 86.2% accuracy has been observed. When principle component was set to two then we got accuracy of 72.4%. 100% accuracy was achieved when we set principle component equal to three.

Table-II: Results of PCA and Random Forest Classifier on dataset

PC=1				
	Precision	**Recall**	**F1-Score**	**Accuracy**
Bajara	0.71	1	0.83	86.2%
Jowar	1	1	1	
Maize	1	0.33	0.5	
Wheat	1	1	1	
PC=2				
Bajara	0.71	1	0.83	72.4%
Jowar	1	1	1	
Maize	0.33	0.33	0.33	
Wheat	1	0.5	0.67	
PC=3				
Bajara	1	1	1	100%
Jowar	1	1	1	
Maize	1	1	1	
Wheat	1	1	1	

Precision and recall values are used for retrieving information about classification. When precision and recall values are nearer or equal to 1 then we say that classification results are good. But if precision and recall values are nearer to zero that means our classification model is giving higher variance.

V. CONCLUSION

Mapping and detecting plant species using hyperspectral remote sensing is quite complicated task because many plant species shows similar reflectance properties. In collection of database we observed that reflectance property of wheat and maize was near about similar. So for discrimination, considering only reflectance properties are not sufficient because it has given us very poor classification result. We have increased our feature set by calculating vegetation indices. It has been observed that NDVI, SR, RVI and RENDVI shows better classification results.

REFERENCES

1. Miglani, A., Ray, S. S., Pandey, R. and Parihar, J. S., “Evaluation of EO-1 Hyperion data for agricultural applications”, J. Indian Soc. Remote Sensing, 36, 255–266, 2008.
2. Carlos Antonio da Silva Junior, Marcos Rafael Nanni , Muhammad Shakir , Paulo Eduardo Teodoro , José Francisco de Oliveira-Júnior , Everson Cezar , Givanildo de Gois , Mendelson Lima , Julio Cesar Wojciechowski, Luciano Shozo Shiratsuchi, “Soybean varieties discrimination using non-imaging hyperspectral sensor”, Elsevier, Infrared Physics & Technology 89, 338–350, 2018.
3. Jan Rudolf Karl Lehmann , André Große-Stoltenberg, MeikeRömer and Jens Oldeland, “Field Spectroscopy in the VNIR-SWIR Region to Discriminate between Mediterranean Native Plants and Exotic-Invasive Shrubs Based on Leaf Tannin Content”, Remote Sens., 7, 1225-1241, 2015.
4. Shreedevi Moharana , Subashisa Dutta, “Hyperspectral Remote Sensing Of Paddy Crop Using In-Situ Measurement And Clustering Technique”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-8, 2014.
5. Jeffrey H. Wilson, Chunhua Zhang and John M. Kovacs, “Separating Crop Species in Northeastern Ontario Using Hyperspectral Data”, Remote Sens., 6, 925-945, 2014.

AUTHORS PROFILE

Miss. Pooja V. Janse, is currently pursuing the Ph.D. degree in Computer Science and Engineering from Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, and Maharashtra, India. She is currently working as BSR Senior Research Fellow sanctioned by UGC to the Computer Science and Information Technology Department, Dr. B.A.M. University, Aurangabad. She is a life member of CSI, ISCA, CSTA and IAEng. She has published more than 24 research papers in reputed National and International Journals / conferences. Her research interest includes the digital speech signal processing, Remote Sensing and Geographical Information System (GIS) Technology.

Dr. R. R. Deshmukh, Professor, M.E., M.Sc. (CSE) Ph.D. FIETE, Presently working as Professor in Department of CS & IT, Dr. B.A.M. University, Aurangabad, (MS), India. He is a Fellow member and working as Chairman, IETE Aurangabad Centre, He has been elected as sectional member of ICT section of Indian Science congress Association. He is life member of ISCA, CSI, ISTE, IEEE, IAEng, CSTA, IDES and a senior member of ACEEE. He was Vice-Chairman of IETE Aurangabad Centre for last 4 years and member of Management Council of Dr. B.A.M. University. He is Member of Academic Council and Senate member of Dr. B.A.M. University, Aurangabad. He published more than 170 research papers in reputed National and International Journals / conferences. He is reviewer and editor of several journals at national & international level. His areas of specialization are Human Computer Interaction, Digital Speech Signal processing, Data Mining, Data Warehousing, Image Processing, Pattern Recognition, Artificial Intelligence, Computational Auditory Scene Analysis (CASA), Neural Networks etc.

Retrieval Number: 100.1/ijitee.I92970710921
DOI: 10.35940/ijitee.I9297.0710921
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication