IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Special Issue-14, May 2019 E-ISSN: 2347-2693

Classification of land cover using Data Analytics for Hyperspectral Imaging

Sakshi Suman[1*], Sanjana Suman[2], Santhosh J[3], Sham Vignesh[4], K Anitha[5]

[1, 2,3,4,5]School of Computing & Information Technology, REVA University, Bangalore, India

Corresponding Author: sakshisuman12@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7si14.345348 | Available online at: www.ijcseonline.org

***Abstract*—** Recent advances in remote sensing technology have made hyperspectral data with hundreds of narrow contiguous bands more widely available. The hyperspectral data can, therefore, reveal narrow differences in the spectral signatures of land cover classes that appear to be similar when viewed by multispectral sensors. If successfully used, the hyperspectral data can yield higher classification accuracies and more detailed class taxonomies. In this approach, we are using deep learning and neural networks to train a model for classifying land cover using data analytics in hyperspectral imaging.

***Keywords*—** *Hyperspectral imaging, Land cover classification, Deep learning, Tensor flow*

I. INTRODUCTION

Hyperspectral Image classification is the process of labelling the different landscape features. In our approach, we are using Deep Learning and Neural Networks to train a model and classify an input hyperspectral image. This type of classification can help to understand the landscape features of a particular area which can be used to predict land usage and also to suggest optimal use of land.

Deep learning is a subfield of machine learning which uses artificial neural networks that are inspired by the structure and function of the biological neural network[1]. In spite of being a new approach, it has become very popular recently. Deep learning has achieved greater success in many applications where machine learning has been successful only at certain rates. Specifically, it is preferred in the classification of big data sets because it can provide fast and efficient results. In this study, we used Tensor flow, one of the most popular deep learning libraries to classify dataset, which is often used in data analysis studies. It is an open source artificial intelligence library developed by Google. Using Tensor flow, we have studied and compared the effects of multiple activation functions on classification results. In this Study, Convolutional Neural Network (CNN) is used as a deep learning artificial neural network.[11]

Here, we are using the Indian Pines data set for training and classification. The Deep learning framework used for this research is Tensor Flow.
The applications of hyperspectral image classification are given below:
A planned strategy for proper utilization of land

- Reduction of unplanned construction
- Identification of different areas to differentiate between agricultural and non-agricultural land
- Fixing of tax policies by the government by knowing the rate of growth
- Get to know the rate of change in population
- Geological exploration from far to near field, Minerals, detection and mapping rare Earth elements and base metals, Mapping and monitoring Mine waste
- Soil characterization and monitoring, digital soil mapping, quantitative soil spectroscopy, soil erosion and land degradation mapping
- Monitoring Dry land areas, Water resource management and early-warning signals of ecosystem shifts
- Mapping of aquatic ecosystems contamination with plastic debris
- Simulation of cereal canopy reflectance

In this project, we are proposing a method to extract features of landscape cover from hyperspectral images. Some of the related works are summarised in the following.

II. RELATED WORK

N-FINDER Algorithm:
An end member is a pure and idealistic signature that can be used to specify a spectral class[5][7].
The hyperspectral unmaking problem is concerned with the decomposition of the hyperspectral image into a product form, where the spectrum in each pixel is represented as a collection of material spectra that are referred to as

endmembers and the mixing proportions of these materials in each pixel that are known as the abundances [4]. One popular approach to this problem is to first estimate the endmembers using endmember extraction algorithms such as N-FINDER. [3]

Convolutional Neural Networks:
CNN can also be effectively employed to classify hyperspectral data if an appropriate layer structure is used. CNN operates on a per pixel basis and that convolution is made along the spectral dimension. The paper presents results from land coverage classification indicating that the method achieves performance rates competitive with SVMs and other DL based methods. In a majority of the papers being reviewed, SVM has been used as a reference method.[6]

Machine Learning:
Machine Learning (ML) is the science of algorithms that can automatically learn from data to find patterns, to define features, and to assign properties. The illustration includes two steps: feature extraction and feature-to-label mapping.

Deep Learning:
Deep learning is part of a larger family of machine learning methods based on learning data characterization. Learning is of 3 types, unsupervised, semi-supervised and supervised.[2] Deep learning uses a cascade of multiple layers of nonlinear preprocessing units for feature extraction and transformation. Each consecutive layer uses the output from the previous layer as input.[8][9]

III. METHODOLOGY

Hyperspectral Image:
A HSI dataset was downloaded. The data was pre-processed and all the noise was removed. We converted the image into .mat form.

Segmentation Approach:
The next step was the literature review and analysis of all the existing methodologies for the segmentation.

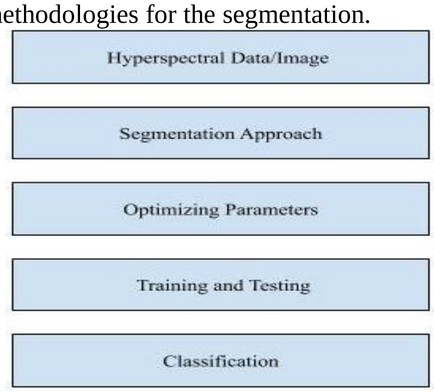


Fig1. System Architecture

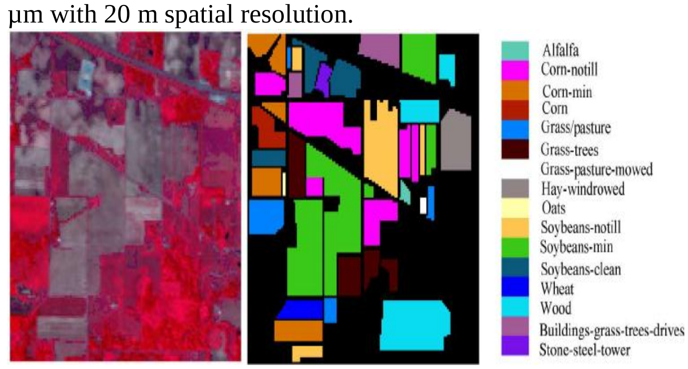
Optimizing Parameters:
The previous step was followed by detailing of the image and optimizing the different parameters of the HSI image including size, number of bands, noise and the respective area of the given satellite image.

Training Data:
Also explained earlier, these steps involve deep learning where the machine is trained to recognize patterns using a training data set. Here we have got a training data set and also a ground truth image of the format .mat.

Framework: An open source tool called Tensorflow is used for the machine training purpose. The output hence is a good segmented image from which buildings may be classified and found out.

IV. RESULTS AND DISCUSSION

A. **Dataset Description:**
We have used the widely used, Indian Pines dataset, which is captured by the visible/infrared imaging spectrometer (AVIRIS) in Northwestern Remote Sens. 2018, 10, 1271 8 of 20 Indiana. It covers the wavelength ranges from 0.4 to 2.5 µm with 20 m spatial resolution.



(a) (b)

Fig2. Indian Pines dataset and corresponding ground truth. (a) False color composite image (R-G-B = band 50-27-17); (b) The ground truth image with 16 land-cover classes.

B. **Classification Results**

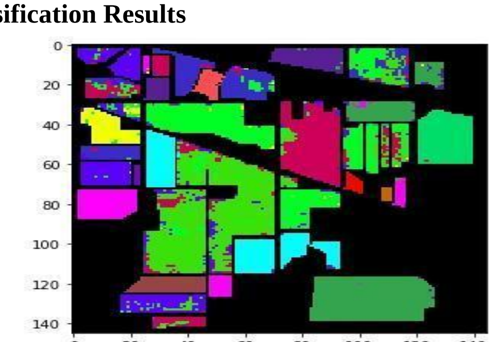


Fig3. Classified image

49.52230623956309 Test loss (%)
82.91065158017948 Test accuracy (%)

	Precision	Recall	f1-score	support
Alfalfa	1.00	1.00	1.00	11
Corn-notill	0.70	0.76	0.73	357
Corn-mintill	0.77	0.78	0.78	208
Corn	0.81	0.86	0.84	59
Grass-pasture	0.95	0.96	0.95	121
Grass-trees	0.99	0.98	0.99	183
Grass-pasture-mowed	1.00	1.00	1.00	7
Hay-windrowed	1.00	1.00	1.00	120
Oats	0.83	1.00	0.91	5
Soybean-notill	0.70	0.83	0.76	243
Soybean-mintill	0.84	0.67	0.75	614
Soybean-clean	0.67	0.82	0.74	148
Wheat	1.00	1.00	1.00	51
Woods	0.98	0.95	0.96	316
Buildings-Grass-Trees-Drives	0.81	0.93	0.87	97
Stone-Steel-Towers	0.96	1.00	0.98	23
micro avg	0.83	0.83	0.83	2563
macro avg	0.88	0.91	0.89	2563
weighted avg	0.84	0.83	0.83	2563

This chart describes how well the model has performed by listing out the measurements of different parameters.
Precision: It is the ratio of correctly predicted positive observations of the total number of predicted positive observations.

Recall: It is the ratio of correctly predicted positive observations to all observations in positive class.
F1 score - It is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into consideration. It is not as easy as accuracy but in the case of uneven class distribution, F1 is more useful.

Support- It is the number of actual occurrences of the class in the specified dataset. If support is not balanced in the training data, it may indicate structural weaknesses in the reported scores of the classifier and could indicate the need for stratified sampling or rebalancing.

After performing the experiment, we have come to the conclusion that the resultant accuracy in prediction is about 82.9%.

V. CONCLUSION AND FUTURE SCOPE

This paper proposes a method for segmentation and classification of land cover through deep learning. Deep learning is used when we have an enormous amount of data and things can go complex. In this study, to tackle this problem of large data and information paving a way to complexities, we used Tensorflow, one of the most popular deep learning libraries to classify dataset, which is often used in data analysis studies. Using Tensorflow, we have studied and compared the effects of multiple activation functions on classification results. In this Study, Convolutional Neural Network (CNN) is used as a deep learning model.

The most important part of this project is its usage of classifying land cover. Depending upon the requirements we can further narrow down or reduce the dimensionality for better efficiency. For instance, eliminating water bodies spectral region was our way to reduce dimensionality as we were concerned with the distinguishing of the different agricultural areas. Henceforth, we can see that a single hyperspectral image with its given ground truth can be put to use in different ways, gathering more information contributing to higher accuracy.
Further improvement in this project could lead to more accurate results. It can be enhanced by:

- Getting the hit ratio up and reducing the time consumed.
- A method for proper pre-processing so that most of the noise may be removed.
- Proper enhancement of the project so that it may be put to industrial use to classify the different land covers.
- Integrate with other projects of similar aims to boost productivity

ACKNOWLEDGEMENT

The completion of any task depends upon the co-operation, coordination and consolidated efforts of several resources of knowledge, energy and time and above all on proper guidance.

We are grateful to the head of C&IT School Dr Sunil Kumar S Manvi for providing us with numerous facilities and opportunities in the form of laboratories and equipment.
We owe this moment of satisfaction, with a deep sense of gratitude to our project guide Prof. Anitha K, for the technical guidance, persistent encouragement, perpetual motivation and everlasting patience, with whom this project would not be successfully completed. Working under their guidance has been a very faithful and valuable experience. We are sincerely thankful to all the members of staff of the department of C&IT, REVA University and all those who

have helped us directly or indirectly throughout the project work.

References

[1]. *"Hyperspectral Image Analysis using Deep Learning"*- a Review Henrik Petersson, David Gustafsson and David Bergström Swedish Defence Research Agency (FOI)Division of C4ISR

[2]. Yushi Chen, Zhouhan Lin, Xing Zhao, Gang Wang, and Yanfeng Gu. *"Deep learning-based classification of hyperspectral data".* Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 7(6):2094–2107, June 2014

[3]. *"Random N-Finder (N-FINDR) Endmember Extraction Algorithms for Hyperspectral Imagery"* Chein-I Chang, Fellow, IEEE, Chao-Cheng Wu, Member, IEEE, and Ching-Tsorng Tsai

[4]. N. Keshava and J. Mustard, *"Spectral unmixing,"* IEEE Signal Process. Mag., vol. 19, no. 1, pp. 44–57, Jan. 2002

[5]. M.Winter, *"Fast autonomous spectral end-member determination in hyperspectral data,"* in Proc. 13th Int. Conf. on Applied Geologic Remote Sensing, vol. 2, Vancouver, pp. 337–344, Apr. 1999.

[6]. J. M. Nascimento and J. M. Bioucas-Dias, *"Vertex component analysis: A fast algorithm to unmix hyperspectral data,"* IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 898–910, Apr. 2005.

[7]. N. Dobigeon, S. Moussaoui, M. Coulon, J. Y. Tourneret and A. O. Hero, *"Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery,"* IEEE Trans. Signal Process., vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[8]. Makantasis, K.; Doulamis, A.; Doulamis, N.; Nikitakis, *"A. Tensor-based classification models for hyperspectral data analysis".* IEEE Trans. Geosci. Remote Sens. 2018, 56, 1–15. [CrossRef]

[9]. Zhang, M.; Li, W.; Du, Q. *"Diverse region-based CNN for hyperspectral image classification".* IEEE Trans. Image Process. 2018, 27, 2623–2634. [CrossRef]

[10]. http://www.escience.cn/people/feiyunZHU/Dataset_GT.html the website for dataset and ground truth.

[11]. *"TensorFlow: A System for Large-Scale Machine Learning"* Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean,Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur,Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker,Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, Google Brain

Authors Profile

Miss. Sakshi Suman currently pursing B.Tech degree in Computer Science and Engineering, School of Computing & Information Technology, REVA University, Bengaluru, Karnataka, India.

Miss. Sanjana Suman currently pursing B.Tech degree in Computer Science and Engineering, School of Computing & Information Technology, REVA University, Bengaluru, Karnataka, India.

Mr. Santhosh J currently pursing B.Tech degree in Computer Science and Engineering, School of Computing & Information Technology, REVA University, Bengaluru, Karnataka, India.

Mr. Sham Vignesh currently pursing B.Tech degree in Computer Science and Engineering, School of Computing & Information Technology, REVA University, Bengaluru, Karnataka, India.

Mrs K Anitha pursed Bachelor Degree in Electronics and Communication Engineering from AMIE, IEI, India in year 2011 and M.Tech degree in Digital Communication Engineering from Visvesvaraya Technological University, Belgaum, Karnataka, India in year 2014. She is currently working as Assistant Professor in School of Computing and Information Technology, , REVA University, Bengaluru, Karnataka, India. She has 5 years of teaching experience and 3 years of Industry Experience. Her area of interest is Data analytics, IoT, Image processing & Digital Communication.