International Journal of Electronics, Electrical and Computational System
IJEECS
ISSN 2348-117X
Volume 6, Issue 10
October 2017

Academic Science

Analyzing the Extraction of Various features from Multispectral Satellite Data and Generation of quick alerts for Emergency Management

Ch. Rajya Lakshmi [1]
[1]Assistant Professor, BVRITN,
Dr. K.Rammohan Rao[2],
[2]Senior Scientist NRSC, Hyderabad
Dr. R.Rajeswara Rao[2]
[2]Professor & Head, JNTUKUCEV

Abstract: *Now a days Geospatial information processing is very critical for effective, collaborative decision making process during emergency situations. Processing the multispectral high resolution satellite imagery data with accuracy is very big challenging issue Because that data includes various Patterns with heterogeneous data items ,those data items classification and Features Extraction is a very Difficult task. This paper focus on how various features like water bodies, vegetation lands, mining lands can be extracted from High Resolution Multi _Spectral satellite image data can be processed with an efficient and scalable manner. The key objective of this paper is how Multispectral and heterogeneous data can be processed for finding of various with accurate and efficient results using decision tree c4.5 classification With Hadoop Google Map Reduce Architecture. After Data processing Completion the resultant data compare with previous Data and also give quick alert messages to the Decision Support Management for Handling of Emergency situations.*

Key Words:GeoSpatial,HadoopMapreduce,C4.5,Decision Support Management

I.Introduction:

The most fundamental challenge for Big Data applications is to explore the large volumes of data and extract useful information or knowledge for future actions [3]. Big Data as an emerging trend and the need for Big Data mining is arising in all science and engineering domains. Satellite multispectral remote sensing imagery data has been used over decades for feature extraction. Satellite images obtained from remote sensing is the only way to detect and study any feature locally or globally that data includes various features like water bodies, vegetation lands ,schools ,hospitals, coastal areas like etc. The growth has lead to an emerging field known as Big Data which is nowadays positioned among top ten strategic technologies. GIS systems handle complex and large amounts of spatial big data [1]. This type of data is generated via satellites, aircraft-mounted cameras, global positioning system devices, etc. Applications of GIScan be spatial search and imagery processing [2]. Satellite image classification is not complex, but the analyst has to take many decisions and choices in satellite image classification process.Currently,operators manually extract cartographic features, such as buildings, roads, and trees, by visual interpretation of satellite imagery and aerial photography. Semi-automatic algorithms to assist cartographic technicians would improve the process. In the previous papers so many authors studied various classification methods for satellite images (Bernardini et al., 2008; Alonso et al., 2007).

2.Literature Review on Decision Tree Classification for various Feature Extraction

The decision tree method includes two main Techniques: learning and classification. The input of decision tree learning algorithm depends up on attributes and attribute values of the training samples, and Now a days, more classification techniques for the extraction of features from satellite data have been reported based on the spectral separability of the features in the visible and infrared regions of the spectrum. We are identified

various adopting the mathematical formulae that deal with the band combinations of the electromagnetic spectrum output is a decision tree. Satellite images generated from remote sensing is the only way to detect and study any feature locally or globally. Generally spatial data can be collected in three formats those are vector data, raster data, network data[5]. Classification refers to the process of partitioning images into homogenous regions. And each regions corresponds to some particular land cover type. There are 3 main image classification techniques :1. Unsupervised image classification. It is based on the pixels clustering technique[6].2. Supervised image classification Technique .It is based on based on the spectral signature defined in the training set [7].3. Object-based image analysis [8].It is based on multi-resolution segmentation it generates homogenous image objects of different shapes and scales based on texture, context and geometry. In the previous research most of the authors extract various features like forest ,agricultural land, water bodies, bare lands depends up on accuracy assessment with Kappa coefficient value that lies between 0 and 1, where0 represents agreement due to chance only. Meanwhile 1represents complete agreement between the two data sets. Negative values can occur but they are spurious. It is usually expressed as a percentage from 83.5% to 89.17%,[10].The below figure2.1 shows various Decision Tree Classification Techniques and their importance for the high spectral image classification.

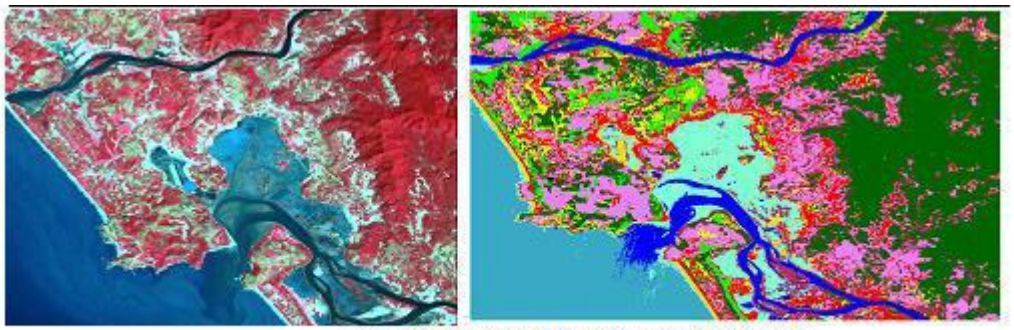
Authors	Decision Tree Classification Techniques
Hunt et al.1996	Classified Objects using Decision Tree
Morgan et al.1963	Inspired by Automatic Detection(AID) Program
Morgan et al.1973	Chi-Square Automatic Interact Detection
Quinlan 1975	Decision Tree ID3
Breiman et al.1984	Classification Regression Tree(CART)
Quinlan 1993	C4.5
Murthy et al.1994	Oblique Classifier(OC1)
Friedl and Brodley 1997	Tested the Classification problems for Remote Sensing Data
Punia et al.(2011)	C5.0
Everendlek and Gulheyar(2011)	Boosted Decision Tree

Fig 2.1 Decision Tree Techniques For Various Satellite Image Classification

3. **Satellite Image Data Processing with Hadoop Technique for Various Feature Extraction**

Hadoop is used to process terabytes of data with Distributed environment.. In the Hadoop architecture, data processing is very important, it affecting Hadoop performance. However ,in a heterogeneous environment, the data required for performing a task is often nonlocal, which affects the performance of Hadoop[11]. MapReduce is a programming model used in clusters that have numerous nodes and use considerable computing resources to manage large amounts of data in parallel. MapReduce is proposed by Google in 2004[11]. In the MapReduce model, Satellite Data Can be processed is called a “job”. A job can be divided into two parts: “map tasks” and “reduce tasks”, in which the map-tasks run the map function and the reduce-tasks run the reduce function. Map function processes input data assigned by the master and produce many intermediate _key, value pairs. Based on _key, value_pairs that are generated by map function processes, the reduce function then merges, sorts, and finally returns the result. The returned result can be classified with best classification algorithm for various feature extraction like schools,hospitals,waterbodied,vegetation lands, coastal lines etc.. By using this technique improve parallel computing for the various feature extraction from satellite image data rather than single processing for single feature extraction from high spectral data. This can shorten the process time of various feature extraction from large data sets of satellite imagery data. In the Hadoop Map Reducing Technique HDFS File system used for the distribution of data in various file formats. which stored large dataset on to various data nodes by dividing it into chunks to achieve parallelism. In the proposed system, we are performing Satellite Data processing including segmentation, computations, and result in merging by Hadoop Technique. At segmentation step, the whole image is segmented into blocks so

that these blocks can be processed in parallel to generate results. In our feature detection scenario, the image blocks are simply dividing into fixed size of blocks, as shown in Fig. 3.1. Each block is represented by the subscripts a and b where a represents the row and b represent the column number of the block. Each segmented block is then processed by computing statistical parameters with reference of Pixel value distribution of several image blocks and machine learning algorithms to extract various features in each block.



Befor and after classification of image

Fig 3.1 Study Area [Multispectral Image(resolution=5.8m)][12]

Pixel value distribution of several image blocks, such as water bodies, vegetation lands ,land cover block with no water bodies, vegetation land, and Land block with water bodies, and two blocks that have only one area comes under water body area inspected. The distribution among pixel values is quite low for all water bodies pixels as shown inFig.3.1.

Table1. Statistical measurements of Water Bodies classified blocks by C4.5 Classification

Block	Block size (no. of values)	No. of Water Bodies in Blocks	Mean value of Water Bodies pixels Mean_W	Mean value of the Water Bodies Pixels Mean_L	|Mean_L−Mean_W|
1	10,000	1	356	2225	1863
2	10,000	2	251	1525	1274
3	10,000	1	1750	3458	1708

Table1 clearly shows that the absolute difference between the mean value of Water Bodies sub block and the mean value of the complete block is very high as shown in the last column of the table. The difference of various blocks fron one feature to another feature can be classified with in a block wise shown in the Fig3.2 .

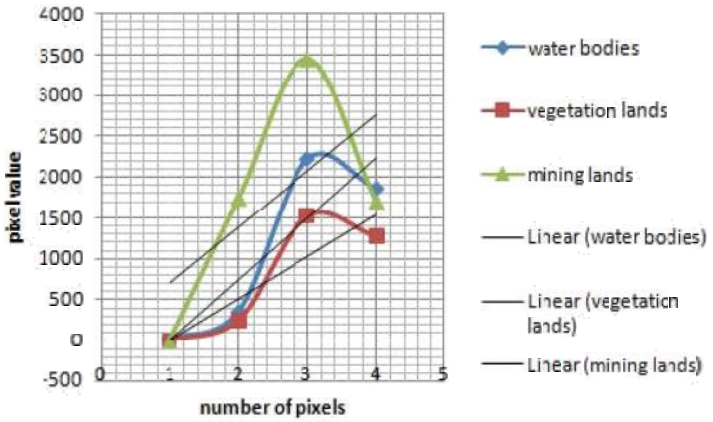
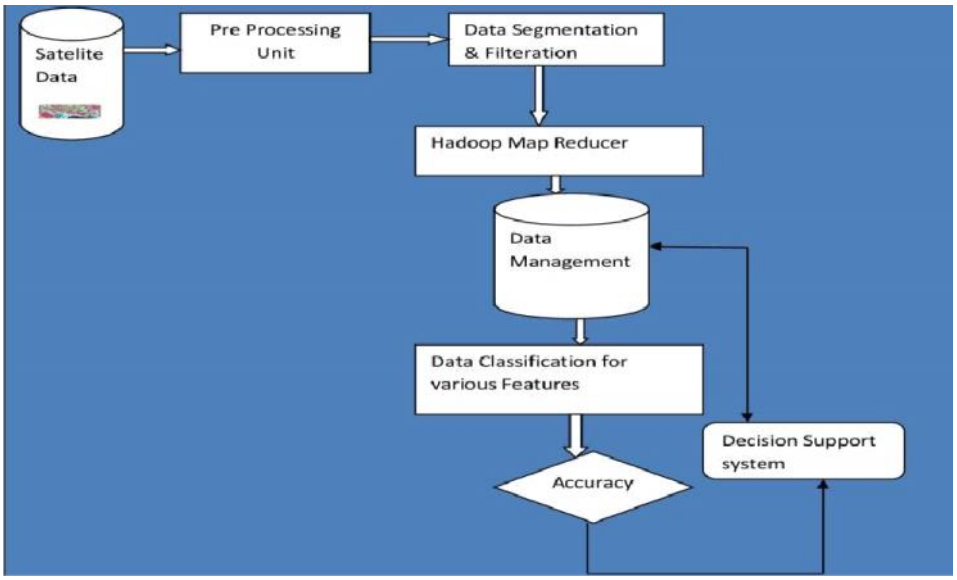


Fig 3.2 Pixel value distribution with classification algorithm

Academic Science

4.Methodology

In this paper, we implement decision trees using the C4.5 algorithm and classify pixels. To overcome the effect of late data processing, we used the method of pruning with decision tree classification with c4.5 algorithm to improve classification accuracy. In this used pixel based continuous feature extraction method[13]. Decision trees represents in the form of parametric nature and classified using pixels values using C4.5 Algorithm[12].The various parameters can be classified by using pixel values like mean,varience, Euclidean distance [13]. In the Proposed Architecture the data processing unit process the satellite image data after processing it will sent to the filtration, the filtration and load balancer server have two basic responsibilities, such as filtration of data and load balancing of processing power. Filtration unit identifies the useful data for analysis since it only allows useful information, whereas the rest of the day using various parameters like mean,varience eta are blocked and are discarded. Hence, it results in enhancing the performance of the whole proposed system. Apparently, the load-balancing part of the server provides the facility of dividing the whole filtered data applied to Hadoop Mapreducer. Hadoop provides the facility of parallel, high-performance computing using a large number HDFS File Systems. Therefore, it is suitable for analyzing large amount of remote sensory image data. The proposed architecture uses a similar mechanism for load balancing; hence, preference is given to Hadoop for sophisticated analysis, algorithm development, and testing architecture. The results generated by each File system sent to the aggregation server for compilation, organization, and storing for further Processing[12]. Aggregation server stores the compiled and organized results into the result's storage with the intention that any server can use it as it can process at any time. The aggregation server also sends the same copy of that result to the decision-making server to process that result for making decision. The decision-making server is supported by the decision algorithm, which inquire different things from the result, and then make various decisions(e.g., in our analysis, we analyze land, sea, and ice, water bodies, swimming pools, where as other finding such as fire, storms, Tsunami, earthquake can also be found). At the end of the classification send these results to the decision management unit that must be strong and correct enough that efficiently produce results to discover hidden things and make decisions. The decision part of the architecture is very significant because any small error in decision-making can degrade the efficiency of the whole analysis and more hazards will comes. If Decision Unit is efficient it can quickly alert the system at the time of emergency situations.



IV.I Proposed Architecture with Satellite Image

Academic Science

5.Conclusion

In this paper, features extraction mechanism from satellite images is proposed by taking the Various features blocks of data. The various features detection with in satellite images is performed by us-ing proposed Architecture model. The implementation model has various units and phases including collection,filtration,segmentation. In this Method Hadoop Mapreducer is a very important data processing model, and c4.5 is very suitable classification algorithm, data can be processed with accuracy and scalable manner.

6.References

[1] Ed. S. Wang and M. GoodchildEnabling Spatial Big Data via CyberGIS: Challenges andOpportunities , under review as a book chapter in CyberGIS:Fostering a New Wave of Geospatial Innovation and Discovery(), Springer , 2014.
[2] Michael R. Evans, Dev Oliver,XunZhou, and Shashi Shekhar SpatialBig Data: Case Studies on Volume, Velocity, and Variety , in Big Data:Techniques and Technologies in Geoinformatics ,isbn 978-1-46-658651-2, CRC Press, 2014.
[3]. A. Rajaraman and J. Ullman, Mining of Massive Data Sets.Cambridge Univ. Press, 2011.
[4]. Alonso, M.C., Sanz, A. and Malpica, J.A., 2007. Classificationof high resolution satellite images using texture. In Bebis G. et al. (Eds.): *Advances in Visual Computing ISVC07*. Part II LNCS4842, pp. 499–508. Springer-Verlag.
[5]. Wadii Boulila and Imed Riadh Farah,Improvement of Satellite Image Classification :Approach Based on Hadoop/Map Reduce.In https://www.researchgate.net/publication/299598117.
[6] Lv, Z., Y. Hu, et al. Parallel K-means clustering of remote sensing imagesbased on mapreduce. Web Information Systems and Mining, Springer:162-170.2010
[7] Li Bo, et al.Parallel ISODATA clustering of remote sensing images based on MapReduce", in Proc. CyberC 2010 Huangshan, pp.380-383, 2010.
[8] Krishna Bahadur K. Improving Landsat and IRS Image Classification:Evaluation of Unsupervised and Supervised Classification through Band Ratios and DEM in a Mountainous Landscape in NepalRemoteSens.Volume1(4), 1257-1272;2009.
[9] T. Blaschke Object based image analysis for remote sensing ISPRSJournal of Photogrammetry and Remote Sensing 65, 216 ,2010.
[10]. Mohd Hasmadi Ismail, and Kamaruzaman Jusoff, Satellite Data Classification AccuracyAssessment Based from Reference Dataset.In World Academy of Science, Engineering and TechnologyInternational Journal of Environmental, Chemical, Ecological, Geological and Geophysical Engineering Vol:2, No:3, 2008.
[11].Chia-WeiLeea, Kuang-YuHsieha, Sun-YuanHsieha,b,∗, Hung-ChangHsiaoa , A Dynamic Data Placement Strategy for Hadoop in Heterogeneous Environments in Big Data Research 1 (2014) 14–22 in www.elsevier.com/locate/bdr
[12].Pooja A P, Jayanth J, Dr Shivaprakash Koliwad, Classification of RS data using Decision Tree Approach in International Journal of Computer Applications (0975 – 8887)Volume 23– No.3, June 2011
[13].M.MazharURathorea,AwaisAhmada,AnandPaula,∗,JiajiWub,Real-time continuous feature extraction in large size satellite images.In Journal of Systems Architecture 000 (2015) 1–11 [m5G;December 14, 2015;21:32]
[14] .Apache HBASE project, http://www.hbase.apache.org
[15].Raghavendra Kune1, Pramod Kumar Konugurthi1, Arun Agarwal2,Raghavendra Rao Chillarige2, and Rajkumar , XHAMI - Extended HDFS and MapReduce Interface for Image Processing Application.
[15] Sarade Shrikant D., Ghule Nilkanth B., Disale Swapnil P., Sasane Sandip R., "LARGE SCALE SATELLITE IMAGE PROCESSING USINGHADOOP DISTRIBUTED SYSTEM", IJARCET, ISSN : 2278 – 1323, Volume-3, Issue-3, March 2014
[17].Wei and dei and Wei and Ji,A Map Reduce Implementation Of C4.5 Decision Tree Algorithm in IJDTA Vol.7,NO.1 (2014),PP.49-60 .htt:p//dx.doi.org/10.14257/ijdta.2014.7.1.05.