Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 5, May 2021: 4958-4963

Land Cover Mapping Using Image Recognition

Kapila Moon, Rushikesh Yadav, Gandharva Tatkare and Saurabh Karambelkar
(kapila.moon@rait.ac.in) Department of Electronics Engineering,
D Y Patil, Deemed to be University
Ramrao Adik Institute of Technology, Sector 7, Nerul, Navi Mumbai

Abstract - Data of the land covered is not only a useful factor to evaluate earth's land change but it is also an essential product for government regarding to decision making and research. The simplest way for an efficient land cover mapping is through remote sensing data. The unique characteristics of these urban targets were taken into account to create the Object- Based Image Analysis (OBIA) in consensus with the traditional pixel-based technique of image analysis, and thus was regarded as a reliable process to compartmentalize image of high resolution. The reason for classification in remote sensing is to ascertain and classify the elements on geography of the earth. It is seen in many real-world applications, such as land use/cover mapping, urban planning, agriculture and geology, etc. The aim of such categorization is to categorize the image (pixel or object) where its category is unknown, and attach it to one of the many categories on the basis of the categorization process and rules.

Keywords: OBIA (Object Based Image Analysis), IEEE (Institute of Electrical and Electronic Engi4neers), RFA (Random Forest Algorithm), ISE (Integrated software environment), I/O (Inputs and Outputs).

I. INTRODUCTION

1.1 Motivation - The need for efficient use of land is growing day by day. There is a need of Technology in this domain so as to find better solutions. The land can be broadly classified into three categories Barren, vegetated and rest the one used by humans for homes/industries. The only way land in a area/region can be seen through technology is by satellite image, this image has the information about all three categories of land. But for person to manually classify the area under the three categories of land is a tiring task. Thus as the main motivation of this project it was though that this tiring task can be done easily and efficiently using Image Recognition. Then as there where challenges into this project, there were also many more advantages. Using various Algorithms, it is possible to distinguish the vegetated land correctly and quickly, not only the vegetated land but also using object-based image recognition various object could be distinguished on a satellite image. The further this idea motivated was that this technique can be used on a large scale so that any surveying body can easily get the idea and estimation about a land to be surveyed. As people tend to realize the need to use land efficiently the need and use of this type of techniques will grow. The cost to development of this project is cheap and also there is not much of a maintenance requirement.

1.2 Objective - The objective of OBIA is to establish and apply theory, processes, and tools for

automating the replication and improvement of human understanding of remotely sensed image images. Picture segmentation, or clustering of pixels into homogeneous objects, and eventual grouping or naming of the objects, as well as modeling based on the features of the objects, are all part of OBIA.

1.3 Problem Definition - The current problem of land scarcity is growing day by day. We need to have a tool to efficiently get data about the land. Without correct data it is not possible to manage land in a region, having data helps in many other ways too. OBIA will efficiently get the land data needed correctly and also quickly, it gives us details from a just a satellite image.

II. Literature Survey

Object-based image analysis (OBIA) incorporate the pixels, where pixels are first being clustered into objects based on either the spectral similarity or an external variable such as ownership, soil or geological unit. Many variables may be decided, categorized as spectral, shape and neighborhood. Some of the common examples of spectral variables are mean value and standard deviation of a specific spectral band; shape variables which involves size, perimeter and compactness; neighborhood variables indicate, for example - the mean difference of an object compared to darker ones. Objects are like hierarchy system where an object is a part of a 'super object', where super object is a combination of various small neighboring objects, where those small objects can be further subdivided into smaller objects: 'sub-objects'. With the help of OBIA, knowledge on a landscape may be included by introducing rules. When few trees or a pond or grass is located near urban area it is generally park or a yard. Or if there are lot of trees or big lake it might belong to forest. And with help of OBIA we can distinguish between them with ease but not with traditional image analysis. With the help of OBIA accuracy and detailfor classification purposes increases

III. System Design

A. Hardware Design - As the project is software based there are no major hardware requirements, basic computer systems with the required software are enough for implementation of the project

B. Software Design - In object-based image analysis each segment represents an object. Objects represent buildings, roads, trees, elds or pieces of those features, depending on how the segmentation is done. Image Segmentation is the process by which various subgroups (of pixels) called Image Objects, of a digital image are partitioned, which can reduce the complexity of the image, and thus analyzing the image becomes simpler.

Different Types of Segmentation:

1. Felzenszwalb's segmentation
2. Quickshift segmentation
3. SLIC - K-Means based image segmentation
4. Compact watershed segmentation of gradient images

C. Random Forest Algorithm - Random Forest Algorithm is a simple and flexible to use

machine learning algorithm which creates accurate results. Random forest is a algorithm which is supervised and constantly learning. The "forest" created by it, are an group of decision trees which are normally trained along with the bagging" method. The broad idea of the bagging process is that a concatenation of learning models increases the overall result. In simple words, to get a more accurate and stable prediction, random forest creates various decision trees and groups them together. One major advantage of the random forest algorithm is its flexibility which can be used for regression and classification functions, and we can also view easily the importance it gives to the input features.

D. Kmeans Unsupervised Algorithm - K means classification of crusting algorithm is unsupervised algorithm, it combines the unlabeled data-set into various clusters. Here K tells us the number of pre-defined clusters which need to be built in the method, for eg. if K=2, cluster count is 2, and for K=3, cluster count is 3, and so on. Kmeans allows us to classify the data into different groups and is a easy way to uncover the categories of groups in the unlabeled data-set, without any training and on its own.

IV. System ImplementationOBIA Algorithm Working

Step1 - Data preparation of image of land

1. Acquiring the desired data
2. The Landsat data and NAIP imagery by changing projection
3. Create a stack raster
4. Create different points for training and testing the model by selecting points on the data created in above steps

Step2 - Processing the data from Step1

1. Read NAIP images as numpy array
2. Segmentation of images
3. Get statistics for image segments
4. Splitting the set points created in step1 in trainingand testing set
5. Assigning statistics of segments to training dataset to train classification algorithm
6. Classify each image segment
7. Perform random forest algorithm to get the output

Kmeans Unsupervised Algorithm Working

1. First of all, we have to decide number of clusters byselecting 'k'.
2. Centroids of k points are selected randomly
3. Pre-defined k clusters are formed by assigning datapoints to their closest centroid

4. New centroids are found by calculating variance foreach cluster
5. Again repeat 3rd step for new centroids
6. If any point is reassigned then perform 4th step againor jump to next step
7. Model is ready

V. Advantages and Application

The major advantages of remote sensing are, the capability to acquire data over large coverage of area; and also classify the natural features or physical objects on the surface; to detect objects on a surface in structured manner and examine their changes as time passes.

1. It is a comparatively cheap, fast process to acquire current data over a large geo-graphical area.
2. It is also cost effective and fast method to construct maps if there are no accurate land surveys available.
3. It is flexible to change according to the computer, and mix with other geographic coverage processes.
4. Data sensed remotely can be easily analyzed with help of fast computers and can be used for a lot of different applications

The primary applications are:

(a) Analyzing the number of rural roads: -The number of Rural Road now can be analyzed using many Remote Sensing techniques with high precision. It is needed as it saves time and leads to overall development of rural areas.

(b) Analyzing land used and land covered: -Various physical properties of land can be determined, analysis of land being used for various reasons can be done through Remote Sensing technologies.

(c) Increasing quality and quantity in farming: -The quantity of farm land available can be determined which can be used increase or decrease farming, the neighboring objects such as lakes or pond/wells and habitats can be analyzed and thus the quality of farming can be increased.

VI. Conclusion and Future work plan

Accuracy depends on which algorithm is used for segmentation and classification, resolution of image. With more segments accuracy will increase but so will computation time With Slic segmentation algorithm and forest image classification algorithm we found the balance between time and accuracy. By using the Lidar data and NDVI as another layer in raster data the accuracy will increase significantly. Can be used very small land area or can scaled up for bigger projects. Gives accurate results and can be used in projects which need precision. Once data is obtained the same techniques are easy to apply. As no hardware is required It is cheap to use.

Kmeans is a unsupervised algorithm, random forest algorithm is supervised algorithms Since we cannot give in how much classes an image should be classified, some of the features gets classified in one single class like road and houses category and some get classified in wrong classes like bare

land and trees. With the help of supervised algorithms, we can define classes before the classification and image gets classified accurately. That's why random forest algorithm which is supervised algorithm is more accurate.

VII. Result and Analysis

Here we have shown our results which is the percentage of land covered, we have shown the percentage of color in an image thereby we can some idea about the percentage of the land covered by either Road, Bare land, Grass, Water, Constructed land. The data(percentage) we get is not accurate to the land if measured but this data helps and is accurate for mapping remotely or through image analysis. The aim was to help get of data of large areas which is not possible or will take years to map/survey.

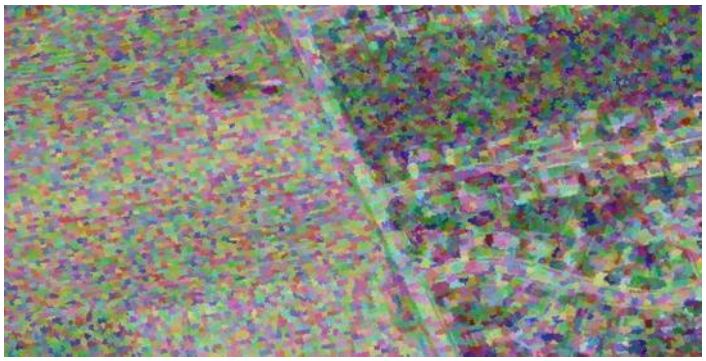

Fig.1 – SLIC segmentaion

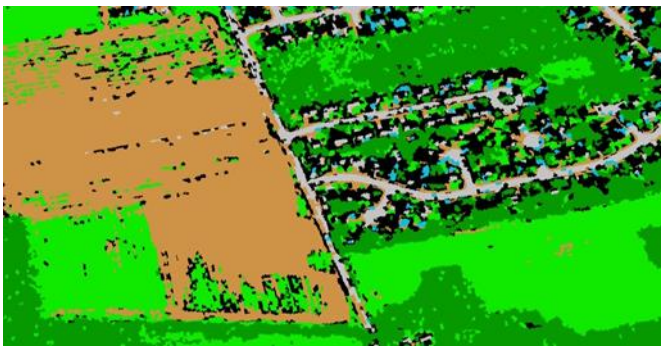

Fig.2 – Classification done using OBIA

Finally, by calculating how much percentage each colour occupies (in Fig.2) we can find how much percentage each land cover type is occupying

IX. REFERENCES

[1] Shafat Khan,Khalid Mohiuddin , \Evaluating the parameters of ArcGIS and QGIS for GIS Applications", International Journal of Advance Research in Science and Engineering, Vol No.7, January 2018.

[2] David G. M. Franca, Rodolfo G. Lotte, Claudia M. de Almeida, Sacha M.O. Siani et al. "Object-based image analysis for urban land cover classification in the city of Campinas - SP, Brazil", 2015 Joint Urban Remote Sensing Event (JURSE), 2015

[3] Steiniger, S. and Bochner, E. (2009), \An overview on current free open source desktop GIS developments" , International Journal of Geographical Information Science.

[4] Mohammed Chachan Younis, Edward Keedwell, Dragan Savic. "An Investigation of Pixel-Based and Object-Based

Image Classification in Remote Sensing", 2018 Inter- national Conference on Advanced Science and Engineering (ICOASE), 2018

[5] Jitendra Kumar Jaiswal, Rita Samikannu,"Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression"IEEE 2017

[6] Youguo Li, Haiyan Wu, A Clustering Method Based on K-Means Algorithm “, International Conference on Solid State Devices and Materials Science.

[7] Hariharan B, Arbelaez P, Girshick R 2, \Simultaneous detection and segmentation", European Conference on Computer Vision,297 - 312.

[8] Long J, Shelhamer E, Darrell T, Fully Convolutional Networks for Semantic Seg- mentation", Conference on Computer Vision and Pattern Recognition, IEEE, 3431 -3440.

[9] Dr Badri Basnet, Dr Ernest Dunwoody, \QGIS for distance learning and teaching".