# Real-Time Data Diagnostic Architecture for Satellite Image Mining using SIFT Algorithm

[1]V.P. Muthukumar
Assistant Professor,
PG and Research Department of Computer Science and Applications, Vivekanandha College of Arts and Sciences for Women(Autonomous), Tiruchengode,
Tamil nadu.

[2] Dr. S. Subbaiah
Assistant Professor,
Sri Krishna Arts and Science College,
Coimbatore, Tamil nadu.

***Abstract*— Remote sensors create an immense measure of information from satellites. These days, there is an enormous interest in ongoing information for remote detecting applications and to separate helpful data from the satellite picture. This paper analyzes distinctive mining approaches in different satellite picture applications. A two-level blending approach is utilized to separate the ocean region. In any case, the presentation is lower because of the absence of highlight extraction. The progressive district mixing approach is utilized to naturally distinct the ocean and land region. The consolidating approach can be better portrayed by directed data joined with the element extraction. This results in bringing out real-time analytical architecture to detect land and sea areas. The sensors are conveyed in the woodland to highlight out the temperature and weight for fire discovery. Sensors are of a significant expense. Sensors can be crushed because of climatic changes and by creatures. These issues can be maintained a strategic distance from by utilizing satellite picture digging applications for fire identification utilizing constant diagnostic engineering. The highlights of the satellite picture can be separated by utilizing the Scale Invariant Component Change (Filter) calculation. The speed of distinguishing fire utilizing the satellite picture can be expanded by the utilization of Hadoop, a parallel handling system.**

***Keywords—LANSAT; Remote Data; HadoopDstributed File System(HDFS); Synthetic Aperture Radar(SAD); Scale Invariant Feature Transform(SIFT); LANDSAT; Remote data.***

## I. INTRODUCTION

Starting late, a great deal of excitement for Information mining has risen. Mining strategies can be executed on new frameworks as existing techniques are updated and new items created. While mining apparatuses are actualized on elite parallel handling frameworks. In this way, clients can examine a huge database in minutes. The fast execution of mining techniques causes clients to break down huge amounts of information. The cutting edge innovation in mining systems gives out gathering, overseeing, breaking down and preparing of remote information. Remote sensors are intended to dissect the earth observatory framework. Numerous works have been done in the various fields of remote detecting information from the satellite, for example, angle based edge identification [1] and change discovery [2]. Ongoing information systems engineering is focused on rapid consistent constant and gigantic disconnected information.

These days, the world gets changed to the advanced world and producing an enormous measure of information. From these informational indexes, to break down information is in danger by utilizing the current innovation. The information is created by gushing information, in this manner the information will land at rapid and the calculation needs to process all the showed up information. In this manner, there is a requirement for a design to dissect both the ongoing and disconnected informational collections. The progression in the PC innovation and remote detecting build the gigantic measure of detecting information. The earth observatory information from the rocket is around 1.7GB, this information is gathered by a single satellite. This huge measure of information needs to procedure to remove the valuable data. Along these lines, there is a requirement for a design for load adjusting, data aggregation, and decision investigation.

Remote detecting is the study of getting data about regions from satellites. The satellites around the earth are creating assortments of the information inconsistently. Detecting pictures are as advanced pictures. The picture should be expertly processed productively. As of late, Orfeo tool stash is utilized to process pictures in enormous volumes [3], joined with MapReduce and Hadoop Conveyed Record Framework (HDFS). Orfeo brings about an proficient and decreased execution time, yet weaknesses in fast information handling. An issue with the large information scientific was the absence of coordination between databases. In this way, experts were blocked by a dull procedure of trading information from the database [4]. The examination of information utilizing Hadoop brings about quick parallel handling of information. Numerous information mining calculations are moving towards Hadoop, yet the accelerate of the parallel k-means algorithm is not linear in type. The principle reason brings about nonlinear were that the correspondence overhead increments as increment the dataset size. These issues can be understood by utilizing real time analytical architecture.

Remote sensors gather a lot of information from the satellite. The gathered information has no significance. The helpful information needs to get the concentrate from the gathered information. Now and again, the gathered

information may be not clear. To resolve such an issue utilizes architecture to analyze offline and real-time data.

## II. DATA MINING TECHNIQUES

Data Mining is the process of extracting knowledge from large amount of data. It offers a platform to do research. The large amount of data crates complexity in acquiring knowledge. Data Mining is the process of knowledge discovery which performs Selection, integration, Transformation and mining of data.


Figure 1: Data Mining Process[5].

The gathered image or data can be utilized for mining from various databases, information distribution center. Mining methods can be executed on existing programming stages just as equipment stages and can be coordinated with frameworks. The different strides in the information mining process are portrayed in figure 1. Data Mining is bolstered by three advancements, to be specific, huge data assortment, multiprocessor Systems, and Data Mining algorithm.

The most common techniques used in Data Mining are:

***Association:*** It is a simple and easy technique to relate two or more items.

***Classification:*** Generally decision trees are used to find classification.

***Clustering:*** It is strategy of dividing data into groups of similar objects.

***Prediction:*** It is a combination of trends, classification, and relations methods.

***Feature Selection and Extraction:*** It is the process of reducing attributes. The information gathered from satellite are then mined to get useful predictions. After that the feature extractions are carried out for next level data processing. In feature extraction method rearranges the huge measure of assets to clarify an enormous arrangement of information precisely.

## III. IMAGE MINING

Finding pieces of information from the gathered information put away in the social database has been significant work in information mining. The huge assortment of the picture can be mined to find new data. The primary issue of picture mining is that it consolidates the field of substance based recovery of pictures, databases and information mining.

The picture mining process has two stages. The first and significant stage is mining enormous measures of gathered pictures. Another part is joining the mining of gathered information and the relating numerical information.

In typical the image contents are a group of blocks, after that use some specific techniques to find feature of the images. Figure 2 illustrates the mining techniques of image. Then the gathered images are processed for feature extraction. After removing feature from the mage mine the content of the image. In the next step the content of the image is evaluated with the stored image dataset. The significant points to notice on mining images are:

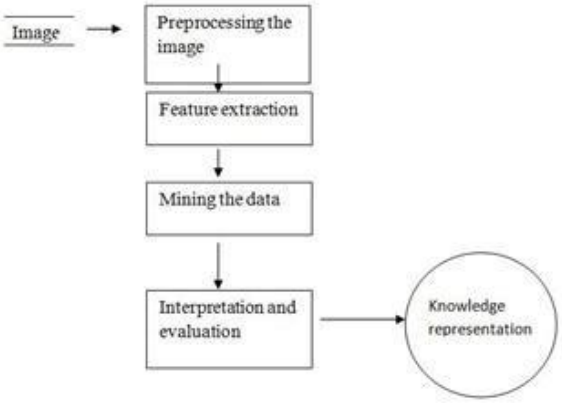


Figure 2: Image Mining[6]

Segment images into the findable regions.

Compare and analyze the segmented image data to the stored dataset. After that apply the data mining algorithm to create the object of association rules.

A picture is gotten to once for the feature extraction. The element extraction brings about picture blob and picture blob descriptors. The picture blob is put away in a document and the blob descriptors are put away in another record. The blob descriptors are utilized to recognize the object of association rules. When a picture is portioned then it isn't important to look through the image content.

## IV. LAND ESTIMATION

A way to deal with land estimation is to enroll the Synthetic Aperture Radar (SAR) image obtained by TerraSAR sensor with the geological map. However, these methodology is not applicable , if map is not accessible or the updation gets flop because of the normal risk. Supersolution and Mean-shift algorithm were acquainted with tackle these issues. K-means is an effective strategy utilized in land cover estimation. Mean-shift system is utilized for optical picture design acknowledgment. The greater part of the picture separating systems utilized for division depended on Mean-Shift and K-means approach. Since the contrast among land and ocean are of slight change and the condition of ocean territory is difficult to decide [7]. These methodology flopped in distinguishing the harbor.

### *A. Sea-Land Segmentation*

From the remotely detected pictures the land and ocean territory can be divided, by utilizing Local Binary Pattern(LBP). LBP is progressively reasonable for remote detecting picture handling. For a land pixel, LBP discovers to constantly zero and to the genuine land pixel

from satellite pictures, LBP isn't zero. It will bring out bogus cautions [8]. Along these lines, need a technique to lessen the bogus caution rate and to portion the land and ocean with high accuracy. Different issues in the current technique to section land and ocean are versatility and the information gathered from remote territories are not prepared for investigation. LANDSAT 5 satellite pictures are utilized for sectioning area and ocean region [9]. LANDSAT pictures don't meet the necessities because of versatility and goals issues. The information gathered from sensors are utilized to isolate the land and ocean territory. It is all the more testing to the coast line extraction and in object discovery. The different articles on the ground, make land forecast more muddled than ocean zone. The various leveled locale blending approach is utilized to naturally remove the ocean and land zone [10]. The blending approach can be better described by managed data joined with the element extraction. This outcomes in drawing out an ongoing diagnostic design to recognize land and ocean territory.

### *B. Data Mining techniques for sea-land detection*

The measure of information rates created in the computerized world is expanding step by step. The current advancements bombs in removing the example dataset. It is important to plan a strategy for investigating both disconnected and ongoing information. The information mining methods include information securing, handling and investigation. The gathered information is preprocessed. At that point, the valuable information is sent to the individual base station for information handling. The filtration and load balancing algorithm channel the gathered information and afterward balance the heap. It builds the framework's effectiveness. The detecting satellite uses SPECAN technique to change over gathered information into the image [11]. Channel the helpful information from the gathered picture. In the information examination process compute mean worth and standard deviation of pictures in each square. At that point figures the most extreme example esteem. The determined incentive for ocean territory is higher than the land region. The primary explanation behind less mean worth is because of the pixel estimation of the land picture is lower. Analyze the pixel estimations of the gathered picture with the information previously set for land and ocean zone. It brings about the recognition of the ocean zone and land zone.

Figure 3 shows the remote information process. The remote detecting information unit in the information procedure gathers the information from the satellite and afterward sent to the base station for additional information procedure. The information handling is of constant and disconnected information process. In the disconnected information process, the information is sent to the capacity server farm and procedure the information later on. The putting away of approaching ongoing information diminishes the exhibition of constant information process. Progressively information preparing the information is legitimately sent to the filtration and burden adjusting server. The filtration server just permits valuable data and disposed of every single undesirable datum. The heap adjusting server isolates the separated information and afterward sent to the preparing server. The heap adjusting server adjusts the intensity of information handling by appropriating continuous information. This server assumes a significant job in expanding the framework proficiency. The preparing server forms the information and sent the outcome to the collection server. Each preparing server makes numerical activities to get the outcome from the picture portion. The servers do the procedure in parallel. The collection server assesses the outcome and afterward sent to the basic leadership server. Every one of the servers is bolstered by calculations. At last, the information examination unit in the remote information processor gives a choice. The decision making server with the help of decision making algorithm makes the decision and evaluates whether the area is land or sea or forest.

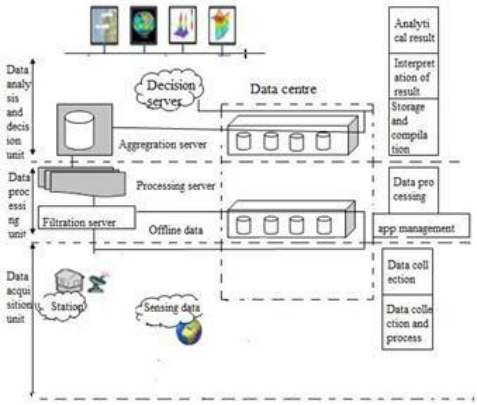


Figure3: Remote Data Process [11]

## V. FIRE DETECTION

Forests are a fundamental asset. The event of fire brings about expanding harm to the biological system. It is important to take measures against the backwoods fire. This requires a productive strategy to identify woodland fire continuously checking. The Aerial vehicle, ground-based techniques were utilized for the discovery. In addition, dangerous fire can abuse the life of aerial pilot and the restricted range is in ground-based measurement.

### *A. A Routing Protocol in Fire Detection*

Specialists discover that the AODV protocol is the best steering protocol utilized for backwoods fire location applications. Vitality utilization is the fundamental isssue in AODV protocol even it has best parcel convey proportion. DSDV, LEACH and APTEEN directing conventions are utilized to conquer the AODV downside. The group hub browsed the sensor hubs for DSDV convention. Bunch hub is picked progressively for Filter and APTEEN steering convention. In case of APTEEN group hub promote delicate edge an incentive to the sensor hubs. Drain isn't appropriate for continuous observing. APTEEN( TEEN and LEACH ) is sufficient for checking [12]. As far as possible the quantity of proactive transmissions and for receptive transmission utilizes edge esteems. The primary favorable position of

APTEEN is that it devours less measure of vitality when contrasted with other steering conventions. These conventions have absence of security and absence of nature of administration in information transmission. Consequently, it is important to advance other method for fire identification.

### *B. Sensor Based Fire Detection*

Temperature and moistness are the fundamental segments in fire discovery. Set a limit an incentive for mugginess and temperature and look at the sensor readings, if the two readings are same produces a caution demonstrating the fire is recognized. The sensors are conveyed to detect the carbon monoxide, stickiness, carbon-dioxide and the temperature. The bunch head sensor hub gathers the information from the relating group hubs. The control station gets the information from the bunch and the information preparing is performed. The principle point of such framework is to diminish the fire independent of significant expense. The slight varieties in temperature and residue particles makes the sensors report fire delay. On the off chance that a hub gets harmed, at that point the hub can interfere with the information procedure. Henceforth creates inaccurate data.

### *C. MAC Protocol for Fire Detection*

Fire identification applications need high transmission reliability. Sensor hubs utilized for fire identification needs unwavering quality due to the cruel condition. It is important to implement dependable strategies. Unwavering quality is portrayed by the quantity of parcels transmitted to the goal during an interim. Execution of MAC protocols brings about constrained dependability. Low dependability of Macintosh conventions were because of the nonattendance of course support. These issues were illuminated by adjusting MAC protocols with the route maintenance. The supports are completed by expanding cost in power utilization and bundle delay. Along these lines, it is important to advance an ease and high solid transmission strategy.

### *D. Visual Techniques for Fire Detection*

Visual detection methods can reduce fire delay and results in fire detection accuracy. Visual techniques combine sequential method and color verification techniques to make easy fire detection. It requires a real-time monitoring system to detect the fire with high precision. Satellite images give real-time information to predict the hazards. Analysis of image by extracting the pixel in real-time results in the prediction of fire detection with the least amount of cost. Camera surveillance systems can be used for the detection of fire. A camera or Optical system can reduce false information about the fire by capturing each movement. Forest-Watch is a camera system for semiautomatic fire detection. The camera needs to install by humans. Capture images within a range, so the camera sensor system results in an inefficient detection method. Camera sensors are of the high cost. It is necessary to provide a system that detects fire in real-time with less cost.

Visual camera sensors take pictures of the forest to detect the fire. A rotating motor is mounted on the camera sensor to get a full image of the forest. The collected image is then processed by using MATLAB to detect smoke [13]. The camera sensors need to install on the communication tower. The tower is placed on a location that faces large areas of the forest.

| Techniques Used | Advantages | Limitations |
|---|---|---|
| AODV Protocol | Best packet ratio for data transmission | Energy Consumption |
| Sensor Nodes | Real-Time Monitoring | Lack of security and reliability |
| MAC Protocol with Route Maintenance | Highly Scalable | Increased in cost of Power Consumption |
| Visual Techniques | Real-Time Monitoring with high precision | Camera Sensors of high Cost |

Table I. Comparison Study of Fire Detection Techniques

The installation of camera sensors on the towers is of great cost. Radio acoustic system is another method helps in the detection of the forest fire. The radio system consists of a radio center and an acoustic source. The radio system has more feeling to change in temperature and provides reliable temperature measurement. The deployment of the radio acoustic system in an optimal place is more complex. Table 1 shows the advantages and limitations of various forest fire detection techniques.

Forests can contribute to high amounts of financial wealth. So, it is necessary to carry out some approaches to detect and extinguish fires. Fire causes environmental damage to human lives. Fire modeling can predict possible fire behavior without getting burn [14]. Most of the fire detection approaches are based on satellite images. Also, the sensors are organized in the forest area to check out the temperature. Each sensor node is well maintained with a temperature sensor. When there occurs a change in temperature, the sensor nodes send packets to the cluster nodes. When the node gets spoiled result in error temperature readings. The image processing techniques are used to overcome the sensor node issues. A rule-based color model for fire pixels, which makes use of RGB color space is used to preserve the environment from the fire [15]. Since the image processing techniques cannot perform well in the real-time fire monitoring system.

The data mining approach uses meteorological data collected from satellite to detect the burned area. Modeled the data to predict the burned area, may result in the error message or lower accuracy for predicting large fires. The fire creates economical, ecological damage and causes harassment to humans. So, it is

important to prevent the spreading of fire. The existing method used meteorological data to predict the fire. It results in the prediction of the burned area and is an offline approach [16]. To overcome this prediction of fire, I proposed a real-time data analytical architecture for remote sensing. Besides, uses various data mining algorithm for data processing and scale-invariant feature transform algorithm (SIFT) for feature extraction of an image.

## VI. CONCLUSION

The data gathered from remote sensors are the test to the coast-line extraction. The absence of highlight extraction in the combining strategies brings about the forecast of land and ocean territories with less exactness. Sensor hubs are conveyed in the woods and are furnished with a temperature sensor to identify the fire. Be that as it may, when the hub gets harmed, it brings about blunder temperature perusing and yields a bogus caution. To diminish these issues, utilizing image preparing systems. These procedures may not perform well continuously checking frameworks. While, the real-time data analytical architecture can break down both constant and disconnected information process. Alongside different algorithms, the architecture is utilized to identify the land just as ocean zone. Notwithstanding ongoing information expository design, a scale-invariant feature transform algorithm (SIFT) can be utilized to discover the fire. The event of fire will get informed to the relating base station.

## REFERENCES

[1] A. Paul, J. Wu, J.-F. Yang, and J. Jeong,"Gradient-based edge detection for motion estimation in H.264/AVC," , IET Image Processing, vol. 5, no. 4, pp. 323327, Jun.2011.

[2] J. Shi, J. Wu, A. Paul, L. Jiao, and M. Gong, "Change detection in synthetic aperture radar image based on fuzzy active contour models and genetic algorithms " , Math. Prob. Eng, vol. 2014, 15 pp, April2014.

[3] Tanuja A, SwethaRamana,VTU Belgaum, India, "Analyzing Big Data using Hadoop" , International Journal for Innovative Research in Science and Technology, Volume 3, Issue 03, August2016.

[4] A. Labrinidis and H. V. Jagadish, "Challenges and opportunities with Big Data" ,in proceedings of 38th International Conference on Very Large Data Bases Endowment, Turkey, Aug. 2731, 2012, vol. 5, no. 12, pp. 20322033.

[5] https://www.google.com/imgres?imgurl=https%3A%2F%2Fimage2.slides erve.com%2F3841931%2Fdata-mining-a-kdd-process-n.jpg&imgrefurl=https%3A%2F%2Fwww.slideserve.com%2Ftuan%2Fdata-mining-a-kdd

[6] A. A. Khodaskar; S. A. Ladhake, "Image Mining: An Overview of Current Research" , Fourth International Conference on Communication Systems and Network Technologies,2014.

[7] J. Martn-de-Nicolas, D. Mata-Moya, P. Jarabo-Amores, N. del Rey-Maestre, J. L. Barcena-Humanes Signal Theory and Communications Department,"Segmentation techniques for land mask estimation in SAR imagery",International Conference on Computational Intelligence,2013.

[8] Yu Xia, Shouhong Wan, Peiquan Jin and LihuaYue, University of Science and Technology of China, "A Novel Sea- Land Segmentation Algorithm Based on Local Binary Patterns for Ship Detection",International Journal of signal processing and pattern recoginition,2014.

[9] Paulo Cortez1 and AnibalMorais, University of Minho, Portugal, "A Data Mining Approach to Predict Forest Fires using Meteorological Data",http://www.dsi.uminho.pt/pcortez

[10] Dongcai Cheng, GaofengMeng and Chunhong Pan,"Sea-land segmen-tation via hierarchical region merge and edge directed graph cut", IEEE International Conference on Image Processing (ICIP),2016.

[11] Muhammad MazharUllahRathore, Anand Paul, Bo-Wei Chen, Bormin Huang, and Wen Ji, "Real-Time Big Data Analytical Architecture for Remote Sensing Application " , IEEE journal of selected topics in applied earth observations and remote sensing,2015.

[12] V. Devadevan, Bhopal and S. Suresh, SRM University, India,"Energy efficient routing protocol in forest fire detection system",IEEE 6th International Conference on Advanced Computing,2016.

[13] Michel Owayjan, George Freiha, Roger Achkar, ElieAbdo, SamyMallah American University of Science and Technology, "Firxio-forest fire detection and alerting system",17thmeditteranean international conference,2014.

[14] QasimSiddque"SurveyofForestfiresimulation",Globaljournalofcomputerscienceandtechnology,1987, Qasim1987@hotmail.com

[15] Vipin V, Asst. Professor, St.Josephs College of Engineering, Palai, India"Image Processing Based Forest Fire Detection", International Journal of Emerging Technology and Advanced Engineering, February2012.

[16] MichaelKostiuk,MAGeographyGeospatialAnalyst,Canada,"usingremotesensingdatatodetectsealevelchange", International conference on ecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS, 2002.