BHOO-MI – BHOONIDHI META INTELLIGENCE

Dr. Charu Vaibhav Verma[1], Vivek Soni[2], Shailja Mishra[3], Suyesha Deshmukh[4], Shobhit Dwivedi[5]

[1]Assistant Professor, Dept of CSE, PIEMR, Indore, RGPV Bhopal (M.P.), India
kharecharu111@gmail.com
[2] Dept of CSE, PIEMR, Indore, RGPV Bhopal (M.P.), India
vsoni2259@gmail.com
[3] Dept of CSE, PIEMR, Indore, RGPV Bhopal (M.P.), India
shailjamishra07091980@gmail.com
[4] Dept of CSE, PIEMR, Indore, RGPV Bhopal (M.P.), India
suyeshasd@gmail.com
[5] Dept of CSE, PIEMR, Indore, RGPV Bhopal (M.P.), India
shobhitdwivedi9@gmail.com

Abstract

Presently satellite data products are supplied to the users in various formats and with required meta data associated to it. In general, this Meta data contains the latitude/longitude information of the area covered. The utilization of satellite imagery in diverse fields necessities efficient methods to add value to the core images. In this research work, we designed a comprehensive methodology to design a system which provides the intelligence to the basic data product by adding various geographic features over that area. Both natural and manmade geographic features were considered in study. The features include post offices, monuments, airports, museums, hospitals, educational institutes (NITs & IITs) etc. All these features are extracted and preserved as a centralized repository/database. Leveraging a combination of Python programming, web scraping techniques, and advanced geospatial libraries such as Geopandas and Shapely, our system undertakes a multi-step approach. Initially, the system reads the metadata of the satellite image to obtain corner latitude and longitude coordinates. Then the system identifies and extracts relevant feature data. These feature's meta information is then compiled into a shapefile, providing a comprehensive overview of the features present within the satellite image. This shapefile serves as a valuable resource for further analysis and decision- making in applications such as environmental monitoring, urban planning, and disaster response.

Keywords:

GIS, Intelligent Satellite Image, Satellite-Image, Data-Preprocessing in Band Data, Point-Data, Geospatial libraries

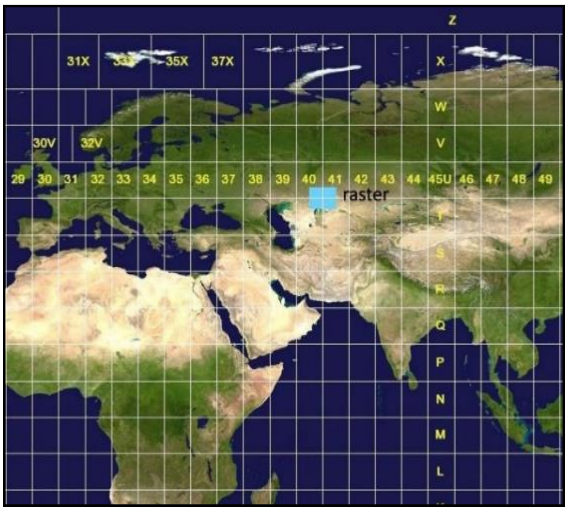
1. INTRODUCTION

Satellite imagery plays a crucial role in various fields like environmental monitoring, disaster assessment, and urban planning. Traditionally, extracting valuable information from these images, such as identifying specific features like post offices, monuments, or educational institutes, has been a manual and time-consuming process. This often requires specialized knowledge of image analysis or geospatial data handling, limiting accessibility and hindering the efficiency of large-scale projects. The primary objective of this work is to design and implement a Python-based system that automates the feature finding process from satellite imagery. This system will specifically target features of high societal interest, such as post offices, monuments, airports, museums, educational institutes, etc. Traditionally, extracting such features from satellite images has been a manual and timeconsuming process. This research aims to eliminate this burden by developing an automated system. The system will operate independently, processing satellite images and identifying features of interest without requiring human intervention.

This research aims to enhance the utilization of satellite imagery by integrating advanced techniques in geographic feature extraction. By leveraging Python programming, web scraping methodologies, and geospatial libraries, it seeks to add value to satellite data products by identifying and extracting natural and man-made geographic features such as post offices, monuments, and educational institutes. The extracted features are compiled into a centralized repository, facilitating further analysis and decision-making in areas like environmental monitoring, urban planning, and disaster response. Through the integration of cutting-edge technology and methodologies, this research contributes to the advancement of spatial intelligence and enhances the accessibility and utility of satellite data in diverse applications. This approach not only enhances the accessibility of satellite data but also adding Intelligence to the data. It facilitates informed resource allocation in various domains reliant on spatial information.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Fig.1. EPSG Projection of the World



2. LITERATURE REVIEW

New digital systems for the processing of photogrammetric and remote sensing images have led to new approaches to information extraction for mapping and Geographic Information System (GIS) applications, with the expectation that data can become more readily available at a lower cost and with greater currency as shown in figure 1. Demands for mapping and GIS data are increasing as well for environmental assessment and monitoring. Hence, researchers from the fields of photogrammetry and remote sensing, as well as computer vision and artificial intelligence, are bringing together their particular skills for automating these tasks of information extraction. The paper will review some of the approaches used in knowledge representation and modelling for machine vision, and give examples of their applications in research for image understanding of aerial and satellite imagery[1].

The importance of road extraction from satellite images arises from the fact that it greatly enhances the efficiency of map generation and thus can be a big help in car navigations systems or any emergency (rescue) system that needs instant maps. Therefore, increasing research is being dedicated and focused on the development of efficient methods to extract topographical meaningful features (like roads) from digital remote sensed images. The work deals with extraction of roads from satellite images. This is a challenging domain compared to extraction from aerial images as satellite images are noisy and of lower resolution. In this method, a Vectorization Approach for the automatic method of road extraction is being used where the image is segmented to identify the road network regions followed by a decision making and continuity procedure to correctly detect the roads and the Vectorization step to identify the line segments or curved segments which represents the road. This method may be employed for obtaining information for feeding large-scale Geographic Information System [2].

Road detection from satellite images can be considered as a classification process in which pixels are divided into road and background classes and can be used as a criterion in road extraction process to discriminate between road and nonroad pixels. Apart from the spectral information, textural parameters and contextual information are usually used by human being in object recognition from images. Contributing texture information in the neural network input parameters seems to be an improving idea for road detection from satellite images. Different texture parameters show different aspects of textural behavior in a defined neighborhood of a given pixel. Artificial neural networks are found to be superior to several previous techniques due in part to their ability to incorporate both spectral and contextual information. In this paper, Neural Networks are applied on high resolution satellite images for road detection [3].

Recently, settlement planning and replanning process are becoming the main problem in rapidly growing cities. Unplanned urban settlements are quite common, especially in lowincome countries. Building extraction on satellite images poses another problem. The main reason for the problem is that manual building extraction is very difficult and takes a lot of time. Artificial intelligence technology, which has increased significantly today, has the potential to provide building extraction on high-resolution satellite images. This study proposes the differentiation of buildings by image segmentation on high-resolution satellite images with U-net architecture. The open-source Massachusetts building dataset was used as the dataset. The Massachusetts building dataset includes residential buildings of the city of 15 Boston. It was aimed to remove buildings in the high-density city of Boston. In the U-net architecture, image segmentation is performed with different encoders and the results are compared. In line with the work done, 82.2% IoU accuracy was achieved in building segmentation. A high result was obtained with an F1 score of 0.9. A successful image segmentation was achieved with 90% accuracy. This study demonstrated the potential of automatic building extraction with the help of artificial intelligence in high-density residential areas. It has been determined that building mapping can be achieved with high-resolution antenna images with high accuracy achieved [4].

Feature extraction techniques are extensively being used in satellite imagery and getting impressive attention for remote sensing applications. The state-of-the-art feature extraction methods are appropriate according to the categories and structures of the objects to be detected. Based on distinctive computations of each feature extraction method, different types of images are selected to evaluate the performance of the methods, such as binary robust invariant scalable key points (BRISK), scale-invariant feature transform, speeded-up robust features (SURF), features from accelerated segment test (FAST), histogram of oriented gradients, and local binary patterns. Total computational time is calculated to evaluate the speed of each feature extraction method. The extracted features are counted under shadow regions and pre-processed shadow regions to compare the functioning of each method. We have studied the combination of SURF with FAST and BRISK individually and found very promising results with an increased number of features and less computational time. Finally, feature matching is conferred for all methods [5].

The huge benefit of the Internet is that it not only connects people from all over the world but also provides information that would otherwise be unavailable. Anyone who needs access to the latest satellite imagery can take it via many platforms provided remote sensing data, either for free from the websites belonging to national research and space centers or for a fee charged by commercial companies. The article offers an overview of the most used services, whose data are available and suitable for processing using EarthExplorer – a platform of the U.S. Geological Survey agency [6].

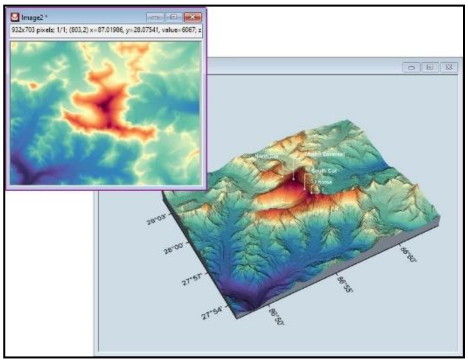


Fig.2. Geotiff image

Many GIS companies, raster data providers, and their clients have requested that the companies concerned with delivery and exploitation of raster geographic imagery develop a publicly available, platform interoperable standard for the support of geographic TIFF imagery as shown in figure 2. Such TIFF imagery would originate from satellite imaging platforms, aerial platforms, scans of aerial photography or paper maps, or as a result of geographic analysis. TIFF images which were supported by the public "geotie" tag set would be able to be read and positioned correctly in any GIS or digital mapping system which supports the "GeoTIFF" standard, as proposed in this document [7].

Scientific articles often contain relevant geographic information such as where field work was performed or where patients were treated. Most often, this information appears in the full-text article contents as a description in natural language including place names, with no accompanying machine-readable geographic metadata. Automatically extracting this geographic information could help conduct meta-analyses, find geographical research gaps, and retrieve articles using spatial search criteria. Research on this problem is still in its infancy, with many works manually processing corpora for locations and few cross-domain studies. In this paper, we develop a fully automatic pipeline to extract and represent relevant locations from scientific articles, applying it to two varied corpora. We obtain good performance, with full pipeline precision of 0.84 for an environmental corpus, and 0.78 for a biomedical corpus. Our results can be visualized as simple global maps, allowing human annotators to both explore corpus patterns in space and triage results for downstream analysis. Future work should not only focus on improving individual pipeline components, but also be informed by user needs derived from the potential spatial analysis and exploration of such corpora [8].

This research explores effective strategies for integrating geographical data into analytical workflows. Three main approaches are investigated: (1) Conversion of Shape Files to CSV or Excel Format, involving the utilization of Geographic Information Systems (GIS) to translate shape files containing polygon data into CSV or Excel files with latitude and longitude coordinates; (2) Transformation of Point Data into Shape Files, which entails using GIS tools to convert point data into shape files for compatibility with existing spatial datasets; and (3) Incorporation of Mapbox Maps as Background Images, allowing the integration of non-interactive Mapbox maps as visual references within analytical outputs. Additionally, the study advocates for the inclusion of these features in future software releases through endorsing relevant Community Ideas and contributing to discussions. Practical guidance is offered through step-by-step instructions and community forums. Ultimately, these methodologies aim to enhance analytical endeavours by enriching them with spatial insights, fostering deeper understanding and informed decision-making [9].

Satellite imagery be composed of images of Earth or other planets which are captured by satellites. A satellite image is an image of the earth taken using artificial satellites. Satellite images available in three main types which are visible light images, water vapor images and infrared images. Imaging of satellites are operated by board associative and businesses or organization around the world. Satellite imaging firms vend images for different under authorize. Satellite Images are authorized to legal board associative and organization such as Apple Maps and Google Maps. Satellite imagery tool which are window - based software is used to record and examine data about the earth. Different satellite imagery tools are used like Bhuvan by ISRO, Google Earth, Earth Alerts, GIS, USGS Global Visualization Viewer, etc [10].

Geospatial data visualization is significantly changing the way we view spatial data and discover information. On the one hand, a large number of spatial data, which carry extremely valuable information, are generated on daily basis. On the other hand, these data are not well utilized due to the lack of free and easily used data visualization tools. This paper describes a way of visualizing massive spatial data at no cost by utilizing publically available visualization tools like Google Earth. We illustrate our methods by visualizing a million global download requests for satellite images maintained by the Earth Resources Observation and Science (EROS) Center of U.S. Geological Survey (USGS) [11].

3. PROBLEM IDENTIFIED

The objective of this research is to address these limitations by developing an automated system for feature extraction from satellite images. By leveraging the power of Python programming and readily available geospatial libraries like Geopandas and Shapely, we propose a solution that streamlines the process. Our system aims to eliminate the need for manual feature identification, saving valuable time and resources.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

4. PROPOSED METHODOLOGY

In this section we present two different modules which are responsible for extracting data from the satellite geotiff image. We have taken IRS-IC-LISS-3 satellites band2, band3 and band 4 data for testing our system. We have tested the product folder in which these data are available and the metadata associated to this tiff file and we have tested them in our module to check the working efficiency of our system.

4.1 DATA ACQUISITION

Collecting the data about the parameters like Post offices, Monuments, Museums, Water Bodies, National Highways, Airports, Educational Institutes and many more from multiple resources, several platforms can be utilized.

4.1.1 **Government Websites:** Many government websites provide public datasets containing information about the various facilities, including post offices, airports and educational institutes. Examples include data portals maintained by national or regional government agencies responsible for transportation, education and tourism.

4.1.2 **Open Data Portals:** Open data portals host a wide range of datasets contributed by government agencies, research institutions, and non-profit organizations. Platforms like inayatdoor.com, Clinic Spots.com, and many more websites are available to display multiple data and can be extracted by using web scrapping.

4.1.3 **Geospatial Data Platforms:** Geospatial data platforms such as OpenStreetMap (OSM) provide crowdsourced geographical data that includes the locations of various amenities worldwide. Users can access and download data from OSM or use its APIs for programmatic access to extract information about post offices, monuments, airports, museums, and educational institutes.

4.1.4 **Research Databases:** Academic databases like Google Scholar, IEEE Xplore, and ResearchGate may contain research papers or studies with spatial data related to public facilities. Researchers often publish datasets or provide access to supplementary materials containing valuable spatial information.

4.1.5 **Commercial Geospatial Databases:** Some commercial geospatial databases offer comprehensive datasets with detailed information about various amenities, including post offices, airports, and educational institutions. These databases may require subscription or payment for access to their datasets.

4.2 MODEL ARCHITECTURE

The proposed architecture is depicted in figure 3.

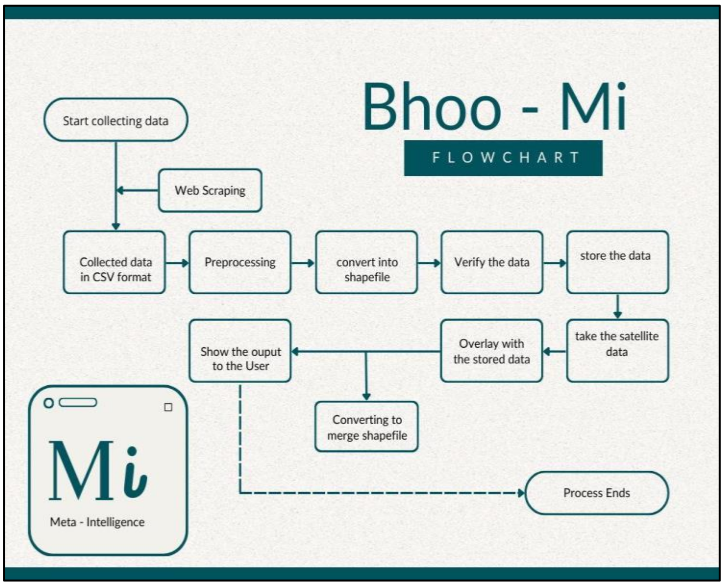


Fig.3. Proposed Model Workflow

4.2.1 Implementation

Implementation of this system has been done using python google colabs environment.

- For collecting the data, we used web scraping for getting the tabular data on web to the csv file format using python request library
- After collecting the data, we need to perform the preprocessing on the data
- And once we are done with preprocessing, we will move to extract latitudes and longitudes from the address we have fetched.
- We have fetched latitudes and longitudes by using geocoders and geopy library in python
- After getting done with it we have stored all the data including Post offices, Monuments, Museums, Hospitals, roads, rivers then we converted into shapefile and verify that whether the library returns valid latitudes longitudes or not.
- If we find any error then we dropped the data or recover it by checking the real value for that.
- Once we are prepared with our data then we will take product folder from the bhoonidhi platform and start checking with our stored data.
- Featured from the satellite image can be extracted by overlaying with latitudes longitudes
- Once it had some extracted features, It will return us a shapefile as output which we have visualized above.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

4.2.2 Proposed Model workflow

Step 1: Start collecting data
Step 2: Web Scrapping
Step 3: Collect data in CSV format
Step 4: Preprocessing
Step 5: Convert into shapefile
Step 6: Verify the data
Step 7: Store the data
Step 8: Take the satellite data
Step 9: Converting to merged shapefile
Step 10: Show the output to Use

4.3 FILTERED SHAPEFILE AS FINAL OUTPUT

After getting the filtered latitudes longitudes from the database we now store the latitudes andlongitudes in a list and then we will convert the list of csv to shapefile based on the attribute names. Once the user gets merged featured shapefile. It further can be visualized as per the user requirements. Here is the filtered shapefile visualization shown in figure 4.

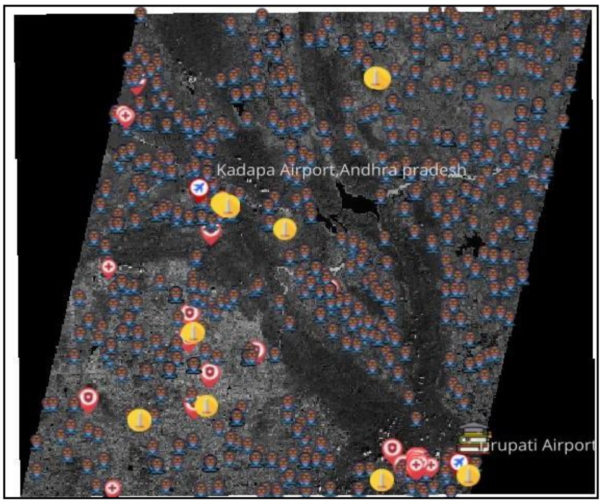


Fig.4.Filtered shapefile

5. RESULT

The implemented system demonstrates robust feature extraction capabilities from satellite imagery. By leveraging Python programming, web scraping methodologies, and advanced geospatial analysis techniques, the system efficiently identifies and extracts pertinent features such as post offices, monuments, airports, museums, and educational institutes. The extracted feature data is compiled into a comprehensive shapefile, offering a valuable resource for further analysis and decision-making across various domains, including environmental monitoring, urban planning, and disaster response. This approach enhances the accessibility of satellite data and facilitates informed decision-making processes reliant on spatial information.

6. CONCLUSION

In this research work we successfully established a Python-based system that automates feature extraction from satellite images, revolutionizing the traditional approach. By leveraging geospatial libraries and a pre-built feature database, the system can pinpoint and extract information about specific features of interest within an image. This eliminates the tedious and error-prone process of manual identification, saving valuable time and resources.

7. FUTURE SCOPE

This work serves as a stepping stone for further exploration and refinement. The future holds exciting possibilities to expand the system's capabilities. Firstly, we envision incorporating a wider range of extractable features. This could involve including historical landmarks, specific infrastructure types, or even detailed classifications of vegetation. Secondly, the system's current scope can be dramatically expanded to encompass global feature extraction. By integrating additional satellite image datasets and enriching the feature database, the system could be adapted to extract data across the entire world map. This global reach would unlock a vast array of potential applications. Finally, the current static parameters could be transformed into dynamic parameters that update automatically based on real-time data or user input. This would introduce a new level of flexibility and adaptability to the feature extraction process.

ACKNOWLEDGEMENT

We extend our heartfelt gratitude to our Director, (PIEMR, Indore) Dr. ManojKumar Deshpande, and our HOD (Dept. of CSE) Dr. Piyush Choudhary, for their unwavering support. We are thankful to Dr. Charu Vaibhav Verma (Project Guide), Mrs. Pooja Pandey, and Professors, for their guidance. Special thanks to Dr. Manju Sarma, Deputy Director at NRSC ISRO, for her exceptional guidance. We also express our gratitude to T. Radhika and Mrs. Jaya Saxena for their invaluable support. Additionally, we support Shanmukha Sir and the NRSC scientists for their insightful contributions. Finally, our thanks to our parents, group members, and all who supported us.

REFERENCES

[1] A. Sowmya, John Trinder, "Modelling and representation issues in automated feature extraction from aerial and satellite images", Vol. 55, Issue 1, February 2000.

[2] J. Hormese, C. Saravanan, "Automated Road extraction from high resolution and satellite images", Vol 24, 2016.

[3] B. Chakrapani, "Automated Feature extraction from high resolution satellite imagery using Ann", Vol 1, 2013.

[4] W. Alsabhan, Turky Alotaiby, "Automatic building extraction on Satellite Images using Unet and ResNet50", Vol. 2022.

[5] S. Karim, Y. Zhang, Muhammad Rizwan Asif, S. Ali, "Comparative analysis of feature extraction methods in satellite imagery", December 2017.

[6] F. Canters, "Advances in small-scale map projection research", Special issue: 29th international geographic congress, 2000.

[7] E. Ghaderpour, "Map Projection", December 2014.

[8] G. Shao, "Satellite Data", November 2016.

[9] I. Salim Hilal Almaqbali, F. Mohabbad Ali Al Khufairi, M. Samiullalla Khan, AZameer Bhat, "Web Scrapping: Data extraction from websites", July 2020.

[10] J. Izvoltova, V. Kriauciunaite-Neklejonoviene, R. Sasik, "Actual Satellite Imagery Sources", IOP Conf. Series, November 2021.

[11] M. Ruth, N. Ritter, E. Grissom, B. Borup, "GeoTIFF format specification GeoTIFF revision 1.0", November 1995.

[12] E. Acheson, Ross S. Purves, "Extracting and modeling geographic information from scientific articles", PLos ONE 16(1):e0244918, January 2021.

[13] Salesforce, "Overlaying Latitude and Longitude Data Points Over a Color-Coded Shapefile on a Map", November 2023.

[14] J. Kulkarni, "Satellite Imagery Tools", February 2017.

[15] G.D. Standart, K.R. Stulken, X.Zhang, Z.L. Zong, "Geospatial visualization of global satellite images with Vis-EROS", Volume 26, Issue 7, July 2011.