Automatic Feature Extraction from Remote Sensing Images

Suresh Kumar M[1], Anitha P[2], Avani K S[3], Gowthami Holla K S[4],Kruthi R Aithal [5]

[1] *Asst. Professor ,Information Science and Engineering, DSCE, Karnataka, India*
[2] *Student, Information Science and Engineering, DSCE, Karnataka, India*
[3] *Student, Information Science and Engineering, DSCE, Karnataka, India*
[4] *Student, Information Science and Engineering, DSCE, Karnataka, India*
[5] *Student, Information Science and Engineering, DSCE, Karnataka, India*

ABSTRACT

The recent years have witnessed significant strides in deep learning technology, particularly in the fields of computer vision, natural language processing, and multimedia. This advancement has created new opportunities for overcoming the challenges associated with traditional methods of collecting geographic data. Traditional approaches often demand substantial resources in terms of manpower and materials, while also encountering obstacles such as object masking and variations in colors, widths, and shapes. To tackle these issues, the proposed project presents a novel approach that harnesses the power of deep learning, specifically the YOLOv5 algorithm based on Convolutional Neural Networks (CNNs), to automatically extract spectral-spatial features from high-resolution multi-spectral images. By utilizing YOLOv5, this method aims to revolutionize the field of geographic data collection by enhancing the effectiveness and efficiency of data extraction.

Keyword: - *YOLOv5 algorithm, CNN , Remote Sensing, Feature extraction, Hyperspectral Images*

1. INTRODUCTION

Remote sensing is the process of detecting and monitoring the physical characteristics of an area, by measuring the radiation reflected from a distance, usually from an aerial vehicle like a drone or from a satellite. Researchers and scientists are currently investigating hyperspectral remote sensing, also known as imaging spectroscopy, which is a relatively new technology. It includes the detection and identification of minerals, terrestrial vegetation, and man-made materials based on their spatial and spectral characteristics. Deep Learning techniques have the advantage of automatically learning features, but this characteristic makes it challenging to observe the system's evolution over time. As the number of layers in a network increases, describing and understanding the model becomes increasingly difficult. Additionally, substantial computing power in terms of Memory, CPU, and GPU is required to train Deep Learning models. This is primarily due to the need for large training samples and the complex architecture of deep neural networks. Deep Learning techniques excel at extracting informative features from raw data through a hierarchical series of layers. This process begins with the initial layers capturing basic features such as textures and edges. As the information passes through subsequent layers, more complex and intricate features can be represented, aided by the knowledge acquired from earlier layers. The beauty of deep learning lies in its automatic learning process, enabling it to adapt to and handle a wide range of situations with diverse characteristics. This flexibility and adaptability make deep learning a powerful approach in handling complex and varied data.

2. LITERATURE SURVEY

[1], deep learning approaches were used as it is more efficient than ML handcrafted features. In this paper, if the land image is given as input, it classifies what part of the land is bare land, tea garden, impervious tree, and other surfaces. The main focus of their work was on fusion at the classification level and employed three multiple deep convolutional neural networks by training three different approaches: CaggeNet, GoogLeNet and ResNet5.

[2], a neural network image interpretation system is designed to efficiently extract Land use or Land cover information from high spatial resolution imagery using self-organizing supervised learning Artificial Neural Network (ANN). Learning Vector Quantization (LVQ) approach was employed in this study to classify four main Land use or Land cover feature types: Buildings, water, vegetation, and roads.

[3], Presented in this paper is a new method that significantly improves the performance of the Scale-Invariant Feature Transformation (SIFT) by achieving a higher rate of true extracted features and accurate matching. The method involves multiple stages, including image band selection, image band compression, image sharpening, automatic feature extraction, and the application of the sum of absolute difference algorithm with an experimental and empirical threshold.

[4], a novel method for extracting roads using an ensemble learning model with a postprocessing stage was proposed. The network weights and biases of our proposed deep learning model are transmitted through the random combination of layers of different sub models during forward and backward propagation. Experiments on two challenging datasets of remote sensing imagery showed that the proposed method has a significantly higher performance than the other models. This model is efficient enough to analyse complex scenes inorder to extract information on road images.

[5], vector field learning was used to extract roads from high-resolution remote sensing imaging. This method is usually used for skeleton extraction in nature images but seldom used in road extraction. The results show that all the vector fields can significantly improve the accuracy of road extraction, no matter whether the field is constructed in the road area or completely outside the road

[6], the authors propose a novel deep residual and pyramid pooling network (DRPPNet) for extracting road regions from high-resolution remote sensing images. The DRPPNet consists of three parts: deep residual network (ResNet), pyramid pooling module (PPM), and deep decoder (DD). Especially, the DResNet uses several residual blocks to extract deep road features from input images, which can enhance the learning ability of DRPPNet and avoid gradients vanish.

3. SYSTEM REQUIREMENTS

In this chapter we discuss the hardware and software requirements for the system.

3.1 Hardware Requirements

This program can run in most systems if all the necessary libraries are installed. A system with good CPU and GPU is ideal.

- Processor : Intel Core i5 or above
- RAM : 16GB
- Hard Disk Space : 100GB
- Operating System: Windows

3.2 Software Requirements

- Language : Python
- Libraries : numpy, pandas, sklearn, scipy, spectral, matplotlib, keras
- IDE : Jupyter Notebook, Google Colaboratory

4. IMPLEMENTATION

The powerful feature extraction ability of Deep Convolutional Neural Networks (DCNN) has made it a prominent hotspot in Computer Vision. Object detection based on DCNN follows a common pipeline, which involves region proposal, CNN feature extraction, region classification, and post-processing. In this context, the *YOLOv5* model stands out by treating object detection as a regression problem. By utilizing a single CNN, *YOLOv5* predicts bounding boxes and class probabilities in an end-to-end manner, resulting in faster predictions.

4.1 About YOLO Algorithm

Adapting the You Only Look Once (YOLOv5) algorithm for feature extraction in remote sensing involves training the model on annotated remote sensing images. The YOLOv5 algorithm is renowned for its speed and accuracy in object detection. It utilizes a single neural network to predict bounding boxes and class probabilities simultaneously for multiple objects within an image. This real-time detection capability makes it particularly suitable for applications that require efficient feature extraction from extensive remote sensing datasets.

By training the YOLOv5 model on annotated remote sensing images, it becomes capable of detecting and classifying various features of interest specific to the remote sensing domain. These features may include buildings, roads, vegetation, water bodies, and other relevant objects. The output of the YOLOv5 model provides precise predictions of bounding boxes and corresponding class probabilities for these identified features.

4.2 Methodology

To achieve a robust YOLOv5 model, we collected a dataset consisting of approximately 8000 images encompassing 43 land cover classes. These classes include Church, Cloud, Beach, Tennis court, Island, Bridge, Snow Berg, Palace, Lake, Airport, and many more. For training purposes, we utilized 7000 images, which translates to roughly 150 images per land cover class.

To facilitate the training process, most annotation platforms support the export of data in the YOLOv5 labeling format. This format generates one annotations text file per image. Each text file contains bounding-box (BBox) annotations for each of the objects present in the respective image.

The first step is training, where utilizing a sufficiently large dataset is crucial. When initializing the weights, passing an empty string (' ') to the weights argument yields the most beneficial results for the model.

Feature extraction is a significant aspect of the model, comprising two main parts: the backbone layers, which serve as a feature extractor, and the head layers, which compute the output predictions.

Once satisfactory training performance is achieved, the model is ready for inference. During inference, we can enhance prediction accuracy by applying test-time augmentations (TTA). The input for inference can take various forms, including images, videos, directories, webcams, or streams.

Now that our model is prepared, we save it with the .pt extension. When the file is executed, a new folder is dynamically created to store the output images. These output images will contain bounding boxes and labels indicating the respective class they belong to.

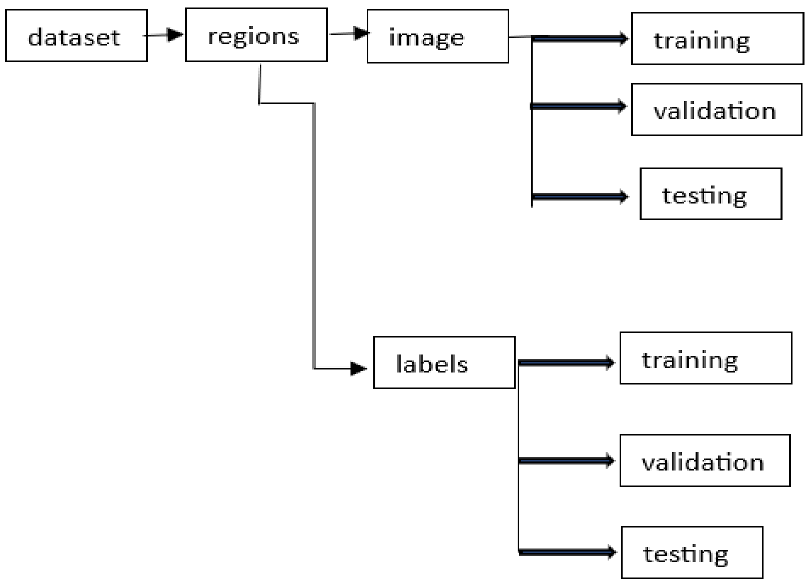
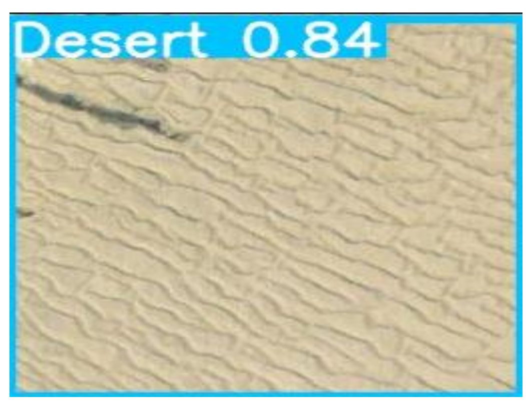


Fig 4.2.1: The data is provided at the following structure

5. RESULTS:



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

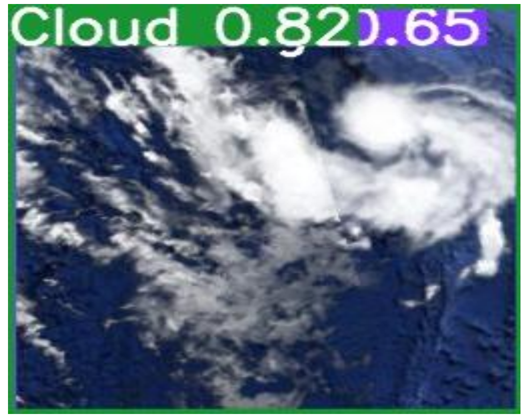


Fig 5.1 : Output images with bounding box and label

6. CONCLUSION:

The traditional supervised machine learning based approaches face a major challenge in feature extraction. This process is not only labor-intensive but may also fail to yield desired results. However, this issue was mitigated by the introduction of Convolutional Neural Networks (CNNs), which automatically extract features. This approach offers a dual advantage: It reduces human effort by automatically extracting features, eliminating the need for manual feature engineering. As the input data progresses through multiple deep layers, highly abstract and robust features are obtained, resulting in improved classification accuracy for the classifier.

7.

REFERENCES

[1]. Akhtar Jamil, Bulent Bayram, "Automatic discriminative feature extraction using Convolutional Neural Network for remote sensing image classification", *2019 Conference Paper October, Istanbul, Turkey,*

[2]. C. Lu, X. Yang, Z. Wang, and Z. Li, "Using multi-level fusion of local features for land-use scene classification with high spatial resolution images in urban coastal zones," *2017 Int. J. Appl. Earth Obs. Geoinf., vol. 70, no. September 2017, pp. 1–12.*

[3]. Stefanski, B. Mack, and B. Waske, "Optimization of Object-Based Image Analysis With Random Forests for Land Cover Mapping*," 2013 Sel. Top. Appl. Earth Obs. Remote Sensing, IEEE J., vol. 6, no. 6, pp. 2492–2504.*

[4]. T. Sasaki, J. Imanishi, K. Ioki, Y. Morimoto, and K. Kitada, "Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data," *2012 Landsc. Ecol. Eng., vol. 8, no. 2, pp. 157–171.*

[5]. D. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks," *2017 IEEE Geosci. Remote Sens. Lett., vol. 14, no. 10, pp. 1685–1689*

[6]. Junjie Li, Yizhuo Meng, Donyu Dorjee, Xiaobing Wei, Zhiyuan Zhang, and Wen Zhang "Automatic Road Extraction From Remote Sensing Imagery Using Ensemble Learning and Postprocessing", *2021 IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 14, 2021 10535*

[7]. M. He, B. Li, and H. Chen, "Multi-scale 3d deep convolutional neural network for hyperspectral image classification," in Proceedings of the IEEE International Conference on Image Processing (ICIP), Sept 2017, pp.3904–3908.

[8]. S. Li, W. Song, L.Fang, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no.9, pp.6690-6709, Sept. 2019, doi: 10.1109/TGRS.2019.290793