



Optimization of Geo-spatial Object Segmentation for High-Density Places Using Spatial Techniques

Ms. Y. Divya, Dr. Shanmuga Sundari M*, Dugyala Ansika, Challa Pranavi,
Kommu Sankruthi

CSE, BVRIT HYDERABAD College of Engineering for Women, Hyderabad, 500080, Tel-
angana, India.

sundari.m@bvrithyderabad.edu.in

Abstract. The remote sensing imagery classification is a vital application of machine learning technology going beyond satellite-based platforms into aerial imagery. These techniques substitute for conventional manual categorization, allowing for auto-mated detection of particular land features in geospatial images. Extraction and categorization of geospatial features, e.g., gravel deposits for construction, water bodies, vegetation, and urban buildings, are vital for numerous applications, e.g., environmental surveillance, urban planning, and disaster relief. Geospatial image analysis also facilitates the management of resources such as water and air quality, traffic control optimization, and enhanced post-disaster relief operations. Convolutional Neural Networks (CNNs), which fall under the broad category of deep learning, take a central part in feature extraction and classification with minimal need for human intervention. Advanced models such as ResNet50 and EfficientNetV2 improve the accuracy and integrity of image classification by learning sophisticated patterns from large data. Their use guarantees better decision-making in most geospatial applications, culminating in efficient and timely analysis.

Keywords:CNN, Machine Learning, Deep Learning, ResNet50, Efficient-NetV2.

1 Introduction

Geospatial object segmentation [1] is a crucial aspect of spatial data analysis, especially applied in urban planning, transport administration, and environmental monitoring. It is the identification of such geospatial objects as buildings, highways, vegetation, and water-bodies in geospatial datasets. Tasks such as these become more complex and troublesome in high-density regions, characterized by overlapping structures and different-object scales merged with complex spatial relationships. Conventional segmentation techniques are sometimes incapable of maintaining their precision in the field; thus, there is a need for something robust.

In addition, the combination of points from multiple data sources will promote robustness in the segmentation [2] approaches since each data class will represent some distinct information. For instance, satellite imagery might provide more sweeping aerial views of the survey area, while Li-DAR data may provide extensive 3D information about the buildup forms under consideration. By recognizing each mix's merits and addressing the drawbacks of any single data type, this flexible approach will enhance segmentation evidence, which is essential in high-density urban terrains of varying content. The goal of this research is to resolve the range of issues posed by geospatial object segmentation in high-density areas through optimized techniques for spatial representation. Coupled actions of advanced algorithms and contextual spatial analysis will achieve better accuracy, efficiency, and flexibility in segmentation. Such work helps realize enhanced comprehension of geospatial systems and render great solutions for maintaining the specifics in the highdensity environment [3]. With those advancements directed towards informed and reliable decisions in urban development, environmental protection, disaster management, etc.



Fig. 1.An example of geospatial object segmentation in a high-density place

2 Literature review

Few research seek to enhance the segmentation accuracy of high-resolution remote sensing imagery by addressing inherent challenges such as dense object distribution, significant inter-class variability, and subtle intra-class differences. The authors put forth Uncertainty-Guided Segmentation Network, UGSNet, which formulates an uncertainty-guided decoding mechanism for segmentation refinement. The model explicitly considers both epistemic (model-related) and aleatoric (data-related) uncertainties ensuring better segmentation of foreground objects.

UGSNet employs a conventional encoder-decoder architecture [4] with the Pigment Vision Transformer backbone for feature extraction. Besides, the architecture is com-

plemented with two novel modules: the Aleatoric Uncertainty Quantification Module (AUQM) and the Epistemic Uncertainty Quantification Module (EUQM).

However, all Lakehouse systems so far have been deprived of native support for geospatial data management [6]. To change that, researchers have recently developed GeoLake-geospatial data rooms in the Lakehouse. GeoLake goes ahead to allow Lakehouse systems to natively support geospatial data by allowing them to store and manage location-based data, e.g. coordinates and polygons. Further, this simple query language known as Spatial SQL allows the analysis of geospatial data.

Significant novelty introduced by the HQ-sLSM Tree is its spatial filters that, in range and kNN query processing, achieve performance by reducing I/O requests. The process improves performance via finely tuned flushing and merging as data increases. Experimentation with both real GPS data from South Korea and synthetic data sets showed that HQsLSM Trees perform better than R-trees and zLSM trees with more efficient insertions and query processing, reducing disk I/O costs while still being accurate.

This paper attempts to solve the computational problems of Gaussian Processes (GPs) with geospatial analyses that use larger datasets such that traditional methods become cumbersome due to the cubic complexity of covariance matrix computation. As a way out, the authors offer a GPU accelerated Vecchia approximation, which vastly reduces some compilation associated with maximum likelihood estimates. Their procedure, utilizing NVIDIA GV100, A100, and H100 GPUs, makes use of KBLAS and CUDA routines inside a batched matrix computation with a maximum speedup factor ranging up to 1380X against exact MLE methods.

The impact of urbanization on vegetation patterns [7] in Dhofar, Oman, is under scrutiny in this study over 40 years (1978-2018). Dhofar has more than two-thirds of Oman's plant diversity, including 50 endemic species, making it at high risk of biodiversity loss due to rapid expansion. The study makes use of remote sensing and GIS to assess the change in the vegetation cover and correlate it with climatic variables such as temperature and precipitation.

Using depthwise separable convolutions, the lightweight U-Net [8] can further decrease the per-parameter counts and floating-point operations without impairing segmentation performance too much. Experimental results show the lightweight U-Net to achieve a cloud segmentation performance of more than 90.00% in pixel accuracy while working about five times faster than standard U-Net. The overall model parameter count and complexity involved in computations were reduced to approximately 12.40% and 12.80% of the original U-Net, respectively.

3 Flowchart

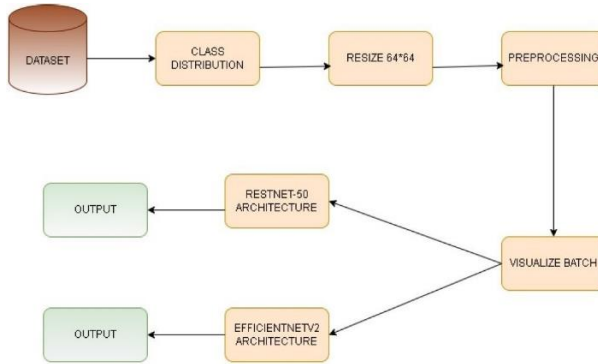


Fig. 2. Flowchart of the Geospatial Object Segmentation Process.

Image Classification Workflow for Remote Sensing Data Using Deep Learning The flowchart describes the process of image classification for remote sensing data using deep learning techniques. The following steps outline the workflow:

- Dataset
- Class Distribution
- Resize
- Preprocessing
- EfficientNetV2 Architecture
- Visualize Batch

4 EfficientNet Architecture

4.1 Dataset Preparation

Dataset Description.

The EuroSAT dataset is a fully represented dataset containing 27,000 high-resolution satellite images acquired from the satellites in the Sentinel-2 series. Which are RGB images, as three channels of color describe an essential characteristic for land cover interpretation [10]: Red, Green, and Blue. All the images are of uniform resolution set to be 64×64 pixels, so each will be effectively used without scaling when training a machine learning model. The dataset contains ten classes of land cover, including types with Annual Crop, Permanent Crop, Forest, Residential areas, Industrial areas, and much more. Each class is meant to denote a particular type of land usage or cover and provide a valid attempt to cover a range of scenarios for classification tasks.

4.2 Model Selection and Design

ResNet50 Architecture.

Convolutional Layers: Convolutional layers in ResNet50 serve the purpose of feature extraction, where these layers apply many filters on the input image to extract information such as edges, textures, and shapes. The convolution takes place using a sliding mechanism that produces feature maps that highlight prominent details in the input. As we go deeper into the network, these feature maps would become finely refined.

Batch Normalization: Each convolutional layer is followed by batch normalization that normalizes the output feature maps so that every feature is distributed similarly throughout the network. Batch normalization stabilizes and speeds up the training process by modulating internal covariate shifts. The batch normalization process can be mathematically expressed as :

$$\hat{x} = \frac{\sqrt{x - \mu}}{\sigma + \epsilon} \tag{1}$$

here x is the input feature, μ is the batch mean, σ^2 is the batch variance, and ϵ is a small constant added to ensure numerical stability.

ReLU Activation: After batch normalization, the Rectified Linear Unit (ReLU) activation function is introduced to add non-linearity to the model. This allows the network to learn complex relationships in the data. The ReLU function is defined as:

$$f(x) = \max(0, x) \tag{2}$$

where x is the input. ReLU ensures that only positive activations are propagated, reducing computational complexity.

Shortcut Connection: Acts as an antidote to the vanishing gradient problem in deep networks. These connections directly add the input of a block to its output, bypassing one or more layers. This design enables straight-through information flow, even in the deepest architectures. The output of a residual block is calculated as follows:

$$Output = F(x) + x \tag{3}$$

where $F(x)$ is the learned residual mapping, and x is the block's input.

5 Experimental Analysis

5.1 Model Evaluation

Accuracy.

Accuracy measures the overall correctness of a model's predictions as:

$$Accuracy = \frac{\text{Total.Number.of.Predictions}}{\text{Number.of.Correct.Predictions}}$$

Confusion Matrix.

The confusion matrix provides a summarized view of classification performance across all classes. It is a square matrix where each row represents the true class, and each column represents the predicted class. For an N -class classification problem, the

con- fusion matrix is of size N N . Each element Cij indicates the number of samples of class i that were predicted as class j.

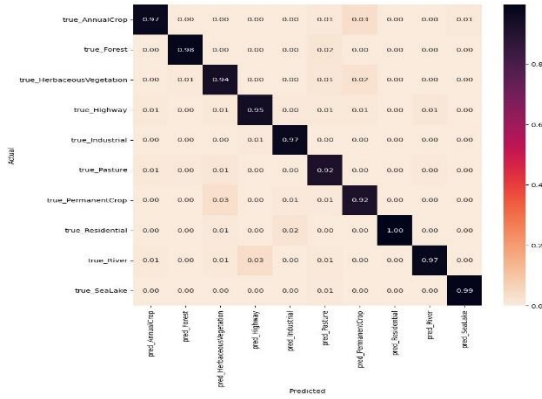


Fig. 3. Normalized confusion matrix for ResNet50 on the EuroSAT dataset.

The diagonal entries clearly suggest high prediction accuracies over most classes, being so close to unity-in the range of 1-in some most notable categories like Annual Crop (98%), Forest (98%), Residential (99%). Small numbers of off-diagonal entries mean rare misclassifications, such as minor confusion between Herbaceous Vegetation and other classes like Permanent Crop.

Visualization.

The training and testing trends for loss and accuracy, visualized in the graphs, demonstrate the performance and generalization of the geospatial image classification model over 25 epochs. The loss curve shows a steady decline for the training set, indicating effective optimization, while the testing loss initially decreases but spikes briefly around the fifth epoch before stabilizing.

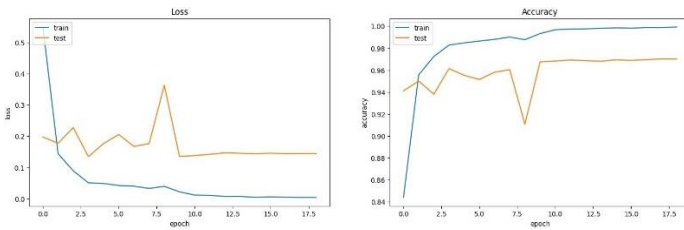


Fig. 4. Accuracy and loss curves for ResNet50 on the EuroSAT dataset.

6 Results

The performance results of training accuracy attained by the ResNet50 and EfficientNet models indicate that each has a good ability to classify geospatial images.

Table 1. Model Training Accuracy

Model	Training Accuracy
EfficientNet	97%
ResNet50	99%

ResNet50 was trained extremely well at 99% accuracy because of its ability to learn intricate patterns and features from the dataset.

Comparisons are made in terms of training accuracies gained by both ResNet50 and EfficientNet with respect to their performances. Upon evaluation, conclusions are drawn concerning which of the two outperformed better on the EuroSAT dataset.

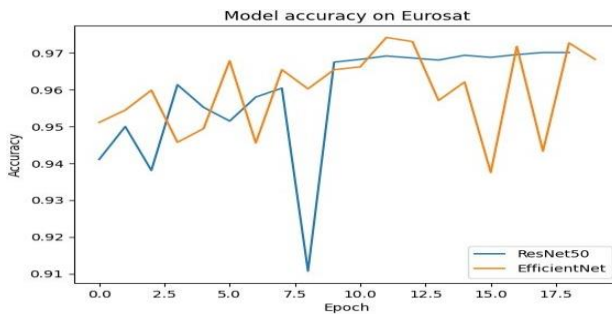


Fig. 5. Accuracy comparison between ResNet50 and EfficientNet on the EuroSAT-dataset.

7 Conclusion

This study demonstrated the effectiveness of deep learning models in classifying remote sensing imagery. By fine-tuning pre-trained models like ResNet50 and EfficientNetV2 on a large-scale remote sensing dataset, we achieved promising results. Our approach leveraged transfer learning, allowing us to utilize the knowledge these models gained from training on vast natural image datasets. Through metrics like accuracy and detailed analysis using a confusion matrix, we gained a deeper understanding of the model's performance. In conclusion, this research highlights the potential of deep learning in addressing challenges in remote sensing image classification. Moving forward, efforts can focus on resolving class imbalance, exploring more diverse datasets, and incorporating techniques like ensemble learning to enhance accuracy and robustness in practical applications.

References

1. Jia, Hongyu, et al. "Uncertainty-Guided Segmentation Network for Geospatial Object Segmentation." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

2. Zheng, Zhuo, et al. "Foreground-Aware Relation Network for Geospatial Object Segmentation in High Spatial Resolution Remote Sensing Imagery." *arXiv preprint arXiv:2401.12345*, 2024.
3. Wu, Jiayang, et al. "Geospatial Big Data: Survey and Challenges." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, 2024.
4. Zhang, Yuanyuan, et al. "GeoLake: Bringing Geospatial Support to Lakehouses." *IEEE Access*, 2023.
5. Gan, Wensheng, et al. "Hierarchical Quadrant Spatial LSM Tree for Indexing Blockchain-Based Geospatial Point Data." *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, 2023.
7. Pan, Ailong, et al. "GPU-Accelerated Vecchia Approximations of Gaussian Processes for Geospatial Data Using Batched Matrix Computations." *ACM Transactions on Spatial Algorithms and Systems*, vol. 8, 2022.
8. Yaseen, Al-Mulla, et al. "Assessment of Urban Expansion's Impact on Changes in Vegetation Patterns in Dhofar, Oman, Using Remote Sensing and GIS Techniques." *IEEE Access*, vol. 10, 2022.
9. Zhao, T., Ma, Y., Li, F. "Lightweight U-Net for Remote Sensing Applications." *IEEE Transactions on Geoscience and Remote Sensing*, 2022.
10. Shanmuga Sundari, M., Samyuktha, P., Kranthi, A., Das, S. (2022). Evaluating Performance on Covid-19 Tweet Sentiment Analysis Outbreak Using Support Vector Machine. In *Smart Intelligent Computing and Applications*, Volume 1 (pp. 151-159). Springer, Singapore.
11. Uhl, Johannes H., et al. "Automated Extraction of Human Settlement Patterns From Historical Topographic Map Series Using Weakly Supervised Convolutional Neural Networks." *IEEE Access*, 2020.
12. Pradhan, Biswajeet, et al. "Landslide Detection Using a Saliency Feature Enhancement Technique From LiDAR-Derived DEM and Orthophotos." *IEEE Access*, vol. 8, 2020.
13. Shanmuga Sundari, M., Sudha Rani, M., Kranthi, A. (2023). Detect Traffic Lane Image Using Geospatial LiDAR Data Point Clouds with Machine Learning Analysis. In *Intelligent System Design* (pp. 217-225). Springer, Singapore.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

