Machine Learning-Based Image Segmentation for Drone Imagery

Table of Contents

I. Introduction

In the current era of digitization and technological advancements, the use of drone imagery (remote sensing) has become crucial in monitoring, managing, and researching various fields such as the environment, agriculture, urban planning, and many others. From collecting wide-ranging and continuous data to analyzing and extracting valuable information from drone images, remote sensing contributes significantly to addressing various social and environmental challenges.

In this context, the topic "Image Segmentation for Drone Imagery" has been chosen to tackle an important and fascinating problem in the domain of drone image processing. The main focus of this project is to partition the pixels in drone images into regions with similar characteristics or belonging to the same object. Image segmentation for drone imagery brings numerous benefits, including precise delineation of objects, evaluation of object scale and boundaries, facilitating informed decision-making in resource management and utilization.

We will start by studying the relevant concepts and methods related to image segmentation, as well as exploring the technologies and tools to be used in the project. Next, we will devise a solution and implement the methods to achieve the set objectives. Finally, we will evaluate the obtained results and draw conclusions, assessing the effectiveness of the proposed solution and suggesting potential avenues for future development.

With dedication, focus, and perseverance throughout the research and implementation process, it is expected that this project will yield positive and meaningful outcomes in applying drone imagery to real-world applications.

II. Related theory

In this section, we will delve into the relevant theories and concepts related to image segmentation for drone imagery. Understanding these foundational principles is crucial for designing an effective segmentation solution.

1. Image Segmentation

Image segmentation is a fundamental task in computer vision that aims to partition an image into multiple meaningful and coherent regions or segments. The goal is to group pixels with similar characteristics together, allowing for a higher-level understanding of the image content. Various techniques are employed in image segmentation, including traditional methods based on color, texture, and intensity, as well as advanced deep learning-based approaches using convolutional neural networks (CNNs).

2. Remote Sensing and Drone Imagery

Remote sensing is the science of acquiring information about the Earth's surface and atmosphere using sensors mounted on satellites or aircraft. Drone imagery, one of the essential sources of remote sensing data, provides detailed visual representations of large areas of the Earth's surface. It captures images in different spectral bands, such as red, green, blue, and near infrared (NIR), which are crucial for extracting valuable information about the Earth's features and conditions.

3. Segmentation Techniques for Drone Imagery

Segmenting drone imagery poses unique challenges due to varying atmospheric conditions, complex terrain, and diverse land cover types. Traditional approaches for drone image segmentation include region-based methods like Watershed transformation and K-means clustering. However, in recent years, deep learning methods, particularly CNNs, have shown superior performance in semantic segmentation tasks. Popular architectures like U-Net, FCN (Fully Convolutional Network), and DeepLab have been successfully applied to drone image segmentation tasks.

4. Spectral Indices and Feature Extraction

Drone imagery often includes multiple spectral bands, capturing different information about the Earth's surface. Spectral indices, such as NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index), play a crucial role in analyzing vegetation and water bodies, respectively. Feature extraction techniques are used to transform the raw pixel values into meaningful representations, allowing for efficient training of segmentation models.

5. Shapefile and GIS (Geographic Information System)

Shapefile is a popular geospatial vector data format used in GIS for representing geographical features like points, lines, and polygons. When performing image segmentation for drone imagery, creating a shapefile containing polygons corresponding to segmented regions is essential for further analysis and interpretation.

6. Training Data and Evaluation Metrics

To train a segmentation model, we need labeled training data, where each pixel in the image is tagged with a specific category or region. To assess how well the segmentation model performs, we use common evaluation metrics like Intersection over Union (IoU), Dice coefficient, and pixel accuracy. These metrics help us measure the accuracy and effectiveness of the model's segmentation results.

Understanding these fundamental theories and concepts will serve as a strong foundation for the subsequent stages of the project, where we will design and implement the image segmentation solution for drone imagery.

III. Technology Used

In this section, we will introduce the technologies employed to build the image segmentation solution for drone imagery.

1. Geospatial Data Abstraction Library (GDAL)

GDAL is a powerful library used for reading, writing, and processing geospatial data. It provides support for various raster and vector data formats, including the commonly used .tif file format for drone images. GDAL enables us to efficiently access and manipulate drone imagery, making it an essential tool in our project.

2. Python Programming Language

Python, with its extensive libraries and ease of use, will be our primary programming language for implementing the image segmentation solution. Python provides a rich ecosystem for data manipulation, machine learning, and geospatial data processing, making it an ideal choice for this project.

3. Geographic Information System (GIS) Software

In this project, we will utilize Geographic Information System (GIS) software, particularly QGIS, as one of the key tools for visualizing and analyzing the segmentation results. QGIS is a popular and open-source GIS software that provides a wide range of functionalities for working with geospatial data, including displaying raster and vector data, performing spatial analysis, and creating thematic maps. With QGIS, we can overlay the segmented polygons on drone imagery, enabling us to interpret and evaluate the segmentation outcomes effectively. QGIS's user-friendly interface and powerful geospatial capabilities make it an excellent choice for this project.

IV. Proposed Solution

In this section, we will present the proposed solution to perform image segmentation for drone imagery at the pixel level. The solution encompasses specific steps and techniques aimed at achieving the objectives of the project.

1. Creating Shapefile from .tif File

The initial step of our proposed solution involves utilizing the QGIS software to create the Shapefile that contains polygons that correspond to different classification areas within the drone image. These polygons act as spatial representations of distinct regions in the image. For each polygon, we will define three important fields:

1. ID: An integer field that denotes the order of the polygon. Each polygon will have a unique ID, which allows for easy identification and referencing.

2. Type: An integer field indicating the category or class of the area represented by the polygon. The type values will be used to classify the regions based on specific features. For example, we may assign values like 1 for water, 2 for grass, 3 for buildings, and so on.

3. Name: A string field that contains descriptive names for each polygon, providing additional information about the classification category.

In this project, we have decided to utilize six distinct types for classifying the polygons. These types will help us effectively label and categorize the various regions present in the drone image. By creating such a shapefile with detailed attributes, we can facilitate the subsequent phases of the image segmentation process, ensuring accurate identification and separation of different features within the image.

Type	Name
0	Water
1	Building
2	Road
3	Tree

4	Grass
5	Other

And we will have two shapefiles(datasets) for training and testing model.

2. Feature Extraction and Index Calculation

Next, we will utilize Python along with relevant libraries to extract and calculate various indices from the .tif file. The calculated indices include R (Red), G (Green), B (Blue), GLI (Green Leaf Index), NRGDI (Normalized Difference Red Green Index), and DSM (Digital Surface Model). These indices play a crucial role in marking essential and valuable features present in the drone imagery.

To extract these indices, we will perform the following steps:

1.Extracting Pixel Values

For each polygon in the shapefile, we will extract the pixel values from the corresponding areas in the drone image. This step will allow us to capture the spectral information (R, G, B) for each location of interest.

2. Calculating Green Leaf Index (GLI)

The GLI is a critical vegetation index used to assess the health and abundance of vegetation. It is computed as $GLI = \frac{2*G-R-B}{2*G+R+B}$. This index will help us identify the vegetation-rich areas in the drone imagery.

3. Calculating Normalized Difference Red Green Index (NRGDI)
The NRGDI is another essential vegetation index that highlights the presence of healthy vegetation. It is calculated as $NGRDI = \frac{G-R}{G+R}$. The NRGDI will further aid in the identification of vegetation regions.

4. Calculating Digital Surface Model (DSM)
The DSM represents the variation in elevation across the drone image area. It is obtained through the comparison of DSM data with the original elevation data. This information will be useful in understanding the topography and terrain characteristics.

By extracting and computing these indices from the polygons in the shapefile, we can effectively capture and quantify vital information related to vegetation, land cover, and elevation. These indices will serve as significant attributes in the subsequent stages of our image segmentation process, contributing to accurate and meaningful classification of the drone imagery.

3. Data Processing and Storage in CSV Files

In this phase, we will utilize the Python programming language to extract information from the shapefile and the calculated indices. The data obtained from these sources will then be processed and stored in CSV (Comma-Separated Values) files, which will be used for training and testing the subsequent machine learning models.

The process involves the following steps:

1. Data Extraction from Shapefile

Using Python, we will read the shapefile and extract relevant information from each polygon. This information includes the ID, type, name, and spatial coordinates of the polygons. Additionally, we will retrieve the boundaries of each polygon to define the regions of interest.

2. Data Retrieval from Calculated Indices

We will retrieve the computed indices, such as R, G, B, GLI, NRGDI, and DSM, from the previous step. These indices hold valuable spectral and elevation information related to the drone imagery.

3. Data Preprocessing and Feature Engineering

The extracted data may require preprocessing to handle any missing or inconsistent values. Additionally, feature engineering techniques may be applied to enhance the representation of the data and improve the performance of the machine learning models.

4. Store data in CSV Format

After the necessary processing and feature engineering, we will store the data in CSV format. Each row in the CSV file will represent a specific polygon, and the columns will contain the relevant attributes, including the ID, type, name, spatial coordinates, and the computed indices (R, G, B, GLI, NRGDI, DSM). This structured format will facilitate data management and accessibility for further analysis.

5. Data Split for Training and Testing
To train and evaluate machine learning models effectively, we will create two CSV files (training and testing sets) from 2 created Shapefile in above step. The training set will be used to train the models, while the testing set will be used to assess their performance.
By processing and storing the extracted information in CSV files, we create a comprehensive dataset that incorporates both spatial and spectral information. This dataset will serve as a valuable resource for training and evaluating machine learning algorithms in the subsequent stages of the project, helping us achieve accurate and reliable image segmentation results.

4. Applying Machine Learning Algorithms

With the prepared data, we will proceed to apply various machine learning algorithms, such as k-nearest-neighbor, support vector machine, and random forest, for the classification of the test data. These algorithms will learn from the labeled training data and then be utilized to predict and classify objects within the satellite image.
1. k-Nearest-Neighbor (k-NN) Algorithm
The k-nearest-neighbor algorithm is a simple and effective classification method. Given a new data point, it identifies the k-nearest neighbors based on their features in the training dataset. The majority class among these neighbors is then assigned to the new data point as its predicted class.

2. Support Vector Machine (SVM)

SVM is a powerful supervised learning algorithm that is commonly used for classification tasks. It finds the optimal hyperplane that best separates the different classes in the feature space. This hyperplane maximizes the margin between the classes, leading to improved generalization and robustness of the classifier.

3. Random Forest

Random forest is an ensemble learning method that combines multiple decision trees to create a strong classifier. Each decision tree is trained on a random subset of the data, and the final prediction is obtained through a majority voting process.

By employing these machine learning algorithms, we aim to leverage their strengths and capabilities to accurately classify objects within the drone imagery based on the extracted features and indices.

5. Result Evaluation

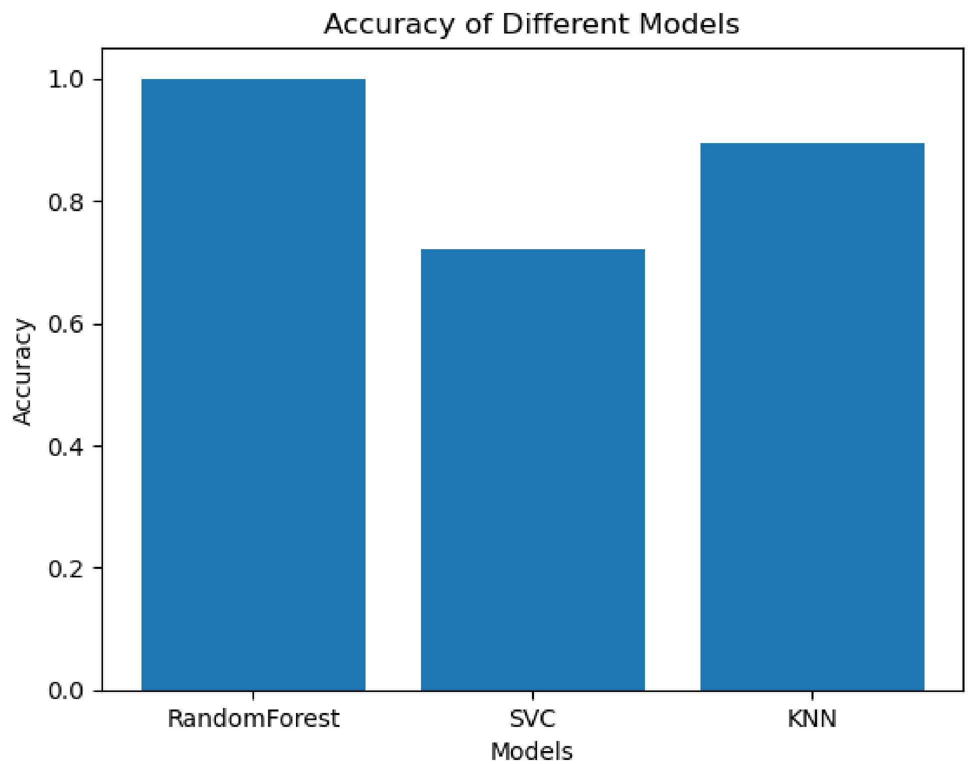
Finally, we will perform testing with the test dataset to evaluate the accuracy of the trained machine learning models. This evaluation will be conducted by calculating various performance metrics, including accuracy, confusion matrix, and other evaluation indices. These metrics will help us assess the precision and effectiveness of the drone image segmentation solution.

The goal of this evaluation is to ensure that the proposed solution achieves high accuracy in identifying and classifying objects within the drone imagery automatically and reliably.

With the execution of the steps, we aspire to build a precise and efficient drone image segmentation solution that can automatically detect and classify objects within the satellite imagery with a high degree of accuracy and reliability. This accomplishment would have significant applications in various fields, such as environmental monitoring, urban planning, and land use analysis.

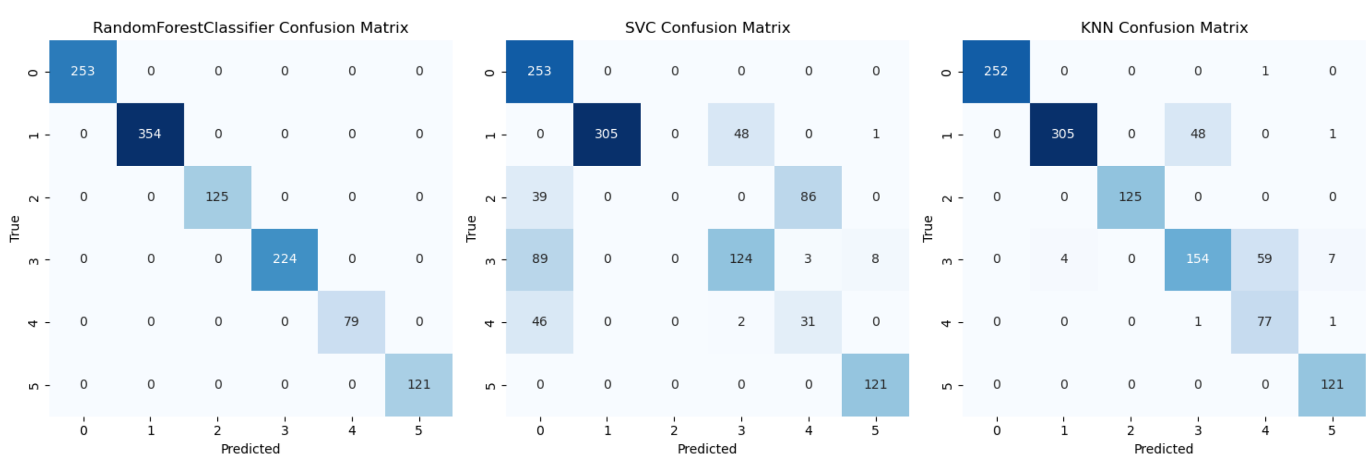
V. Achieved Result

Based on the above steps, utilizing the three algorithms: K-Nearest Neighbors, Support Vector Machine, and Random Forest, we can observe that all three algorithms achieve difference levels of accuracy, and RandomRorest might be the best model among the three models applied to this problem.



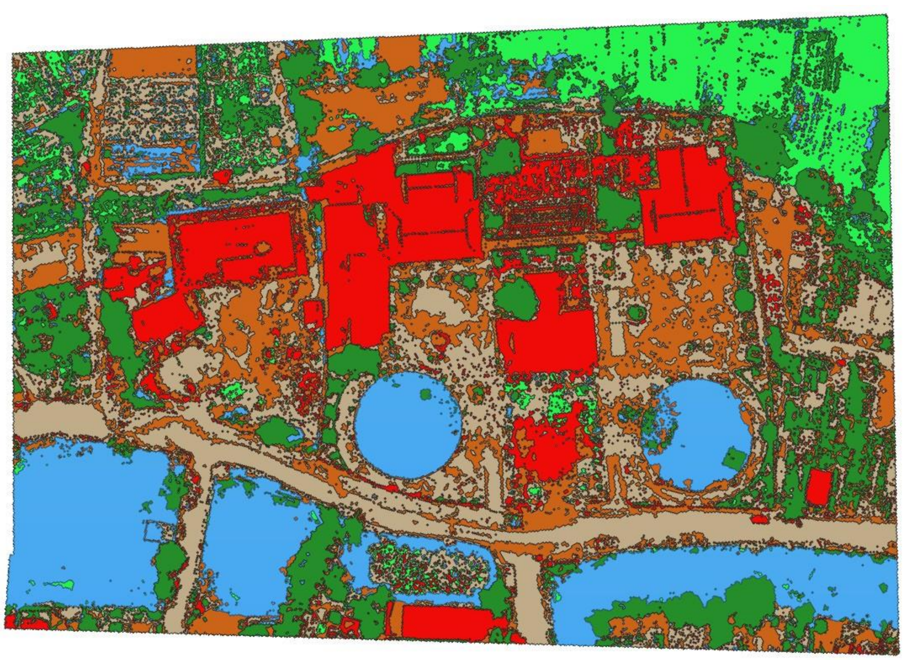
However, this level of accuracy may not be optimal due to certain factors, such as inadequate sampling, for example, having an overrepresentation of certain types of polygons (bias), or insufficient data for training purposes.

Below are the confusion matrices for each algorithm:

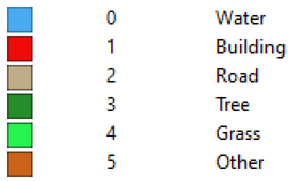


After applying the RandomForest model to the task of classifying the entire large image (scene Chua_Kham_Son), here are the classification results we obtain compared to the original image.

Original image (scene Chua_Kham_Son)



Classified image



VI. Conclusion

In this project, we successfully proposed a solution for drone image segmentation based on K-Nearest Neighbors, Support Vector Machine, and Random Forest algorithms. The aim was to automatically detect and classify objects within drone images, providing valuable insights for various applications.

Throughout the implementation, we followed a systematic approach, including preprocessing the data, extracting relevant features, and training the machine learning models. The results demonstrated that all three algorithms achieved accuracy levels ranging from 70% to 97%. However, we also identified certain limitations that could have impacted the accuracy, such as sampling bias and insufficient training data.

Despite the challenges faced, our solution serves as a solid foundation for further improvements. To enhance the accuracy and robustness of the model, increasing the diversity and volume of training data, along with fine-tuning the model's hyperparameters, will be crucial.

Drone image segmentation has extensive real-world applications, ranging from environmental monitoring to urban planning and disaster management. By continuously refining the proposed solution and exploring advanced techniques, we aim to contribute to the advancement of image analysis in these domains.

In conclusion, our project has laid the groundwork for an effective drone image segmentation system. With continued efforts in addressing the limitations and incorporating cutting-edge methods, we envision a more accurate and efficient solution that can significantly impact various fields, facilitating better decision-making and resource allocation in the future.