NASA MODIS IMAGERY CLOUD IDENTIFICATION

Dongwei Fu
University of Colorado, Boulder
Boulder, CO 80309
dofu3785@colorado.edu

ABSTRACT

Clouds play an important role in moderating the Earth's climate. Satellite imagery, with its advantage of high spatial coverage is one of the most effective methods for studying clouds. In this project, I plan to investigate the problem of classification and segmentation of shallow cloud features in satellite imagery using Convolutional Neural Network (CNN) technique.

1. Introduction

Climate change is one of the most pressing challenges human-beings are currently facing. Cloud covers approximately 70% of the Earth's surface, playing a critical role in regulating the planet's climate by reflecting sunlight and trapping heat, which affects weather patterns and the global energy balance. Understanding cloud formations is essential for accurate climate modeling and predictions. Satellite imagery is the most effective approach for studying shallow cloud patterns, structures, and formation because it provides extensive spatial coverage. In this project, we will be using dataset curated by researchers from Max Planck Institute for Meteorology [1][2]. The dataset is provided as an online Machine learning competition posted on Kaggle [1]. It consists of hand labelled RGB satellite imagery taken from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard NASA's two polar-orbiting satellites, Terra and Aqua. The cloud labels are divided into four classes: Fish, Flower, Sugar and Gravel. Each class (label has its distinct cloud features.

2. Related work

Satellite image segmentation can be viewed as combining classification and localization. Through image segmentation, it partitions the image into smaller segments, and then tries to understand what is given at a pixel level and identifies the shapes and boundaries of the object. The final output of image segmentation is a mask where each element indicates which class the pixel belongs to. One of the most widely used image segmentation model architecture is U-Net [3]. Originally developed for biomedical image analysis, U-Net gets its name from its architecture, a U-shaped structure consists of two main components: an encoder network and a decoder network. The encoder compresses the input image by downsampling, and the decoder then upsamples the compressed features and restores the original spatial resolution. This design ensures that both contextual and detailed information are preserved, making it highly effective for tasks requiring precise localization of objects in images.

3. Proposed work

3.1 Dataset Exploratory Data Analysis

We will start by performing EDA analysis on the training and test dataset. This allows us to familiarize with the satellite imagery dataset and experiment with ideas of data preprocessing to prepare for model training.

3.2 Data Preprocessing

Data will be preprocessed for model training purposes. There are available modules that can help streamline the data preprocessing, such as DataGenerator using the Tensorflow.Keras package.

3.3 Model Architecture and Setup

Currently the plan is to use a U-Net model architecture that can be implemented using Python Tensorflow package. Segmentation Models [4] has been used in the past for similar image segmentation tasks.

3.4 Proposed Timeline

09/01/24 - 09/15/24: Identify the problem assess project feasibility. Select corresponding dataset to perform exploratory data analysis.

09/16/24 – 09/30/24: Prepare data for model training. Data preprocessing and data visualization.

10/01/24 – 10/10/24: Choose baseline model and perform model training and initial model evaluation. Fine tune model.

10/11/24 – 10/15/24: Finalize analysis and write final project report and presentation.

4. Model Evaluation and Discussion

Dice coefficient is used for this project to evaluate model performance. It is often used to compare the pixel-wise agreement within a predicted segmentation and corresponding ground truth, using the following equation:

$$\frac{2 * |X \cap Y|}{|X| + |Y|} \tag{1}$$

Here X defines the set of pixels predicted, whereas Y defines the ground truth of the training set. When X and Y are empty, the dice coefficient is defined as 1. Unlike the more popular "accuracy" metric, Dice coefficient is calculated by the overlap of the two segmentations divided by the total size of the two objects (images).

5. Conclusion

This section will be updated once all model runs, and analyses have been carried out.

6. References

[1] Rasp S., H. Schulz, R. Walter, and M. Demkin. 2019. Understanding Clouds from Satellite Images. Kaggle. https://kaggle.com/competitions/understanding_cloud_organization

[2] Rasp, S., H. Schulz, S. Bony, and B. Stevens. 2020. Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. Bull.

Amer. Meteor. Soc., 101, E1980–E1995, https://doi.org/10.1175/BAMS-D-19-0324.1.

[3] Ronneberger, O., Fischer, P., & Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

297–302, https://doi.org/10.2307/1932409.

[4] Iakubovskii, P. 2019. Segmentation Models. GitHub repository. Retrieved from https://github.com/qubvel/segmentation_models