CS4490 2021-Summer Thesis Project Proposal

USING CONVOLUTION NEURAL NETWORK TO LABEL DIFFERENT PARTS OF REMOTE SENSING IMAGE

Ziqingqing Ye

Department of Computer Science

Western University

2021/07/10

Project supervisor:
Prof. Roberto Solis-Oba Department of Computer Science

Course instructor:
Prof. Nazim H.Madhavji

1. Abstract:

Context and motivation:
The classification of remote sensing images plays a key role in several areas. For example, Adapt the measurements to the local conditions: land capability maps, vegetation cover maps and other maps can be obtained through remote sensing image classification, which can be used as the basic map for the next studies on the environment, land use and greening. Traditional remote sensing classification methods have the disadvantages of poor robustness and slow speed and when a particular data classification requires a particular algorithm, the accuracy is relatively low. In recent years, deep learning has achieved success in many fields. This method can capture the high-dimensional nonlinear characteristics of data, so it is of great significance to try to apply it to the classification of remote sensing.

Principal ideas and Anticipated results:
This project requires a large dataset to train a network, the data sent into network should be multimodal, prior knowledge of the accuracy of each individual label, and the data can't contain too much noise. The network can use a convolution-based network which could better capture the image properties.

Anticipated impact of results
In contrast to existing technologies, we can handle different situations with just one model. This model has the advantage of high inference speed and high accuracy.

2. Detailed proposal:

⇨ **2.1 Research Objectives**:

GOAL:
The goal is to using deep learning method to build a model which is trained by a large remote sensing dataset, this model could be used to classify various remote sensing image.

(1) **OBJECTIVE :** Preprocess the dataset, denoise and normalize data, and then split dataset into training dataset, validation dataset and test dataset.
(2) OBJECTIVE : Build a model and train it, then finetuning the hyperparameters to get best performance.

SIGNIFICANCE:
The original image we received was large and full of noise. If we do not use super pixel segment the image, the training will not success. We will have to decompose the original image into many small images and transform them to make the data suitable for training.

The first thing we get is original images. Because the size of each image is different, it is necessary to segment each image by super-pixel. The purpose of remote sensing image classification is to enable us to easily and successfully identify each feature in the map. Therefore, it is necessary to distinguish these features in training data and then we can build an effective model.

⇨ **2.2 Research Approach and Methodology**:

- Technical issues: We need to acquiree a series of remote sensing images, and then label it manually and preprocess it, after that, we should build a convolutional neural network for training. Finally use the dataset train this model and verify it.

- Technologies, components: This project developing, implementing and validating remote sensing classification algorithms. We also need a GPU server and a data set to support this research.

- Identify key issues in the research methodology
First, split training dataset, validation dataset and test dataset from the data set. These datasets will be used to optimize the model and verify the results in future experiments. And I also need to use image process technology, such as image transform super-pixel to make these data suitable for training. After that, use deep learning algorithm to build the model.

- Technology/Solution (results): The result is that we can perfectly classify most of unlabeled remote sensing images.

- Challenges anticipated: Some objects of the image are too small to detect. For a particuller feature on the map, the number of training samples is very limited. For example, if the feature we want to recognize is the road, but there are few training pictures containing the road feature.

- Mitigation of challenges: try resampling dataset. we can change the dataset that you use to build your predictive model to have more balanced data.[1] You can add copies of instances from the under-represented class or delete instances from the over-represented class. For example, if there are only a few pictures containing the road, we can train several more times on these pictures to ensure that each feature of the map is learned and trained about the same number of times.

- Validation strategy
In order to ensure the applicability of our results, I split the data into training set, validation set and test set. First, use the training set to optimize model and then use validation set to verify the result after each training epoch. Finally, we use test set to verify results after training process compleated.

3. Value of the results and industrial relevance:

-Derivation of technology and Application: At the academic level, our research integrates the almost whole process of remote sensing image classification using deep learning, so that future researchers can derive from our research to other fields.

-Unlimited range of possible uses: The ideal result is that we'll make a complete application that includes the technology we discussed earlier. It helps the unmanned vehicle sense the ground environment, at the same time, it helps to support agriculture, meteorology and other fields of research

4. Reference:

[1] Brownlee, J. (2020, August 15). *8 Tactics to Combat Imbalanced Classes in Your Machine Learning Dataset*. Machine Learning Mastery. https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/.