

Detecting Damaged Buildings based on Post- Hurricane Satellite Imagery

Problem & Context:

After a hurricane, damage assessment is essential to emergency workers and first responders so resources can be planned and designated accordingly. One way to gauge the damage extent is to detect and quantify the number of damaged buildings, which is typically done by ground survey methods. This process can be labor intensive and time-consuming. With satellite imagery readily available today, the goal of this study is to improve the efficiency and accuracy of building damage detection with image classification algorithms.

Target Clients:

The primary clients of this project targets emergency workers, first responders, and politicians in charge of allocated disaster relief resources. Not only will this project, if successful, save time and money for disaster workers, It will be beneficial for those affected by the natural disaster.

Data:

The data are satellite images from Texas after Hurricane Harvey divided into two groups (damage and no_damage). The dataset is structured in the following format:

1. train_another : the training data; 5000 images of each class
2. validation_another: the validation data; 1000 images of each class
3. test_another : the unbalanced test data; 8000/1000 images of damaged/undamaged classes
4. test : the balanced test data; 1000 images of each class

All images are in JPEG format. The dataset is available for download via the following link:

<https://www.kaggle.com/kmader/satellite-images-of-hurricane-damage>

Constraints & Scope:

- Satellite imagery resolution is not as high as various benchmark datasets commonly used to train neural networks with respect to the objects of interest.
- There are some inconsistencies in image quality. Since the same region can be captured multiple times on different days, the same coordinate may have multiple images of different qualities
- The ground truth labels could be incorrect in some cases.

Approach:

Multiple steps will be taken to build a predictive model for this project as well as to analyze the resulting predictions.

1. The dataset will be loaded into a pandas dataframe and any preprocessing of the images will be completed.
2. I will develop an imagery classification model by building a Convolutional Neural Network (CNN) from scratch and a pre-trained model using transfer learning.

3. Grad-Cam, a technique for understanding what features are most important by using the gradients of the convolution layer will be employed for better model interpretability.

Deliverables:

The final draft of the project will be presented in the form of a slide deck and formal project report . Jupyter Notebooks will be delivered detailing each step taken and code written for the analysis of the project. A Github repository for the project will be created as well.