mattlegro / **SAR-OceanImage-CNN**

☆ **0** stars   ⑂ **0** forks

☆ Star	🔔 Notifications

<> **Code**   ⊙ Issues   ⇅ Pull requests   ▷ Actions   ▥ Projects   📖 Wiki   ⚠ Security

⑂ master ▾   **Go to file**

mattlegro Finalizing project files, preparing for review.   ...   1 minute ago   ⟲ **11**

View code
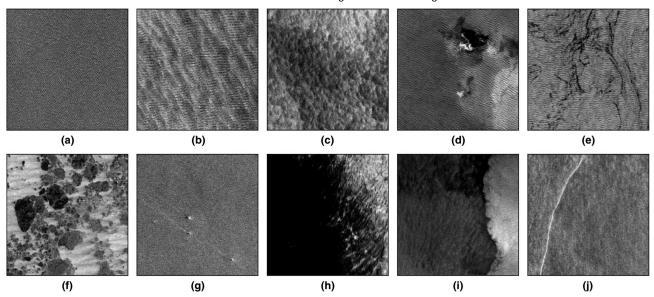
≡ README.md

Project Overview

For this project, we will train a fully connected convolutional neural network to predict, from manually labeled Sentinel-1 Synthetic Aperture Radar (SAR) imagery, whether an image contains one of ten primary open ocean phenomena.

The Data

This project uses a manually labeled dataset presented in a paper published by Wang et al. in 2019. It contains 37000 images roughly equally distributed into ten classes, all acquired by the Sentinel-1 satellite in 2016: Pure Ocean Waves, Wind Streaks, Micro Convective Cells, Rain Cells, Biological Slicks, Sea Ice, Icebergs, Low Wind Area, Atmospheric Front, and Oceanic Front. Descriptions for each of the classes and in depth descriptions of the preprocessing techniques used to prepare the images for machine learning can be found in the source paper by following the link above. Below is a set of example images, containing one from each class, where letters (a)-(g) are respective to the list of categories provided above.
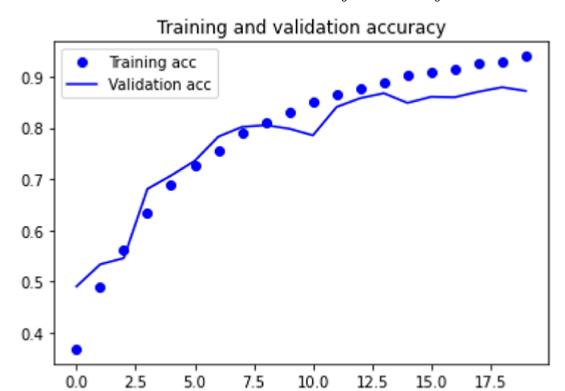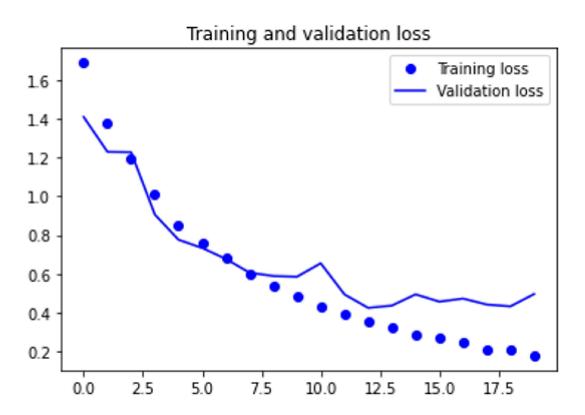
Business Problem

Oceans cover 70 % of the Earth's Surface, with air-sea interactions playing a crucial role in climate and weather projections. Sentinel-1 and Sentinel-2 satellites launched in 2014 and 2016 take hundreds of thousands of images every month that are largely unprocessed, making these images an untapped source of potential for improving weather and climate models. Developing classifiers and other models to sort through these images makes possible the otherwise unmanageable task of garnering information from such a huge dataset. This project is thus a demonstration of the capacity for CNNs to learn how to identify images collected as a step toward using the images for meteorological benefit.

Results

In the end, two CNNs were independently trained, taking only a single night on a single GPU home desktop computer, yielding accuracies of 84 and 86 % on a set of images reserved for test evaluation. On the x-axis is the number of training epochs.

Training and validation accuracy

Training and validation loss

Areas for Improvement

It is notable that both networks struggled in identifying certain phenomena: Ocean and Atmospheric Fronts. From the paper, "Periodic signatures of ocean waves can coexist with [both] Ocean Front [and] Atmospheric Front." Additionally, there is some class imbalance in the dataset. Ocean Fronts are least represented at about a quarter of the frequency of the other images. While there are two other categories that were underrepresented and performed fine, it is possible that a combination of imbalance and similarity to / presence of other classes in images primarily labeled with these categories caused this inaccuracy. In the future, augmenting the dataset with rotated or flipped images to relieve imbalance issues and artificially increase the number of training images could improve performance.

Conclusions

Based on our results, we can assert that CNNs are certainly capable of learning to identify useful features of SAR images and can absolutely benefit researchers in combing through the ever-increasingly massive collections generated by Sentinel-1 and other earth satellites.

Releases

No releases published

Packages

No packages published

Languages

- **Jupyter Notebook** 100.0%