Machine Learning Engineer Nanodegree

Capstone Proposal

Nicholas Condo
April 17, 2020

Understanding the Amazon from Space

Domain Background

The Amazon is the largest and most biodiverse tropical rainforest in the world, covering an area of 2.1 million square miles. It is comprised of an estimated 390 billion individual trees divided into 16,000 species. The Amazon has been referred to as the "lungs of the planet", as it helps to stabilize the earth's climate and slow global warming by fixing CO2 and producing 20% of the world's oxygen.

Since 1978 over 289,000 square miles of Amazon rainforest have been destroyed across Brazil, Peru, Columbia, Bolivia, Venezuela, Suriname, Guyana, and French Guiana. Every minute, the world loses an area of forest the size of 48 football fields. There is concern that the destruction of the forest will result in loss of biodiversity, habitat loss, and the release of the carbon contained within the vegetation, which could accelerate global warming. Better data about the location of deforestation and human encroachment on forests can help governments and local stakeholders respond more quickly and effectively.

Problem Statement

While considerable research has been devoted to tracking changes in forests, it typically depends on coarse-resolution imagery from satellites such as Landsat (30 meter pixels) or MODIS (250 meter pixels). This limits its effectiveness in areas where small-scale deforestation or forest degradation dominate. Furthermore, these existing methods generally cannot differentiate between human causes of forest loss and natural causes.

The goal of this project is to track changes in the Amazon rainforest due to deforestation using satellite image data. The data – provided by Planet – has a ground-sample distance (GSD) of 3.7m and an orthorectified pixel size of 3m. The task is a multi-label classification problem where each image should be labeled with an appropriate atmospheric condition, common land cover/land use phenomena, and rare land cover/land use phenomena. Each image will have one and potentially more than one atmospheric label and zero or more common and rare labels. Correctly labeling these images can help identify deforestation and limit the damage being done by illegal human encroachment on the ecosystem.

Datasets and Inputs

The data for this project is taken from a Kaggle competition, [Planet: Understanding the Amazon from Space](). The images were derived from Planet's full-frame analytic scene products using their 4-band satellites in sun-synchronous orbit and International Space Station orbit. The set of chips for this competition use the GeoTiff format and each contain four bands of data: red, green, blue, and near infrared. Also included is a separate set of JPG chips which were processed using the Planet visual product processor and then saved as jpg chips.

The labels for this task were chosen in collaboration with Planet's Impact team and represent a reasonable subset of phenomena of interest in the Amazon basin. The labels can broadly be broken into three groups: atmospheric conditions, common land cover/land use phenomena, and rare land cover/land use phenomena. Each image will have one and potentially more than one atmospheric label and zero or more common and rare labels.

Solution Statement

Over the past few years, computer vision has taken giant leaps due to advances in deep learning and convolutional neural networks. At this point, deep leaning models are able to outperform humans on many image classification tasks, such as [ImageNet](). One great benefit that has come from researchers striving to develop the best performing model on ImageNet, has been the advent of transfer learning. Transfer learning is when a model developed for one task is reused for a model on a second task. In short, a model trained on ImageNet can then be "fine-tuned" on a different dataset, achieving state-of-the-art results while requiring minimal training time. This technique has become prevalent in computer vision applications, and popular deep learning frameworks today all provide built-in support for using popular models trained on ImageNet. Harnessing the power of transfer learning, I plan to develop a deep learning model to classify the satellite images using the PyTorch deep learning framework.

Benchmark Model

The winner of the Kaggle competition held in 2017, was user "bestfitting", producing an F2 score of 0.93317 on the private leaderboard. He provided a brief overview of his solution here, where he describes his method of ensembling 9 different deep learning models, in addition to some data augmentation and postprocessing, to achieve this result. Using his private leaderboard result as a benchmark, I will attempt to improve on his F2 score of 0.93317.

Evaluation Metrics

Submissions to this competition are evaluated based on their mean F2 score. The F score, commonly used in information retrieval, measures accuracy using the precision p and recall r. Precision is the ratio of true positives (tp) to all predicted positives (tp + fp). Recall is the ratio of true positives to all actual positives (tp + fn). The F2 score is given by:

$$(1+\beta^2)\frac{pr}{\beta^2 p + r} \text{ where } p = \frac{tp}{tp+fp},\ r = \frac{tp}{tp+fn},\ \beta = 2.$$

Project Design

The workflow I plan to follow will first involve exploring the data to get an idea of the distribution of classes and view the associated images. I then plan to create a resnet-50 model pretrained on ImageNet to classify the images and produce a submission file to upload to Kaggle. I will utilize the fastai library for developing and training the models, which is a helper library that sits on top of the PyTorch deep learning framework. Once I have a working pipeline – from input images to submission file – I will then begin to optimize the model by experimenting with different hyperparameters such as image sizes, batch sizes, learning rates, and data augmentation techniques. I will then experiment with ensembles and post processing techniques, such as creating thresholds for the different labels. I will also experiment with other popular computer vision networks to potentially include in a final ensemble based solution.