Understanding Deforestation in Amazon Basin using Convolutional Neural Networks

RIT | Rochester Institute of Technology

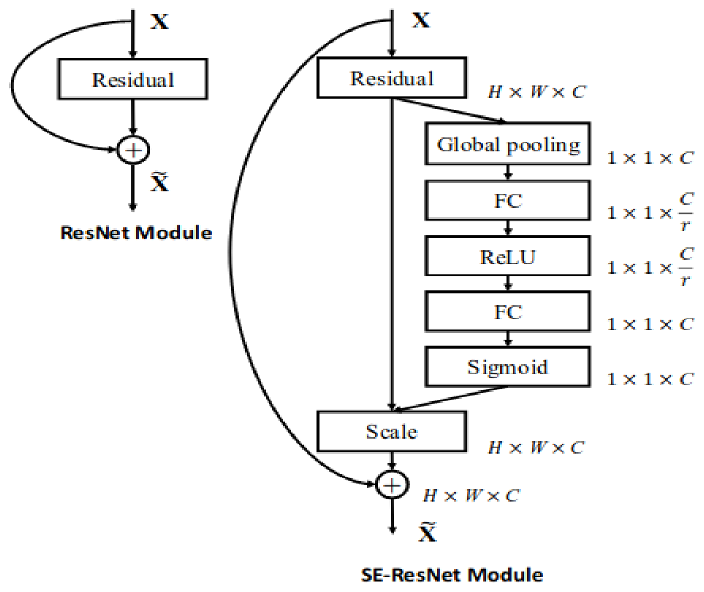
Aayush Soni, Department of Computer Science, RIT.

Introduction

Every year, an estimated **18 million** acres of forest cover is lost from various parts of the world, according to UN Food and Agriculture Organization. Deforestation in Amazon Basin accounts for the largest share among them, leading to reduced biodiversity, habitat loss, climate change and other detrimental effects.
Planet Labs Inc. has built and deployed the largest constellation of Earth-imaging satellites, collecting daily imagery of the entire land surface at 3-5 meter resolution, providing the highest quality of earth-imaging ever produced. Planet Labs has provided satellite image chips of the Amazon Basin, both with and without labels corresponding to various atmospheric conditions and land cover/land use.
Having an efficient and reliable algorithm to automate the labeling of new satellite images using existing data, can help the global community better understand where, how, and why deforestation happens all over the world - and ultimately how to respond.

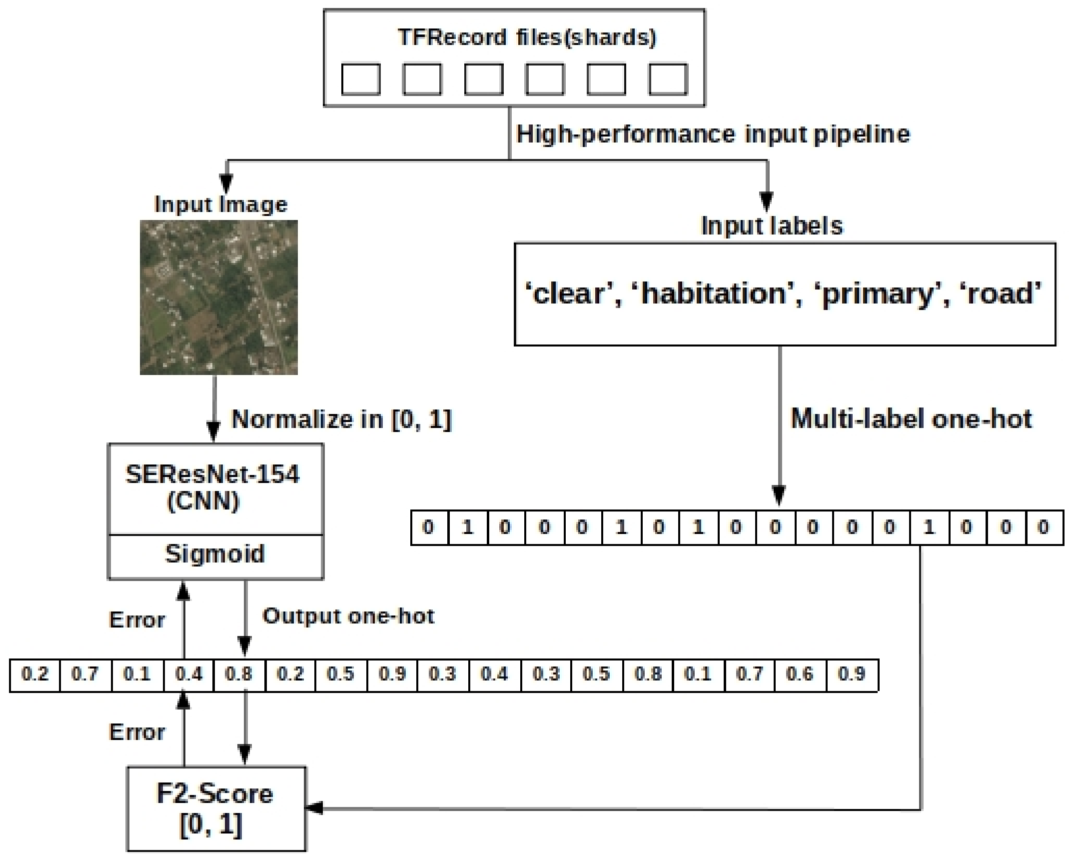
Methodology

This problem can be framed as a **Multi-label, Multi-class** problem, where each image can have more than one label. The training process can be divided into three parts:
Data Pre-processing: The images were normalized and the labels were one-hot encoded. Both were serialized and then written to multiple TFRecord shards for fast reading.
Data Reading and Augmentation: The image-labels pairs are read parallelly from TFRecord shards using a high-performance input pipeline. While reading, data is augmented by performing random horizontal and vertical flipping of images with a probability of 0.5 each, and shuffling the data after ever epoch. Finally, this stream of data is randomly split into 80-20 training-validation ratio.
Feature Extraction and Training:
In this model, **Squeeze-and-Excitation ResNet-154**[2], [3] is used as an image feature extractor. It modifies ResNet-154 using SE blocks as shown on the right, and has achieved state-of-the art results on ImageNet data set. For the last dense layer, the Sigmoid activation is used to keep individual label probabilities independent of each other. Finally, F2-score is used to determine error between output one-hots and ground-truth one-hots, which is propagated back through the network for training. The Schematic to the right clearly describes the steps of the training process.



-[2]

Schematic of Training Process



The Dataset

The data set used in this project was provided by Planet Labs on the Kaggle platform[1]. The total size of data is approximately **53 GB**, consisting of Images and labels as described below:
Images: There are around 40k(21 GB) multi-labeled and 61k(32 GB) unlabeled images in TIFF format, each of shape **(256, 256, 4)**, where the first three channels are B, G and R channels respectively, and the fourth channel corresponds to Near-Infra-Red(NIR) image data.
Labels: The data set consists of 17 unique labels, that describe various aspects of the image such as type of Cloud Cover, Haze, Primary Rain Forest, Water (Rivers & Lakes), Habitation, Agriculture, etc.

Below is a complete list of labels in the data set:

1. 'slash_burn'
2. 'cloudy'
3. 'partly_cloudy'
4. 'blooming'
5. 'primary'
6. 'haze'
7. 'conventional_mine'
8. 'clear'
9. 'road'
10. 'selective_logging'

11) 'agriculture'
12) 'water'
13) 'habitation'
14) 'artisinal_mine'
15) 'blow_down'
16) 'cultivation'
17) 'bare_ground'

Some of the labeled samples from the data set are given below.

Results

Training Setup: The model was trained for 50 epochs with a batch size of 24, and learning rate ranging from 1e-3 to 1e-5. Training was performed on Google Colaboratory, which consists of 12 GB of RAM and a Tesla T4 GPU with 16 GB Video Memory.
Evaluation Metric: F2-score, as recommended by the challenge authors, is used for evaluation on this task. Its described using below equations:

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

-[1]

This metric ranges from [0, 1], with higher score indicating better accuracy.
Testing Results: As discussed before, there is a separate Test data, the accuracy on which is calculated by Kaggle servers based on the uploaded results. The model in this project converges on the Train/Validation data as expected, and achieves an **F2 Score of 90.3%** on the Test data.
Examples: Two examples of ground-truth vs output labels of Validation images are shown to the right.

Actual: agriculture, partly_cloudy, primary, water
Output: agriculture, partly_cloudy, primary, water, slash_and_burn

Actual: primary, road, habitation, clear
Output: primary, road, blow_down, habitation

Conclusion

This project demonstrates the efficacy of modern Computer Vision algorithms -specifically Convolutional Neural Networks- in the analysis of complex earth satellite imagery with high accuracy. It also demonstrates the large positive impact CV algorithms can make on climate change and the protection of nature and biodiversity.

References

[1] Kaggle Challenge: www.kaggle.com/c/planet-understanding-the-amazon-from-space/overview/data
[2] Squeeze-and-Excitation Networks: arxiv.org/pdf/1709.01507.pdf
[3] SEResNet-154 Keras Implementation: www.github.com/titu1994/keras-squeeze-excite-network