Impact report

Estimating deforestation in Bolivia with image segmentation

Fanni Varhelyi, Sanha Tahir, Sukriti Mahajan

1. Introduction

Global warming is a critical issue facing humanity. While there are many contributing factors, one of particular interest is the issue of deforestation. Deforestation refers to intentional repurposing of forest land (National Geographic). It is concerning for two reasons. First of all, deforestation in itself is a carbon emitting activity: trees emit carbon when cut down. Two, trees are actively removing carbon from the air. Thus, the removal of forests is doubly harmful. Countering this trend has seen some progress and public awareness, but some of the reforestation or protection programs have dubious quality as well (Economist). Thus, understanding and countering this trend, and monitoring both deforestation and efforts to stop or counteract this on a detailed level will be crucial in the coming years. In particular, the deforestation of the Amazonian rainforest in areas such as Brazil or Bolivia has been alarming and worth investigating (e.g., WWF).

Our team set out to investigate this phenomenon, and see if we can contribute to the effort that is already being made to monitor and quantify deforestation. Monitoring is traditionally done through the analysis of satellite images: from image classifiers, methods have shifted mostly towards utilizing convolutional neural nets, such as VGG or ResNet, and applying semantic segmentation (Bragagnolo et al., 2022). More recent research in the last 2-3 years seems to utilize the convolutional encoder-decoder architecture, U-Net, developed in 2015 for medical imaging (Ronneberger et al., 2015). One paper compared U-Net, VGG, and ResNet, and found that an Attention based U-Net model could outperform all others and had an accuracy of ~94% (Bragagnolo et al., 2022).

Based on all the above, we formulated our concrete research question as follows:

> *How serious is the recent deforestation trend in Bolivia, and can a U-Net based image segmentation approach be used to monitor changes on a detailed scale?*

2. Data collection and initial analysis

There are numerous publicly available sources for satellite imagery, most notably the Landsat collection published by USGS. We opted for an alternate data source, provided by Planet through Norway's International Climate & Forests Initiative, aimed to make tropical imagery publicly accessible and thus advancing research and monitoring efforts (Planet). This dataset was attractive for two reasons. First, each image covers a smaller and thus more detailed area than the publicly accessible Landsat images, with pixel resolution of 4.77m, 4096x4096 resolution per image, GeoTIFF format (Planet).

Second, the imagery provided is already pre-processed: the images are comprised of multiple satellite passings and cleaned of clouds or other obstructions where possible. This dataset enabled us to focus on a small area within Bolivia where we identified increased deforestation when visually inspecting Global Forest Watch's map and database. We selected the location 15.69° - 16.48° S, 60.83° - 61.81° W, which resulted in 36 satellite images for our test data. The publicly available dataset covers the period 2016-2023, and we decided to use the normalized images for summer months for each year between 2017 and 2023. This resulted in 252 images for analysis. All of them were acquired through the Planet API in both full scale (specifications outlined above), and as 3-band 256x256 png files.

To train our model, we needed Amazon rainforest imagery that already had masks. For this purpose, we leveraged three datasets published by Bragagnolo et al. (2021). One had 45 3-band, 512x512 rgb GeoTIFF images with corresponding 1-band masks. The others had 619 and 605 4-band, 512x512 GeoTIFF images and corresponding 1-band masks in png format. Of these, we selected 250 images each for computational reasons. Finally, we acquired a base version of the model itself from John & Zhang (2022).

Most of our exploratory analysis was aimed to consolidate and preprocess the images acquired from different sources and to ensure the trained model can provide predictions on the test dataset. We also leveraged a base semantic segmentation model to provide a baseline for our predictions. Training images were normalized and changed to a float 255 data type. To improve accuracy and enlarge our training data, we leveraged data augmentation: performed rotations, added noise, and cropped into images. Then, we converted the acquired png images to GeoTIFF format for our test data. Once this preparatory work was done, we looked into different options to train our models.

3. Methodology and modeling

3.1 Model Architecture

The U-Net architecture, called as such because of its characteristic U-shaped structure, follows a traditional encoding-decoding structure, The encoding path sequentially reduces spatial resolution, increases feature channel and extracts hierarchical features, through a block of two convolutions and a Max Pooling layer. Each encoding block corresponds to a decoding layer that restores the resolution through upsampling and convolutional layers. This design is especially effective for capturing both global context and fine-grained details in images (Ronneberger et al., 2015).

The distinctive feature of U-Net is the incorporation of skip connections, which connect corresponding layers in the encoding and decoding paths. These skip connections facilitate the flow of information between different scales, preserving the spatial details during the upsampling process (Ronneberger et al., 2015).

An illustration of our architecture is shown on Figure 1.

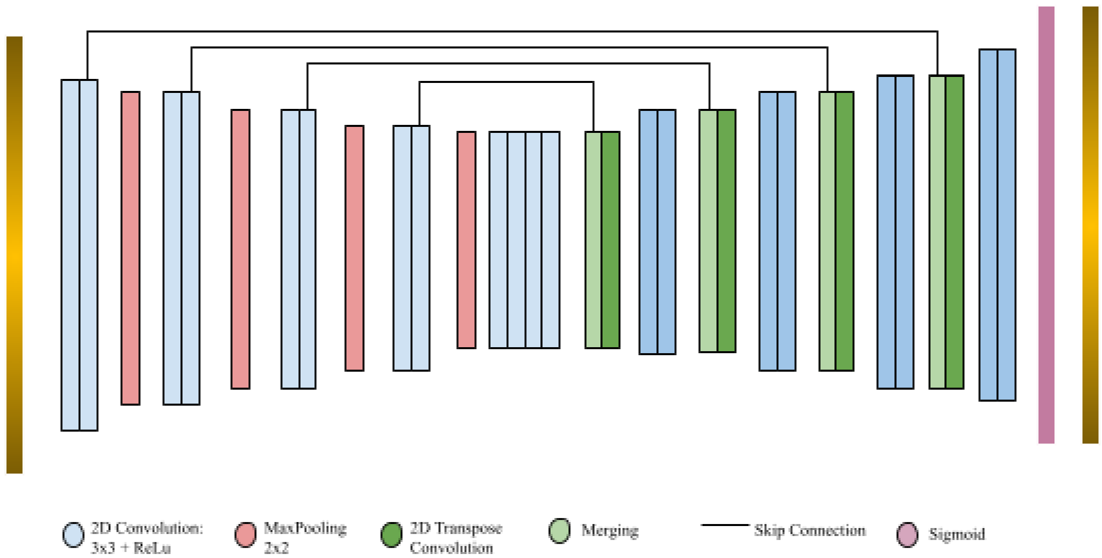


Figure 1: U-Net Architecture (self-authored)

We used an Adam Optimizer, with a learning rate of 0.0001 and a loss function of binary cross entropy. It is important to note that the shape of our input layer and the corresponding output layer were dynamic, allowing us to experiment with different sizes and different numbers of bands (i.e. the number of color channels) within the images.

We also experimented with the addition of attention in this model, by adding gates in the skip connections, before we did the concatenation. This attention mechanism allows the model to dynamically focus on specific regions of the input, enhancing our model's ability to capture and utilize relevant information of forest and non-forest areas for our image segmentation task (Bragagnolo et al., 2022).

3.2 Evaluation Metrics

We also experimented with our monitoring metrics.The choice of a monitoring metric during experimentation can have a significant impact on how we assess and optimize different models' performance. Changing the monitoring metric will change the way our model optimization works and how well it generalizes to unseen data. This was particularly important in our case, since our goal was for our final model to work well on image sets acquired from Planet; images for which we do not have masks.

In this vein, while we used the traditional metric of validation accuracy, we also used validation Mean Intersection over Union (Mean IoU). Mean IoU is commonly used to evaluate the performance of image segmentation models. Also known as the Jaccard Index, IoU is a measure of the overlap between the predicted segmentation mask and the ground truth mask. It is calculated as the intersection of pixels belonging to that class (in our case, either forest or non-forest) divided by the union of pixels belonging to that class (Rosebrock, 2023). The reason we chose this metric to experiment with is because it provides a more nuanced measure than accuracy. We wanted our models to score high mean IoU values, which would indicate good segmentation.

4. Results:

In total, we ran eight models, with different configurations that have been specified in the table Appendix 2. Before we dive into quantifiable metrics, we predicted the mask for our unseen Bolivia data:

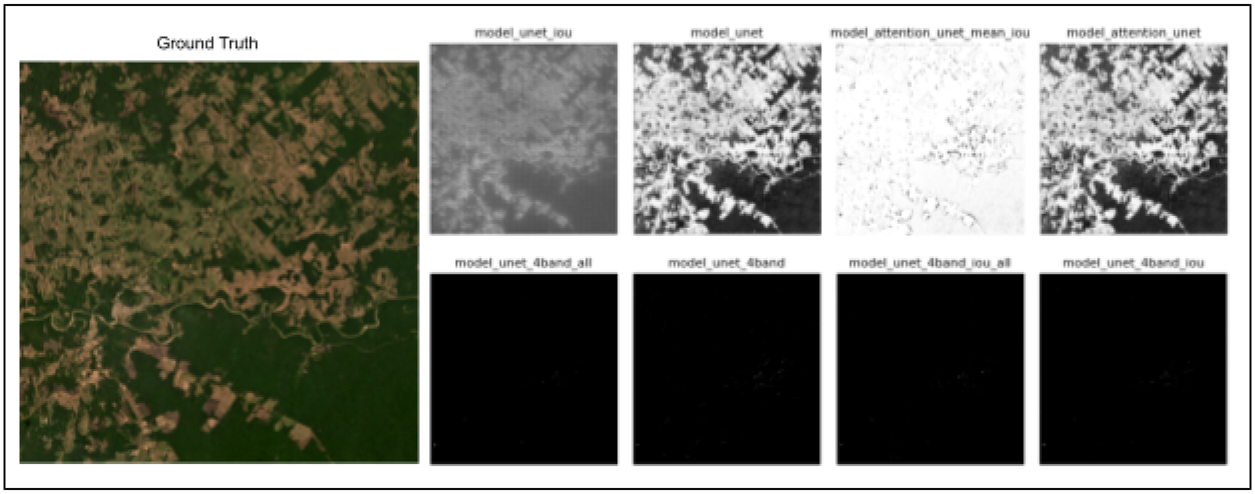


Figure 2: Predictions on a test file from Bolivia using different models.

The model failed to produce predictions when trained on 4 band images. This failure persisted even after different tries of preprocessing of the 4-band test images. We theorize this was due to differences in the data structures between the training and test data. When using 3-band regular RGB images, the model had no issues predicting high quality masks. We attempted to strip the 4-band images of their last channel and train the model thus, but it yielded similar unsatisfactory results. As a result, we decided to focus our work on the 4 models trained on 3-band images. Evaluation metrics for these models are given below:

Model	Accuracy Score	Precision	F1
model_unet_iou	0.53	0.74	0.62
model_unet	0.83	0.91	0.87
model_attention_unet_mean_iou	0.94	0.95	0.94
model_attention_unet	0.94	0.95	0.94

Table 1: Evaluation metrics of select models.

While model_attention_unet has higher metrics across the board, model_unet performed best for unseen Bolivia data. The results also indicate that mean IoU did not particularly help during the training, and it's even more apparent in the monitoring graphs, where the metric oscillates wildly between excellent values and poor values. This may be because our hyperparameters needed further refinement, for example learning rate or loss choice. To perform analysis on the rate of deforestation, we selected the outputs of model_unet, and worked with the masks predicted by this model.

As we analyzed the amount of predicted land area over the time period of 2017-2023, we saw an increase in non-forest areas in Bolivia over the years, which corresponds to our initial hypothesis. A summary of our results can be seen on Figure 3.

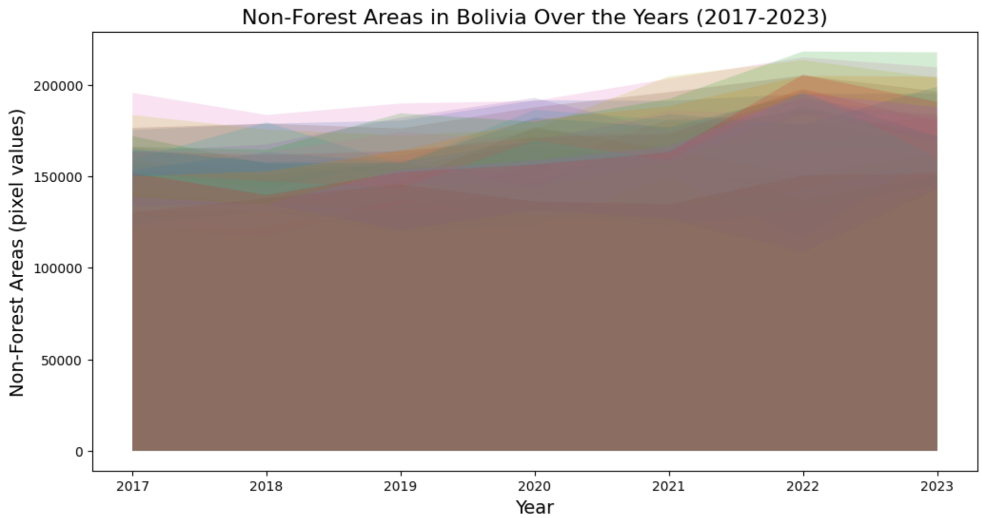


Figure 3: Non-forest areas over the years 2017-2023

Our findings indicate an overall increase of non-forest areas, about 17.6% over the investigated period. Based on the visual inspection of the images, most are agricultural. We also perceived a higher intensity of deforestation between 2020 and 2022, which somewhat aligns with the literature (e.g., the Guardian reports intensification of recent deforestation in 2023). We would also expect a similar intensity between 2022 and 2023, but our results indicate similar forested areas in 2022 as in 2023.

Conclusion

To conclude, we have managed to successfully implement the U-Net model to generate masks for a previously unexplored domain. By leveraging this model, we were able to quantify the extent of deforestation, providing a robust framework for systematic analysis and enabling the longitudinal tracking of this dire phenomenon. Our results indicate that a U-Net-based model, trained on a small sample of training images from the same region, can perform well on unseen new training data. We quantified deforestation for 2017-2023 in a Bolivian area with known deforestation, and saw an increase of 17,6% of land coverage.

Our limitations also highlight the difficulty of working with satellite images from different sources, and how beyond the modeling itself, a deep domain knowledge expertise is critical when transforming images from one format to another, and deciding which bands to work with. Our final methodology still has its merits. Using 3-band images could mean leveraging smaller and simpler input data that more people are familiar with. A model trained on more diverse data could theoretically give potential researchers accurate predictions on any image from the region it was trained on, even a screenshot. This could expand potential users of monitoring tools. We hope that our unique contribution can foster a more informed and data-driven approach to address environmental concerns.

References

Deforestation. National Geographic. URL: https://education.nationalgeographic.org/resource/deforestation/ (Accessed: 06 December 2023)

Deforestation in Bolivia has jumped by 32% in a year. What is going on? (2023) Thomas Graham. Guardian: https://www.theguardian.com/global-development/2023/oct/12/deforestation-in-bolivia-has-jumped-by-32-in-a-year-what-is-going-on (Accessed: 06 December 2023)

Forests burn, soils dwindle and people suffer. (2023) WWF. URL: https://wwf.panda.org/discover/knowledge_hub/where_we_work/amazon/amazon_threats/ (Accessed: 06 December 2023)

Trees alone will not save the world. (2023) The Economist: https://www.economist.com/special-report/2023/11/20/trees-alone-will-not-save-the-world (Accessed: 06 December 2023)

L. Bragagnolo, R.V. da Silva, J.M.V. Grzybowski, Amazon forest cover change mapping based on semantic segmentation by U-Nets, Ecological Informatics, Volume 62, 2021, 101279, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2021.101279

Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

Ahmad Alzu'bi, Lujain Alsmadi, Monitoring deforestation in Jordan using deep semantic segmentation with satellite imagery, Ecological Informatics, Volume 70, 2022, 101745, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2022.101745.

L. Bragagnolo, R.V. da Silva, J.M.V. Grzybowski, Towards the automatic monitoring of deforestation in Brazilian rainforest, Ecological Informatics, Volume 66, 2021, 101454, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2021.101454.

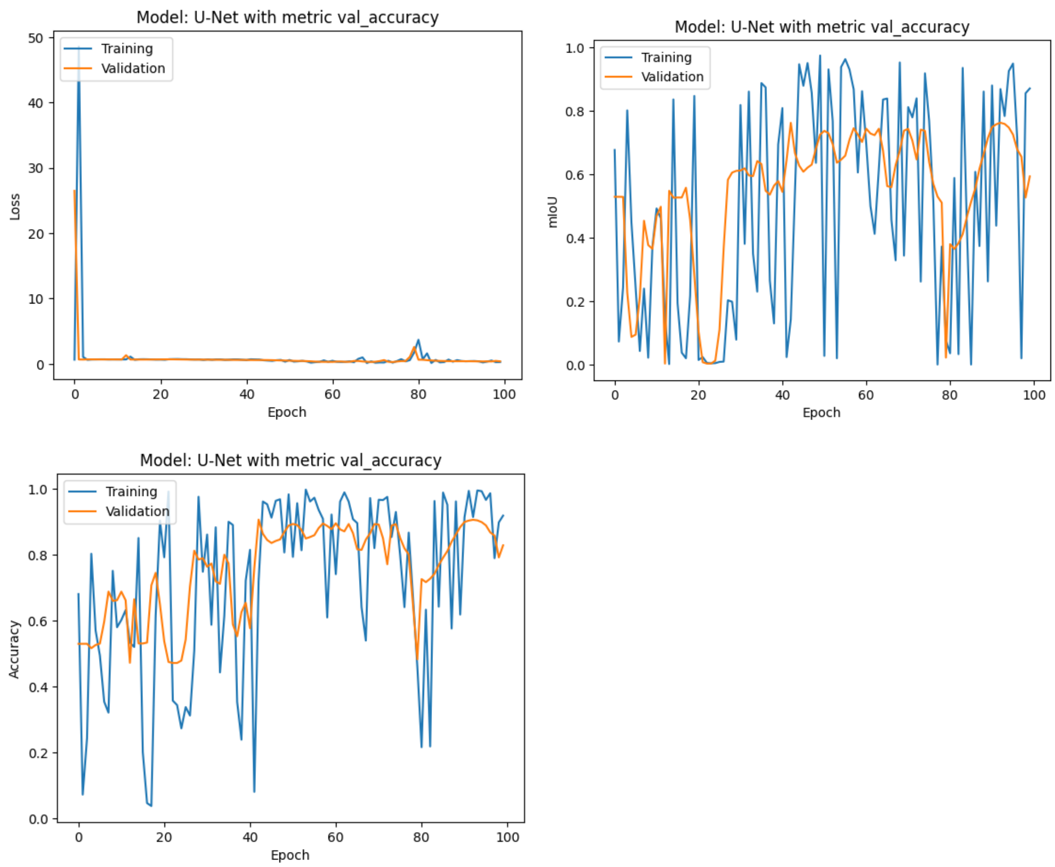
David John, Ce Zhang, An attention-based U-Net for detecting deforestation within satellite sensor imagery, International Journal of Applied Earth Observation and Geoinformation, Volume 107, 2022, 102685, ISSN 1569-8432, https://doi.org/10.1016/j.jag.2022.102685.

Rosebrock, A. (2023) Intersection over union (IOU) for object detection, PyImageSearch. Available at: https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/ (Accessed: 06 December 2023).

Data sources

Planet: https://developers.planet.com
USGS Landsat: https://landsatlook.usgs.gov
Amazon and Atlantic Forest image datasets for semantic segmentation.
https://zenodo.org/records/4498086 and https://zenodo.org/records/3233081

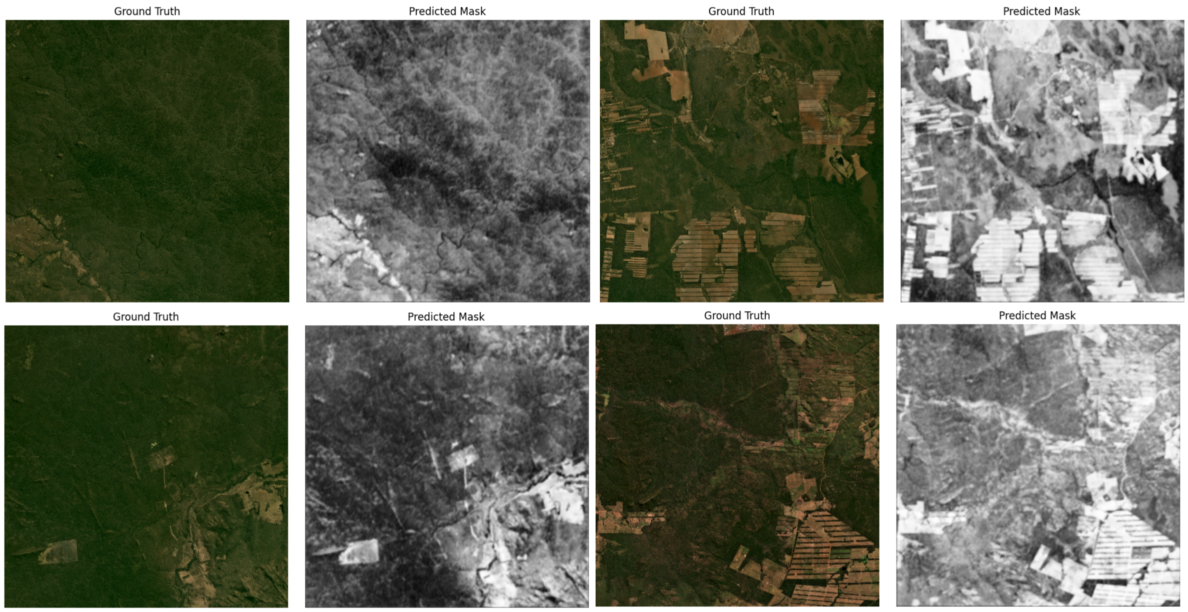
Appendix:



Appendix 1: Monitoring metrics for U-Net Modeling

Model	Training Image Band	Monitoring Metric
U-Net	30 Amazon 3-band	Accuracy
U-Net	30 Amazon 3-band	Mean IoU
Attention U-Net	30 Amazon 3-band	Accuracy
Attention U-Net	30 Amazon 3-band	Mean IoU
U-Net	250 Amazon 4-band	Accuracy
U-Net	250 Amazon 4-band	Mean IoU
U-Net	500 Amazon + Atlantic Forest 4 band	Accuracy
U-Net	500 Amazon + Atlantic Forest 4 band	Mean IoU

Appendix 2: Table for different model configurations



Appendix 3: Predicted masks for Bolivia images using model_unet, trained on 3-band Amazon Forest images