Deforestation Detection in the Brazilian Amazon Using Transformer-based Networks

Mariam Alshehri
EE & Computer Science
University of Missouri, USA

Anes Ouadou
EE & Computer Science
University of Missouri, USA

Grant Scott
EE & Computer Science
University of Missouri, USA

***Abstract*—Deforestation is a critical environmental issue that has far-reaching impacts on climate change, biodiversity, and the livelihoods of local communities. Conventional methods such as field surveys and map interpretation are not feasible, especially in vast regions like the Brazilian Amazon. In this paper, we adapt ChangeFormer, a transformer-based change detection model, to detect deforestation in the Brazilian Amazon, leveraging the attention mechanism to capture spatial and temporal dependencies in bi-temporal satellite images. To evaluate the model's performance, we implemented a rigorous methodology to create a deforestation detection dataset using Sentinel-2 images of selected conservation units in the Brazilian Amazon during 2020 and 2021. The model achieved a high accuracy of 94%, demonstrating the potential of transformer-based networks for accurate and efficient deforestation detection.**

***Index Terms*—Change detection, deep learning, deforestation**

I. Introduction

Deforestation poses a severe threat to the natural ecosystem as it causes the depletion of biodiversity, instability of ecosystems, and contributes significantly to climate change. The Brazilian Amazon, the world's largest rainforest, is vital for regulating climate and carbon levels. However, rapid deforestation rates have led to a plethora of issues, including increased greenhouse gas emissions, reduced carbon storage capacity, and more frequent forest fires [1].Given the catastrophic implications, it is imperative to promote effective policies to prevent it based on accurate and timely data, and deforestation detection (DD) is the primary source of such data.

While DD is recognized as an essential task in preserving the Brazilian Amazon, a significant obstacle arises due to its vast size, covering approximately 5.2 million km^2 of land.Remote sensing imagery (RSI) has emerged as an effective approach for detecting deforestation due to its wide area coverage, cost-effectiveness, and ability to provide consistent and repeatable data.Sentinel-2, launched by the European Space Agency (ESA) with a 10 m resolution and 5-day revisit rate, is an ideal open-access data source for DD. In RSI change detection (CD) applications, quantitative analysis is done to identify surface changes from images of the same place captured at different timestamps [2].

From the literature, there is evidence that deep learning (DL) methods outperform traditional machine learning techniques in CD applications. For instance, [1] implemented a convolutional neural network (CNN) to predict annual changes in vegetation cover in the Brazilian Amazon, but the technique incurs high computational costs due to redundant operations.

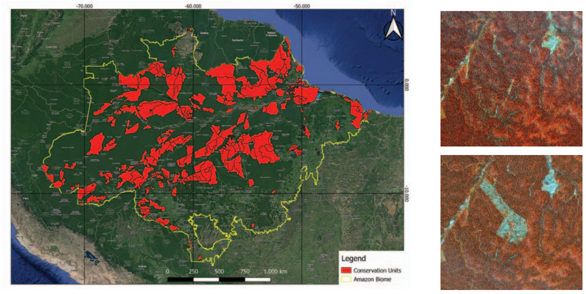

Fig. 1. The Amazon Biome (left) is outlined by the yellow boundary and the conservation units are highlighted in red. Color-shifted infrared from chips for 2020 (top) and 2021 (bottom) highlight deforestation changes up close.

TABLE I
THE TOP FIVE CONSERVATION UNITS BY SIZE OF DEFORESTATION

Area ID	Area(km^2)	Deforestation(km^2)	Deforestation(%)
59	16792	528.33	3.14
89	13017	214.59	1.63
55	1974	108.92	5.51
291	20395	114.70	0.56
165	9315	91.31	0.98

The transformer model is one of the DL models that introduce the attention mechanism to accommodate multi-temporal images. The model can easily be scaled, captures long-range sequence features, and facilitates efficient parallel processing. Despite its benefits in computer vision, the transformer model has received limited attention in DD research. The objective of this study is to investigate the application of the transformer-based network, ChangeFormer [3], in deforestation detection in the Brazilian Amazon. It is expected that the model will offer better deforestation detection accuracy in comparison to CNNs.

II. Methods and Data

We used Sentinel-2 Level-2A Surface Reflectance images with a maximum cloud cover percentage of 20%. We obtained ground truth polygons from the PRODES project datasets, provided by the Brazilian National Space Agency (INPE). Figure 1 (left) displays the polygons of these conservation units, as well as the boundary of the Amazon Biome.

We selected the top-5 conservation units with the highest deforested land areas for 2020 and 2021 (see Table I). To ensure image dates align closely with labeling dates, we identified the earliest and latest labeling dates for each area in PRODES, as set image acquisition dates to $\pm$30 days.We selected the visible and near-infrared bands at 10 m resolution and Scene Clas-

2023 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-3984-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAI54212.2023.00130

979-8-3503-3984-0/23/$31.00 ©2023 IEEE
DOI 10.1109/CAI54212.2023.00130

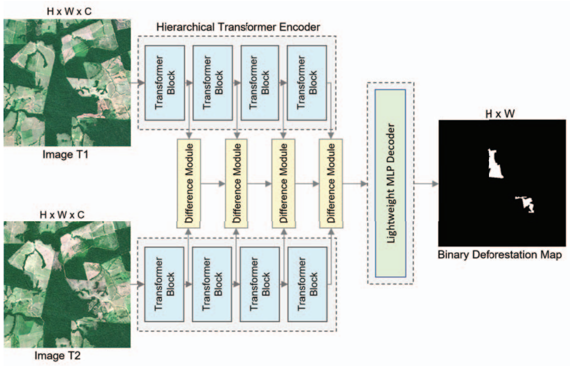

Fig. 2. A simplified overview of the ChangeFormer architecture, showing its three main components: a siamese hierarchical transformer encoder, four difference modules, and a lightweight MLP decoder.

sification Layer (SCL) at 20 m resolution. The near-infrared band distinguishes vegetation from other features, while the SCL band acted as a mask to eliminate cloud or cloud-shadow areas. We employed three band combinations, Color-shifted Infrared (NIR-R-G), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI), to highlight vegetation spectral signatures. Figure 1 (right) highlights the deforestation in two NIR-R-G images taken in 2020 and 2021.

We applied filtering at both the raster and chip levels. High-quality rasters were selected manually, while at the chip level, we removed single-class chips and those with less than 10% "change" class area.At the end of the chip creation process a total of 7,734 pairs of chips in 256×256 size and 1,406 pairs of chips in 512×512 size were generated.The dataset then was split into three subsets: (60%, 20%, 20%) for training, testing, and validation, respectively.

III. Experimental Results

We employed ChangeFormer [3], a transformer-based network specifically designed for the task of change detection, which has demonstrated outstanding performance in detecting building and general changes in urban areas, see Fig. 2. We initialized the model randomly and optimized its performance by tuning hyperparameters. We trained multiple model configurations for 200 epochs and evaluated their performance using overall and change-class-specific metrics such as F1 score, IoU score, precision, and recall.

We found that the 256×256 chip size outperformed the 512×512 chip size in all metrics, likely due to the availability of more training chips, with the top three results from the former surpassing the best result from the latter. We further determined that the optimal combination for the ChangeFormer model was a learning rate of 0.0001, CE loss, AdamW optimizer, and the NIR-R-G band. This achieved an overall accuracy of 94%/92% (256×256/512×512) with mIoU scores of 83%/80% and F1 scores of 90%/88%, respectively.

A previous study [4] also applied deep learning for deforestation detection by comparing various fully convolutional network architectures. Similar to our work, they used the PRODES dataset for ground truth, and they utilized both

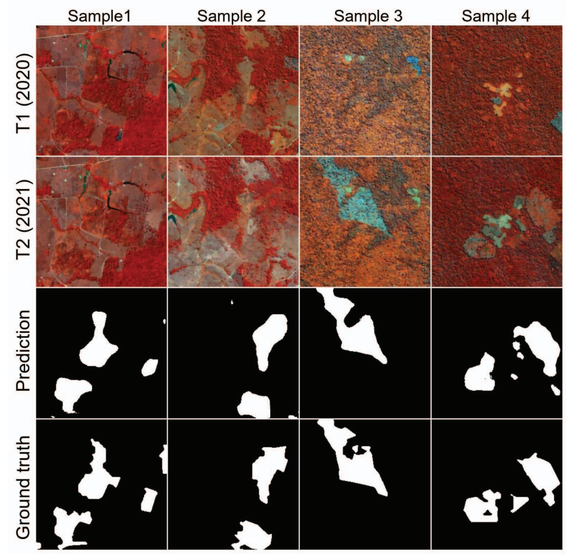

Fig. 3. Four image samples with corresponding predicted and ground truth deforestation maps. The top two rows show images from 2020 and 2021, respectively. The third and fourth rows display the predicted and ground truth maps, respectively.

Sentinel 2 and Landsat-8 satellites imagery. The best result for their study was achieved by FC-DenseNet with an F1-score of 70.7%. However, ChangeFormer model obtained at least 81% F1-score for the change-class in both chip sizes. Similar trends were observed in recall and precision, where we achieved 80% and 86% in our training process compared to 75.1% and 78% reported in the previous study.

IV. Conclusion

We obtained Sentinel-2 satellite imagery and ground truth data for deforested areas in the Brazilian Amazon rainforest in 2020 and 2021. We conducted a thorough hyperparameter search, exploring various configurations to identify the best settings for our task. Our investigation showed that color-shifted infrared composite and cross-entropy loss with AdamW optimizer yielded the best results. In comparison to existing research in the field of deforestation detection, our findings suggest that transformer-based networks are capable of achieving significantly improved levels of accuracy.

References

[1] P. P. de Bem, O. A. de Carvalho Junior, R. Fontes Guimarães, and R. A. Trancoso Gomes, "Change Detection of Deforestation in the Brazilian Amazon Using Landsat Data and Convolutional Neural Networks," *Remote Sensing*, vol. 12, no. 6, p. 901, Jan. 2020.

[2] L. Khelifi and M. Mignotte, "Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis," *IEEE Access*, vol. 8, pp. 126 385–126 400, 2020.

[3] W. G. C. Bandara and V. M. Patel, "A Transformer-Based Siamese Network for Change Detection," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2022, pp. 207–210.

[4] D. L. Torres, J. N. Turnes, P. J. Soto Vega, R. Q. Feitosa, D. E. Silva, J. Marcato Junior, and C. Almeida, "Deforestation Detection with Fully Convolutional Networks in the Amazon Forest from Landsat-8 and Sentinel-2 Images," *Remote Sensing*, vol. 13, no. 24, p. 5084, Jan. 2021.