INSTITUT D'ENSEIGNEMENT SUPÉRIEUR DE RUHENGERI

B.P. 155, Ruhengeri, Rwanda

T : +250 788 90 30 30, +250 788 90 30 32, W : www.ines.ac.rw, E : info@ines.ac.rw

DEEP LEARNING-BASED ANALYSIS OF SATELLITE IMAGE TIME SERIES FOR MAPPING FOREST REGENERATION IN AMAZON RAINFOREST

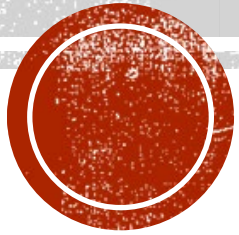
Presented by:

Fulgence HATANGIMANA - S2493209

Supervisors

Dr Raian Maretto,

Prof Claudio Persello



November 2024

THESIS ASSESSMENT BOARD:

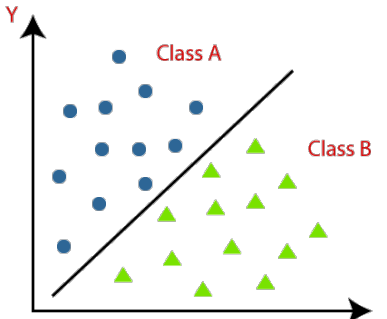
dr. Mariana Belgiu

dr. Marcos Adami

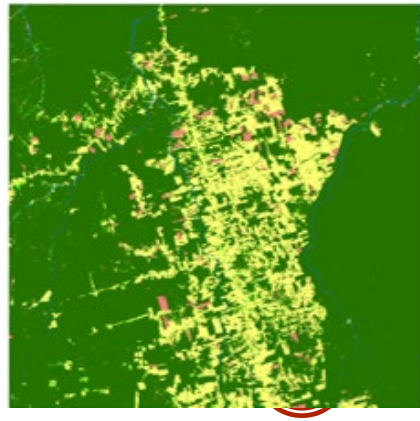
Fédération des géomètres francophones

INTRODUCTION

- The Amazon rainforest, particularly the Pará region, experiences significant deforestation.
- Accurate satellite image classification is crucial for environmental monitoring, and urban planning among others.
- However, current methods rely on traditional approaches, which often lack precision, and accuracy, making them time-consuming.
- **This Research Gap -** introducing a deep-learning approach to improve accuracy in mapping forest regeneration



Classification Algorithm



RESEARCH OBJECTIVES

To develop and implement a deep learning-based model for analyzing satellite image time series for mapping forest regeneration in the Amazon rainforest.

Specific objectives

1. To explore deep learning-based models, specifically the transformer and hybrid transformer, to accurately identify and distinguish regenerated areas from other classes in the Amazon rainforest.

2. To Assess the model's capability by analyzing satellite image time-series data over different time lengths, ensuring improved accuracy and a more realistic representation of the regenerated area.

3. Evaluate the best-performing model for mapping regenerated areas from the developed models.

4. To quantify the area of secondary forest across the entire study area based on the classification generated by the proposed method.

RESEARCH QUESTIONS

1. Which of the two deep-learning architectures, transformer, and hybrid transformer, is the most suitable for accurately differentiating regenerated areas from other classes in the Amazon rainforest?

2. What are the capabilities of the model to effectively analyze time-series data across varying time steps to achieve a more accurate and realistic representation of regeneration in the Amazon rainforest?

3. What is the performance of the best-performing model for mapping regenerated areas from the developed models?

4. To what extent is there a secondary forest across the entire study area based on the classification generated by the proposed method.?

STUDY AREA AND DESCRIPTION

- **Amazon Rainforest**: Covers approximately 5.5 million km², making it one of the largest biodiverse ecosystems on Earth.
- **Pará State**: A major Brazilian state largely covered by the Amazon, with high deforestation and regeneration rates.
- A smaller impacted region in Pará was chosen to optimize the modelling process

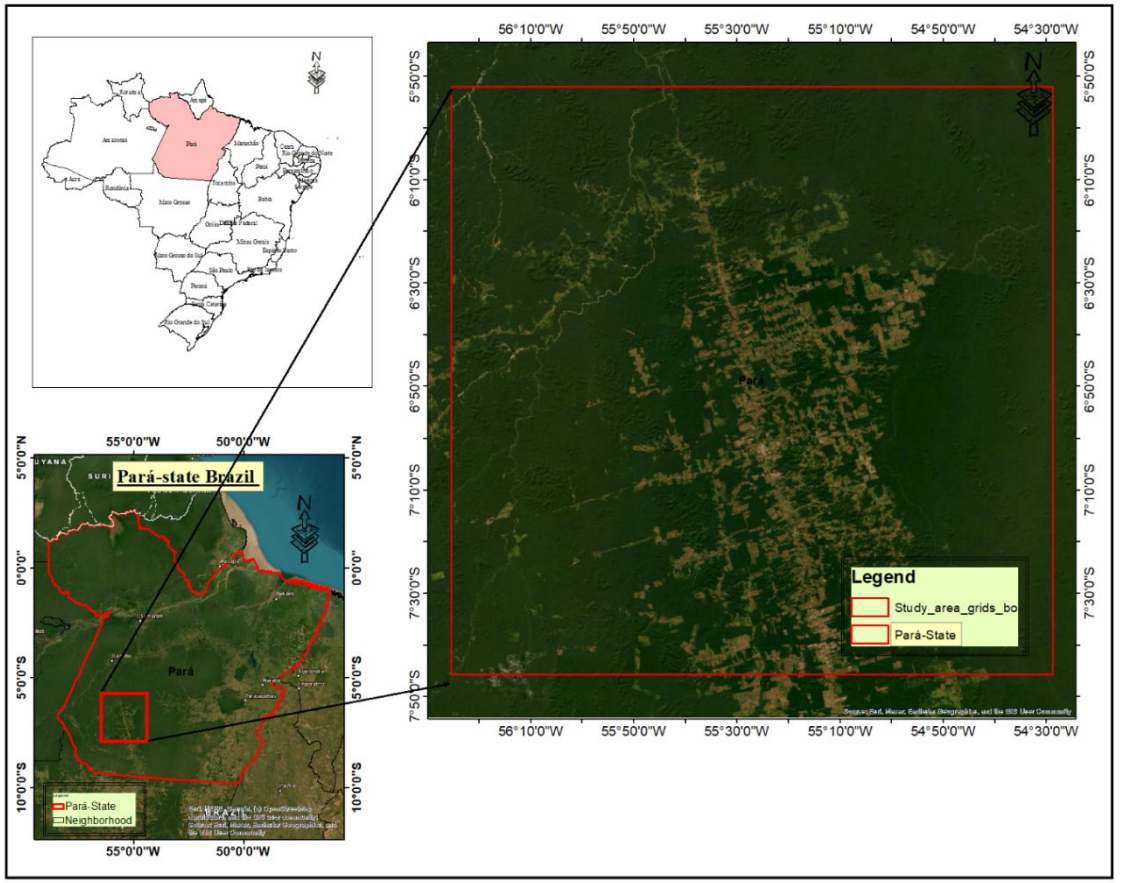


Figure1:The study area is the Amazon rainforest in Pará State, Brazil. The large map highlights the study grid (red outline).

DATASETS USED FOR THIS STUDY

	Dataset	Source	Temporal
1	Land Use and Land Cover Map	https://brasil.mapbiomas.org	2021
2	Deforestation Data from PRODES	https://terrabrasilis.dpi.inpe.br	2021
3	Satellite images	USGS Earth Explorer	2012-2021
4	Brazil tiles	Brazill data Cube	2022
5	Administrative boundaries	https://www.diva-gis.org	2022

Software and tools

python™

ArcGIS® Pro

Google Earth Engine

Software packages

NumPy
Mathematical functions

PyTorch
Machine Learning

Keras
Deep Learning

Tensorflow
Machine Learning

Pandas
Data analysis and manipulation

Matplotlib
Data visualisations

METHODOLOGY

❑Data pre-processing steps :

Atmospheric correction and cloud masking:

- ✓ Could an obscure land surface, making the data from those pixels unusable for tasks like land-cover classification.
- ✓ Correction – masking out the cloud from images that don't represent the land surface.

❑ Input data **: (samples (S), channels (C), time steps(t_i)**

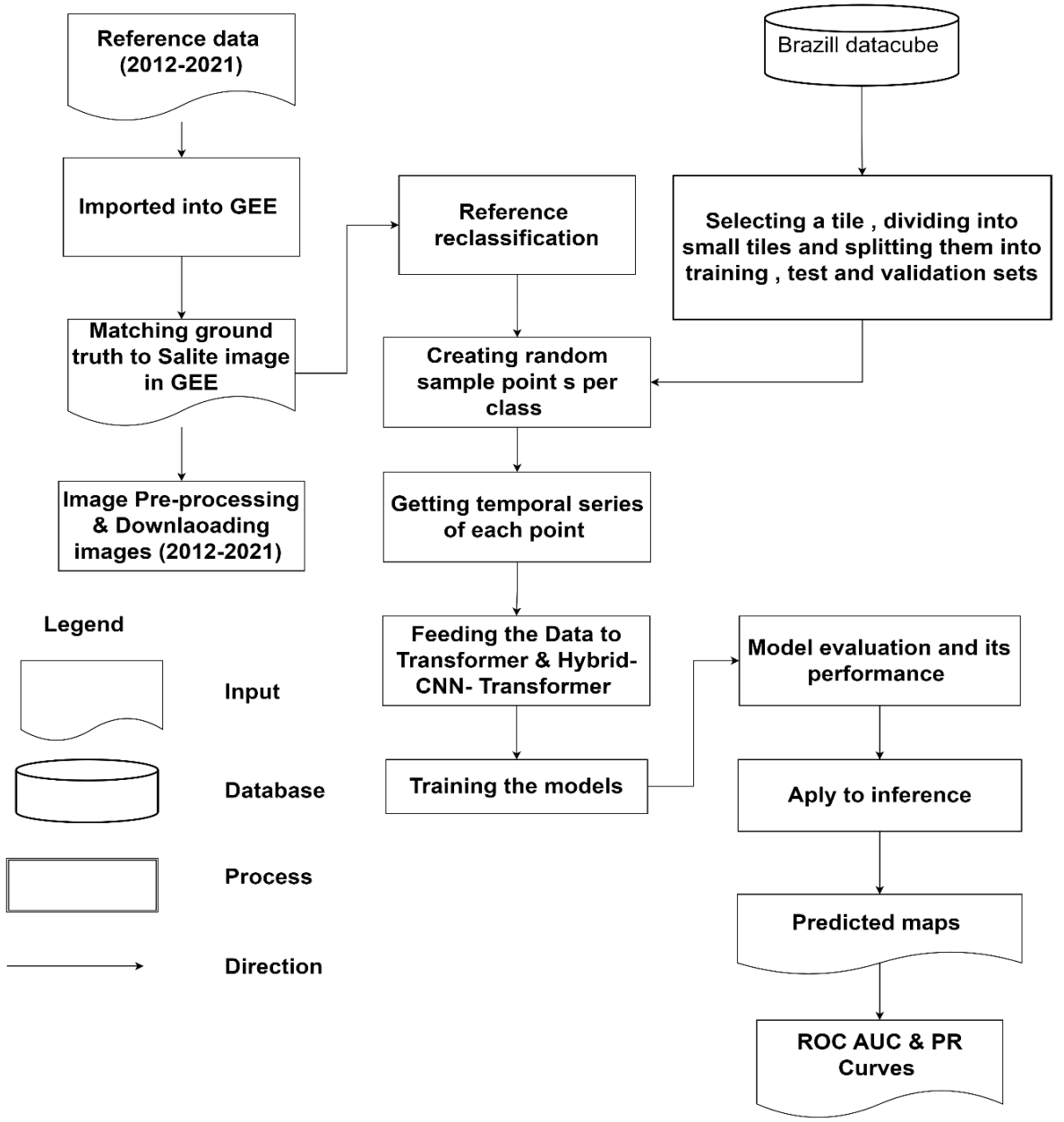
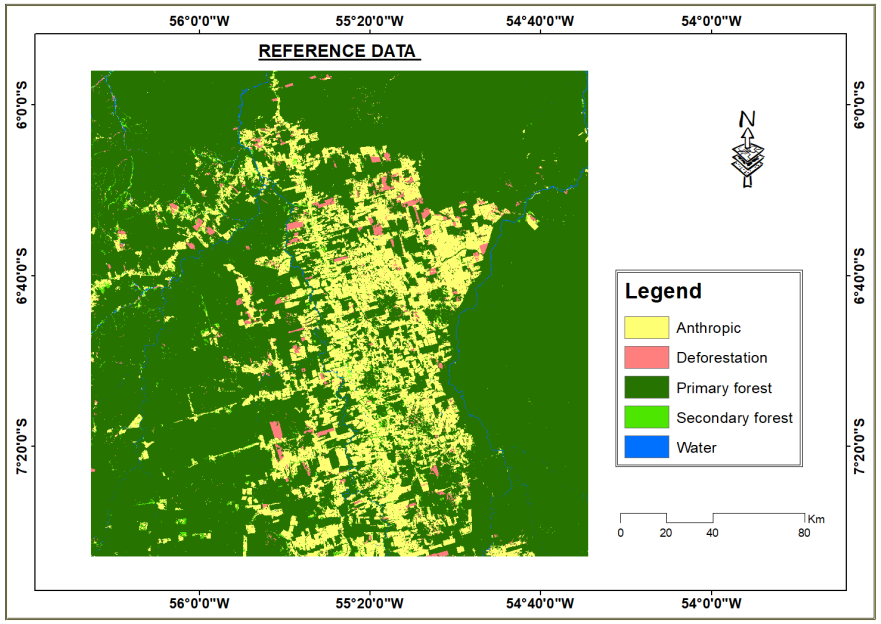


Figure 2: Workflow for Land Cover Classification using Satellite Imagery and Deep Learning Models

DATASET PREPARATION AND MODEL SETUP



https://brasil.mapbiomas.org/en/colecoes-mapbiomas/

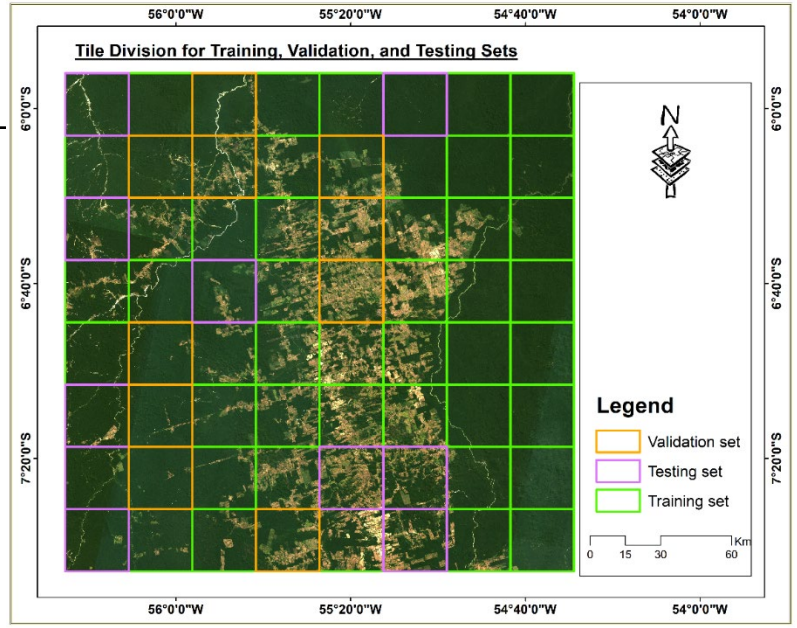


Figure 3: Division of large tiles into 64 smaller tiles and splitting them into training (green), validation (orange), and testing (yellow) sets.

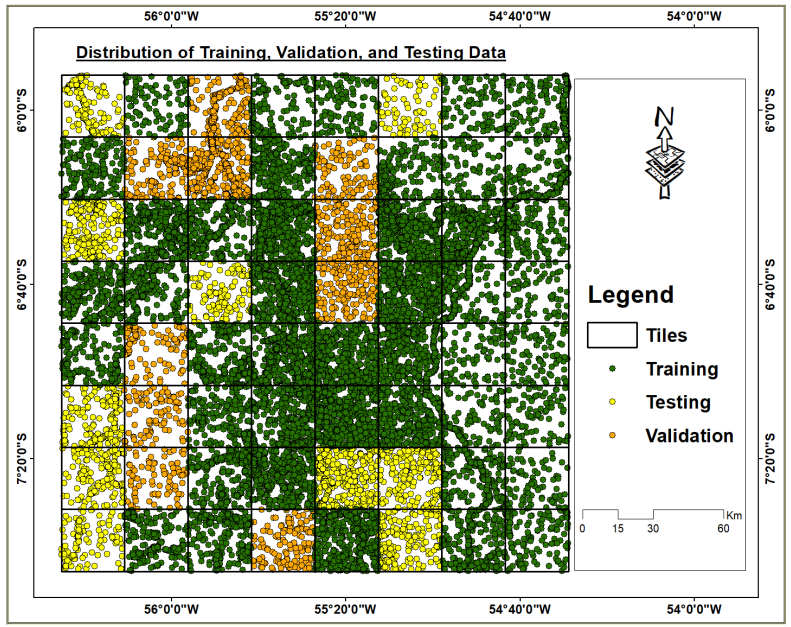
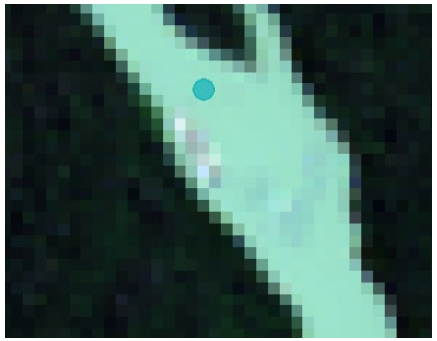
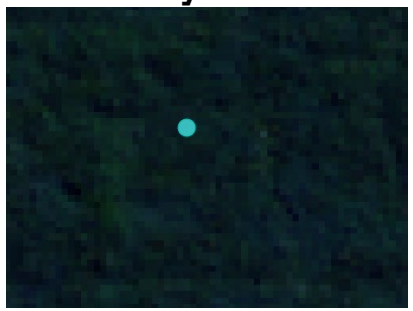


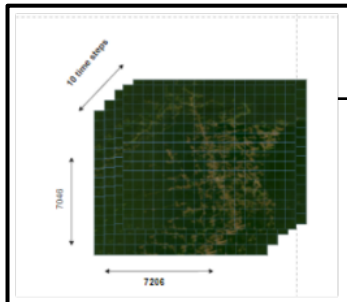
Figure 5: Distribution of random points for the training (green), testing (yellow), and validation (orange) sets.



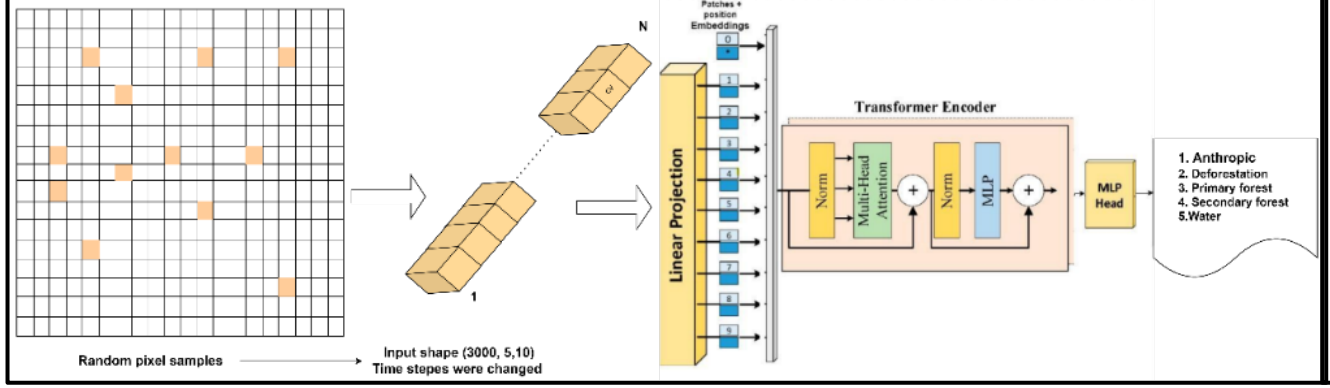


MODEL ARCHITECTURE AND IMPLEMENTATION –OBJ (1)

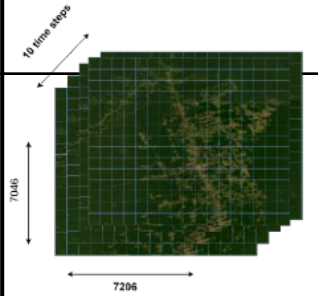
Standard transfomer



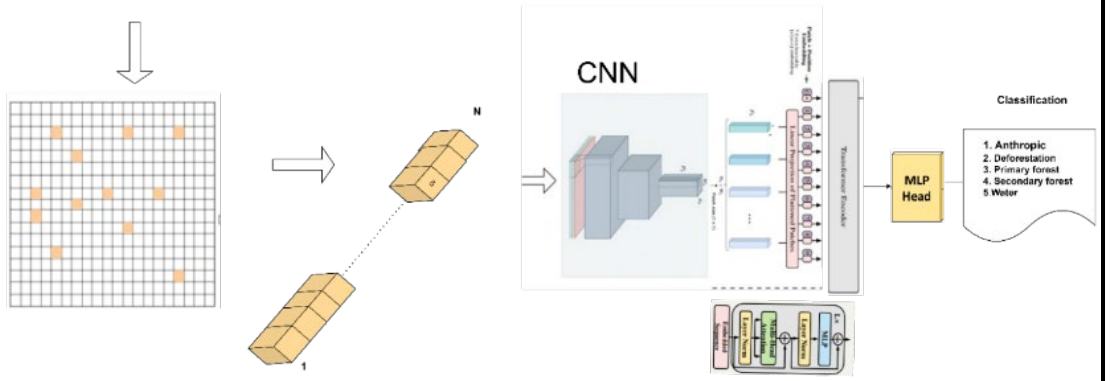
The transformer model uses self-attention mechanisms to capture local & global dependencies within the satellite image time series data.



Hybrid transfomer



The hybrid transformer model combines CNN with transformer layers



Input Data: Both models were trained on **1D sequence of pixels across five bands over different time lengths i.e. (samples (S), channels (C), time steps(t_i)),** which were labelled into various classes such as anthropic, deforestation, primary forest, secondary forest, and water bodies

RESULTS

ASSESSING THE MODEL'S CAPABILITY BY ANALYZING TIME-SERIES DATA OVER DIFFERENT TIME LENGTHS, (OBJ-2)

Research Question 1: Which of the two deep-learning architectures, transformer, and hybrid transformer, is the most suitable for accurately differentiating regeneration areas from other classes in the Amazon rainforest? **This answer Research Q1 & Q2.**

Table 7: Model capabilities with time series data on performance metrics over different time lengths (Using testing **points**)

Time length (t_i)	Model	Precision	Recall	F1-score	Overall accuracy (%)
1	Transformer	0.71	0.71	0.71	70.80
	Hybrid Transformer	0.74	0.74	0.74	73.88
3	Transformer	0.81	0.81	0.8	80.74
	Hybrid Transformer	0.83	0.82	0.88	81.76
5	Transformer	0.84	0.83	0.83	83.70
	Hybrid Transformer	0.85	0.84	0.84	85.50
10	Transformer	0.86	0.85	0.85	85.48
	Hybrid Transformer	0.86	0.86	0.86	86.36

RESULTS

THE PERFORMANCE OF THE BEST-PERFORMING MODEL FOR MAPPING REGENERATED AREAS –OBJ (3)

Q3. What is the performance of the best-performing model for mapping regenerated areas from the developed models?

Table 8: Performance of the hybrid model to map regenerated areas using test sample points.

	hybrid transformer				Transformer			
Class	**Precision (User)**	**Recall (Producer)**	**F1-Score**	**Overall Accuracy**	**Precision**	**Recall (Producer)**	**F1-Score**	**Overall Accuracy**
Anthropic	87.98	87.8	87.89		84.41	88.8	86.55	
Deforestation	84.79	78.6	81.58		94.49	68.6	79.49	
Primary Forest	79.94	86.5	86.5		78.56	85.4	81.84	
Secondary Forest	81.65	81	81.32		76.22	85.9	80.77	
Water	97.8	97.9	97.85		97.92	98.7	98.31	
Overall Accuracy				**86.36**				**85.48**

Table 9: Performance of the hybrid model for mapping regenerated areas.

	Hybrid transformer				Transformer			
Class	**Precision (User)**	**Recall (Producer)**	**F1-Score**	**Overall Accuracy (%)**	**Precision (User)**	**Recall (Producer)**	**F1-Score**	**Overall Accuracy (%)**
Anthropic	99.42	87.56	93.02		89.98	89.12	90	
Deforestation	51.21	71.82	59.75		38.97	30.74	51.45	
Primary Forest	94.8	94.98	94.89		95.48	98.63	97.61	
Secondary Forest	64.95	67.73	66.32		61.24	59.69	61.91	
Water	55.42	94.29	69.85		88.81	54.1	67.82	
Overall Accuracy				**81.087**				**79.346**

QUALITATIVE ANALYSIS OF THE RESULTS

I. Predicted maps made by Hybrid and Transformer Models

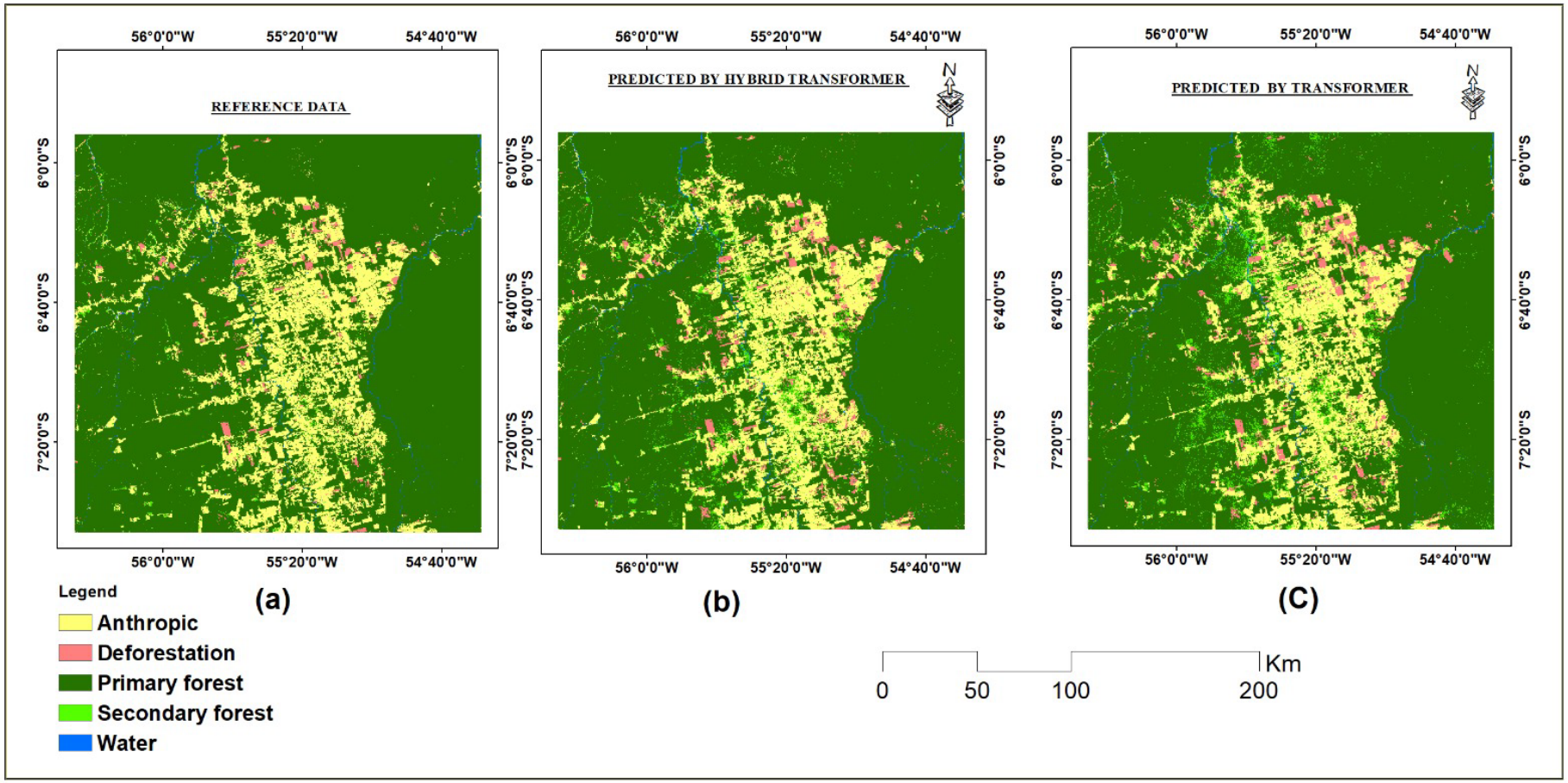


Figure 8: Depicts predicted and reference maps for forest regeneration

II. IDENTIFYING AREAS OF ACCURATE PREDICTION

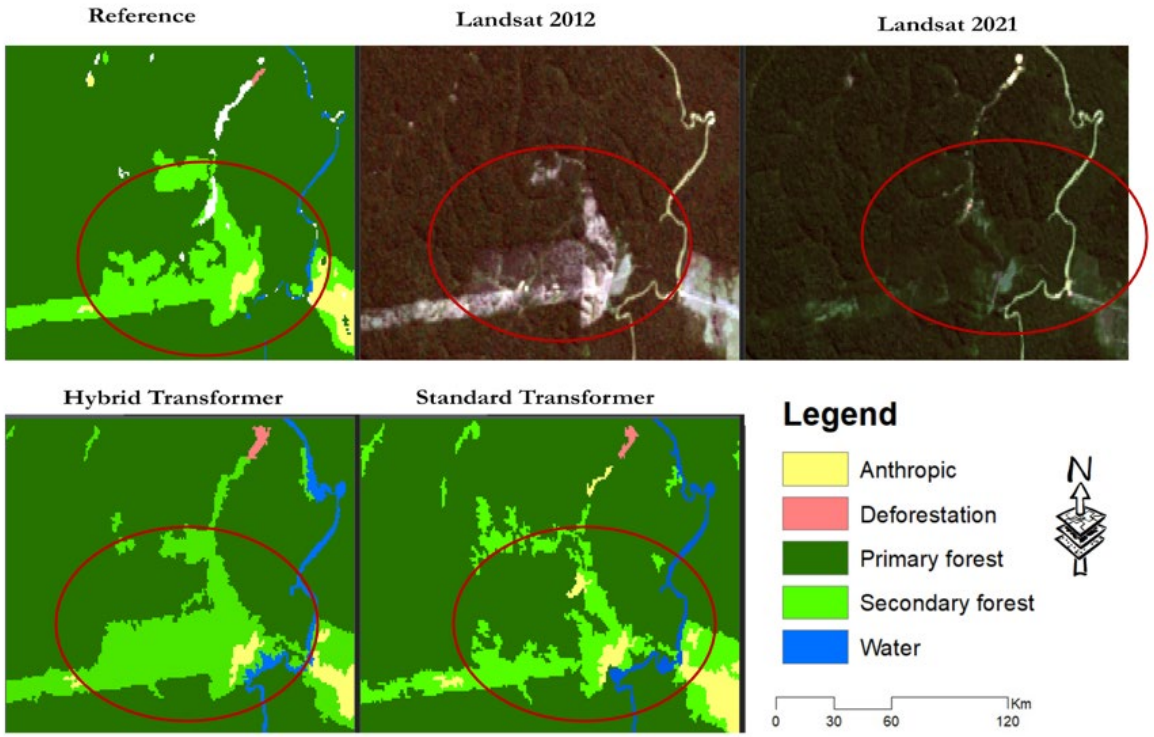


Figure 11: Comparison of model predictions with reference data and historical Landsat images (a)

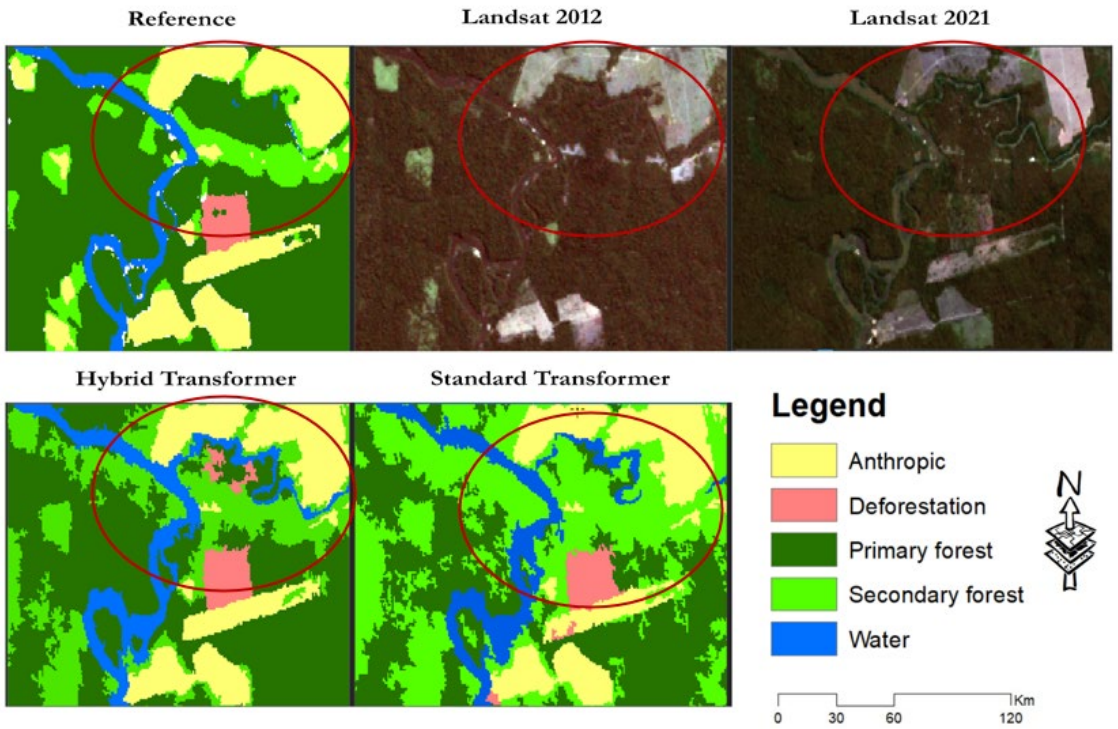
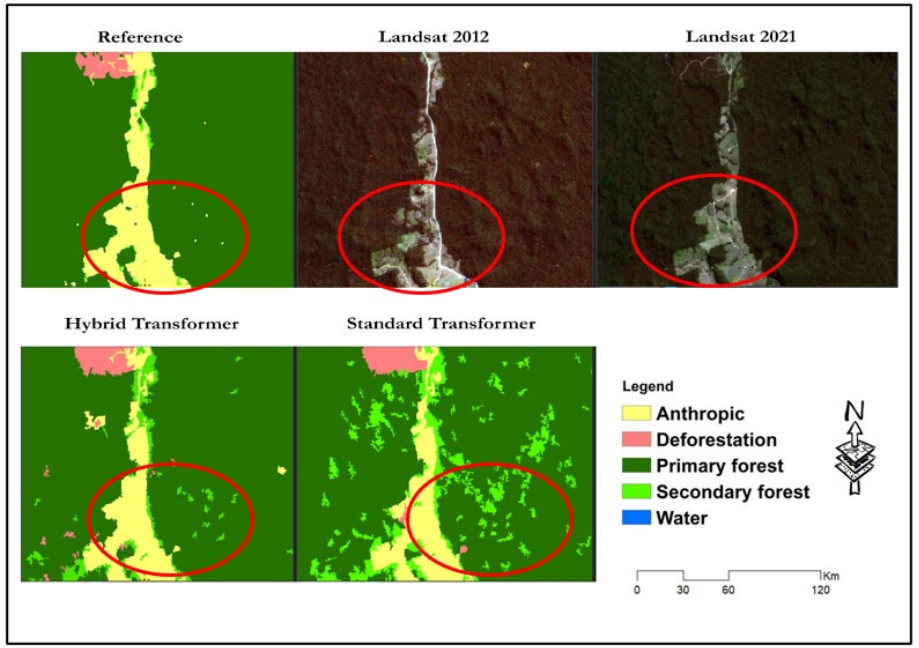
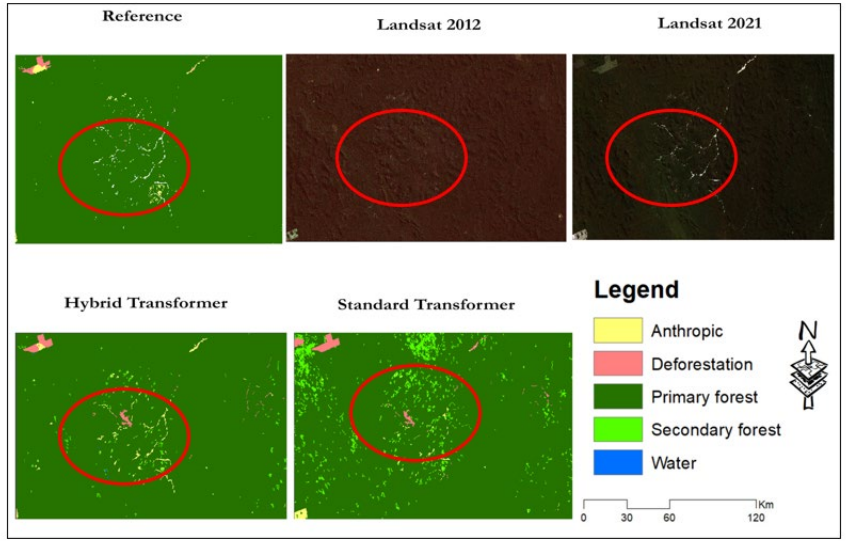


Figure 12: Comparison of model predictions with reference data and historical Landsat

III. INACCURATE PREDICTIONS:

- Models struggled in regions with uncertainties in reference data, etc.
- Hybrid Transformer outperformed the traditional Transformer but still faced difficulties in certain areas.
- Figure 10 shows noisy predictions (third column) versus accurate reference data (fourth column), highlighting misclassification issues.
- The model misclassified cloud cover as forest in the third and fourth rows.
- Understanding these inaccuracies is crucial for refining model training and improving overall accuracy



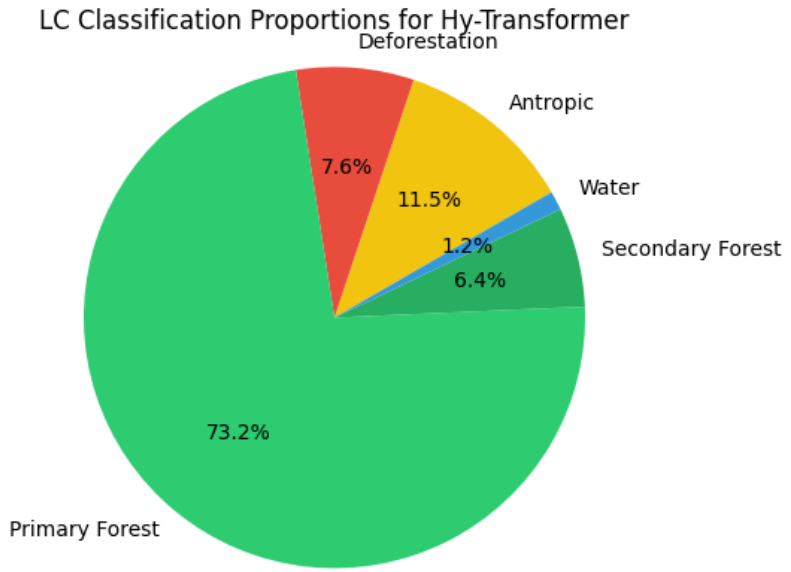


fied

QUANTIFICATION OF THE EXTENT OF SECONDARY FOREST –OBJ (4)

Q4. To what extent is there a secondary forest across the entire study area, based on the classification generated by the proposed method?.

Model Accuracy: Best-performing model achieved an overall accuracy of 86.38%.



The model effectively identified and quantified various land cover classes, aiding in environmental monitoring and forest management.

CONCLUSIONS AND RECOMMENDATIONS

The main objective of this research was to develop and implement a deep learning-based model for the analysis of satellite image time series to map forest regeneration areas.

- Successfully developed a hybrid transformer model that accurately identifies and distinguishes regeneration areas in the Amazon, surpassing the traditional transformer with an accuracy of 86.36%.
- **Objective 2 Achieved**: Demonstrated the model's capability to analyze time-series data, with significant accuracy improvements at longer time lengths, providing a more realistic and detailed representation of forest regeneration dynamics.
- **Objective 3 Achieved**: The hybrid model outperformed in mapping regenerated areas, showing superior performance in accuracy, precision, recall, and F1-score metrics.
- **Objective 4 Achieved**: Quantified the extent of secondary forest across the study area, providing valuable data for monitoring forest recovery and guiding reforestation efforts.
- This research offers a robust tool for large-scale environmental monitoring, with significant implications for forest conservation and climate change mitigation.

RECOMMENDATIONS

- **Future research should keep exploring hybrid transfomer** that explicitly incorporate space-time dimensions to enhance the accuracy of forest regeneration mapping.

- **Extend Time Series Analysis**: Future studies should continue to explore the potential of hybrid transformer models for handling extended time-series data

- **Integrate Additional Data**: Combine satellite imagery with NDVI and DEM data to enhance model accuracy.

- **Conduct Field Surveys**: Owing to the current limitations imposed by inaccuracies in the reference data, it is advised that future research includes comprehensive field surveys.

Thank you