Deforestation Detection using Satellite images

2024 Math to Power Industry Workshop

Pacific Institute for the Mathematical Sciences

Math INDUSTRY

Team 3 (Finite Carbon)
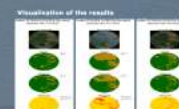
Pipeline

1. Data Acquisition

2. Preprocessing

3. Splitting the Data

4. Cross Validation and Hyperparameter Optimization

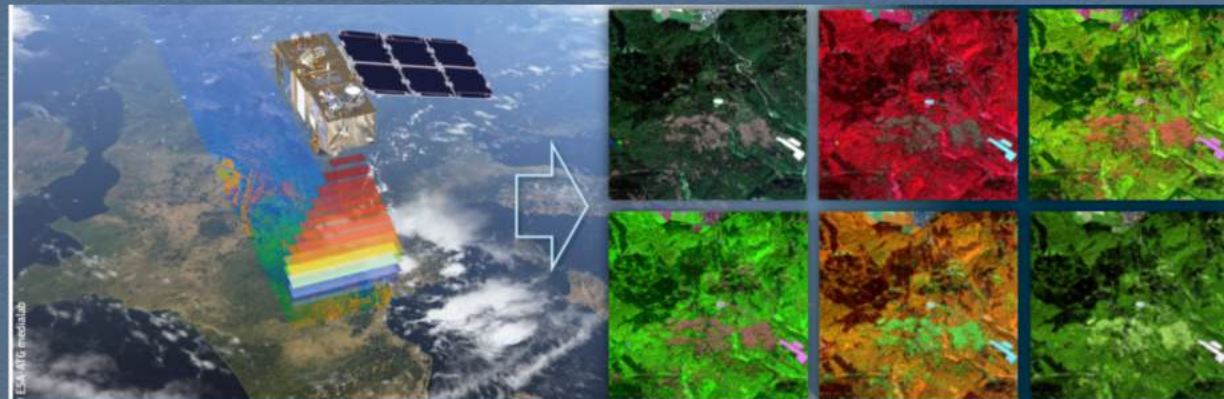### Visualization of the results

Team 3 (Finite Carbon)

- **Mentor**
  - Bahareh Yekkehkhany (Applied Remote Sensing Scientist, Finite Carbon)

- **Members**
  - Isaac Asamoah (University of Saskatchewan)
  - Patrik Coulibaly (University of British Columbia)
  - Arman Jahangiri (University of Calgary)
  - Yasaman Shahhosseini (University of Victoria)
  - Brian Zambrano (University of Alberta)

Problem Statement

The objective of this project is to implement a machine learning algorithm for distinguishing between forest and non-forest regions in Canada using satellite imagery.

Bulkley-Nechako District, Central B.C.

Our approach

- **Image source**:
  ESA Sentinel-2 Satellite images from
  the Google Earth Engine (GEE)
  (higher resolution, revisiting frequency)

- **Methodology**:
  Training a Random Forest Classifier
  based on F1 and Recall scores

Pipeline

1. Data Acquisition

3. Splitting the Data

2. Preprocessing

4. Cross Validation and
   Hyperparameter Optimization

1. Data Acquisition

1. Data Acquisition

- Time (May-September 2021, 2022, and 2023)
- Manually Selected Region

Pipeline

1. Data Acquisition

- Time (May-September 2021, 2022, and 2023)
- Manually Selected Region

2. Preprocessing

3. Splitting the Data

4. Cross Validation and Hyperparameter Optimization

2. Preprocessing

2. Preprocessing

- Cloud Percentage Selection
- Feature Engineering (NDVI, NDMI)
- Cropping the area of interest
- Cloud Masking
- Creating reference label
- Normalization (MinMax Scaling)

Pipeline

1. Data Acquisition

- Time (May-September 2021, 2022, and 2023)
- Manually Selected Region

3. Splitting the Data

2. Preprocessing

- Cloud Percentage Selection
- Feature Engineering (NDVI, NDMI)
- Cropping the area of interest
- Cloud Masking
- Creating reference label
- Normalization (MinMax Scaling)

4. Cross Validation and Hyperparameter Optimization

3. Splitting the Data

3. Splitting the Data

- Create Train/Test Dataset

Pipeline

1. Data Acquisition

- Time (May-September 2021, 2022, and 2023)
- Manually Selected Region

2. Preprocessing

- Cloud Percentage Selection
- Feature Engineering (NDVI, NDMI)
- Cropping the area of interest
- Cloud Masking
- Creating reference label
- Normalization (MinMax Scaling)

3. Splitting the Data

- Create Train/Test Dataset

4. Cross Validation and Hyperparameter Optimization

Results

Metrics	Train	Test
Precision	0.98	0.97
Recall	0.99	0.98
F1	0.99	0.98
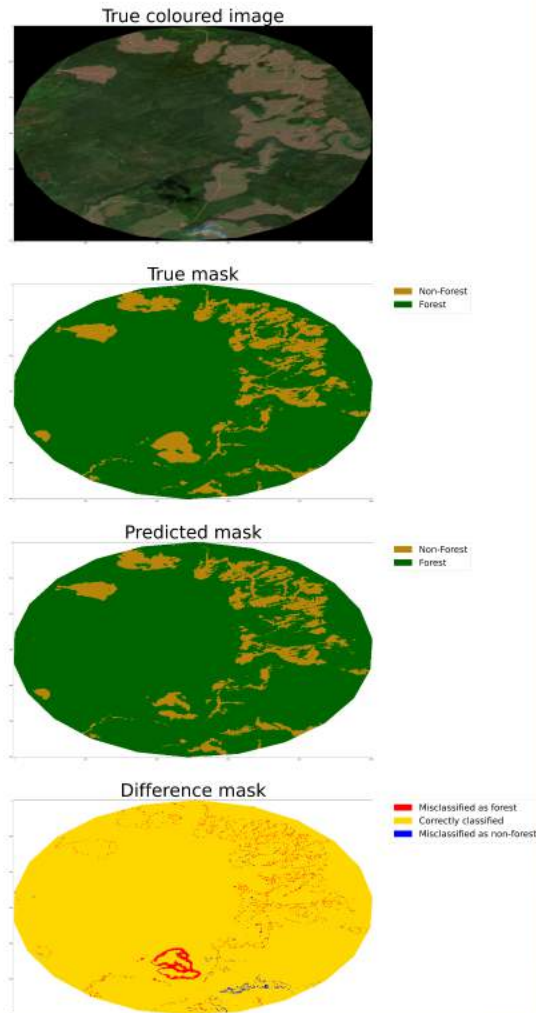
Validation Table: Train size: 20% (3,896,557). Test size: 80% (15,586,228)

Feature	NDVI	RED	BLUE
Feature Importance	28%	14%	12%

Feature Importance Table

Visualization of the results

Location: 53°26'50"N 124°38'13"W (3km radius)
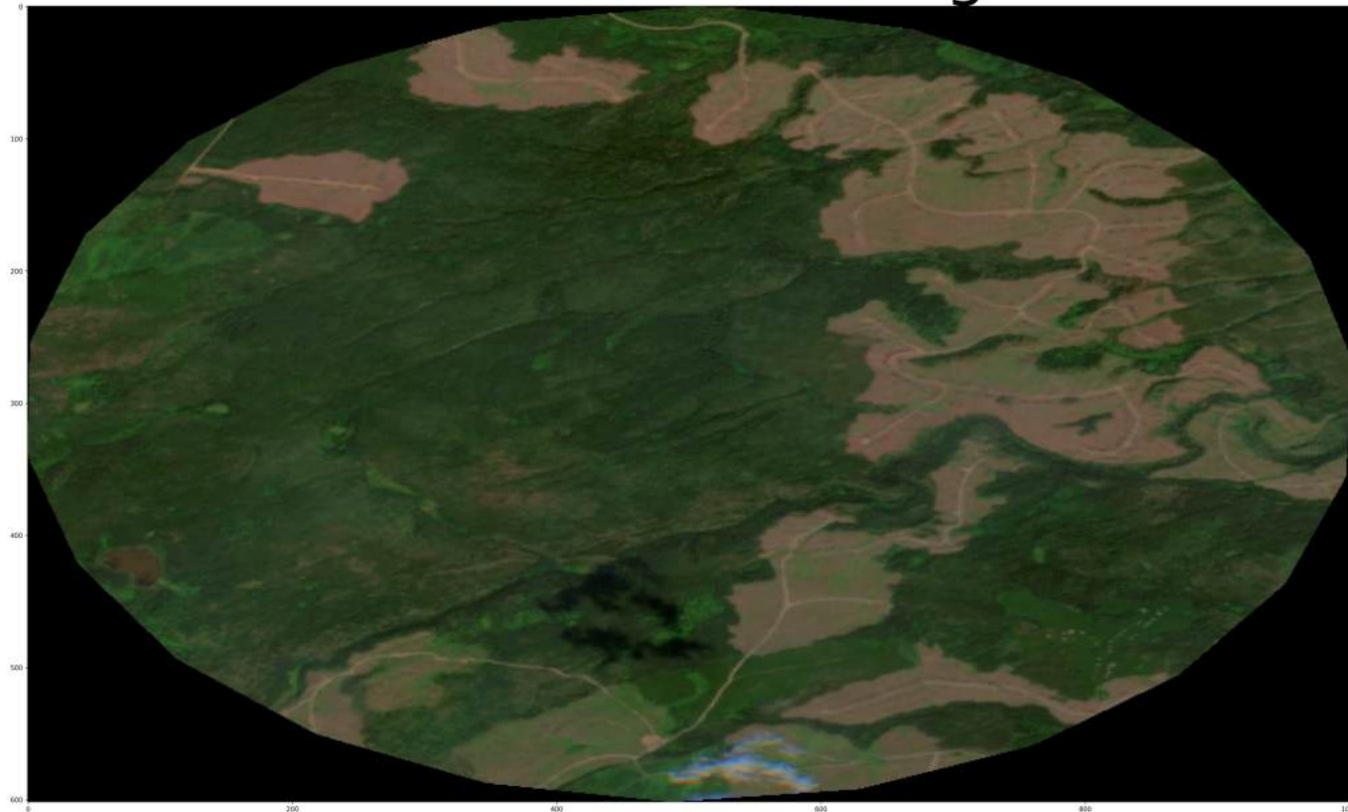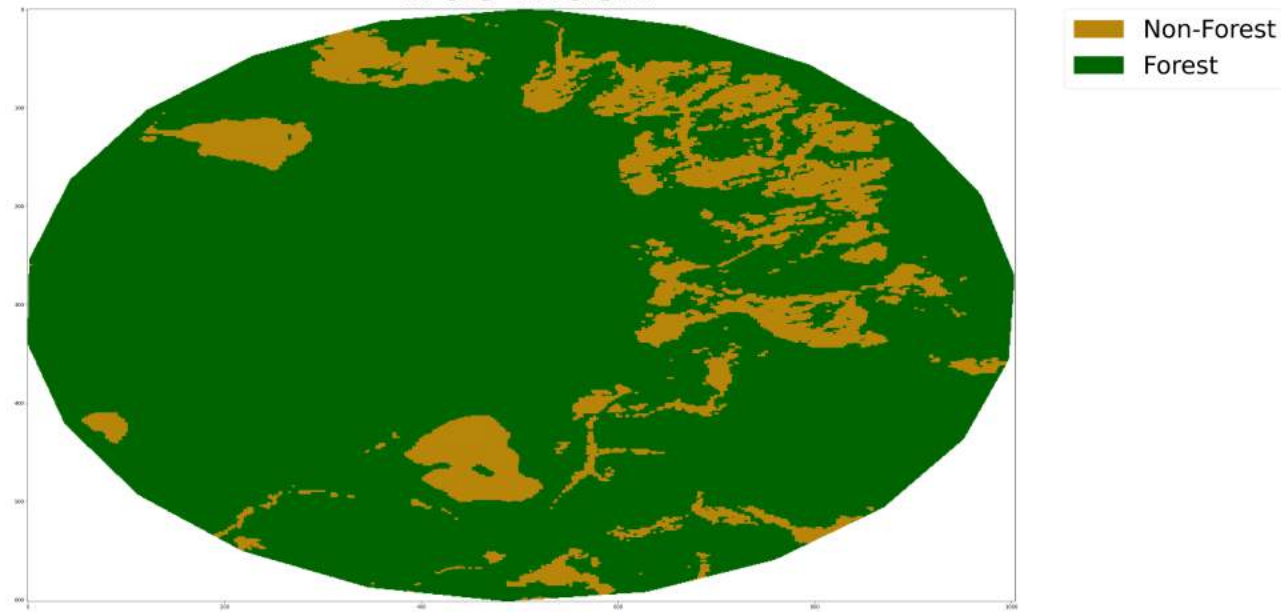Acquisition date: 07/20/2023

Location: 53°26'50"N 124°38'13"W (3km radius)
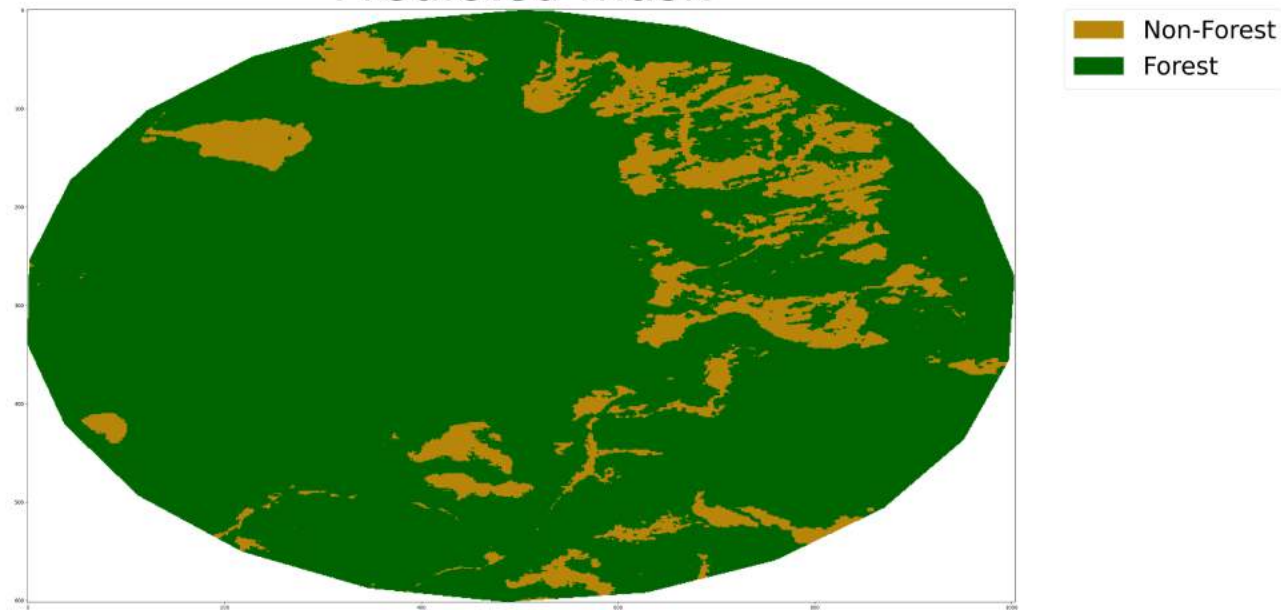Acquisition date: 07/12/2023

Location: 53°26'50"N 124°38'13"W (3km radius)
Acquisition date: 09/10/2023

Location: 53°26'50"N 124°38'13"W (3km radius)
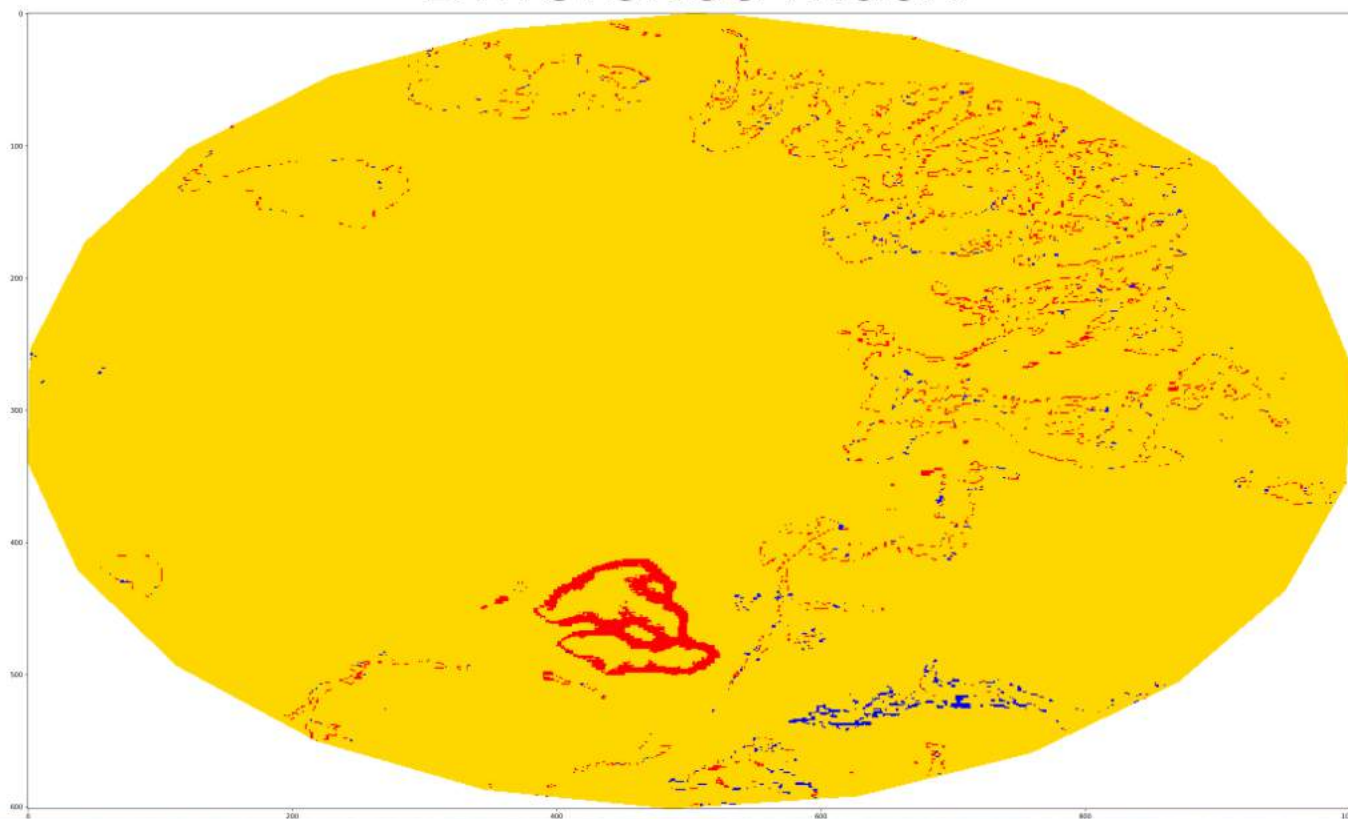## Acquisition date: 07/20/2023

True coloured image

True mask

Non-Forest
Forest
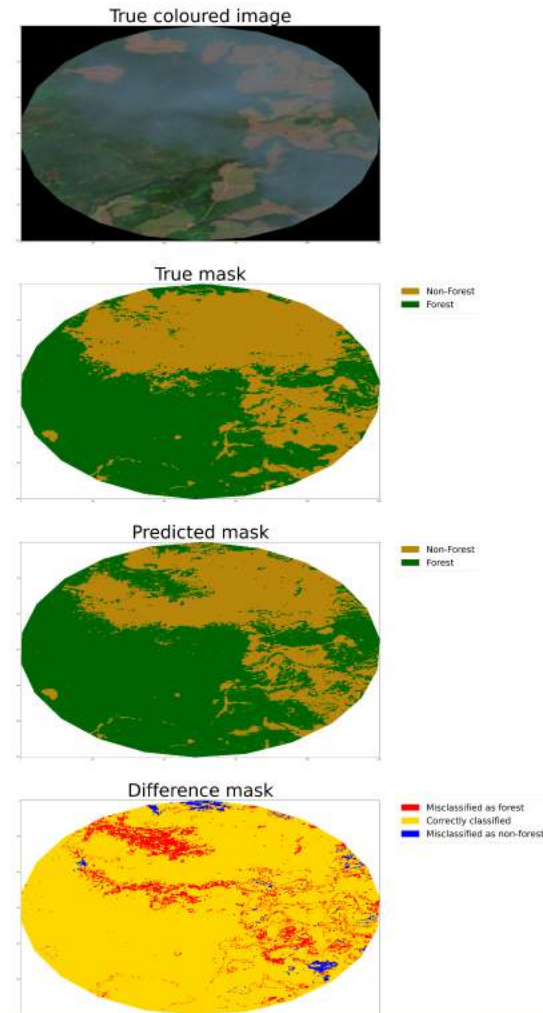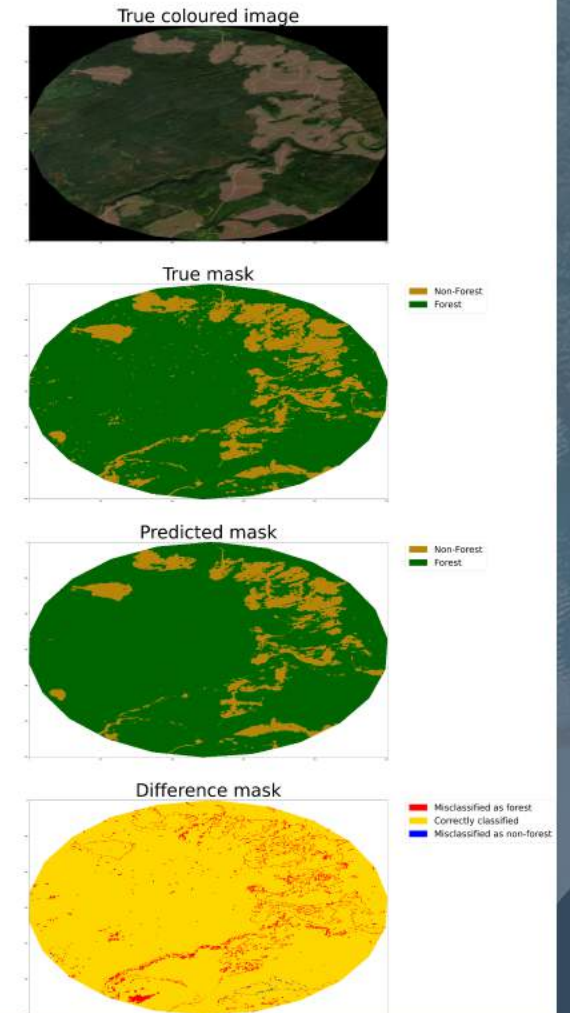
Predicted mask

Non-Forest
Forest

Difference mask

Misclassified as forest
Correctly classified
Misclassified as non-forest

Visualization of the results

Location: 53°26'50"N 124°38'13"W (3km radius)
Acquisition date: 07/20/2023

Location: 53°26'50"N 124°38'13"W (3km radius)
Acquisition date: 07/12/2023

Location: 53°26'50"N 124°38'13"W (3km radius)
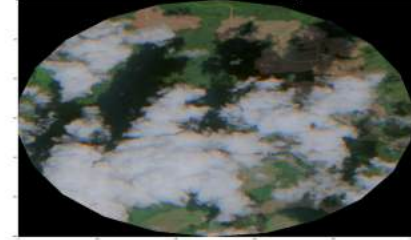Acquisition date: 09/10/2023
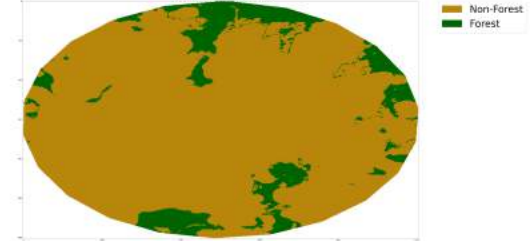
Possible Improvements/Future Works

- Forest Change Detection
- Convolutional Neural Networks (CNN)
- Satellite Integration
  - Landsat 8/9
  - MODIS
- Improve Cloud Cover
- Classify to human/nature-based deforestation

Location: 53°26'50"N 124°38'13"W (3km radius)
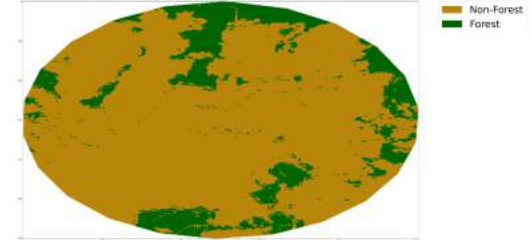Acquisition date: 06/27/2023

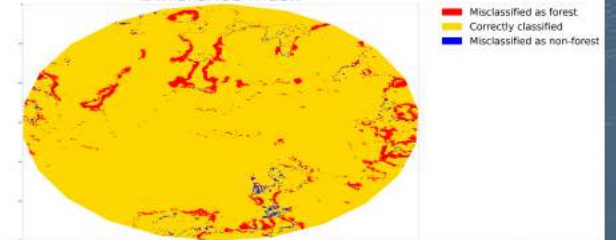# Thank you for your attention!