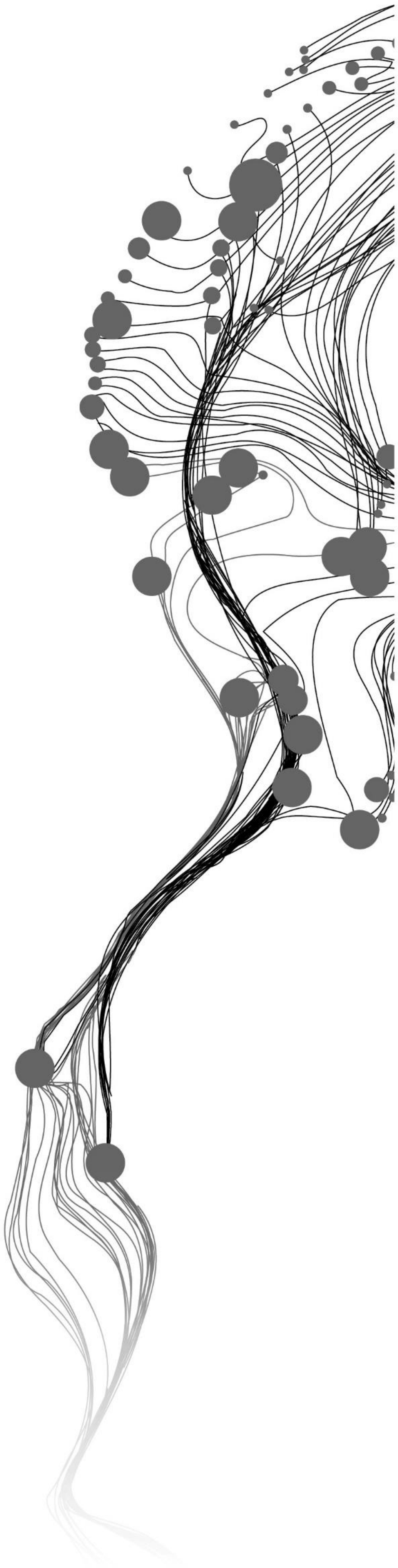
Deep Learning based Multi-Source Data Fusion to Map Deforested Areas in Amazon Rain Forest

AKHIL CHHIBBER
JUNE, 2022

SUPERVISORS:
Dr. Raian Vargas Maretto, ITC
Dr. Mariana Belgiu, ITC

EXTERNAL ADVISOR:
Dr. Sugandh Chauhan, Space4Good



Deep Learning based Multi-Source Data Fusion to Map Deforested Areas in Amazon Rain Forest

AKHIL CHHIBBER
Enschede, The Netherlands, June 2022

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geo-Informatics

SUPERVISORS:
Dr. Raian Vargas Maretto, ITC
Dr. Mariana Belgiu, ITC

EXTERNAL ADVISOR:
Dr. Sugandh Chauhan, Space4Good

THESIS ASSESSMENT BOARD:
Prof. Dr. ir. A. Stein (Chair)
Dr. Marcos Adami (External Examiner, INPE

ABSTRACT

Accurate deforestation mapping can provide useful information for efficient forest management. Frequent cloud cover often hampers deforestation mapping in tropical forests when only using an optical image. Since optical remote sensing is ineffective in cloudy weather conditions, a possible alternative is the use of all-day and all-weather Synthetic Aperture Radar (SAR). This study aims to overcome this limitation of the optical image by fusing optical (Sentinel-2) and SAR (Sentinel-1) images. With that, we aim to improve deforestation detection through Deep Learning (DL) based late fusion, using as a test site an area in Pará State, Brazil. We compared the accuracies of the deforestation maps generated for the year 2020 from standalone optical and SAR images with maps predicted using late fusion which includes both Sentinel-1 and Sentinel-2 sensor data as input. Results showcased that deforestation mapping using the combination of optical and SAR sensor data improved the overall classification accuracy which was also verified using McNemar's statistical significance test. For cloud-free image, Sentinel-1/Sentinel-2 based late fusion provided an overall accuracy of 0.97, 0.94, and 0.91 on the full image, test set-1, and test set-2 respectively, while in the cloudy image, Sentinel-1/Sentinel-2 based late fusion provided an overall classification accuracy of 0.95, 0.91 and 0.88 respectively. In the case of a cloud-free image, the overall accuracy using Sentinel-1/Sentinel-2 based late fusion was +3%, +3%, and +3% higher for full image, test set-1, and test set-2 respectively than Sentinel-1 image and +2% and +1% higher for full image and test set-2 than Sentinel-2 image. In case of cloudy weather condition, the overall accuracy of late fusion using both Sentinel-1/Sentinel-2 image was +1%, +2% and +1% higher for full image, test set-1 and test set-2 respectively than Sentinel-1 image and +10%, +2% and +10% respectively higher than Sentinel-2 cloudy image. The presented approach using late fusion showed the advantage of fusing Sentinel-1 and Sentinel-2 sensor data for deforestation mapping compared to the standalone data source. Also, the results show significant benefits of fusing both Sentinel-1 and Sentinel-2 images even in the cloudy weather condition where 22-48% of the study area was covered with clouds in Sentinel-2 data.

Keywords: Sentinel-1, Sentinel-2, Deforestation, Late Fusion, Semantic Segmentation, Deep Learning

ACKNOWLEDGEMENTS

To start with, I would like to acknowledge the Faculty of Geo-Information Science and Earth Observation (ITC) for providing me with ITC Excellence Scholarship which made my journey for my M.Sc. with Specialization in Geo-Informatics possible for me. I would like to convey my thankfulness to both my first and second supervisors Dr. Raian Vargas Maretto and Dr. Mariana Belgiu respectively who encouraged me and helped me get back on track when my Thesis progress went a little slowed down during my journey.

I would also like to be grateful to Space4Good for offering me a chance to pursue an Internship with them and also for supporting me with a possibility of a Thesis Collaboration with such a talented Advisor, Dr. Sugandh Chauhan who helped me out with her expertise in Synthetic Aperture Radar (SAR) for mapping deforestation.

As this M.Sc. was quite a long journey of 2 years, I am highly indebted to my beautiful family including my loving parents and my brother for believing in me throughout my journey and providing me with unconditional support whenever I was feeling discouraged.

This M.Sc. was not just a result of a few people but enormous support from hundreds of individuals and volunteers in the Deep Learning and Remote Sensing community, as well as our ITC Alumni who helped me in gaining skills in Machine Learning and also helped me whenever I got struck. In the end, I would like to thank every individual who made this long journey of two years a lifetime experience for me which is going to be a forever memory.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background and Problem Statement

The availability of forests on our planet is essential for the lives of all human beings (Lee et al., 2020). Among all the reasons, one of the most important contributions of forests in mitigating climate change through the exchange of carbon dioxide with oxygen. They also help in preventing soil erosion and protecting the riverine ecosystems. Despite the benefits of forests, even now, deforestation activities are taking place on a large scale. Researches have estimated that deforestation is the second leading cause of emission of Greenhouse Gases (GHG) after emissions caused by the burning of fossil fuels (de Bem et al., 2020). Bem et al. (2020) estimated that this continuity of deforestation activities can lead to a decrease in seasonal rainfall globally. This makes early detection of deforestation caused by human-induced activities essential (Ortega et al., 2019).

The Amazon rainforest occupies approximately 5.5 million km² (Andrade et al., 2020), contains half of all the tropical forests in the world, and generates more than 20% of the oxygen in our planet (Ortega et al., 2019). For decades, the Amazon biome has been facing deforestation caused by human activities, like forest fires, illegal logging, the development of informal settlements, etc. World Wildlife Fund mentioned that, if the current amount of deforestation in the Amazon continues, by 2030 more than one-fourth of its forest will get vanished. This makes it urgent to develop policies to protect its forest, considering the necessary resources it provides for the protection and preservation of our planet (Ortega et al., 2019). Therefore, it is essential to have regular deforestation maps with a superior level of accuracy to help in formulating public policies for battling the fight against deforestation (R. V. Maretto et al., 2020).

Looking at the issue of combatting illegal deforestation activities, the Brazilian government has developed some initiatives for monitoring deforestation with the Brazilian National Institute of Space Research (INPE) (Ortega et al., 2019). *Firstly,* the Amazon Deforestation Monitoring Project namely the Program for monitoring deforestation through satellite imagery (PRODES) (Amazônia, 2022) started in 1988 the estimation of the annual deforestation rates. *Secondly,* Near Real-Time Deforestation Detection (DETER) (INPE, 2019), was initiated to inform and reinforce the actions of controlling the illegal deforestation activities, for helping in maintaining policies. The aim of PRODES is to map the anthropic disturbances caused through the clear-cut of primary forests in the Amazon biome in the regions covered by vegetation forest formations as defined by the classification proposed by RADAM, (1978). PRODES map includes four classes (i) deforestation, (ii) non-forest, (iii) hydrography (iv) forest (INPE, 2019). Forest class includes all areas of the primary forest as defined by RADAM, (1978). Water bodies and rivers are part of the "hydrography" class, and the class "deforestation" includes areas of primary amazon forests which were deforested by clearcut having areas bigger than 6.25 ha. As the main goal of PRODES is to detect the disturbances in primary forests, the areas which were deforested in past and abandoned, presenting a regeneration of vegetation and became secondary forests, they are still identified under the "deforestation" class. The "deforestation" class is divided into sub-classes based on the deforestation which is detected every year, since the inception of this project. Lastly, the "non-forest" class includes all the other areas that do not belong to "forest", "hydrography", and "deforestation".

To come up with the final training image based on PRODES, the three classes namely "hydrography", "non-forest" and "forest" were categorized together as "non-deforestation". Every year based on the temporal series, the "deforestation" class is prepared including all the deforested areas until that year. In this research, for producing the labels of deforestation for the year 2020, the areas mapped under the class "deforestation" up to the year 2020 were counted as deforestation. It represents the accumulated disturbances in primary forests from 1988 until the year 2020. The drawback of these projects is their dependency on well-trained experts for manually mapping deforested areas based on their expertise, making them dependent on human interpretation (Adarme et al., 2020). This dependency makes deforestation detection costly and, time-consuming (R. V. Maretto et al., 2020). Hence, there is a demand for automatic methods for the detection of illegal deforestation, which may reduce the need for specialists and consequently the costs of the process (Andrade et al., 2020).

Recently, studies have shown that Deep Learning (DL) has achieved a strong potential in several fields, and also becoming state-of-the-art in the remote sensing community (de Bem et al., 2020). It shows countless possibilities, including image fusion, semantic segmentation, object detection, classification of land cover and land use, etc. In DL, it is possible to express various levels of representations of data to extract abstract and robust features which provide significant information (Ortega et al., 2019). Studies related to mapping deforestation using DL have accomplished a classification accuracy of as high as 95% (R. V. Maretto et al., 2020). Most studies on deforestation detection using DL algorithms, like Isaienkov et al. (2021), Maretto et al. (2020), Shumilo et al. (2020), Matosak et al. (2020), Ortega et al. (2019), Andrade (2020) and de Bem et al. (2020) have been based on using only optical sensors. As stated by Adarme et al., (2020), some regions in the Amazon are covered by clouds for almost the entire year, which restricts the use of optical sensors. The annual mean cloud cover of the Brazilian Legal Amazon (BLA) is roughly 74% (Doblas et al., 2020). This limits the use of optical sensors for deforestation detection in the Amazon. Synthetic Aperture Radar (SAR), on the other hand, is a type of sensor which has an advantage of an all-day and all-weather capability to deliver land information, differently from what can be extracted using optical sensors (Nicolau et al., 2021). This includes moisture, structure, etc. SAR sensors have presented then a good potential for deforestation monitoring, alerts, and preparation of land use land cover maps.

Optical sensors deteriorate from consequences related to illumination, shadows, and clouds (Hughes et al., 2020). Likewise, SAR images also face challenges while classifying different applications of land (Joshi et al., 2016). Some challenges refer to geometric deformations including foreshortening, microwave-radar shadow, and layover (Hughes et al., 2020). One problem faced in all SAR imagery is the occurrence of speckle noises, which increase the ambiguity in measurements and also result in reduced overall accuracy, which may need to be pre-processed by using speckle reduction filters (Joshi et al., 2016). In brief, the individual limitations of SAR and optical data pose challenges but, most of these limitations do not overlap between both the sensors (only exception of topography which is a weakness in both). Because of this, complementarity is possible and both sensors can compensate for their limitations. Therefore, the synergy of the information extracted from both sensors can be utilized for improving deforestation detection.

Image fusion is defined as the *"combination of two or more different images to form a new image by using a certain algorithm"* (Belgiu & Stein, 2019, p. 2). Optical-SAR data fusion has grown in remote sensing (Meraner et al., 2020). In image classification, convolutional neural networks (CNN) directly extract features from Optical and SAR images (Han et al., 2021). However, currently, the state-of-the-art DL-based architecture used for detecting deforestation use only single data sources, as shown by Wahab et al. (2021), Hasret et al. (2018), Isaienkov et al. (2021), Maretto et al. (2020), de Bem et al. (2020), Andarme et al. (2020) and Lee et al. (2020). As mentioned by Adrian et al., (2021), combining the different yet complementary information of Optical and SAR images is an encouraging direction for DL in remote sensing. This fusion of complementary features has also been explored in the literature for many applications including crop-type mapping (Adrian et al., 2021), coastal wetlands monitoring (Wu, 2021), and wildfire monitoring (Rashkovetsky et al., 2021), etc. This study proposes a late fusion-based DL model which will be detailed out in Chapter 2 and also in Chapter 4. This late fusion model fuses two standalone DL models (using Sentinel-1 and Sentinel-2 data as input) for performing multi-modal data fusion for deforestation mapping.

1.2. Objectives and Research Questions

Main Objective
To exploit the complementarity of Optical and SAR sensors for mapping deforested areas through Deep Learning based data fusion.

Sub-Objectives
- To design a multimodal deep learning algorithm that explores complementarity among Sentinel-1 and Sentinel-2 images for mapping deforested areas.
- To compare the classification results obtained using data fusion and using standalone Sentinel-1 and Sentinel-2 images.

Research Questions
- To what degree can SAR data support deforestation mapping when atmospheric conditions affect the optical image?
- To what extent can the fusion of Optical-SAR data improve deforestation mapping relative to using standalone sensor images?

1.3. Hypothesis

Fusion of Optical-SAR data will improve deforestation mapping in comparison to standalone SAR and Optical data.

1.4. Thesis Structure

This Thesis report is divided into six chapters which are structured as follows:

Chapter 1: Introduction
This chapter details the background information and also the justification of why this research is essential, clarifying what is the research problem, objectives, and research questions.

Chapter 2: Literature Review
This chapter includes the details of previous research work conducted for similar applications along with the conceptual explanation of the methodologies and technologies used in this research for Multi-Modal Data Fusion.

Chapter 3: Study Area and Images
This chapter introduces our study area used for deforestation mapping and images that have been used in this research for performing multi-modal data fusion.

Chapter 4: Methodology
This chapter details the methodology followed to accomplish the research objectives and answer the research questions. It gives information about the data pre-processing and methods used for analysis and evaluating the models.

Chapter 5: Results and Discussion
This chapter depicts the results of deforestation using a standalone image as well as multi-modal data fusion on a cloudy and cloud-free image, followed by a detailed discussion on the results obtained.

Chapter 6: Conclusion and Recommendation
This chapter completes the research with some closing remarks on the entire research, its outcomes, and some possible future directions that this study can take.

2. LITERATURE REVIEW

2.1. Image Fusion

Image fusion methods can be clustered into three main categories (Figure 1), namely, pixel-level fusion, feature-level fusion, and decision-level fusion, as stated by Joshi et al., (2016). In pixel-level fusion, original pixels of different images are directly fused, while in feature-level fusion, features obtained from individual sensors are fused. Lastly, in decision-level fusion, the output of classification completed by individual sensors is fused. Decision-level fusion (Figure 1) deals with the identity declaration (ID) as was mentioned by Z. Lio et al., (2018). In the case of decision level fusion, the IDs are obtained from multiple input images which would be Sentinel-1 and Sentinel-2 images in our case by performing semantic segmentation using the DL architecture. In the process of fusion, a joint ID is generated from the input IDs. Fusion can be performed in various ways, averaging is one of the easiest forms of late fusion. As mentioned by Mahyoub et al., (2019), pixel-level fusion is considered unsuitable for Optical-SAR fusion because of the high occurrence of speckle noise in SAR data. This noise leads to issues like a layover, shadowing, and foreshortening (Mahyoub et al., 2019). Registration is a process to geometrically align imagery from various sensors, e.g., Sentinel-1 and Sentinel-2 (Kulkarni & Rege, 2020). Optical and SAR data acquired are geo-referenced, but they suffer sometimes from incorrect alignment. The alignment of both sensors is relatively simpler in feature-level fusion (Figure 4) in comparison to pixel-level fusion (R.Pandit & J. Bhiwani, 2015).

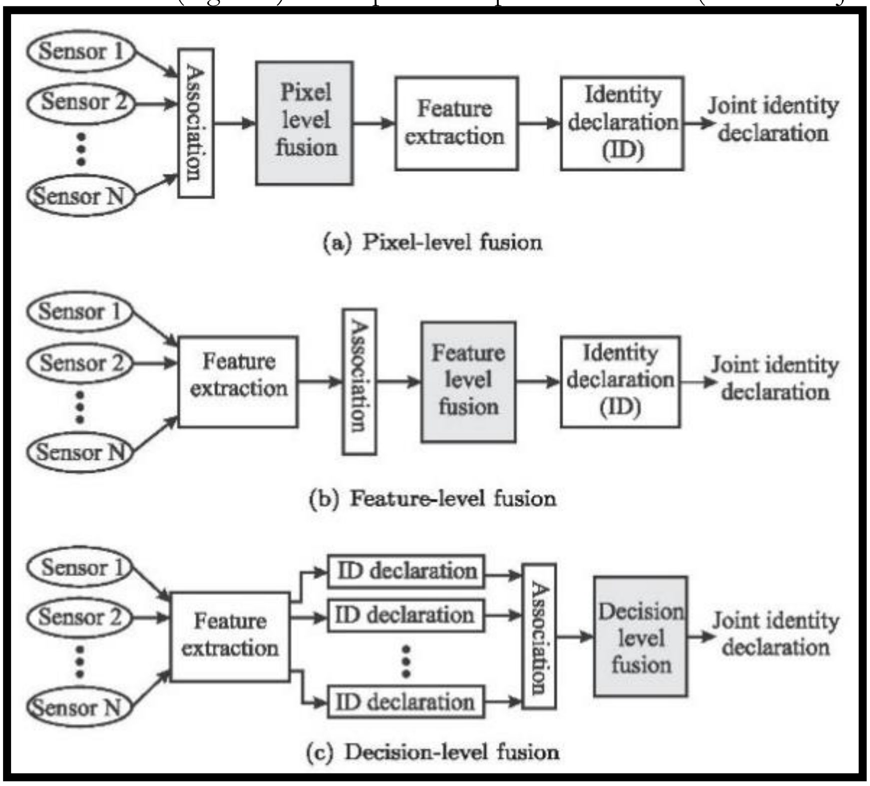


Figure 1: Image Fusion Techniques, (Z. Liu et al., 2018)

One limitation of using the Feature Level Fusion (Figure 4) as was used by Audebert et al., (2018) is that both layers of encoders are supposed to be compatible to fuse the encoders after every convolution. This restricts the DL Architecture in the encoder part to be consistent and use exactly similar architecture. In this research, we proposed a different technique of fusion that depend on the late features in the last decoder. Those networks may or may not be topologically the same. In the case of late fusion, numerous standalone networks are used as the beginning for the preparation of DL-based architecture (Gadzicki et al., 2020). The standalone networks can be heterogeneous like one separate network for optical image and a separate network for SAR image, fitting only the image which they are specifically designed for. The real fusion is then negligibly requiring only the merging of the individual results of each network which was performed by concatenation of convolution layers in the last decoders in this research (Figure 19).

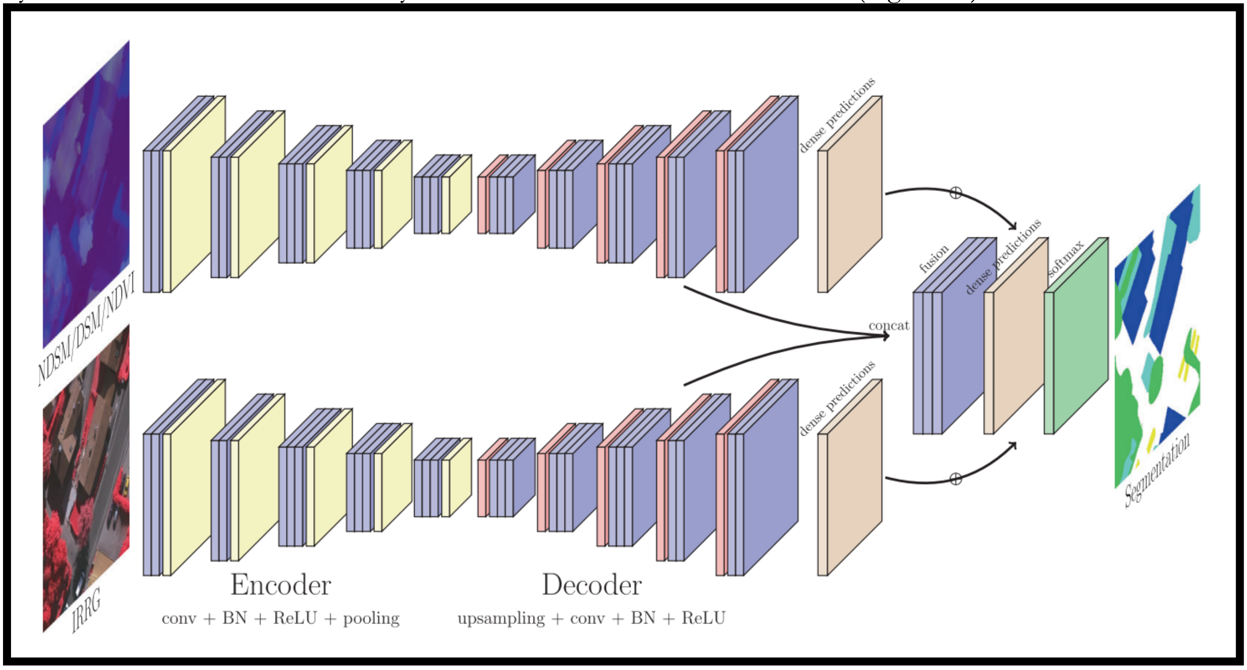


Figure 2: Late Fusion, (Audebert et al., 2018)

2.2. Deep Learning and Fusion

DL has been explored for multi-modal data processing (Audebert et al., 2018). Initially, audio and video data have been fused effectively using two branches in the DL architectures, where one branch was used for audio and the other branch was used for video. Both branches were posteriorly merged in the center of the DL architecture (Ngiam et al., 2011). FuseNet also used by Hazirbas et al., (2017) expanded this concept of fusing two branches to fully convolutional networks (FCN) using Red, Green, Blue, and Depth (RGB-D) data for performing semantic segmentation by using an early fusion method within SegNet Architecture (Figure 3). With the advances in computer vision projects using DL algorithms, the remote sensing community also embraced and utilized these techniques for different applications in Earth Observation. The very first effective patch-based Convolutional Neural Network (CNN) architecture was used within the remote sensing community for the extraction of buildings and roads (Mnih & Hinton, 2010). Vakalopoulou et al., (2015) expanded the method to multi-spectral satellite imagery which contains visible and infrared bands. More recently, a lot of tasks related to semantic segmentation have been shifted to FCN Algorithms (Volpi & Tuia, 2017). This is because FCN models like the SegNet (Figure 3) perform semantic segmentation of satellite imagery. It can also capture the spatial dependencies between different classes, with no pre-processing methods e.g. super-pixel segmentation resulting in the prediction of classes in high resolution. Further, the SegNet architecture is composed of an encoder-decoder structure where feature maps are up-sampled in the decoder which is in line with the encoder to match it back to the input resolution. This results in performing pixel-wise prediction of labels specifically at 1:1 resolution. In the past, the multi-modal fusion of complementary sensors has been explored within the remote sensing community to determine different properties of an area in the semantics being segmented (Audebert et al., 2018). Paisitkriangkrai et al., (2015) utilized LiDAR and Optical data by fusing features using the Random Forest algorithm. Liu et al., (2017) integrated features from auxiliary images which included LiDAR and NDVI data along with their higher-order Conditional Random Field (CRF) for improving the classification of optical data in DL networks.

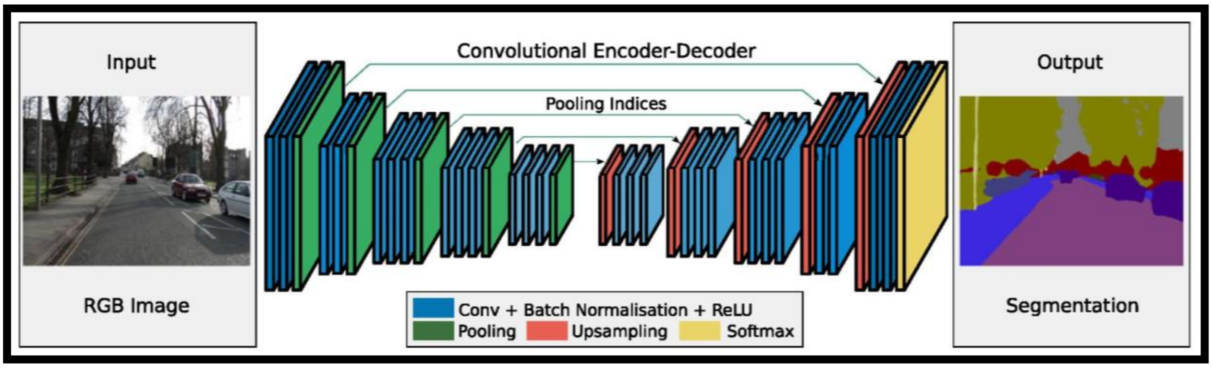


Figure 3: SegNet Architecture, (Badrinarayanan et al., 2017)

Further, Audebert et al., (2017) explored the late fusion (Figure 2) technique for fusing optical and LiDAR data for performing semantic segmentation through fusing the results of standalone classifiers. The use of a multi-modal network was also explored by Audebert et al., (2017) for fusing optical data with OpenStreetMap data, used as auxiliary data, for performing semantic segmentation through the FuseNet Architecture (Figure 4). This multi-modal data fusion has been explored in DL for many remote sensing applications including crop type mapping (Adrian et al., 2021), land cover mapping (Ienco et al., 2019), sea ice classification (Han et al., 2021), wildfire detection (Rashkovetsky et al., 2021), etc. But, it is yet to be explored for deforestation detection using DL with Sentinel-1 and Sentinel-2 images. Considering the works already done in the literature, this study uses a late fusion technique by fusing the features on the last decoder based on the SegNet architecture (Figure 3). This was done to perform a multi-modal data fusion of Sentinel-1 and Sentinel-2 images for mapping deforested areas in the Brazilian Amazon.

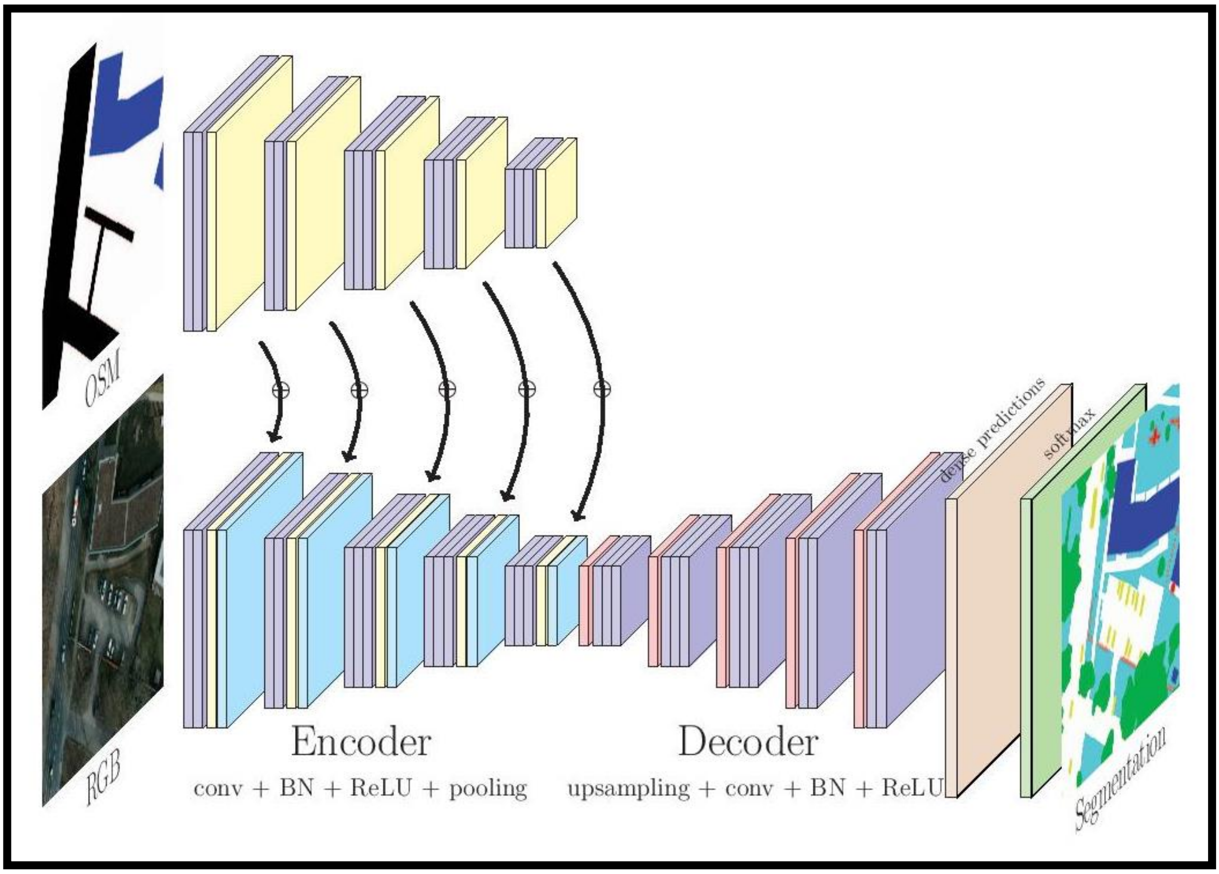


Figure 4: FuseNet using Optical and OSM Data, (Audebert, Saux, et al., 2017)

3. STUDY AREA AND IMAGES

3.1. Study Area

The high occurrence of clouds is an issue for monitoring forests in tropical regions (Andrade et al., 2020). This scenario is also observed in the Brazilian Amazon for nearly an entire year, thereby challenging the use of optical sensor data and making SAR data a favorable possibility (Adarme et al., 2020). In recent decades, within the Brazilian Legal Amazon (BLA), the majority of the deforestation activities have happened within Pará state (152,475 km^2) followed by Amazonas (26,972 km^2), Mato Grosso (146,159 km^2), and Rondônia (61,677 km^2) (Brovelli et al., 2020). Pará State (Figure 5) shows a variety of locations with varying characteristics, including coastal zones, protected areas, and riverside areas (R. V. Maretto, 2020). Because of the active deforestation, presence of protected areas, and also the high frequency of cloud coverage, Pará state seemed to be appropriate for this research. The area selected is positioned on coordinates of 03°17’23” South and 050°55’08” West (Carrero et al., 2020). Within Pará State, a subset was selected based n the availability of both cloudy and cloud-free Sentinel-2 images. The subset is a square of dimensions 112.64 x 112.64 km. The total deforested area within the subset is 4925.14 km^2, from 1988 until the year 2020 based on data from PRODES.

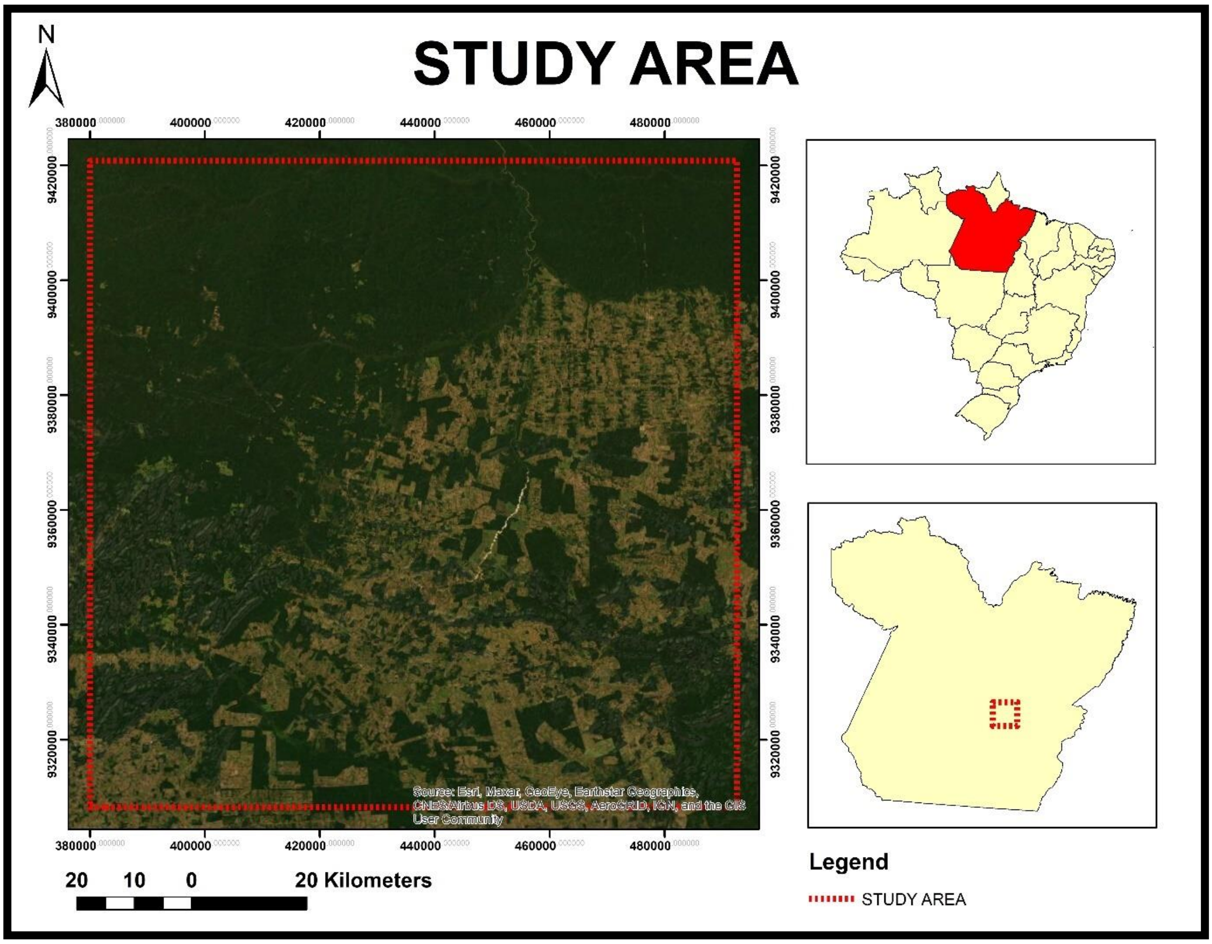


Figure 5: Selected Study Area within Para State

3.2. Images

3.2.1. Sentinel-1

This study used freely available data from the Sentinel-1 platform, which comprises 2 parallel satellites namely Sentinel-1A and 1B (Kumar, 2021) released on the 3rd of April 2014 and the 25th of April 2016 respectively (Shrestha, 2018). Both satellites carry a C-Band SAR sensor providing images between the 2 different orbits. This reduces the temporal resolution to 6 days, instead of 12 days for each satellite individually. It flies at an altitude of 693 km with a swath width of 400 km (Kumar, 2021). It delivers single and dual polarization data in C-Band in two specific formats namely Single Look Complex (SLC) and Ground Range Detected (GRD). Sentinel-1 collects data in 4 acquisition modes with different swath and spatial resolutions with a single (Horizontal Horizontal (HH) or Vertical Vertical (VV)) and dual-polarization (Vertical Vertical (VV) + Vertical Horizontal (VH) or Horizontal Horizontal (HH) + Horizontal Vertical (HV)) (Shrestha, 2018). Table 1 shows the characteristics of the four modes. Further, Sentinel-1 data has three levels: level-0 raw data, level-1 geo-referenced time-tagged data (Single Look Complex, Ground Range Detected), and level-2 ocean-use data (European Space Agency, 2012). For this research, Sentinel-1 Ground Range Detected (GRD) with Interferometric Wide Swath Mode (IW) was used as also used in literature by Wahab et al., (2021) and Hasret et al., (2018) for mapping deforestation. This image was selected for two specific dates including 22nd July 2020 (Figure 6a) and 26th October 2020 (Figure 6b), which were the closest Sentinel-1 images available, to match with the cloud-free and cloudy Sentinel-2 images respectively. To avoid any confusion, from now on in this research, we will use the terminology cloud-free Sentinel-1 image for the image acquired on 22nd July 2022 when weather condition was cloud-free in the Sentinel-2 image. Similarly, we will use the terminology cloudy Sentinel-1 image, for the image acquired on 26th October 2022 when weather condition was cloudy for the Sentinel-2 image.

Table 1: Characteristics of 4 Acquisition Modes, (European Space Agency, 2012)

Modes	Properties
Strip map Mode (SM)	80 km Swath
Interferometric Wide Swath Mode (IW)	240 km Swath
Extra Wide Swath Mode (EW)	400 km Swath
Wave Mode (WV)	20 km * 20 km Vignettes

Data Pre-Processing

Pre-Processing steps were implemented for Sentinel-1 sensor data in the Sentinel Application Platform (SNAP), developed by ESA. The specific pre-processing steps are shown as follows:

1. **Apply Orbit File:** In SAR source products, the orbit information is generally inaccurate. Thus, the accurate orbit information auto-downloaded by SNAP needs to be applied.
2. **Thermal Noise Removal:** During the acquisition of the Sentinel-1 satellite image, the background energy created by the imaging received instruments gets incorporated as thermal noise in the backscatter signals of radar (Phuntsho, 2020). SAR products are highly influenced by thermal noise, specifically in the case of cross-polarization. The thermal noise removal step helps to reduce thermal noise.
3. **Radiometric Calibration:** Calibration is a process that transfers the digital pixel values to calibrated SAR backscattered signals. In this step, backscatter signals are saved in Sigma format.
4. **Speckle Filtering:** Speckle noise caused by coherent processing of backscattered signals makes it difficult to interpret images. To reduce the influence of speckles, the “Lee Sigma” filter was applied using a 3 x 3-pixel moving window through the tool “Single Product Speckle Filtering” in SNAP.
5. **Geometric Correction:** Distortions caused by the side-view characteristic of this sensor (overlapping and shadow) may reduce the quality of SAR images. The Range-Doppler method was chosen for image registration.

Further, a median filter of 3x3 kernel was run on the geometrically corrected raster for reducing speckle noise. After this, the images were re-sampled to the spatial resolution of Sentinel-2 raster (10m) and converted to TIFF format for further processing by using it as input in the DL algorithm, and also for its fusion with Optical data.

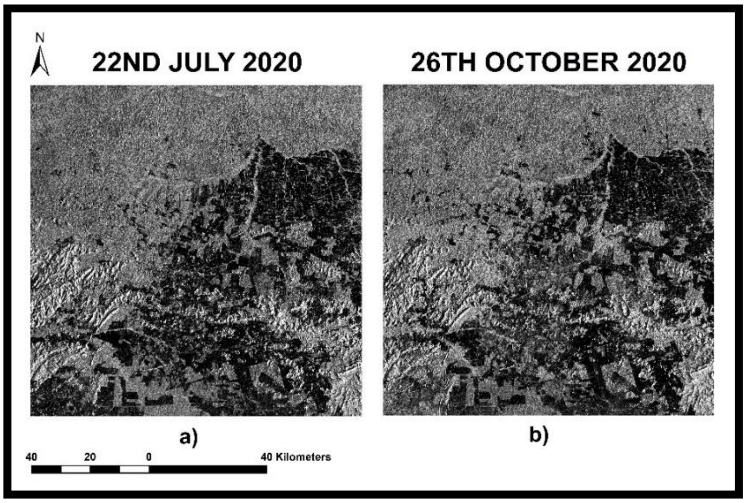


Figure 6: Sentinel-1 VHVV stacked bands for the year 2020

3.2.2. Sentinel-2

The sentinel-2 platform contains two parallel satellites, namely Sentinel 2A and 2B, released on the 23rd of June 2015 and the 7th of March 2017 respectively. It provides us, together, with a revisit time of five days with 10m, 20m, and 60m spatial resolution and 13 spectral bands (Table 2), available from the visible spectrum to short-wave infrared (European Space Agency, 2015). Both images were taken from the open-source Copernicus hub (Sentinelhub, 2022) by the European Space Agency (ESA). Though our SegNet-based DL architecture was able to absorb all the available Sentinel-2 bands. In this research, the highest spatial resolution (10m) bands were used including Red, Green, Blue, and NIR to map the deforested areas. These 4 bands were also selected by Pacheco et al., (2022), Torres et al., (2021), and John et al., (2022) for deforestation mapping. Also, studies related to deforestation mapping using Landsat-8 usually use all bands as also used by Torres et al., (2021). However, due to the larger size of Sentinel-2 images compared to Landsat-8, the processing time taken for Sentinel-2 was higher than Landsat-8. Further, Sentinel-2 images are nine times larger than Landsat-8 as was mentioned by Torres et al., (2021). This was also one of the reasons of taking four bands as more bands take more computational resources and is also a limitation of this study. For our study, we downloaded both cloudy image (Figure 7b) dated 25th October 2020 and cloud free image (Figure 7a) dated 27th July 2020. The Area of Interest (AOI) was extracted from different tiles with cloud coverage ranging from 22 to 48%. Since Sentinel-2 Level-2 (S2L2) data products remove atmospheric errors and provide Bottom of Atmosphere (BoA), the pre-processing on Sentinel-2 was not required (Rahimi, 2020).

Table 2: Sentinel-2 Dataset including Selected four Bands, (European Space Agency, 2015)

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)	Bandwidth (nm)
Band 1: Coastal Aerosol	0.443	60	27/45 (2A/2B)
Band 2: Blue	**0.490**	**10**	**98**
Band 3: Green	**0.560**	**10**	**45/46 (2A/2B)**
Band 4: Red	**0.665**	**10**	**38/39 (2A/2B)**
Band 5: Vegetation Red Edge	0.705	20	19/20 (2A/2B)
Band 6: Vegetation Red Edge	0.740	20	18
Band 7: Vegetation Red Edge	0.783	20	28
Band 8: NIR	**0.842**	**10**	**115**
Band 8A: Narrow NIR	0.865	20	20
Band 9: Water Vapor	0.945	60	20
Band 10: SWIR-Cirrus	1.375	60	20
Band 11: SWIR	1.610	20	90
Band 12: SWIR	2.190	20	180

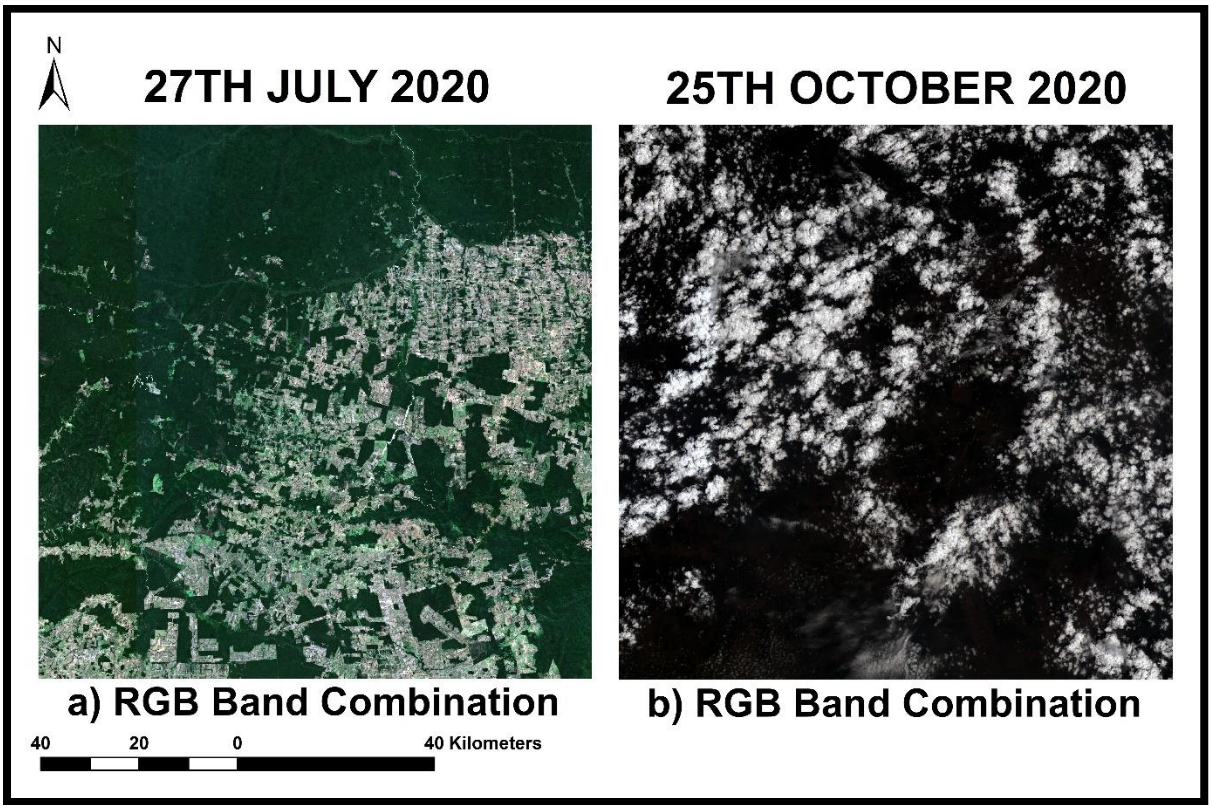


Figure 7: Sentinel-2 cloud free on the left side and cloudy on the right side for the year 2020

3.2.3. Deforestation Data

For deforestation data to be used for training and validation of the late fusion model, images generated by PRODES were used as ground truth until the year 2020. INPE uses satellite data to monitor annual deforestation since 1988, through the PRODES program (INPE, 2019). Together with the DETER program, PRODES has proved to be of great importance for informing public policy actions and planning on amazon (PRODES, 2021). Recent results, from analyzers carried out with independent experts, indicate a level of accuracy of PRODES close to 95%. In PRODES methodology, due to the resolution of the images used and to keep the consistency of the temporal series, the minimum mapping unit is 6.25 ha below which the deforestation activity is not mapped. As the deforestation image was downloaded in shapefile format, it was first converted into raster format using the "Vector to Raster" tool in QGIS (Figure 11). Further, the deforestation data which was available at 30m spatial resolution was re-sampled to 10m to match it with Sentinel-1 and Sentinel-2 images, before feeding in the DL model.

4. METHODOLOGY

4.1. Overall Workflow

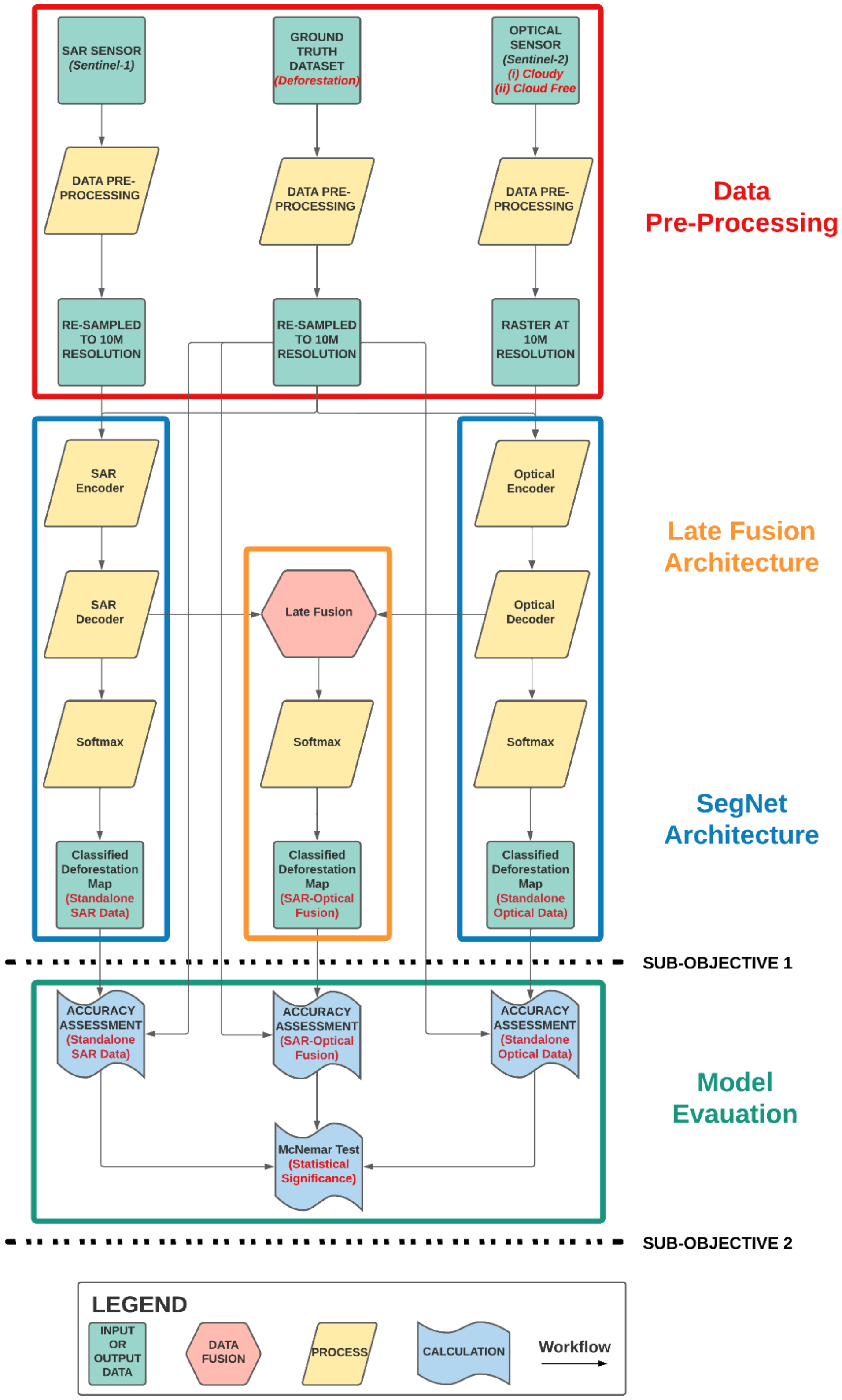


Figure 8: Overall Workflow

The methodology flowchart as can be seen in Figure 8, was divided into three sub-sections namely (i) Data Pre-Processing, (ii) DL Architecture, and (iii) Model Evaluation. The pre-processing steps on Sentinel-1 and ground truth deforestation image are detailed in chapters 3.2.1 and 3.2.3 respectively. Apart from this, the ground truth image was divided into small patches of the size 256 x 256 for using it as an input in the DL architecture. Further, the methodology flowchart includes three separate DL models for deforestation mapping. One Optical-SAR fusion model using late fusion (Figure 19) was proposed for this research and two separate models for standalone SAR and optical data using SegNet Architecture (Figure 14). All the three-segmented maps using Optical-SAR fusion and standalone SAR and Optical sensors were evaluated using accuracy assessment for comparison. To further validate the results of deforestation mapping, between the standalone sensor and late fusion, McNemar's test was performed using Python. This was performed for every scenario including deforestation mapping using standalone sensors, as well as late fusion in both cloudy and cloud-free weather conditions. McNemar's test which is a non-parametric statistical significance test is appropriate to compare the performance of classifications based on machine learning as was mentioned by Mcnemar et al., (1947). The output of this test was used to answer the research questions.

4.2. Data Preparation

Before feeding Sentinel-1, Sentinel-2, and ground truth deforestation images in the DL architecture, data preparation was done to ensure getting the images in a consistent format and spatial resolution.

4.2.1. Ground Truth for Deforestation

Deforestation reference data was taken from PRODES. Since 1988, PRODES have generated maps including yearly deforestation in the region. These maps were then utilized by the Brazilian government to create public policies to fight against deforestation (PRODES, 2021). PRODES uses Landsat images with a spatial resolution of 20 to 30m. In past, PRODES also used other similar satellite images including Disaster Monitoring Constellation (DMC), China-Brazil Earth Resources Satellite program (CBERS), LISS-3 images

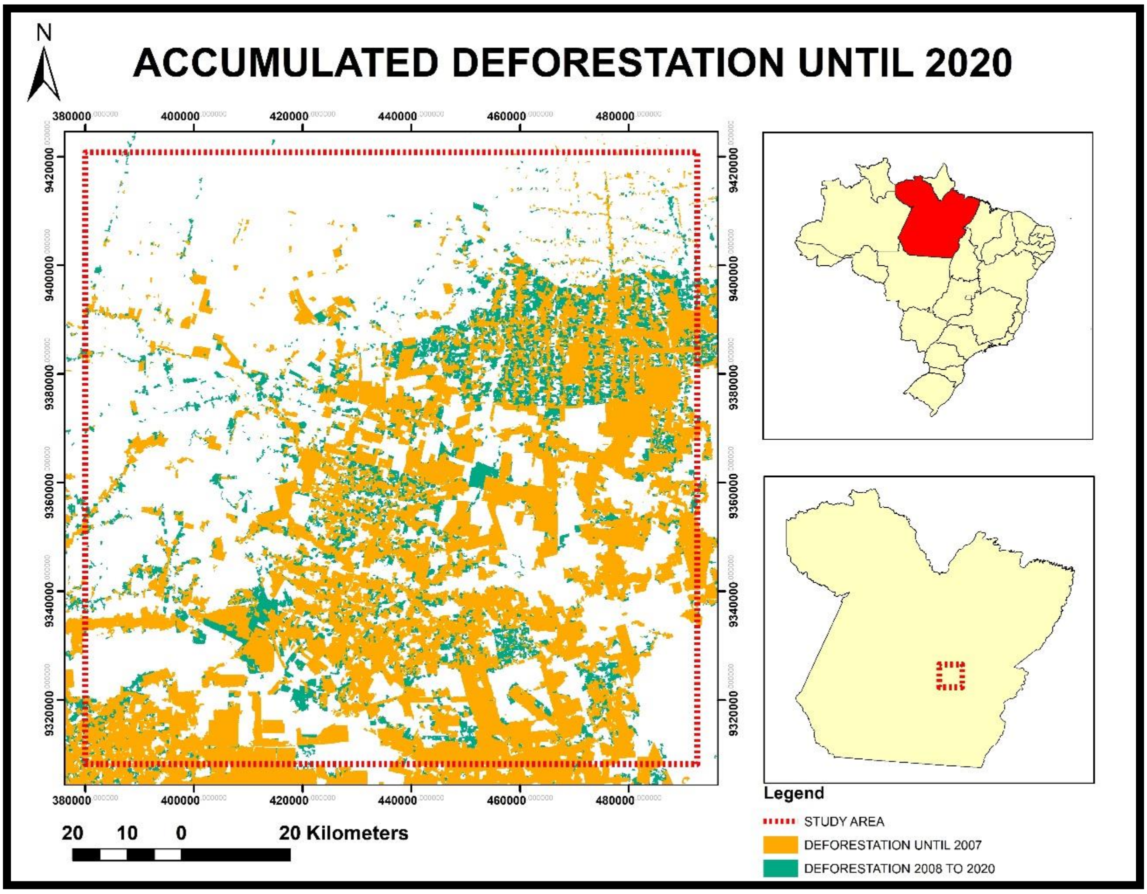


Figure 9: Accumulated Deforestation until 2020

from the Indian satellite IRS-1, etc. (PRODES, 2021) to record and quantify deforested areas which are greater than 6.25 hectares. The data was available as two different shapefiles. The first shapefile included the deforestation from 1988 until 2007 (Figure 9), and the second shapefile included the yearly increase in deforestation from 2008 to 2020 for this study (Figure 9). Both the shapefiles were merged to make one image of deforestation from 1988 until 2020 (Figure 10). Further, this vector image was converted to raster (Figure 12) using the "Rasterize" tool in QGIS (Figure 11) in the "Byte" data type which is acceptable by the DL model. As the spatial resolution of the PRODES image is 30m, it was re-sampled to 10m to make it similar to Sentinel-1 and Sentinel-2 images for using it as input in the DL model.

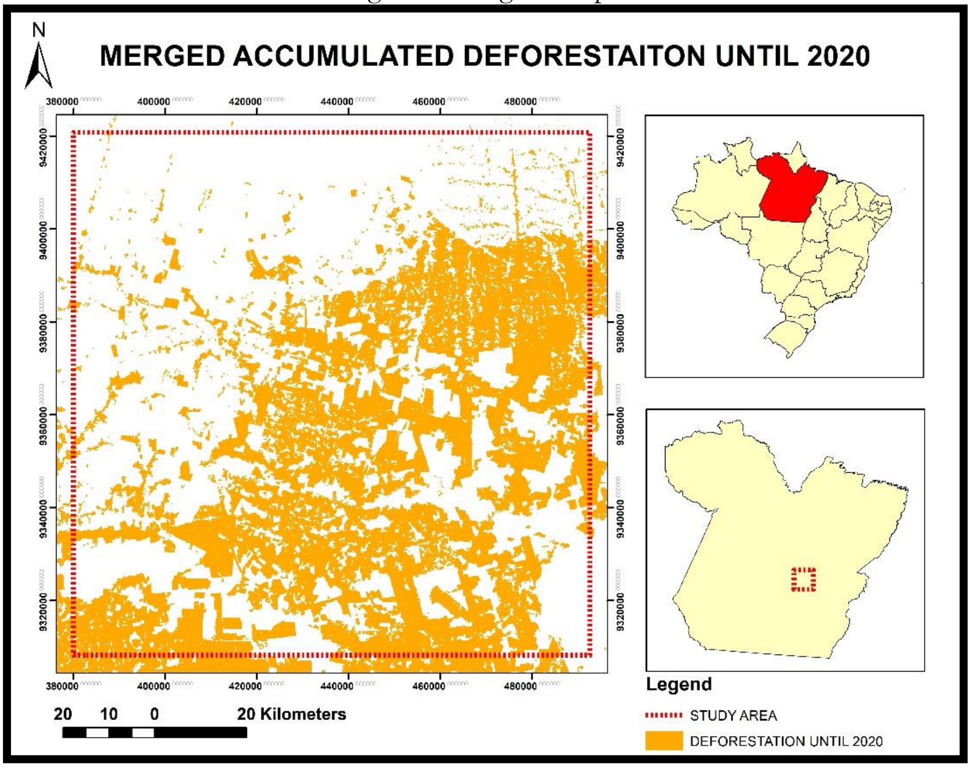


Figure 10: Merged Accumulated Deforestation until 2020

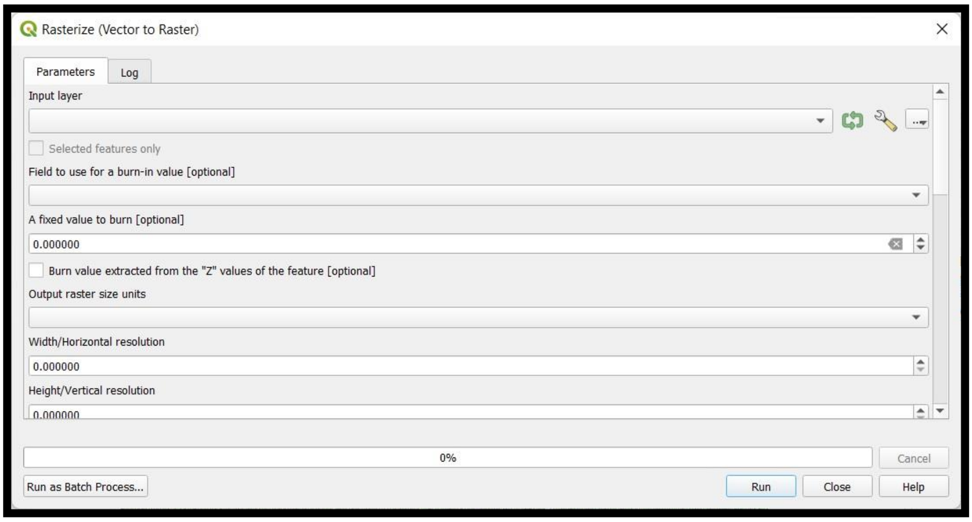


Figure 11: Rasterize Tool

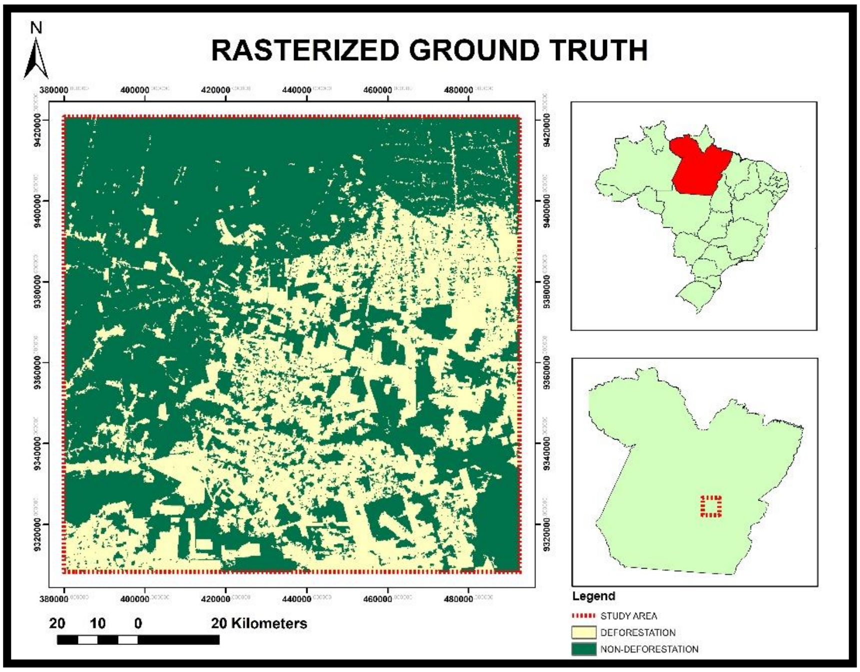


Figure 12: Rasterized Ground Truth

4.2.2. Dividing Study Area into Tiles

The raster image for our AOI was of the dimension 11264 x 11264 pixels which were divided into 16 tiles of 2816 x 2816 pixels each (Figure 13). This was done to get a visual idea of predictions on the test set of size 2816 x 2816 pixels rather than visualizing a small test set of size 256 x 256 pixels. In the study area, 74% of the area was used as a training set, followed by 13% each for the validation set and test set approximately. The validation set was randomly selected using 18% of the training set making it approximately 13% of the entire image. This was to ensure that the proportion of the validation set is somewhere similar to the test set. The 16 tiles of the size 2816 x 2816 pixels were further subdivided into 1936 small patches of 256 x 256 pixels out of which 1439 patches were used for the training set, 255 for the validation set, and 242 (2 tiles) for the test sets.

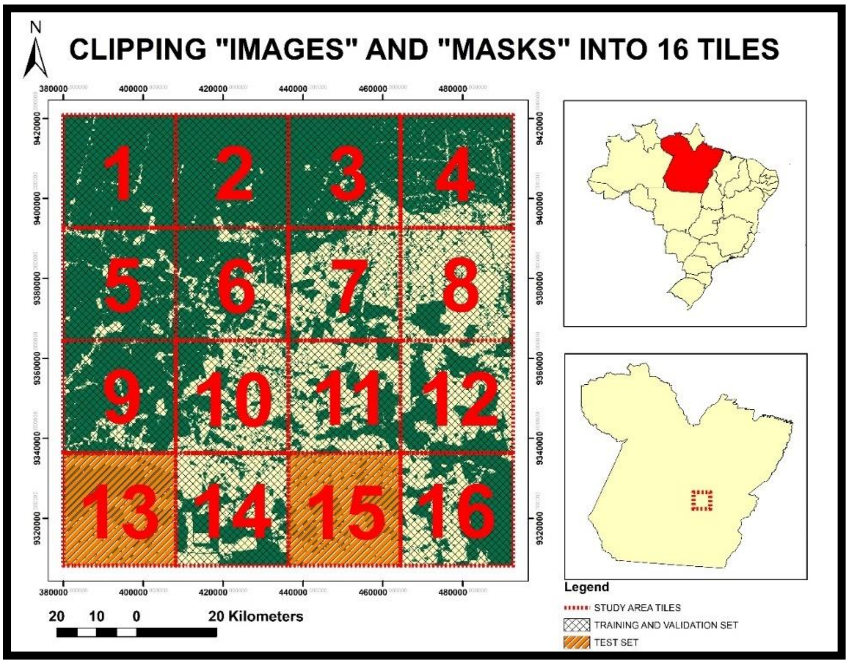


Figure 13: Study Area divided into 16 Tiles

4.3. Deep Learning for Deforestation Mapping

4.3.1. Single Sensor Deforestation Mapping

The input layer in the SegNet architecture, as shown in Figure 14, is the raster which includes raw pixel values of the patches having a dimension of W x H x D. Here, W x H is the patch size and "D" is the depth of that specific patch equivalent to the number of bands e.g. 13 bands in case of Sentinel-2 raster (Musyoka et al., 2018). In this study, for multi-sensor deforestation mapping using late fusion (Figure 19), there are two input layers, one for processing optical data and the other for processing SAR data. For single sensor deforestation mapping using standalone Sentinel-1 and Sentinel-2 images, SegNet (Figure 14), which constitutes the backbones of the late fusion network, was used.

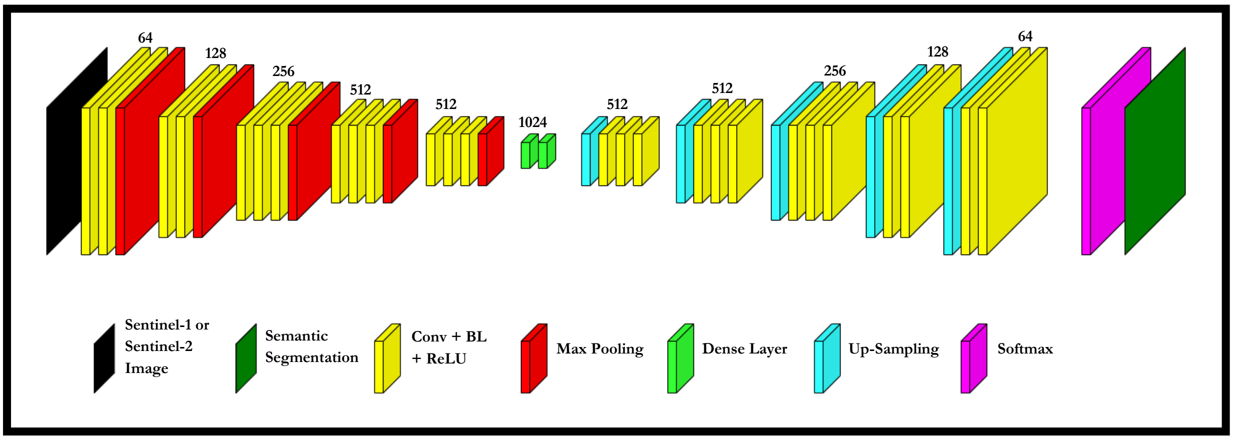


Figure 14: SegNet for Standalone Sensor used in this research

4.3.1.1. Convolutional Layer

The main unit of a Convolutional Neural Network (CNN) architecture is the convolutional layer (Mohammadimanesh et al., 2019). It contains a collection of convolutional filters (Figure 15), where each filter is convolved across the feature maps to generate two-dimensional activation maps (Yan, 2019). The dimension of the output feature map being generated through convolution can be denoted with the following equation:

$$\left[\frac{W - K + 2p}{s} + 1\right] \text{ x } \left[\frac{W - K + 2p}{s} + 1\right], \text{(Yan, 2019)} \text{ ------------------ } 1$$

Where "W" refers to the size of the input image, "K" refers to the size of the convolutional kernel. and "s" and "p" refers to the size of stride and padding respectively. A sliding window was used for applying the convolutional filter.

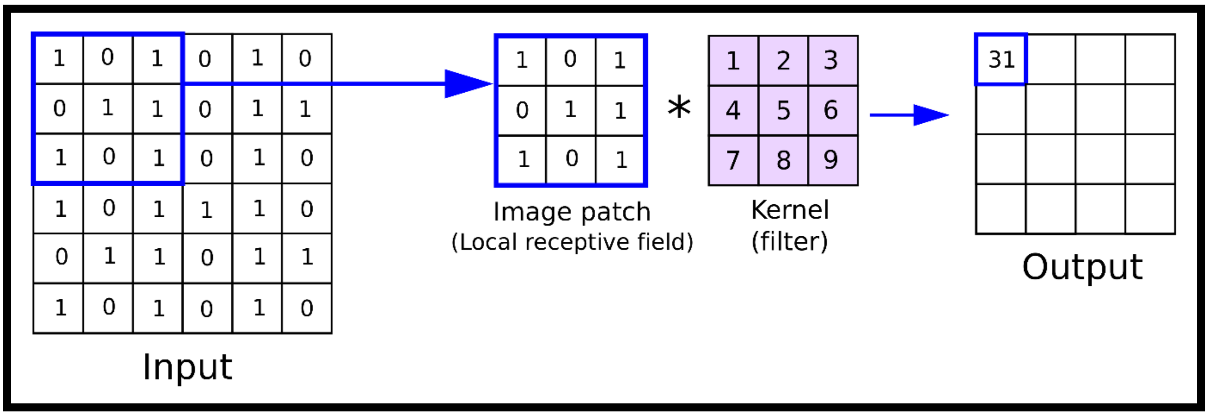


Figure 15: Example of a Convolutional Layer, (Reynolds, 2022)

4.3.1.2. Batch Normalization Layer

During the training process, in case the parameters of prior layers are modified, this results in affecting the distribution of the next layer input (Yan, 2019). This may reduce the training efficiency by requiring careful parameter initialization and lower learning rates. Hence, to solve this issue, Ioffe et al., (2015) used a technique that performs the normalization of each training mini-batch. In both the architectures including SegNet and late fusion used in this study, batch normalization was implemented after the convolutional layers.

4.3.1.3. Activation Function

The activation function, or non-linear layer, improves the capacity of the network to convey complicated non-linear mapping (Mohammadimanesh et al., 2019). Sigmoid (σ) and hyperbolic tangent (tanh) which can be seen in Figure 16, are frequent activation functions used mostly in neural networks (Rizaldy, 2018). The curve of hyperbolic tangent and sigmoid (Figure 16) are related, but the only difference is that in the case of "sigmoid function", the range is between [0, 1] while in the case of "hyperbolic tangent", the range is between [-1, 1]. This makes the derivative of "hyperbolic tangent" greater making its gradient greater than that of the sigmoid. There is an alternative activation function named rectified linear unit (ReLU), introduced by Vinod et al., (2017). ReLU activation function was used in this study, g(z) = max(0, z). It performs threshold operation on every input element. As suggested by the study performed by Rizaldy et al., (2018), replacing the conventional functions including "logistic sigmoid" or "hyperbolic tangent" with ReLU gave a relatively superior outcome. ReLU function was used in this research in both late fusion and single sensor SegNet architecture.

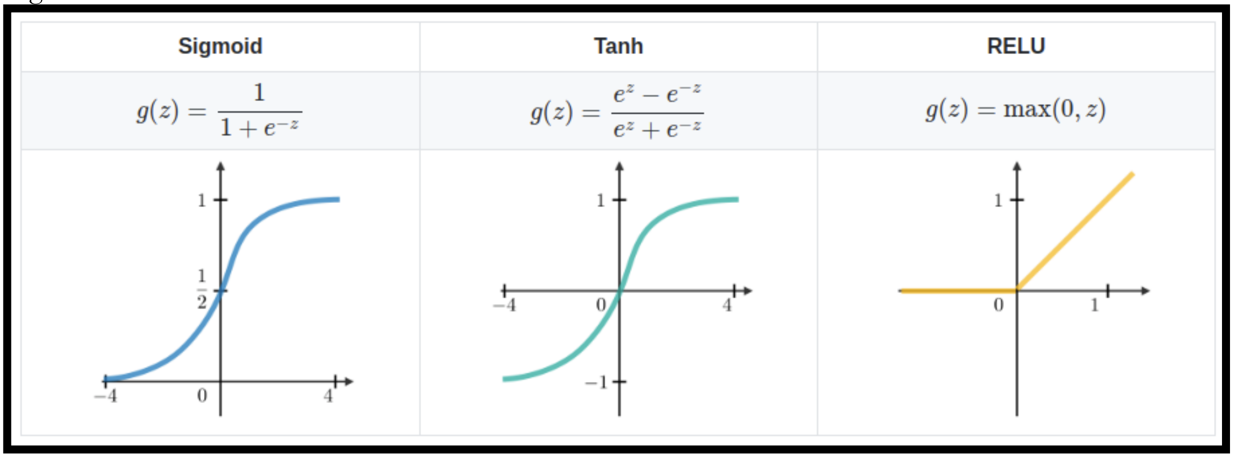


Figure 16: Different Activation Functions, (Jayawardana, 2021)

4.3.1.4. Pooling Layer

Pooling is a strategy for downsampling, usually stacked after the convolutional layer and Activation Function to summarise the output feature map (Rizaldy, 2018). It reduces the dimensionality thereby reducing the computational cost. Max pooling strategy takes a maximum value (Figure 17) which usually outperforms other pooling strategies (Rizaldy, 2018).

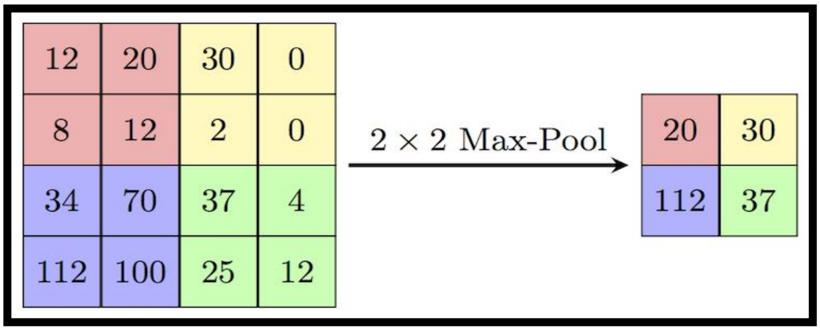


Figure 17: Example of 2x2 Max Pooling with Stride 2, (ComputerScienceWiki, 2018)

4.3.1.5. Softmax Layer

The softmax layer is included as the finishing layer in the FCN architecture used in this study to perform the semantic segmentation (Musyoka et al., 2018). This will help to classify the input dataset into the desired number of output classes which is two, namely "Deforestation" and "Non-Deforestation". It takes a value that is equivalent to the classes of interest z_p, p = 1, 2, 3,…, p which is two in our case. Further, the value of the individual class is assigned with a probability between 0 and 1 (Figure 18). The total sum of these values should be equal to one.

$$\sigma(z)j = \frac{e^{zj}}{\sum_{p=1}^{P} e^{zp}} \; for\ z = 1, 2, 3, \dots, p, \text{ (Musyoka et al., 2018) ----------------- } 2$$

where "z" represents a vector of the class input. In this research, we had two classes which were inside "z" making the output classes, j = 1, 2. "p" is the probability that a pixel belongs to a given class which is "Deforestation" and "Non-Deforestation" in this study.

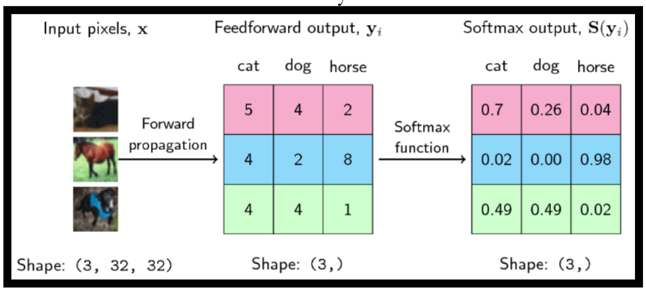


Figure 18: Example of Softmax Layer, (MIRANDA, 2017)

4.3.2. Multi-Sensor Deforestation Mapping

As was also mentioned in section 2.1, one limitation of using the feature level fusion specifically the one performed by Audebert et al., (2018), was that both encoders were expected to be consistent or symmetrical. This was to perform fusion at every convolution which might not always be the case in reality (Audebert et al., 2018). Therefore in this research, we used the fusion method in which we fused the features of the convolutional layers in the last decoder (Figure 19) of both the standalone DL models. This is called late fusion from now on in this research to avoid any confusion. This late fusion was inspired and adapted from Audebert et al., (2018) where instead of concatenating the dense predictions, the convolutions in the last decoder were fused. This concatenation was followed by Batch Normalization, ReLU activation function, 2D convolution, and finally the softmax classifier.

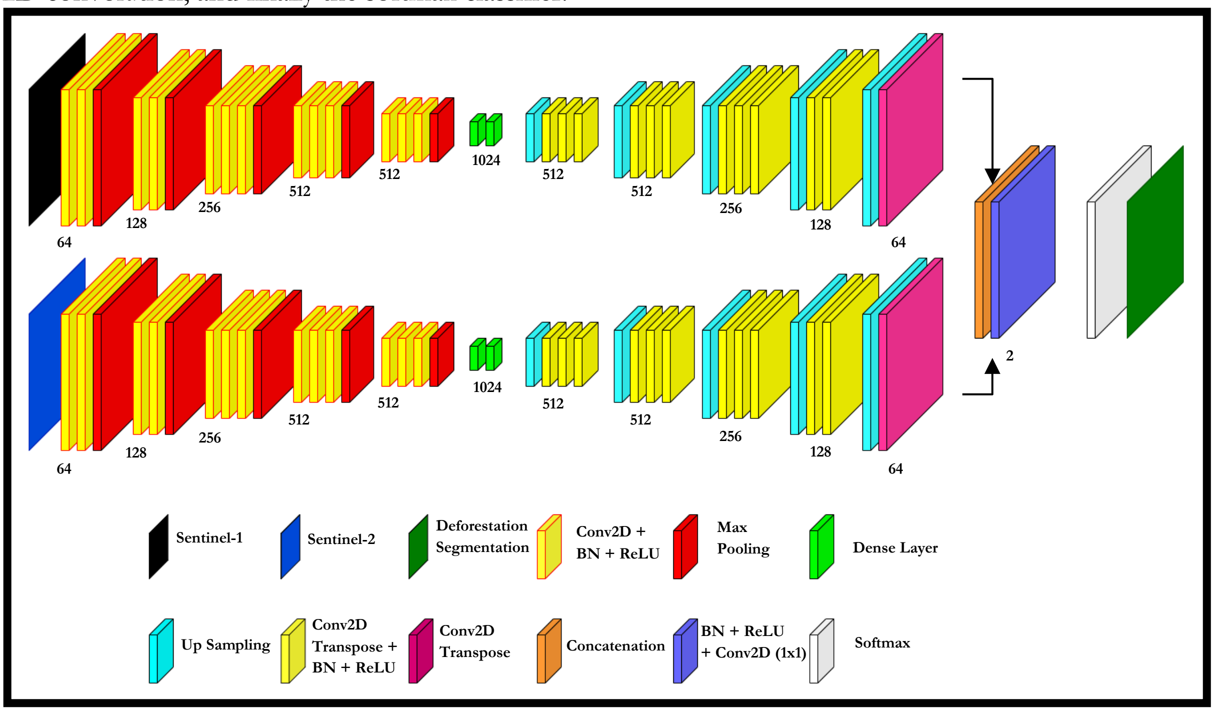


Figure 19: Late Fusion Architecture used in this research

The late fusion approach used in this research uses the SegNet architecture (Figure 14) as its backbone, with two different standalone SegNet Architectures from the encoder to the decoder for both Sentinel-1 as well as Sentinel-2 images. The entire structure of SegNet Architecture was explained in section 4.3.2. Between the encoder and decoder, two dense layers were used to change the dimension. After the last up-sampling followed by a transposed convolution in the last decoder layer of both the standalone Sensors, both the convolutional layers of Sentinel-1 and Sentinel-2 images were concatenated using the Keras function namely "Conv2D Transpose". After this, the concatenated layer was passed through batch normalization, ReLU activation function, a 1x1 convolution, and finally the Softmax function for performing the semantic segmentation. With this, we combine complementary information from both the Sentinel-1 and Sentinel-2 sensor images.

4.3.3. Training Parameters

To obtain an accurate model, a manual optimization strategy was utilized for selecting the optimum combination of hyper-parameters (Table 3). The initial values to start the optimization were based on previous experience as well as the hit and trial method for getting the best possible results. In the case of SegNet implemented using a standalone sensor and late fusion which involved two images including both Sentinel-1 and Sentinel-2, a batch size of 10 was selected to successfully run the model. Other hyper-parameters include the early stopping method from the Keras library to monitor the validation accuracy with the patience of 30 epochs. This means that the model will stop training if there is no increase in Validation Accuracy for 30 epochs. This avoids the model to overfit, saves a considerable amount of computational power, and reduces the time taken to train the model from scratch.

Table 3: Selected Hyper-Parameters

PARAMETER	SegNet	Late Fusion
Batch Size	10	10
Number of Epochs	150	150
Learning Rate	0.01	0.01
Momentum	0.9	0.9
Loss Function	Binary Cross Entropy	Binary Cross Entropy
Activation Function	ReLU	ReLU
Early Stopping Patience	30	30
Evaluation Metrics	F1 Score, User Accuracy, and Producer Accuracy	F1 Score, User Accuracy, and Producer Accuracy

4.4. Model Evaluation

4.4.1. Accuracy Assessment

Accuracy assessment of the segmented output represents the level of details that are segmented correctly in the map (Foody, 2002). The segmented map was compared with reference deforestation data from PRODES (INPE, 2019) for the year 2020 which was assumed to be our ground truth. This was performed in two stages, described below:

1. *Firstly,* as part of research question 1, an accurate assessment of the deforestation map was performed on two different scenarios. Initially, the segmentation was produced by the fusion of Optical-SAR data by using cloudy Sentinel-2 imagery and then by using cloud-free Sentinel-2 imagery.
2. *Secondly,* as part of research question 2, accuracy assessment was performed on three different scenarios, which are the deforestation maps predicted by (i) Optical-SAR fusion, (ii) standalone SAR sensor data, and (iii) standalone optical sensor data.

For performing the accuracy assessment, we used several metrics derived from the confusion matrix (Foody, 2002). It shows the relationship between reference data and the segmented data in a tabular format. We used the following measures for performing the accuracy assessment:

Overall Accuracy (OA): OA indicates the total proportion of accurately segmented pixels relative to ground truth data, with the result coming out in percentage. Apart from OA, we also calculated the accuracy for individual classes by calculating its User Accuracy (UA) and Producer Accuracy (PA) (Figure 20).

User Accuracy: UA refers to the probability that a value predicted to be in a specific class is actually in that class (L3HARRIS, 2022). This is calculated by dividing the number of correctly predicted pixels by the total number of pixels that were classified.

Producer Accuracy: PA refers to the probability that a value in a given class was classified correctly as per the reference dataset (L3HARRIS, 2022). This is calculated by dividing the number of correctly classified pixels by the number of reference pixels in a specific class.

F1 Score: It gives information about the robustness and precision of the segmentation results, (Mayasari, 2019). It is calculated from User Accuracy (precision) and the Producer Accuracy (recall) using the following equation:

$$\text{F1 Score} = 2 \times \frac{\text{UA x PA}}{\text{UA+PA}}, \text{ (Mayasari, 2019) ------------------- } 3$$

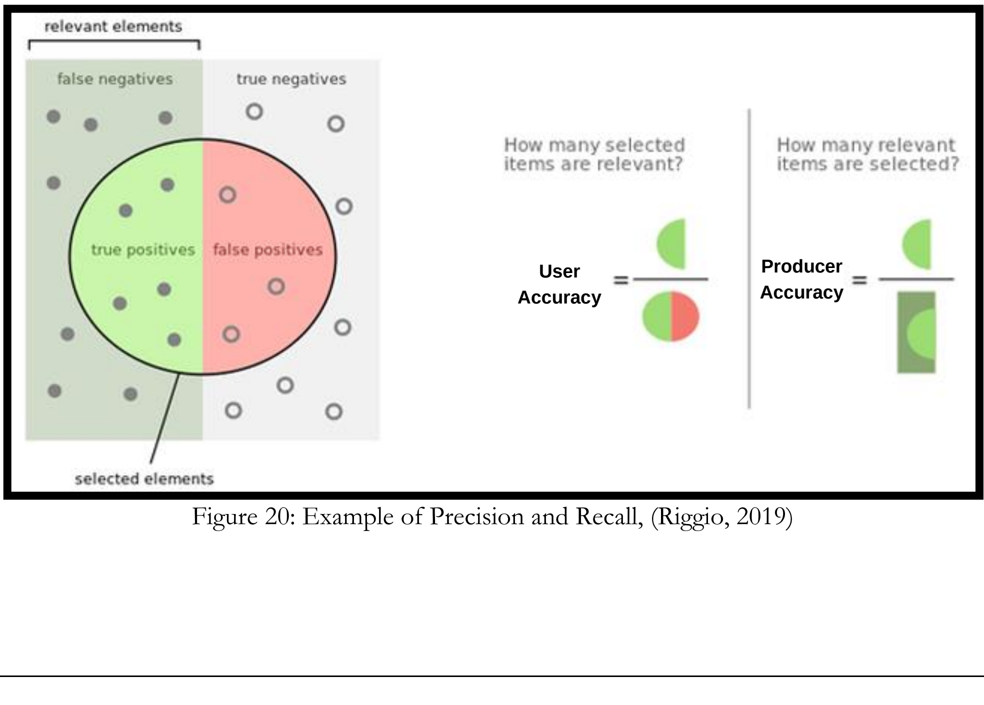


Figure 20: Example of Precision and Recall, (Riggio, 2019)

4.4.2. McNemar's Test

DL models are often computationally expensive and consist of an extremely large number of input images (Huber & Müller-Stach, 2017) as also in our study area. Our study area includes a total of 1936 small patches of the 256 x 256 pixels out of which 1439 patches were used for the training set and 255 for the validation set which is input to the DL model. In the case of a large image, model training can take many hours or even days depending on the availability of computational resources. With the advancement in machine learning models, there is strong attention to the statistical significance test which can compare and assess the predictions based on DL models using just a single test set. McNemar's test may be a suitable test for the assessment of the predictions of DL models with just one test set. This test looks at binary variables which show correct (including both true positive and true negative) and incorrect (including both false positive and false negative) pixels which are of the dimension of 2×2 (Table 5). If we have two trained classifiers, their predictions can be assessed using McNemar's test, an example of which is elaborated in Table 4, by randomly comparing classifications of 10 pixels.

Table 4: Example of a Summary of two Classifications

PIXELS	CLASSIFIER 1 CORRECT	CLASSIFIER 2 CORRECT
1	YES	NO
2	NO	NO
3	NO	YES
4	NO	NO
5	YES	YES
6	YES	YES
7	YES	YES
8	NO	NO
9	YES	NO
10	YES	YES

McNemar's test is performed by preparation of a contingency table based on a comparison of the correct and incorrect prediction of pixels of different classifiers as can be seen in Table 5. The contingency table depends on the assumption that in the case of both the standalone DL classifiers using Sentinel-1 and Sentinel-2 as inputs, the prediction was done on the same test set. The contingency table for Table 4 comparing the predictions of 2 classifiers using 10 pixels as an example can be seen in Table 6, prepared based on Table 5.

Table 5: Example of a Contingency Table

	CLASSIFIER 2 CORRECT	CLASSIFIER 2 INCORRECT
CLASSIFIER 1 CORRECT	Yes/Yes	Yes/No
CLASSIFIER 1 INCORRECT	No/Yes	No/No

Table 6: Example of Contingency Table for Table 4

	CLASSIFIER 2 CORRECT	CLASSIFIER 2 INCORRECT
CLASSIFIER 1 CORRECT	4	2
CLASSIFIER 1 INCORRECT	1	3

McNemar's test checks for the availability of differences in the predictions from two classifiers. It informs if the two models agree or disagree in a similar manner i.e. if the false positives and false negatives of two different DL models match or not. By no means this test gives information about the accuracy or availability of error in the predictions of the DL model. McNemar's test is a non-parametric test based on the confusion matrices (Table 6) and the size of the Table is always 2x2. The equation for McNemar's test is given as follows:

$$z = \frac{f12 - f21}{\sqrt{f12 + f21}}$$, (Foody, 2004)-------------------- 4

where f_{ij} signifies the occurrence of sites in the elements i, j of the confusion matrix. In the literature, there are some ongoing dialogues about this technique, including its use in remote sensing. This is for the comparison based on the assessment of chi-square (X^2) distribution, which can be calculated by just performing a square on Equation 4 with one degree of freedom. In such scenarios, the updated equation is as follows:

$$X^2 \text{ Statistics} = \frac{(f12 - f21)^2}{f12 - f21}$$, (Foody, 2004)-------------------- 5

McNemar's test has a null hypothesis which states that two classifiers disagree with each other equally. In a scenario where the assumed null hypothesis is rejected in the test set, it gives a clear indication that the two DL models disagree with each other in different ways. Based on the selection of significance level, which was selected as 0.05 in this study, the calculation of the p-value has the following interpretation:

- **(i) If p > alpha:** fail to reject H0, the two models do not disagree (e.g. no difference was observed using two different DL models).
- **(ii) If p <= alpha:** reject H0, the two models do disagree (e.g. difference has been observed using two different DL models).

5. RESULTS AND DISCUSSION

5.1. Results

5.1.1. Deforestation Mapping on Cloud-Free Images

For the assessment of the model performance, the standalone and late fusion models were evaluated using three specific examples from (i) the full image, (ii) test set-1, and (iii) test set-2. The full image contains the entire study area of which 74% is used for training. The reason why the full image was also taken along with the test set for evaluation is to make a visual assessment of the performance of the models in the entire study area. The final binary prediction of "Deforestation" and "Non-Deforestation" was then compared to the PRODES map, considered as ground truth in this study, and also evaluated using the metrics mentioned in 4.4. The performances of the model with Sentinel-1, Sentinel-2, and late fusion for deforestation mapping are outlined in Table 7. In the case of cloud-free images, the overall accuracy of Sentinel-1 images on the three examples was 0.94, 0.91, and 0.88 (Table 7), which was lower than Sentinel-2 and late fusion, but still reasonably close. Further, a late fusion of Sentinel-1 and Sentinel-2 presented the best results with an overall accuracy for the three examples as 0.97, 0.94, and 0.91 (Table 7). Figure 21, 23, and 25 show the satellite imagery, and ground truth image along with the predictions in the examples of three evaluations. In terms of the trend observed in the model performance relative to each other, the performance was observed as a minimum for Sentinel-1, in-middle for Sentinel-2, and the maximum for late fusion. This trend was noticed across all three examples. The confusion matrix for the three examples can be seen in Figures 22, 24, and 26 respectively. Apart from this, Table 8 and 9 showcase the user and producer accuracies respectively, in which Sentinel-1 performs lower than Sentinel-2 and late fused.

Table 7: Evaluation Metrics on Cloudy and Cloud Free Image

WEATHER CONDITION	SENSOR / TEST SET	OVERALL ACCURACY	F1 SCORE
Cloud Free	Sentinel-1 Full	0.94	0.95
	Sentinel-1 Test 1	0.91	0.9
	Sentinel-1 Test 2	0.88	0.78
	Sentinel-2 Full	0.95	0.96
	Sentinel-2 Test 1	0.94	0.93
	Sentinel-2 Test 2	0.9	0.81
	Late Fusion Full	**0.97**	**0.97**
	Late Fusion Test 1	**0.94**	**0.94**
	Late Fusion Test 2	**0.91**	**0.81**
Cloudy	Sentinel-1 Full	0.94	0.95
	Sentinel-1 Test 1	0.89	0.88
	Sentinel-1 Test 2	0.87	0.75
	Sentinel-2 Full	0.85	0.88
	Sentinel-2 Test 1	0.89	0.89
	Sentinel-2 Test 2	0.78	0.68
	Late Fusion Full	**0.95**	**0.95**
	Late Fusion Test 1	**0.91**	**0.9**
	Late Fusion Test 2	**0.88**	**0.79**

Table 8: User Accuracy on Cloudy and Cloud Free Image

WEATHER CONDITION	SENSOR / TEST SET	USER ACCURACY	
		DEFORESTATION	NON-DEFORESTATION
Cloud Free	Sentinel1 Full	0.94	0.95
	Sentinel1 Test 1	0.92	0.91
	Sentinel1 Test 2	0.9	0.85
	Sentinel2 Full	0.95	0.96
	Sentinel2 Test 1	0.95	0.93
	Sentinel2 Test 2	0.91	0.9
	Late Fusion Full	**0.95**	**0.98**
	Late Fusion Test 1	**0.94**	**0.95**
	Late Fusion Test 2	**0.91**	**0.92**
Cloudy	Sentinel1 Full	0.92	0.95
	Sentinel1 Test 1	0.89	0.9
	Sentinel1 Test 2	0.88	0.85
	Sentinel2 Full	0.73	0.93
	Sentinel2 Test 1	0.91	0.88
	Sentinel2 Test 2	0.88	0.62
	Late Fusion Full	**0.94**	**0.95**
	Late Fusion Test 1	**0.92**	**0.89**
	Late Fusion Test 2	**0.9**	**0.84**

Table 9: Producer Accuracy on Cloudy and Cloud Free Image.

WEATHER CONDITION	SENSOR / TEST SET	PRODUCER ACCURACY	
		DEFORESTATION	NON-DEFORESTATION
Cloud Free	Sentinel1 Full	0.94	0.95
	Sentinel1 Test 1	0.92	0.91
	Sentinel1 Test 2	0.9	0.85
	Sentinel2 Full	0.95	0.96
	Sentinel2 Test 1	0.95	0.93
	Sentinel2 Test 2	0.91	0.9
	Late Fusion Full	**0.95**	**0.98**
	Late Fusion Test 1	**0.94**	**0.95**
	Late Fusion Test 2	**0.91**	**0.92**
Cloudy	Sentinel1 Full	0.92	0.95
	Sentinel1 Test 1	0.89	0.9
	Sentinel1 Test 2	0.88	0.85
	Sentinel2 Full	0.87	0.85
	Sentinel2 Test 1	0.89	0.9
	Sentinel2 Test 2	0.88	0.62
	Late Fusion Full	**0.94**	**0.95**
	Late Fusion Test 1	**0.92**	**0.89**
	Late Fusion Test 2	**0.9**	**0.84**

Further, though Sentinel-1 performs lower than Sentinel-2 and late fusion, still the difference in overall accuracy, user accuracy, and producer accuracy was not more than 4% in any case. As demonstrated in Table 8, the user accuracy of test set-2 using Sentinel-1 image was 0.9 compared to 0.91 for Sentinel-2 and late fusion. Similarly, as demonstrated in Table 9, the producer accuracy on test set-2 was 0.9 for Sentinel-1 compared to 0.91 for both Sentinel-2 and late fusion. In the overall accuracy and f1 score, late fusion performed relatively better than Sentinel-1 and Sentinel-2 images, but that is not the case for the user and producer accuracy. In Table 8, we can observe that the user accuracy of deforestation of the full image and test set-2 was the same for both Sentinel-2 and late fusion which was 0.95 and 0.91 respectively. Apart from this, in the case of test set-1, the user accuracy of Sentinel-2 (0.95) was higher than late fusion which was 0.94. Though the user accuracy on "deforestation" for full image and test set-1 was the same for Sentinel-2 and late fusion. But, in the case of user accuracy of "non-deforestation", late fusion performs better than Sentinel-2. Sentinel-2 outperformed late fusion in test set-1, but the difference was minimal. This difference can also be observed in the rate of false positives (Table 10), i.e. Sentinel-2 on test set-1 has a false positive rate of 0.06, which is better than late fusion, which is 0.07. However, the false negatives of test set-1 of late fusion were 0.05, which was relatively better than Sentinel-2, with 0.06. considering all the evaluation metrics, the overall performance of late fusion was the best for a cloud-free image in comparison to the use of standalone sensors.

Though late fusion performed relatively better in terms of overall accuracy, f1 score, user accuracy, and producer accuracy, the false positive rate (Table 10) for test set-2 were quite high (0.21). Late fusion not only improves the model performance in terms of overall accuracy. It also reduces the false positives and false negatives relative to Sentinel-1 and Sentinel-2 standalone classifiers in cloud-free conditions. Sentinel-1 image again performs the lowest, in terms of estimation of the false positive and false negative which can be seen in Table 10. The confusion matrix can be seen visually in Figures 22, 24, and 26 below, and the metrics derived from it can be seen in Table 10.

Table 10: False Positives and False Negatives on Cloudy and Cloud Free Image

WEATHER CONDITION	SENSOR/TEST SET	FALSE POSITIVE	FALSE NEGATIVE
Cloud Free	Sentinel1 Full	0.04	0.09
	Sentinel1 Test 1	0.1	0.08
	Sentinel1 Test 2	0.22	0.06
	Sentinel2 Full	0.03	0.07
	Sentinel2 Test 1	0.06	0.06
	Sentinel2 Test 2	0.21	0.04
	Late Fusion Full	**0.03**	**0.03**
	Late Fusion Test 1	**0.07**	**0.05**
	Late Fusion Test 2	**0.21**	**0.02**
Cloudy	Sentinel1 Full	0.05	0.08
	Sentinel1 Test 1	0.13	0.09
	Sentinel1 Test 2	0.27	0.06
	Sentinel2 Full	0.07	0.27
	Sentinel2 Test 1	0.12	0.09
	Sentinel2 Test 2	0.22	0.23
	Late Fusion Full	**0.04**	**0.08**
	Late Fusion Test 1	**0.08**	**0.1**
	Late Fusion Test 2	**0.21**	**0.07**

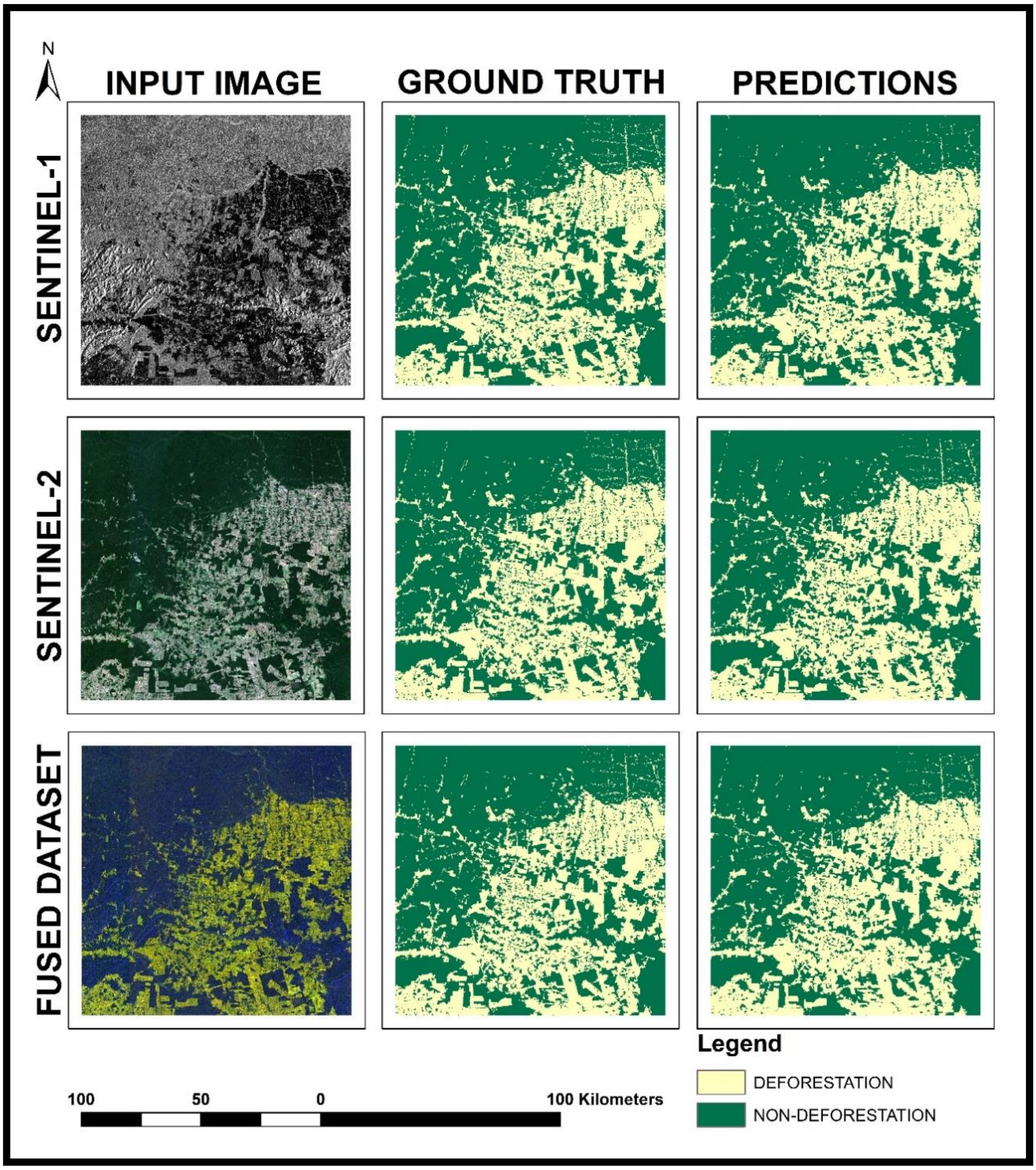


Figure 21: Cloud Free Prediction on Full Image

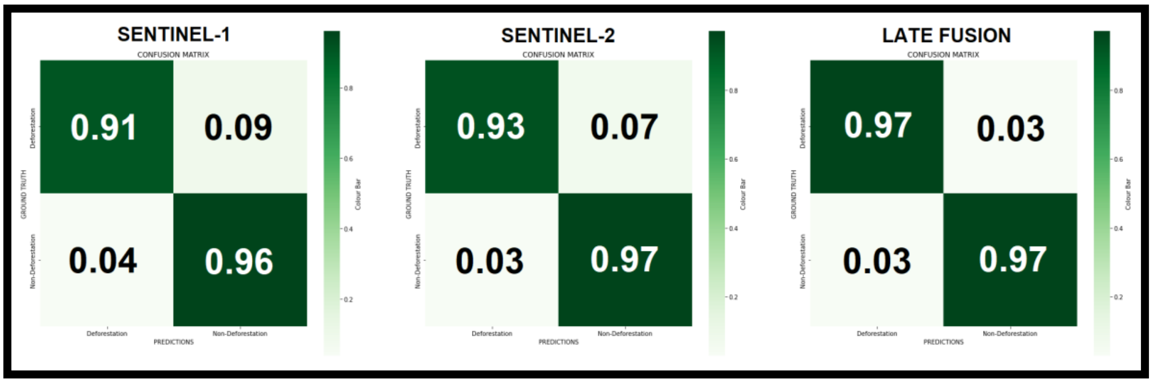


Figure 22: Confusion Matrix of Cloud Free Dataset on Full Image

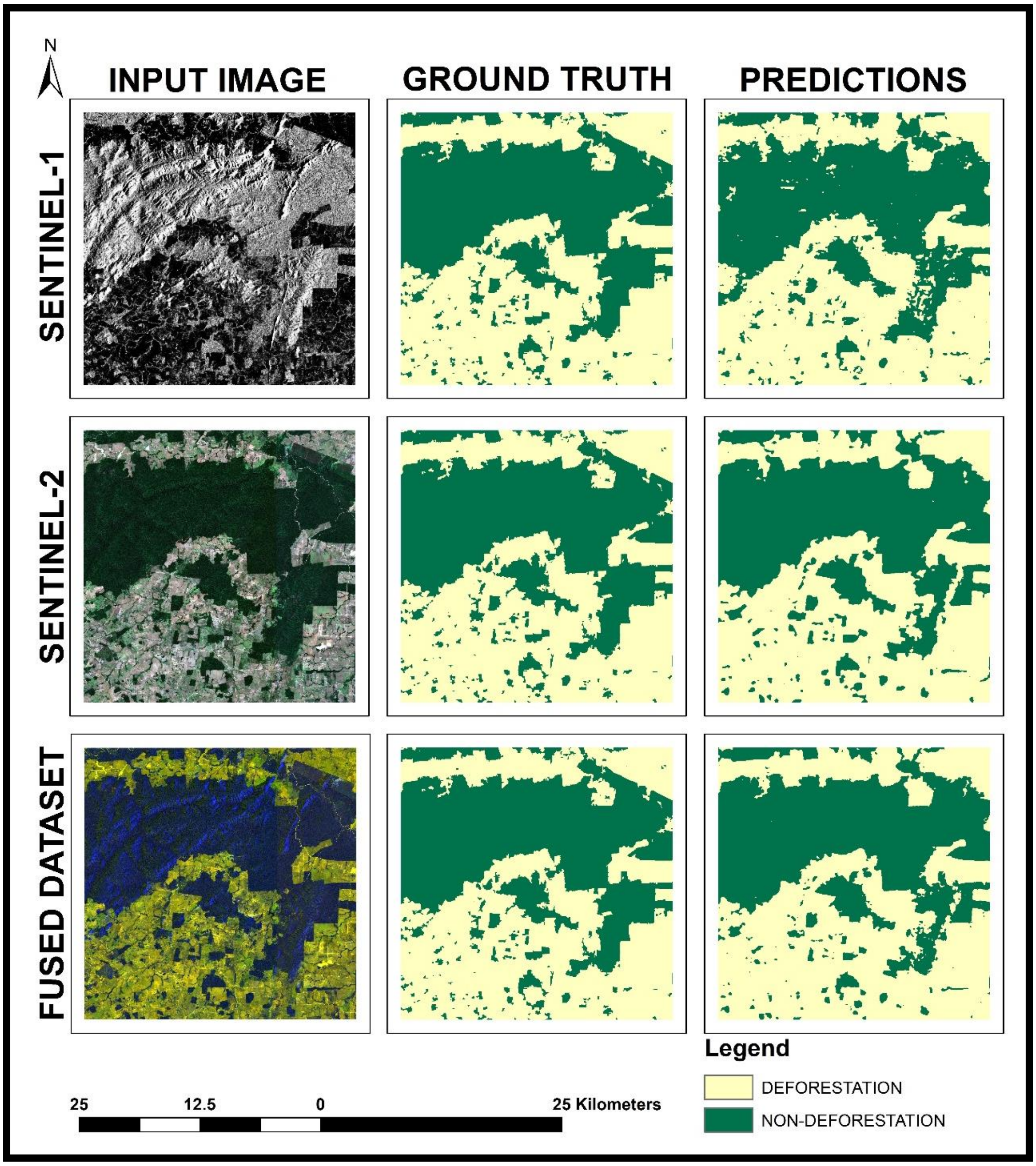


Figure 23: Cloud Free Prediction on Test Set - 1

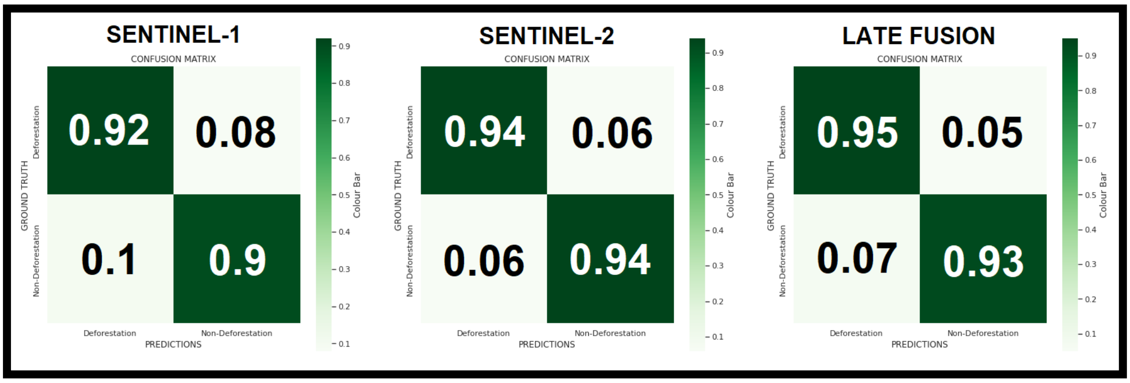


Figure 24: Confusion Matrix of Cloud Free Dataset on Test Set - 1

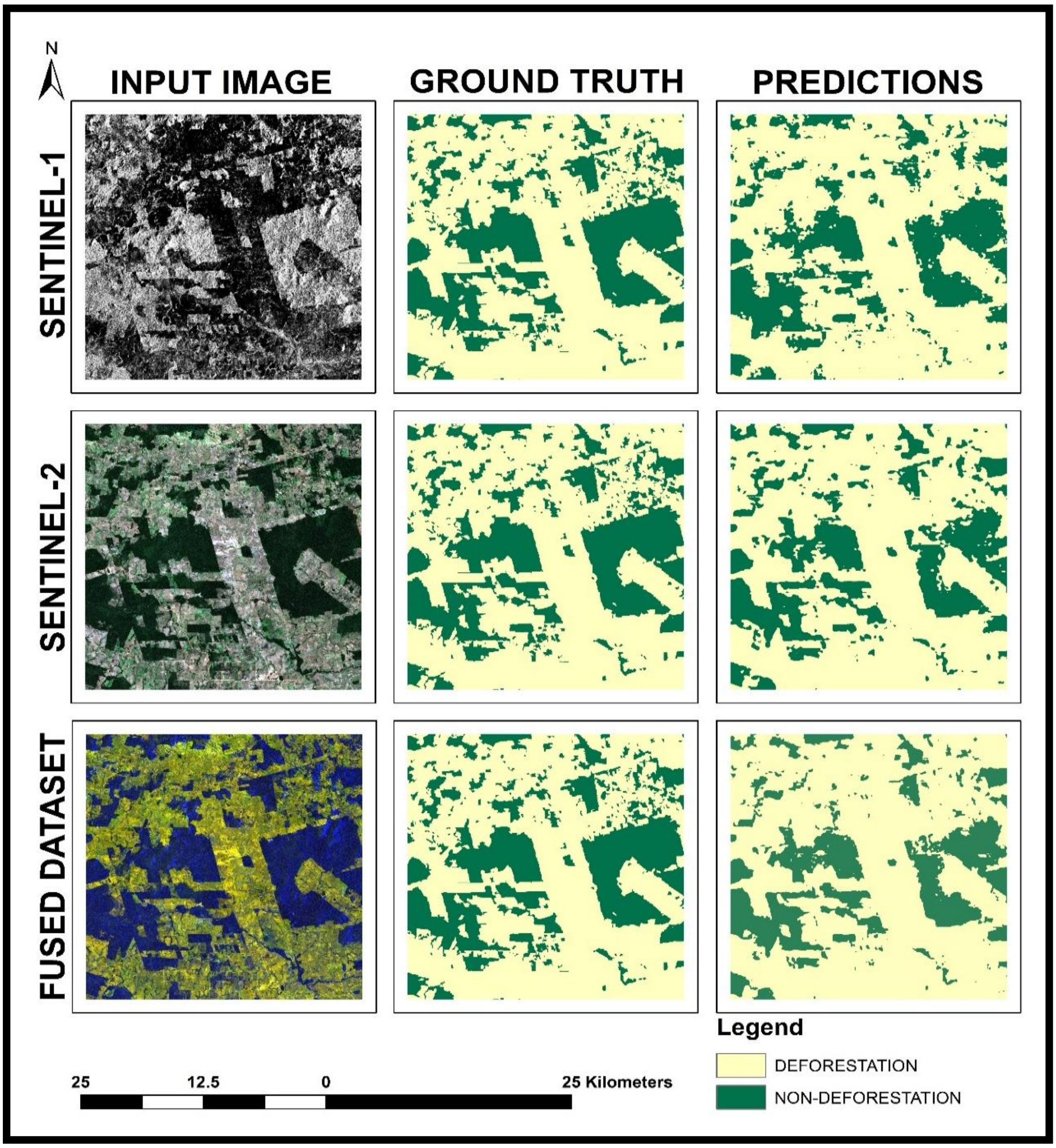


Figure 25: Cloud Free Prediction on Test Set - 2

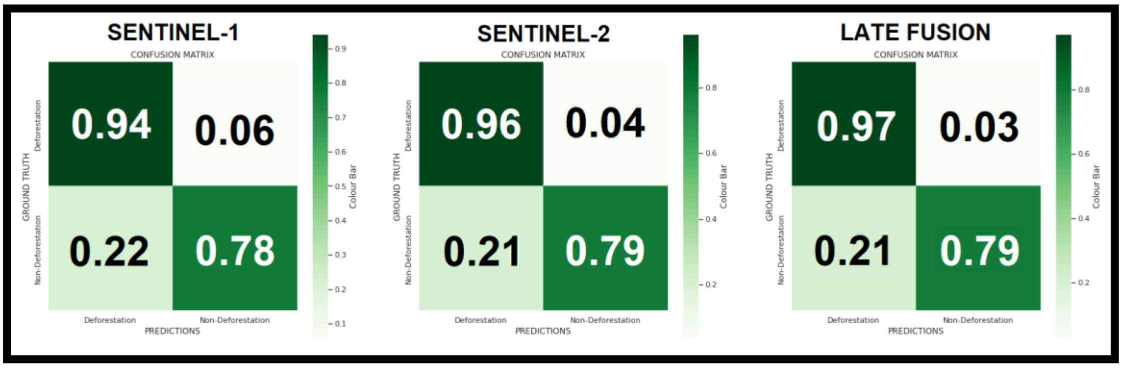


Figure 26: Confusion Matrix of Cloud Free Image on Test Set -2

5.1.2. Deforestation Mapping on Cloudy Images

In the previous section, the performance of the models was assessed on cloud-free Sentinel-2 images. In this section, similar assessments will be done on cloudy Sentinel-2 images acquired on 25th October 2020 and Sentinel-1 images acquired on 26th October 2020. As the Sentinel-2 dataset includes 22-48% clouds, so it is expected that the model will not be able to predict accurately between "deforestation" and "non-deforestation" in the areas obstructed by clouds. The performances of Sentinel-1, cloudy Sentinel-2, and the late fusion model for deforestation detection are outlined in Table 7. The model predictions for Sentinel-1, cloudy Sentinel-2, and late fusion models can be seen in Figures 27, 29, and 31 respectively and their confusion matrix can be found in Figures 28, 30, and 32 respectively for the three examples. As expected, the overall accuracy of the Sentinel-2 image was lower than Sentinel-1 and late fusion. Looking at the overall accuracy of 0.89 in the case of test set-1, Sentinel-2 is performing equivalent to Sentinel-1. However, when observed, it was found that though the overall accuracy of the Sentinel-2 test set-1 image is equivalent to Sentinel-1, visually (Figure 29) the predictions look less accurate than both Sentinel-1 and late fusion. Though the overall accuracy in the case of late fusion for the three examples was 0.95, 0.91, and 0.88, this was slightly lower than the cloud-free image which was 0.97, 0.94, and 0.91. Nevertheless, even in the case of cloudy images, late fusion was performing the best in the overall classification accuracy as well as F1 Score relative to the use of standalone Sentinel-1 and Sentinel-2 images.

A similar trend was observed in the user accuracy of deforestation. However, in the case of the user accuracy of "non-deforestation", Sentinel-1 was performing slightly better than late fusion for test set-1 and test set-2 which is 0.9 and 0.85 respectively against 0.89 and 0.84 late fusion. Despite Sentinel-1 performing relatively better than late fusion in user accuracy, Sentinel-1 presented a higher rate of false positives for the "deforestation" class. This was 0.13 and 0.27 for test set-1 and test set-2 respectively, against 0.08 and 0.21 for late fusion, confirming the improvement in the performance using late fusion. Sentinel-1 performed slightly better than late fusion in terms of False Negatives, which were 0.09 and 0.06 for test set-1 and test set-2 against 0.1 and 0.07 for late fusion. The difference was relatively small and, considering deforestation being the main class of interest, late fusion still performs better than standalone images.

5.1.3. Analysis of Best Performing Model

It was observed in the previous section that late fusion performs relatively better than both standalone models for Sentinel-1 and Sentinel-2 for deforestation mapping in both cases of cloudy and cloud-free images. Comparing the results of late fusion, cloud-free images perform better than cloudy images in terms of overall accuracy which was expected due to the availability of 22-48% clouds. The overall accuracy in cloud-free images was 0.97, 0.94, and 0.91 against cloudy imagery which was 0.95, 0.91, and 0.88 respectively. Considering 22-48% cloudy Sentinel-2 image, the performance of the late fusion-based cloudy image was very close to the cloud-free image. Late fusion on cloud-free data performed relatively better than cloudy data also in terms of the user accuracy (Table 8), and producer accuracy (Table 9). Further, as expected, in terms of false negatives, again the cloud-free image performs better than the cloudy image. Though the cloudy image contains 22-48% clouds in the Sentinel-2 image, still the false-positive detected in test set-2 is similar, which is 0.21. This shows how Sentinel-1 can benefit and supplement cloudy Sentinel-2 data for deforestation mapping. Though the value of false positives is considerably high but was observed similarly in both the cloudy and cloud-free images.

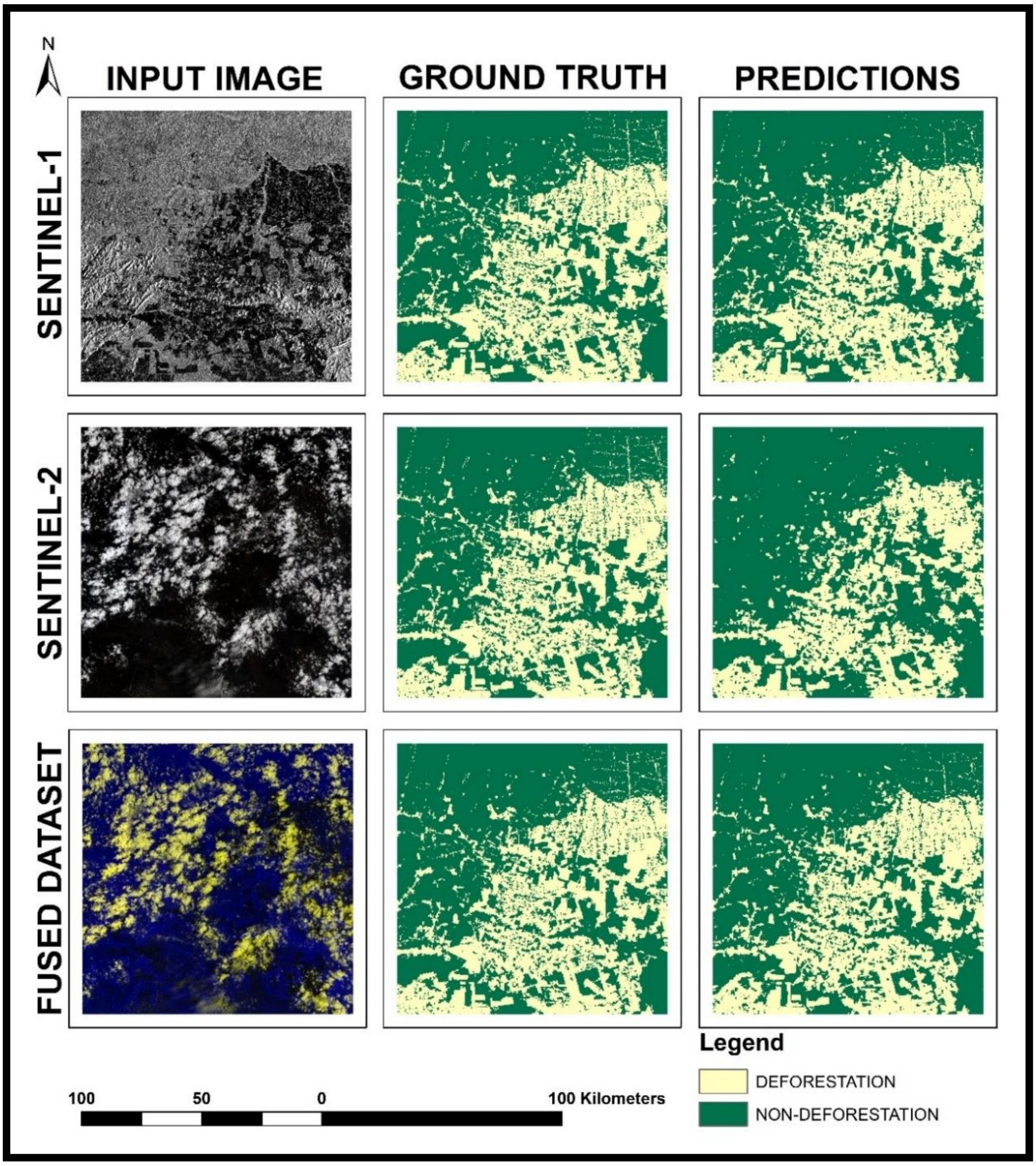


Figure 27: Cloudy Prediction on Full Image

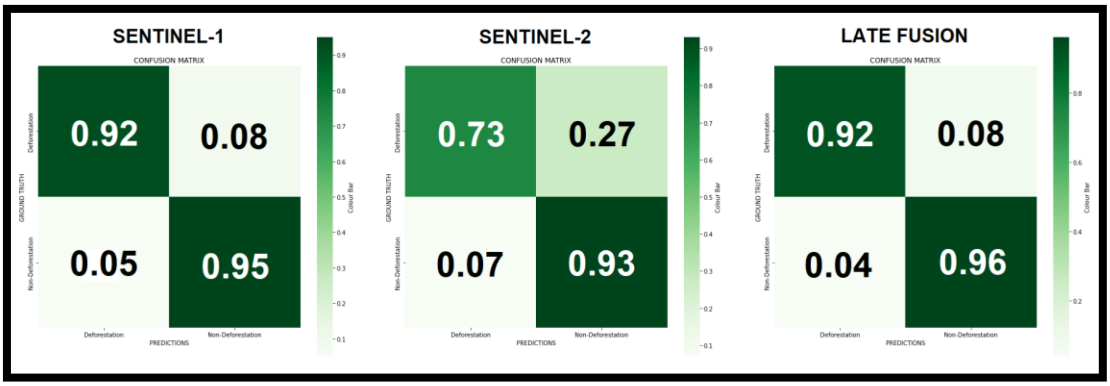


Figure 28: Cloudy Prediction on Full Image

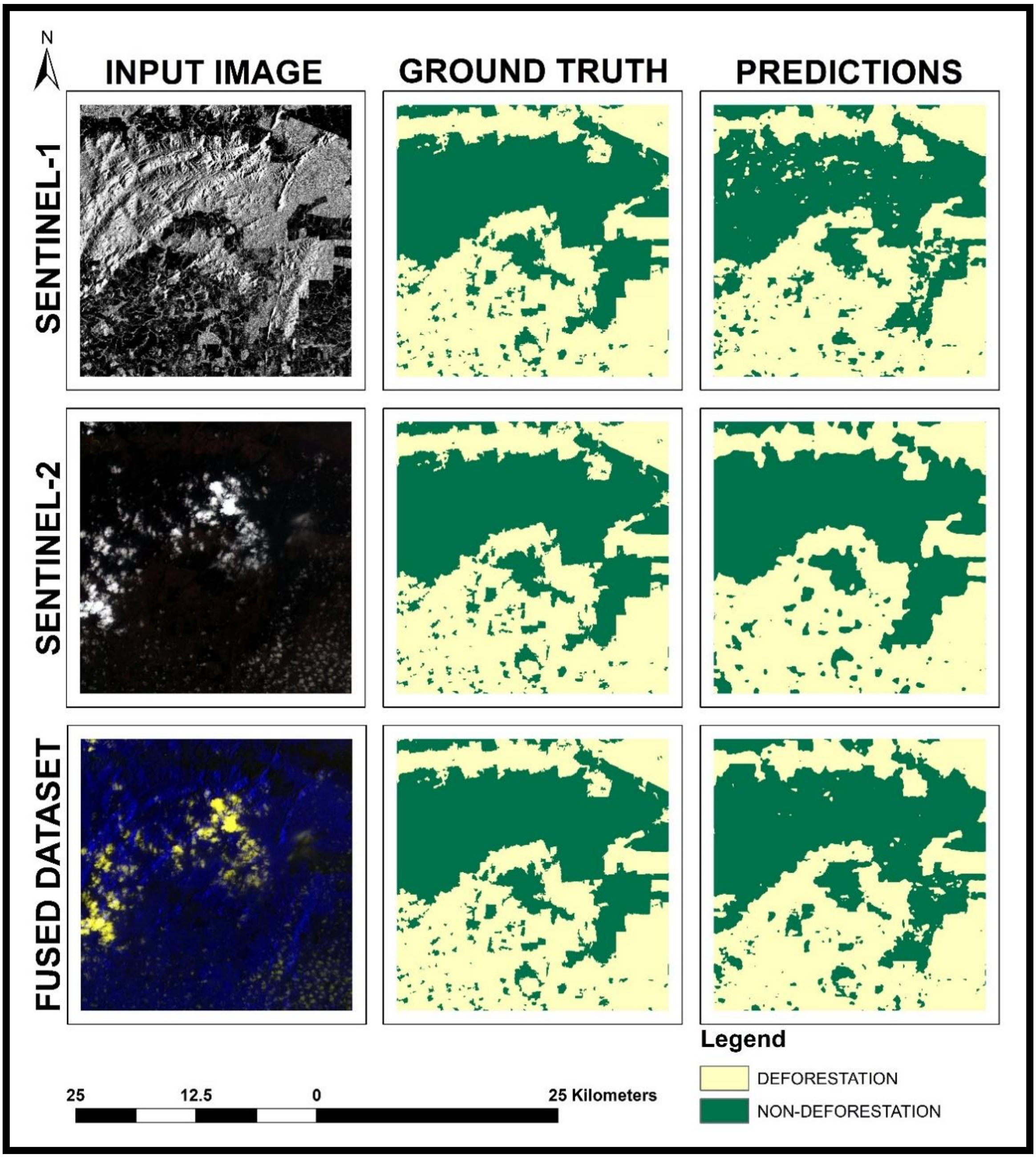


Figure 29: Cloudy Prediction on Test Set - 1

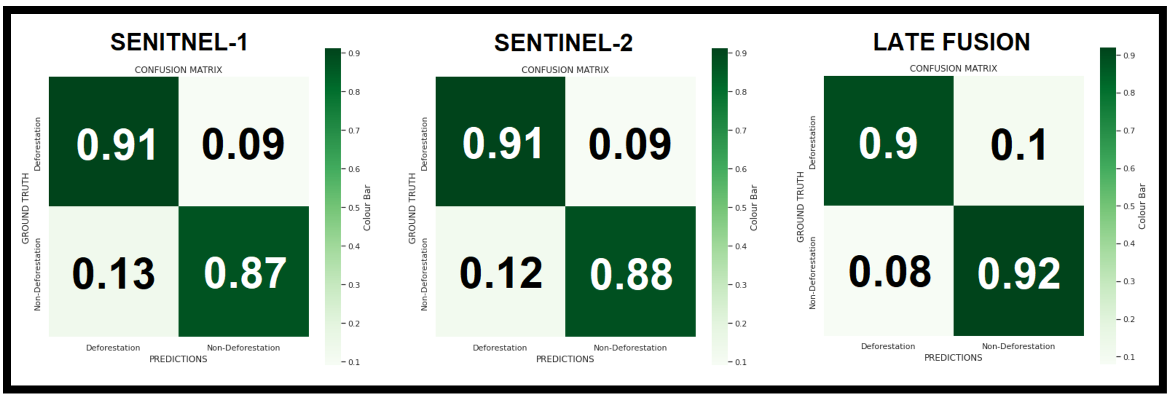


Figure 30: Confusion Matrix of Cloudy Dataset on Test Set-1

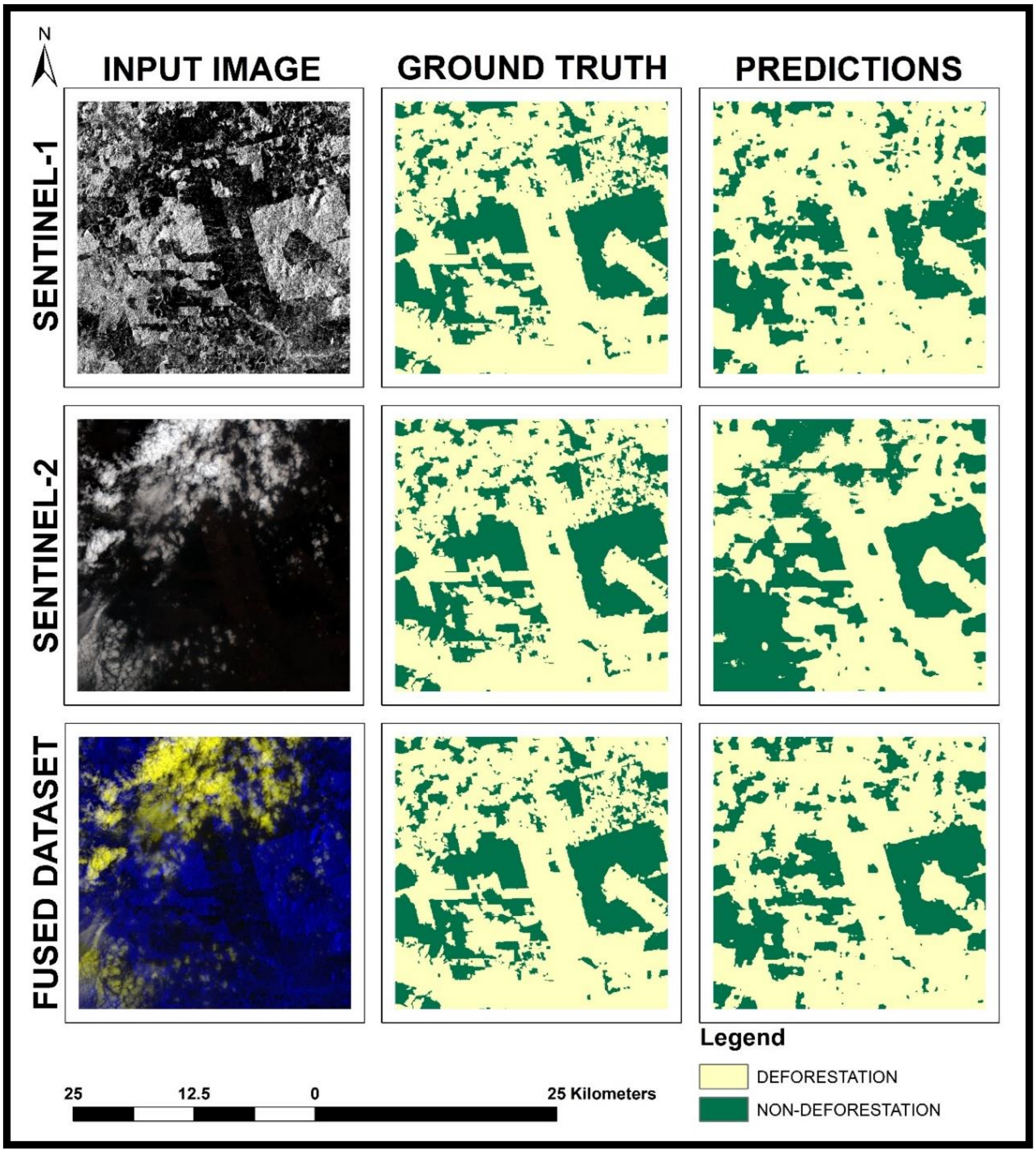


Figure 31: Cloudy Prediction on Test Set - 2

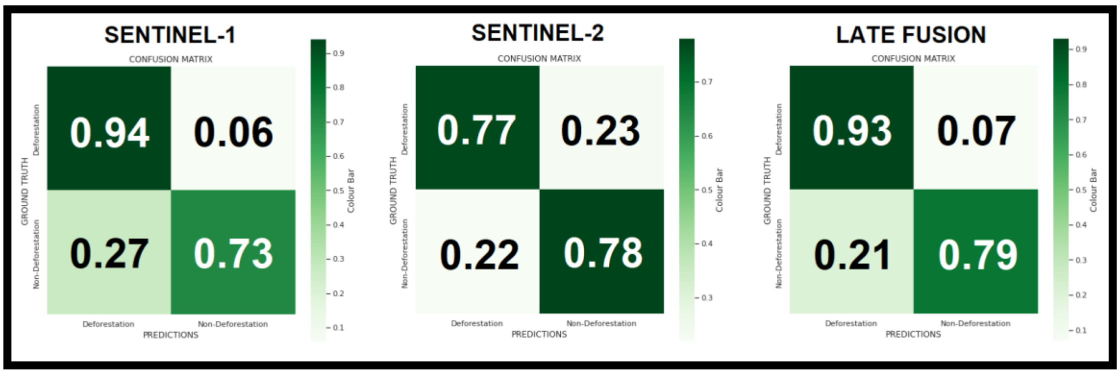


Figure 32: Confusion Matrix of Cloudy Dataset on Test Set - 2

5.1.4. Test of Homogeneity between Classification Distribution

For the assessment of the results of the segmentation using standalone models and late fusion, we performed McNemar's test between Sentinel-1, Sentinel-2, and late fusion-based DL models for the three test scenarios. The results of the McNemar's test (Table 11) showcased that all the predictions differ significantly from each other at one degree of freedom and a significance level of 0.05. The only exceptions were Sentinel-2 and late fusion as given in Table 11. In the case of cloud-free images, classification using late fusion for the full image was exhibiting the greatest differences in the X^2 Statistics (Table 11), relative to the other two sensors. However, in the case of test set-1 and test set-2 things were a little different. In test set-1, Sentinel-2 and late fusion were not significant which was also observed in the overall accuracy (Table 7) which was 0.94 (same for both). Although slight improvement was observed in the F1 score (Table 7) i.e. 0.94 for late fusion-based test set-1 against 0.93 for Sentinel-2 test set-1. The difference between Sentinel-2 and late fusion in test set-1 was statistically insignificant at the 5 percent level of significance with an X^2 Statistical difference of 1271.22. In the case of test set-2, the difference between Sentinel-2 and late fusion is statistically significant. The difference is quite low with an X^2 Statistical difference of only 3374.03 against 55362.51 between Sentinel-1 and late fusion.

In the case of cloudy images, the results were a little different. The results showcased that all the distributions differ significantly from each other. Taking the same degree of freedom and the same significance level, even test set-1 was statistically significant which was found insignificant for cloud-free images. In the case of the full image, a clear X^2 Statistical difference of 14226133.13 and 16041183.54 for Sentinel-1/Sentinel-2 and Sentinel-2/late fusion respectively has been observed. Similar X^2 statistical difference was also observed in test set-2 with a difference of 424994.39 and 528035.21 for Sentinel-1/Sentinel-2 and Sentinel-2/late fusion respectively. Test set-1 gave somewhat different results in which the statistical difference was higher between Sentinel-1/late fusion which was different from the differences observed in test set-2 and full image.

Table 11: McNemar's Test Results

WEATHER CONDITION	TEST SET	PRODUCT	X2	p-Value
Cloud Free	Full Image	Sentinel-1 vs Sentinel-2	190604.81	<0.05
		Sentinel-1 vs Late Fusion	1684811.99	<0.05
		Sentinel-2 vs Late Fusion	985819.78	<0.05
	Test Set - 1	Sentinel-1 vs Sentinel-2	79262.11	<0.05
		Sentinel-1 vs Late Fusion	94182.46	<0.05
		Sentinel-2 vs Late Fusion	**1271.22**	**2.02**
	Test Set - 2	Sentinel-1 vs Sentinel-2	34548.12	<0.05
		Sentinel-1 vs Late Fusion	55362.51	<0.05
		Sentinel-2 vs Late Fusion	3374.03	<0.05
Cloudy	Full Image	Sentinel-1 vs Sentinel-2	14226133.13	<0.05
		Sentinel-1 vs Late Fusion	279315	<0.05
		Sentinel-2 vs Late Fusion	16041183.54	<0.05
	Test Set - 1	Sentinel-1 vs Sentinel-2	3307.06	<0.05
		Sentinel-1 vs Late Fusion	35610.46	<0.05
		Sentinel-2 vs Late Fusion	9262.51	<0.05
	Test Set - 2	Sentinel-1 vs Sentinel-2	424994.39	<0.05
		Sentinel-1 vs Late Fusion	12441.35	<0.05
		Sentinel-2 vs Late Fusion	528035.21	<0.05

5.2. Discussion

Instead of considerable drop in the deforestation rate in the amazon, forests are still being cleared (R. V. Maretto et al., 2020). Despite the forest clear-cuts at such a large scale, such events are hard to track in time (Isaienkov et al., 2021) and lead to the loss of huge areas of forests that are slowly and steadily being cut down. This gives a clear idea about the need for consistent and efficient monitoring of deforestation activities on our planet, and generating precise deforestation maps for updating and facilitating public policies aimed at combating deforestation (R. V. Maretto et al., 2020). In this study, deforestation mapping was explored using Sentinel-1, Sentinel-2, and late fusion models in both cloudy and cloud-free weather conditions. This study showcased a new methodology for deforestation mapping by using a late fusion technique by performing a fusion of features in the last decoder. The model was used for both Sentinel-1 and Sentinel-2 images as input, to perform a multi-modal data fusion. Standalone SegNet models were separately prepared for mapping deforestation using Sentinel-1 and Sentinel-2 images separately. Based on the results in section 5.1.3, as expected late fusion using a cloud-free image performed relatively better than a cloudy image in terms of overall accuracy. In cloud-free weather conditions, Sentinel-2 had a higher user and producer accuracy and performed relatively better than Sentinel-1, as C-Band SAR is usually less sensitive to the forest structures, which are mostly in the "Non-Deforestation" class in our study as also mentioned by Sirro et al., (2018) and Sinha et al., (2015). This can also be quantified by the fact that the user and producer accuracy for "Non-Deforestation" detected using Sentinel-1 for all the test sets was lower than Sentinel-2 image (Table 8 and 9) in a much higher proportion than the "Deforestation" class. The results in the case of the cloudy dataset were quite as were expected which was lower than the cloud-free dataset.

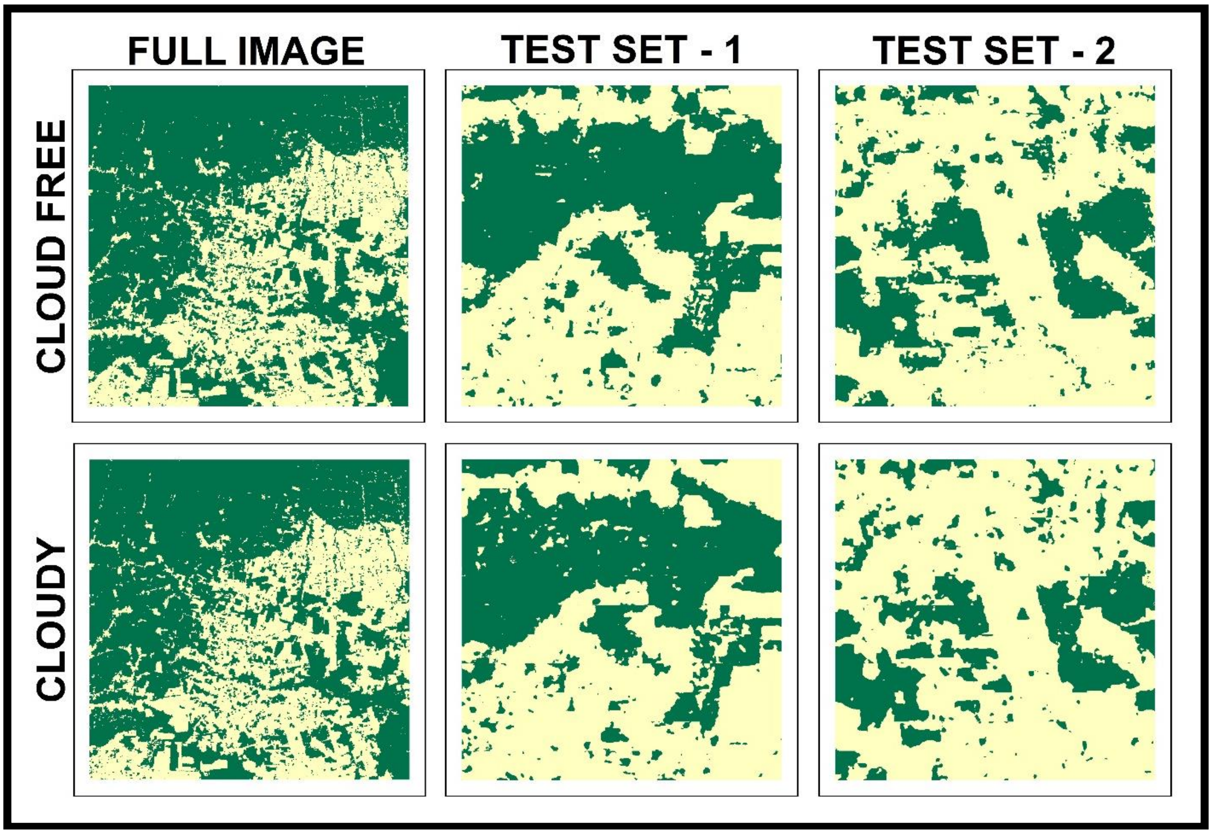


Figure 33: Predictions on Cloudy and Cloud Free Sentinel-1 Image

As Sentinel-1 can capture information in all-weather conditions, so not much difference was observed in the overall accuracy in cloud-free and cloudy Sentinel-1 images which can also be seen visually in Figure 33. But, in the case of the Sentinel-2 image, a clear difference was observed in the overall accuracy between the cloudy and cloud-free images. The difference can also be observed visually in Figure 34. Though there was a huge decrease in the overall accuracy of cloudy Sentinel-2 images, fusing it with Sentinel-1 using late fusion improved the performance of the model for all three examples. This directly shows the benefit of fusing images from different modalities in the case of cloudy images.

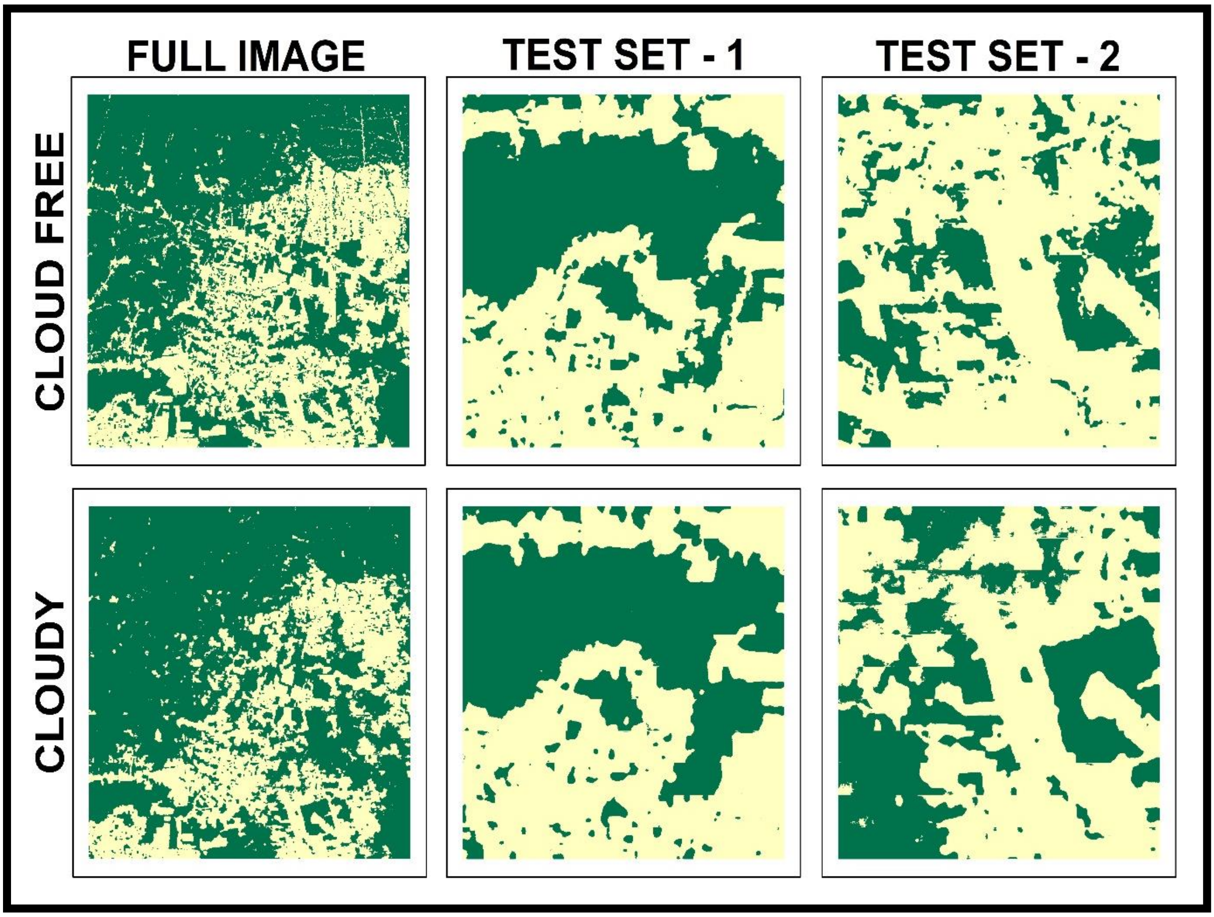


Figure 34: Predictions on Cloudy and Cloud Free Sentinel-2 Image

In the case of cloud-free images, Sentinel-2 performed the best in standalone sensors which is also a reason why the late fusion in cloudy-free images performed relatively better than the cloudy image in terms of overall accuracy. The trends observed in overall accuracy were in-line with the trends observed in the F1 Score (Table 7).

In this study, late fusion was explored for improving the deforestation mapping for both cloudy and cloud-free datasets. As expected, the fusion improved the classification accuracy for deforestation mapping for both cloudy and cloud-free datasets. Looking at Figure 35 classified by cloudy Sentinel-2 test set-2, the lower left part (highlighted in a red dotted circle) was misclassified as "Non-Deforestation" since, with the obstruction by clouds, it ends up confusing the DL architecture using cloudy Sentinel-2 image for predicting between the "Deforestation" and "Non-Deforestation" class. However, late fusion manages to correctly classify Deforestation on the lower left side (Figure 35) in the lower-left part of the image due to the presence of Sentinel-1 Image. This was similar to the predictions on a full image which can be seen in Figure 36, on the upper right side of the Sentinel-2 predicted image and a few other locations (highlighted in a red dotted circle), where small patches of deforestation were predicted as "Non-Deforestation" in green colour. Late fusion recovered the small patches of "Deforestation" which can be seen in the upper right side and other parts (highlighted in a red dotted circle) of Figure 36. Thanks to the availability of the Sentinel-1 sensor which was not affected by clouds. Further, looking at past studies done for deforestation mapping based on optical images by Ortega et al., (2019), Isaienkov et al., (2021), Andrade et al., (2020), Bem et al., (2020), Adarme et al., (2020) specific dates of the image were selected where cloud cover was minimum. Also, in the study done by Maretto et al., (2020), clouds were masked out during deforestation mapping. This is because of the availability of clouds, the DL-based model was not able to perform accurate predictions. This research tried to fix this issue of clouds in the Sentinel-2 dataset specifically by using late fusion and mapping deforestation in cloudy Sentinel-2 image along with Sentinel-1 image.

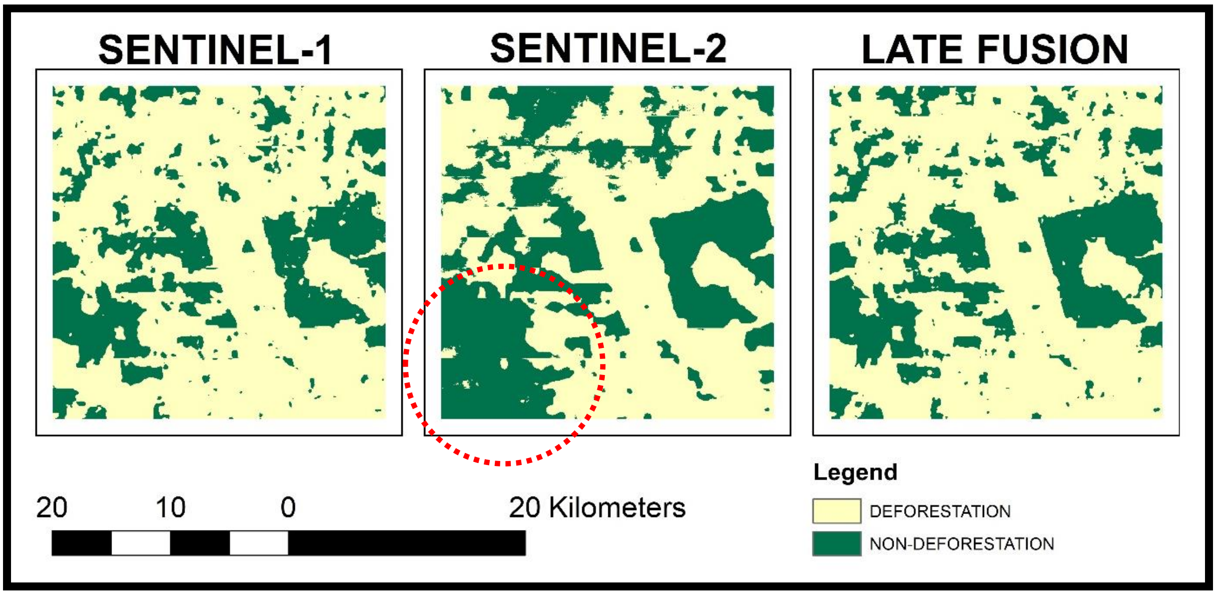


Figure 35: Predictions on Cloudy Image on Test Set-2

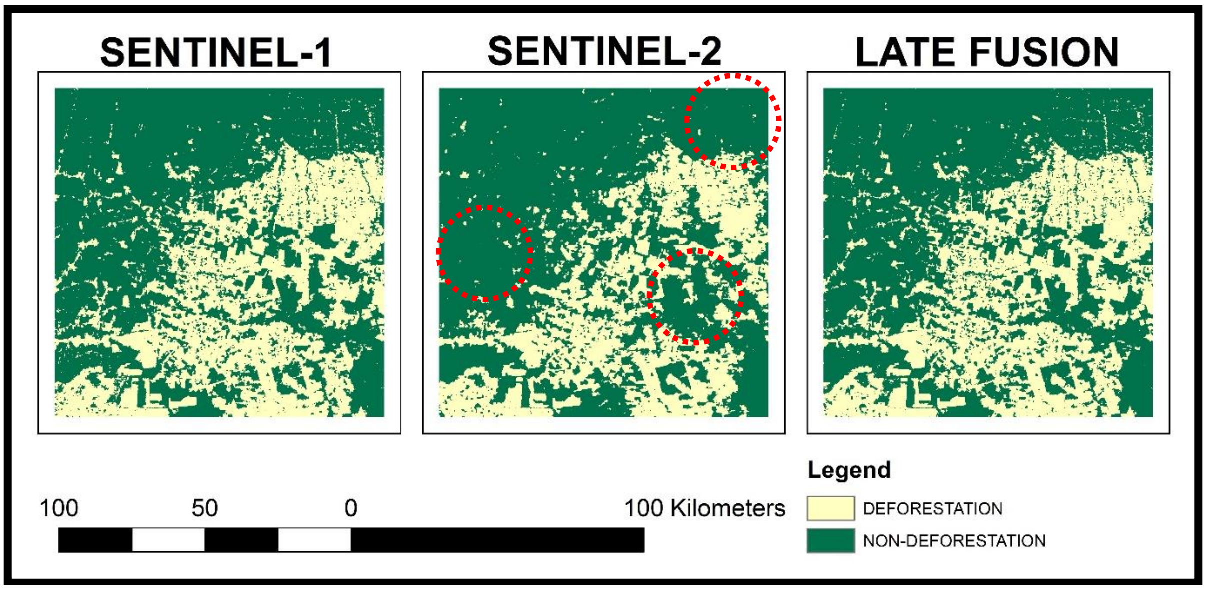


Figure 36: Predictions on Cloudy Image on Full Image

Combining both the images using late fusion improved the accuracy of deforestation mapping, relative to using standalone Sentinel-1 or Sentinel-2 images. Sentinel-2 produced superior results than that Sentinel-1 in the case of a cloud-free situation. While in the case of a cloudy scenario, as expected, Sentinel-1 produces relatively better results than that Sentinel-2. Although only little improvement in deforestation mapping was observed using late fusion, this improvement was observed in terms of all the evaluation metrics including overall accuracy (Table 7), f1 score (Table 7), user accuracy (Table 8), producer accuracy (Table 9), false positives and false negatives (Table 10). This improvement was also observed in both cloudy and cloud-free weather conditions in comparison to the DL model using standalone Sentinel-1 and Sentinel-2 images. The performance of deforestation mapping was relatively better in cloud-free conditions using late fusion than in the cloudy Sentinel-2 image. Based on the results, in the case of deforestation mapping, if the availability of images was not an issue, cloud-free imagery should be considered over cloudy imagery. As standalone images, both Sentinel-1 and Sentinel-2 provided sufficient accuracies, and fusing both optical and SAR data does improve the overall classification accuracy for deforestation detection as was also discussed in 5.1. Our research has shown the advantage of Optical-SAR data fusion in both cloudy and cloud-free conditions.

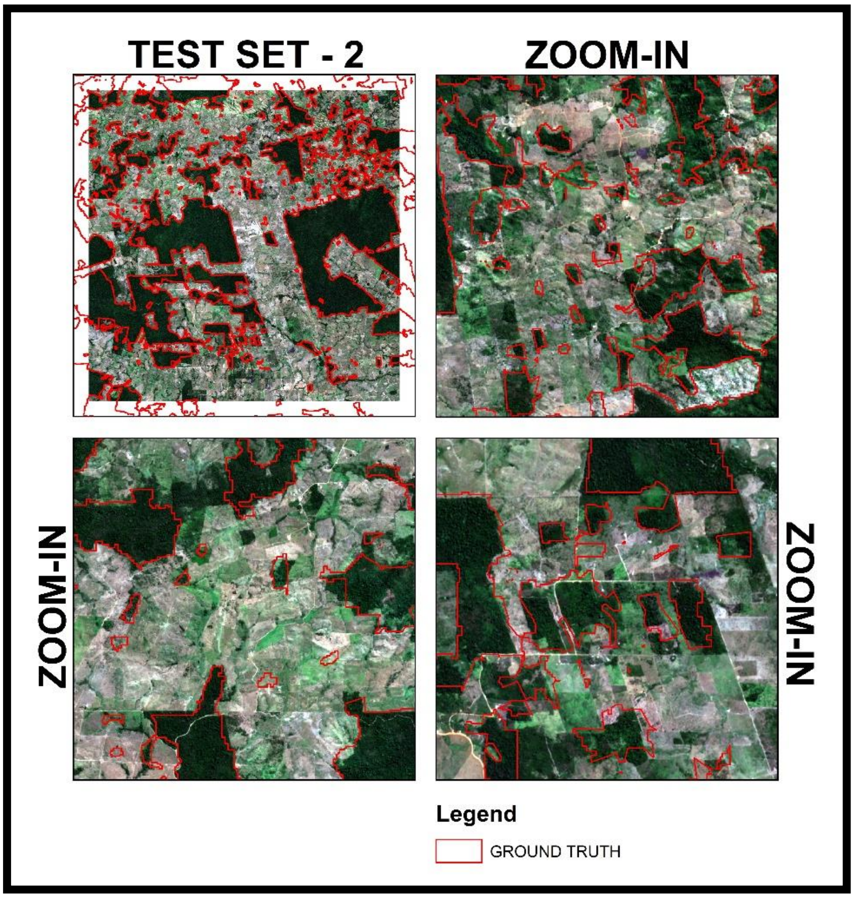


Figure 37: Ground Truth Image of Test Set - 2

In the case of test set-2, late fusion performed relatively better than standalone sensors. But, it was observed that the rate of false positives was considerably higher for deforestation i.e. 0.22, 0.21, and 0.21 using Sentinel-1, Sentinel-2, and late fusion respectively for a cloud-free scenario. For the cloudy dataset, the false positives were 0.27, 0.22, and 0.21 using Sentinel-1, Sentinel-2, and late fusion (Figure 38) respectively. Further, the quantification of False Positives (Table 10) for test set-2 was relatively higher than for test set-1. This can be an issue in a scenario where a DL algorithm was used for decision-making by the government officials, wrong fines can be imposed on a location where no deforestation ever happened on the ground. One of the reasons for a high rate of false positives in test set-2, was that a high number of small patches of "Non-Deforestation" relative to test set-1 were observed, with a size less than 0.1 hectares. This is not quantified but can be seen visually in Figure 38. The DL model usually smoothes the edges in case of extremely small patches as can be seen visually in predictions in cloudy and cloud-free datasets (Figure 21, 23, 25, 27, 29, and 31) thereby leading to the wrong classification. Another reason for the high rate of false positives in test set-2 is the time lag between the PRODES dataset and the acquisition of cloud-free satellite imagery used in this research. As can be seen in Figure 37 in test set-2, due to this lag extremely small areas are wrongly trained to lead to the wrong classification. Another possible reason for the high false positives is the difference in the spatial resolution of PRODES (30m) and the Sentinel-1 and Sentinel-2 images which is 10m.

Further, the performance of the semantic segmentation using DL architecture was also affected by the quality of ground truth images. This is because PRODES is also not 100% accurate, but indicates a level of accuracy close to 95% (PRODES, 2021). In future work, using reference images with a higher spatial resolution like MapBiomas with 10m spatial resolution can be explored for deforestation mapping using DL Architecture prepared in this research. The overall classification accuracy for the cloudy image was close enough to that of the cloud-free image. Improvements can be explored in the fusion model which has been suggested in section 6.3 of this research. Although Sentinel-2 alone did perform the best in terms of deforestation mapping in cloud-free conditions, and also Sentinel-1 in cloudy conditions, the combination of both improved the accuracy of deforestation mapping in both cloudy and cloud-free conditions.

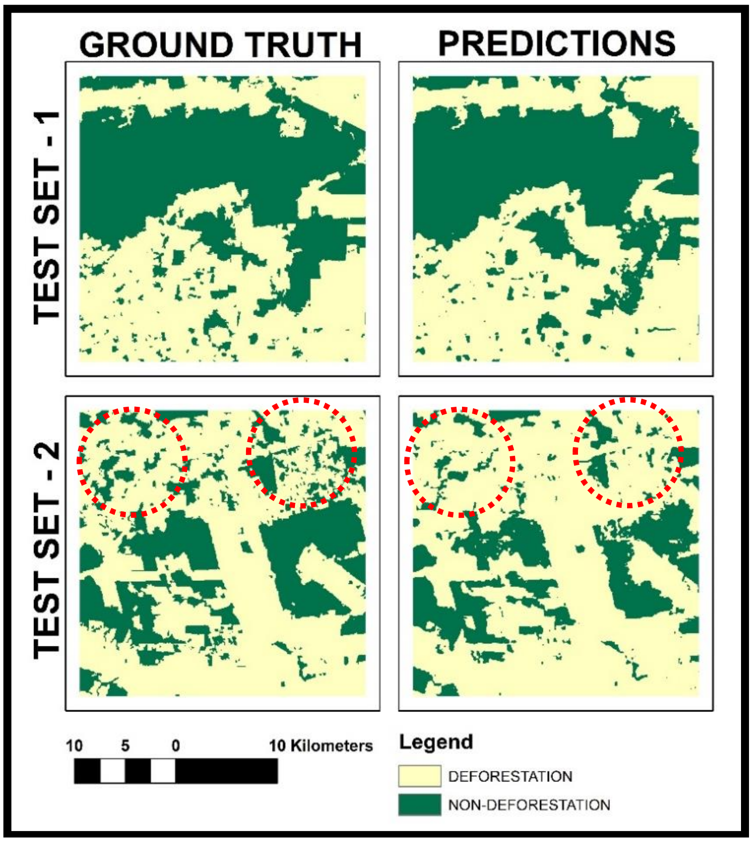


Figure 38: Predictions on Cloudy Image using Late Fusion

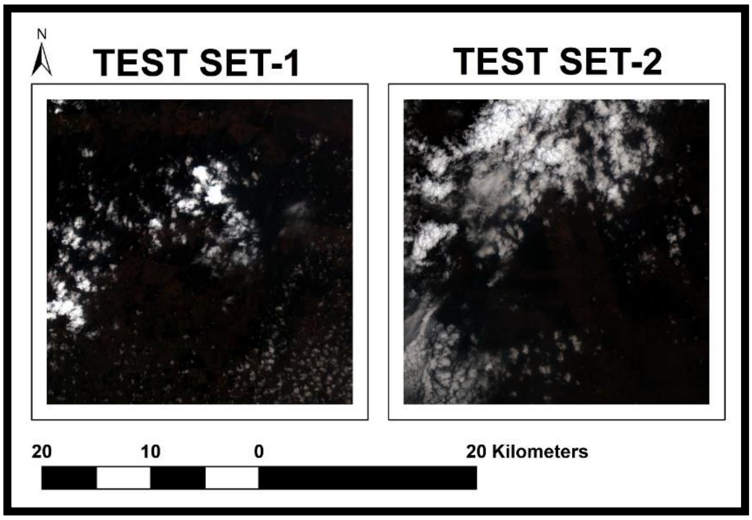


Figure 39: Sentinel-2 Cloudy Test Set-1 and Test Set-2

As per McNemar's test for cloud-free images, all the examples were statistically significant apart from Sentinel-2 and late fusion which was statistically insignificant for test set-1 at a 5% level of significance (Table 11). As per the interpretation of McNemar's test, the level of difference between Sentinel-2 and late fusion is very minimal. But, though being statistically insignificant, still, a minimal level of X^2 Statistical difference was observed though only 1271.22 (Table 11).

In the case of cloudy images, in the full image and test set-2, the statistical difference was observed lower between Sentinel-1/late fusion than between Sentinel-2/late fusion. As also mentioned in 4.4.2 that McNemar's test informs if the two models agree (true positive) or disagree (true negative) in a similar manner. It does not give information about the accuracy of the model. A high rate of difference was observed in X^2 statistics of Sentinel-2/late fusion relative to Sentinel-1/late fusion. This statistical significance test was in line with the overall accuracy, where the accuracy of Sentinel-1 was close to late fusion. But, there was a relatively high difference in the overall accuracy between Sentinel-2 and late fusion. As expected, this high statistical difference was also observed in Sentinel-1/Sentinel-2 in both full image and test set-2. This is because the predictions of Sentinel-2 were quite different (with a high level of differences) than Sentinel-1 and late fusion. It was because of wrong predictions of the DL model in the Sentinel-2 image in the areas obscured by clouds. Although the difference between Sentinel-1 and late fusion was low for the full image and test set-2, relative to the difference from the Sentinel-2 image. But, the difference was statistically significant showing a level of difference between the predictions though small. This verifies that the results of late fusion/Sentinel-1 and late fusion/Sentinel-2 were statistically significant and different, which is in line with the results in the overall classification accuracy.

The observations for test set-1 for the cloudy image were a little surprising and contradicted the full image and test set-2. This is because though results were statistically significant for Sentinel-1, Sentinel-2, and late fusion with each other, the differences observed in Sentinel-2/late fusion were smaller than the difference observed between Sentinel-1/late fusion. Similar trends were also observed in the overall accuracy as though the Sentinel-2 imagery was cloudy, but the overall accuracy on test set-1 was the same (0.89) for both Sentinel-1 and Sentinel-2. In the case of Sentinel-1 test set-1, the wrong classification of small patches of "Deforestation" (Figure 40 highlighted with red dotted line) over the "Non-Deforestation" (green color) areas, makes more false positives in the case of Sentinel-1 predictions relative to Sentinel-2 predictions which were 0.13 and 0.12 respectively. The false positives are reduced to 0.08 using late fusion. Hence, the availability of more false positives by the small Deforestation patches in Sentinel-1 (Figure 40), is one of the reasons for the statistical difference in the case of test set-1 higher between Sentinel-1/late fusion relative to the difference between Sentinel-2/late fusion. The results of McNemar's statistical significance test are in-line with the overall classification accuracy, proving our hypothesis correct.

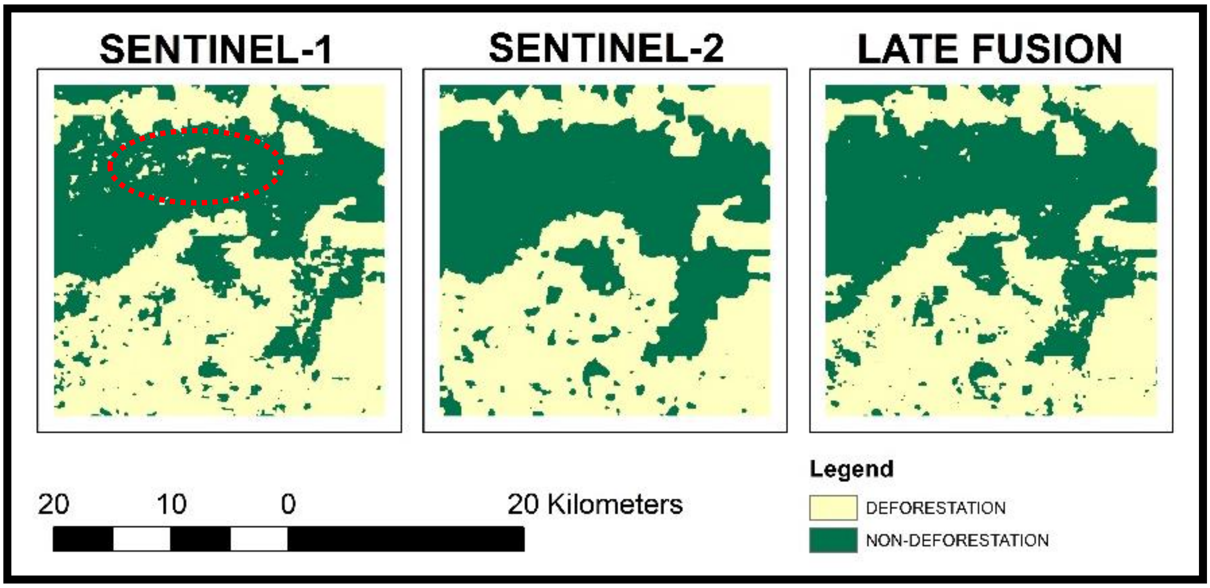


Figure 40: Predictions on Test Set-1 on Cloudy Image

6. CONCLUSION AND RECOMMENDATION

6.1. Conclusions

This study describes a novel approach for the late fusion of Sentinel-1 and Sentinel-2 images for deforestation mapping for the year 2020. Our study confirms that the combination of optical and SAR data significantly improves the accuracies for deforestation mapping, which was also validated using McNemar's statistical significance test. In the case of a cloud-free image, the overall accuracy using Sentinel-1/Sentinel-2 input used in late fusion increased by +3%, +3%, and +3% for full image, test set-1, and test set-2 respectively relative to Sentinel-1 image and increased by +2% and +1% for full image and test set-2 respectively relative to Sentinel-2 image. In case of cloudy weather condition, the overall accuracy of Sentinel-1/Sentinel-2 inputs used in late fusion increased by +1%, +2% and +1% for full image, test set-1 and test set-2 respectively relative to Sentinel-1 image and increased by +10%, +2% and +10% relative to Sentinel-2 cloudy image. With this research, we have confirmed that DL shows potential for multi-sensor data fusion for mapping deforestation specifically using freely available Sentinel-1 and Sentinel-2 imagery. We also have been able to show that data fusion is beneficial for deforestation mapping in case of both cloudy and cloud-free scenarios. In future studies, data fusion can be explored with fusion using more than the two modalities used in this research for improving deforestation mapping in both cloudy and cloudy-free weather conditions.

6.2. Answer to Research Questions

1. **To what degree can SAR data support deforestation mapping when atmospheric conditions affect the optical image?**

Sentinel-1 together with Sentinel-2 using late fusion supports deforestation detection when atmospheric conditions affect optical image by +10%, +2% and +10% for full image, test set-1 and test set-2 respectively.

2. **To what extent can the fusion of Optical-SAR data improve deforestation mapping relative to using standalone sensor images?**

- In the case of cloud-free data, using late fusion improved the overall accuracy by +3%, +3%, and +3% for full image, test set-1, and test set-2 respectively relative to Sentinel-1 Image. The overall accuracy increased by +2% and +1% for full image and test set-2 respectively relative to Sentinel-2 Image.

- In the case of cloudy weather conditions, using late fusion improved the overall accuracy by +1%, +2%, and +1% for full image, test set-1, and test set-2 respectively relative to the Sentinel-1 image. The overall accuracy improved by +10%, +2% and +10% relative to Sentinel-2 cloudy image.

6.3. Recommendations for further studies

Even though the approach presented in this research produced encouraging results, it still has numerous limitations which can be overcome in future works, as a continuation of this research which is as follows:

(i) This research mapped the accumulated deforestation from the year 1988 until 2020 using a single temporal image of the year 2020. In future works, the use of time-series images should be explored instead of a single temporal image used in this research for mapping deforestation.

(ii) Other future directions which can be taken are, fusing Sentinel-1 and Sentinel-2 images for change detection which could not be undertaken in this research due to time constraints.

(iii) In this research only one fusion approach i.e. late fusion was utilized for fusing Sentinel-1 and Sentinel-2 images. Future works can observe the results of deforestation detection using other levels of fusion which can be more of a comparative study for understanding the performance of different fusion methods in both cloudy and cloud-free atmospheric conditions.

(iv) In this research, we only used the highest-resolution (10m) bands including Red, Green, Blue, and NIR as more bands take more computational resources. In future works, more bands of Sentinel-2 images can be explored for deforestation detection.

(v) In the future transferability of the existing trained model in a different area in the amazon or even in a different part of the planet can be explored.

(vi) In this study, Sentinel-1 images were highly beneficial for mapping deforestation in cloudy conditions. Future satellite missions including NISAR which is a joint project between NASA and ISRO will include “S” and “L” Bands which can be explored for data fusion with Sentinel-2 and also for deforestation mapping using standalone “S” and “L” Bands.

(vii) We used late fusion by using the same backbone architectures for both Sentinel-1 and Sentinel-2 images which were SegNet architectures. Future research can explore late fusion by using different backbone architectures before fusion, which could utilize the full potential of both Sentinel-1 and Sentinel-2 images for mapping deforestation in both cloudy and cloud-free weather conditions.

(viii) Due to time constraints, this study was only limited to using standalone sensors and late fusion for mapping deforested areas and evaluating the model using accuracy assessment and McNemar’s test. In future studies, specific features can be identified which were correctly detected on Sentinel-1 and wrong on Sentinel-2 and vice-versa. This can help in the identification of the exact complementary features for deforestation mapping using fusion relative to using a standalone image.

LIST OF REFERENCES

Adarme, M. O., Feitosa, R. Q., Happ, P. N., Almeida, C. A. De, & Gomes, A. R. (2020). Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and cerrado biomes from remote sensing imagery. *Remote Sensing*, *12*(6). https://doi.org/10.3390/rs12060910

Adrian, J., Sagan, V., & Maimaitijiang, M. (2021). Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, *175*, 215–235. https://doi.org/10.1016/j.isprsjprs.2021.02.018

Amazônia, P.-. (2022). *PRODES deforestation estimates in Brazilian Amazon*. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes

Andrade, R. B., Costa, G. A. O. P., Mota, G. L. A., Ortega, M. X., Feitosa, R. Q., Soto, P. J., & Heipke, C. (2020). Evaluation of semantic segmentation methods for deforestation detection in the amazon. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *43*(B3), 1497–1505. https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1497-2020

Audebert, N., Le Saux, B., & Lefèvre, S. (2017). Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10111 LNCS*, 180–196. https://doi.org/10.1007/978-3-319-54181-5_12

Audebert, N., Le Saux, B., & Lefèvre, S. (2018). Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 20–32. https://doi.org/10.1016/j.isprsjprs.2017.11.011

Audebert, N., Saux, B. Le, & Lefevre, S. (2017). Joint Learning from Earth Observation and OpenStreetMap Data to Get Faster Better Semantic Maps. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2017-July*, 1552–1560. https://doi.org/10.1109/CVPRW.2017.199

Belgiu, M., & Stein, A. (2019). Spatiotemporal image fusion in remote sensing. In *Remote Sensing* (Vol. 11, Issue 7, p. 818). MDPI AG. https://doi.org/10.3390/rs11070818

Bem, P. P. de, Junior, O. A. de C., Guimarães, R. F., & Gomes, R. A. T. (2020). Change Detection of Deforestation in the Brazilian Amazon Using Landsat Data and Convolutional Neural Networks. *Remote Sensing 2020, Vol. 12, Page 901*, *12*(6), 901. https://doi.org/10.3390/RS12060901

Brovelli, M. A., Sun, Y., & Yordanov, V. (2020). Monitoring forest change in the amazon using multi-temporal remote sensing data and machine learning classification on Google Earth Engine. *ISPRS International Journal of Geo-Information*, *9*(10), 1–21. https://doi.org/10.3390/ijgi9100580

Carrero, G. C., Philip, •, Fearnside, M., Denis, •, Do Valle, R., Cristiano De, •, & Alves, S. (2020). Deforestation Trajectories on a Development Frontier in the Brazilian Amazon: 35 Years of Settlement Colonization, Policy and Economic Shifts, and Land Accumulation. *Environmental Management*, *66*, 966–984. https://doi.org/10.1007/s00267-020-01354-w

ComputerScienceWiki. (2018). *Maxpool*. https://computersciencewiki.org/index.php/File:MaxpoolSample2.png

de Bem, P. P., de Carvalho, O. A., Guimarães, R. F., & Gomes, R. A. T. (2020). Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks. *Remote Sensing*, *12*(6). https://doi.org/10.3390/rs12060901

Doblas, J., Shimabukuro, Y., Sant'Anna, S., Carneiro, A., Aragão, L., & Almeida, C. (2020). Optimizing Near Real-Time Detection of Deforestation on Tropical Rainforests Using Sentinel-1 Data. *Remote Sensing 2020, Vol. 12, Page 3922*, *12*(23), 3922. https://doi.org/10.3390/RS12233922

European Space Agency. (2012). *Sentinel-1*. Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services. https://sentinel.esa.int/documents/247904/349449/S1_SP-1322_1.pdf

European Space Agency. (2015). *Sentinel-2 User Handbook*. https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

Foody, G. M. (2004). Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy Giles. *Photogrammetric Engineering & Remote Sensing*, *70*(5), 627–633. https://doi.org/10.14358/pers.70.5.627

Gadzicki, K., Khamsehashari, R., & Zetzsche, C. (2020). Early vs late fusion in multimodal convolutional neural networks. *Proceedings of 2020 23rd International Conference on Information Fusion, FUSION 2020*. https://doi.org/10.23919/FUSION45008.2020.9190246

Han, Y., Liu, Y., Hong, Z., Zhang, Y., Yang, S., & Wang, J. (2021). Sea ice image classification based on heterogeneous data fusion and deep learning. *Remote Sensing*, *13*(4), 1–20. https://doi.org/10.3390/rs13040592

Hasret, G., & Wegner, J. D. (2018). *Detecting deforestation using Sentinel-1 SAR data and deep learning Supervisors*.

Hazirbas, C., Ma, L., Domokos, C., & Cremers, D. (2017). FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10111 LNCS*. https://doi.org/10.1007/978-3-319-54181-5_14

Huber, A., & Müller-Stach, S. (2017). Approximate Statistical Tests for Comparing Supervised Classification. *Ergebnisse Der Mathematik Und Ihrer Grenzgebiete*, *65*, 117–133. https://doi.org/10.1007/978-3-319-50926-6_6

Hughes, L. H., Marcos, D., Lobry, S., Tuia, D., & Schmitt, M. (2020). A deep learning framework for matching of SAR and optical imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *169*(September), 166–179. https://doi.org/10.1016/j.isprsjprs.2020.09.012

Ienco, D., Interdonato, R., Gaetano, R., & Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, *158*(February), 11–22. https://doi.org/10.1016/j.isprsjprs.2019.09.016

INPE, I. N. D. P. E.-. (2019). *Metodologia Utilizada nos Projetos PRODES e DETER*. 33. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes/pdfs/Metodologia_Prodes_Deter_revisada.pdf

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015*, *1*, 448–456.

Isaienkov, K., Yushchuk, M., Khramtsov, V., & Seliverstov, O. (2021). Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem with Sentinel-2. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 364–376. https://doi.org/10.1109/JSTARS.2020.3034186

Jayawardana, R. (2021). *ANALYSIS OF OPTIMIZING NEURAL NETWORKS AND ARTIFICIAL INTELLIGENT MODELS FOR GUIDANCE , CONTROL , AND NAVIGATION of Modernization in Engineering ANALYSIS OF OPTIMIZING NEURAL NETWORKS AND ARTIFICIAL INTELLIGENT MODELS FOR GUIDANCE , CONTROL , AND NAVIGATI. April.*

John, D., & Zhang, C. (2022). An attention-based U-Net for detecting deforestation within satellite sensor imagery ☆. *International Journal of Applied Earth Observation and Geoinformation*, *107*(January), 102685. https://doi.org/10.1016/j.jag.2022.102685

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T. A., Reiche, J., Ryan, C. M., & Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing*, *8*(1), 1–23. https://doi.org/10.3390/rs8010070

Kulkarni, S. C., & Rege, P. P. (2020). Pixel level fusion techniques for SAR and optical images: A review. *Information Fusion*, *59*(January), 13–29. https://doi.org/10.1016/j.inffus.2020.01.003

Kumar, D. (2021). Urban objects detection from C-band synthetic aperture radar (SAR) satellite images through simulating filter properties. *Scientific Reports 2021 11:1*, *11*(1), 1–24. https://doi.org/10.1038/s41598-021-85121-9

L3HARRIS. (2022). *Calculate Confusion Matrices*. https://www.l3harrisgeospatial.com/docs/calculatingconfusionmatrices.html

Lee, S. H., Han, K. J., Lee, K., Lee, K. J., Oh, K. Y., & Lee, M. J. (2020). Classification of landscape affected by deforestation using high-resolution remote sensing data and deep-learning techniques. *Remote Sensing*, *12*(20), 1–16. https://doi.org/10.3390/rs12203372

Liu, Y., Piramanayagam, S., Monteiro, S. T., & Saber, E. (2017). Dense Semantic Labeling of Very-High-Resolution Aerial Imagery and LiDAR with Fully-Convolutional Neural Networks and Higher-Order CRFs. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2017-July*, 1561–1570. https://doi.org/10.1109/CVPRW.2017.200

Liu, Z., Blasch, E., Bhatnagar, G., John, V., Wu, W., & Blum, R. S. (2018). Fusing synergistic information from multi-sensor images: An overview from implementation to performance assessment. *Information Fusion*, *42*(October 2017), 127–145. https://doi.org/10.1016/j.inffus.2017.10.010

Mahyoub, S., Fadil, A., Mansour, E. M., Rhinane, H., & Al-Nahmi, F. (2019). Fusing of optical and synthetic aperture radar (SAR) remote sensing data: A systematic literature review (SLR). *International Archives of* *the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4/W12), 127–138. https://doi.org/10.5194/isprs-archives-XLII-4-W12-127-2019

Maretto, R. V. (2020). *AUTOMATING LAND COVER CHANGE DETECTION : A DEEP LEARNING BASED APPROACH TO MAP DEFORESTED AREAS*.

Maretto, R. V., Fonseca, L. M. G., Jacobs, N., Korting, T. S., Bendini, H. N., & Parente, L. L. (2020). Spatio-Temporal Deep Learning Approach to Map Deforestation in Amazon Rainforest. *IEEE Geoscience and Remote Sensing Letters*, *18*(5), 1–5. https://doi.org/10.1109/lgrs.2020.2986407

Matosak, B. M., Maretto, R. V., Korting, T. S., Adami, M., & Fonseca, L. M. G. (2020). Mapping Deforested Areas in the Cerrado Biome through Recurrent Neural Networks. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 1389–1392. https://doi.org/10.1109/IGARSS39084.2020.9324019

Mayasari, R. (2019). *Convolutional Networks for the Classification of Multi-Temporal Satellite Images Convolutional Networks Multi-Temporal Satellite Images*. *February*.

Mcnemar, Q. (1947). *NOTE ON THE SAMPLING ERROR OF THE DIFFERENCE BETWEEN CORRELATED PROPORTIONS OR PERCENTAGES*. *12*(2), 153–157.

Meraner, A., Ebel, P., Zhu, X. X., & Schmitt, M. (2020). Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. *ISPRS Journal of Photogrammetry and Remote Sensing*, *166*(January), 333–346. https://doi.org/10.1016/j.isprsjprs.2020.05.013

MINERAL, D. N. D. P. (1978). Programa de Integração Nacional. Levantamento de Recursos Naturais. *V. 18 (Manaus) – RADAM (Projeto) DNPM, Ministério Das Minas e Energia. Brasil. p. 626.*, *18*, p.626.

MIRANDA, L. (2017). *Understanding softmax and the negative log-likelihood*. https://ljvmiranda921.github.io/notebook/2017/08/13/softmax-and-the-negative-log-likelihood/

Mnih, V., & Hinton, G. E. (2010). Learning to detect roads in high-resolution aerial images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6316 LNCS*(PART 6), 210–223. https://doi.org/10.1007/978-3-642-15567-3_16

Mohammadimanesh, F., Salehi, B., Mahdianpari, M., Gill, E., & Molinier, M. (2019). A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem. *ISPRS Journal of Photogrammetry and Remote Sensing*, *151*(March), 223–236. https://doi.org/10.1016/j.isprsjprs.2019.03.015

Musyoka, G. M., Tolpekin, V. ., & Persello, C. (2018). *Automatic Delineation of Small Holder Agricultural Field Boundaries Using Fully Convolutional Networks*. Unpublished. https://webapps.itc.utwente.nl/librarywww/papers_2018/msc/gfm/musyoka.pdf

Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, 689–696.

Nicolau, A. P., Flores-Anderson, A., Griffin, R., Herndon, K., & Meyer, F. J. (2021). Assessing SAR C-band data to effectively distinguish modified land uses in a heavily disturbed Amazon forest. *International Journal of Applied Earth Observation and Geoinformation*, *94*, 102214. https://doi.org/10.1016/j.jag.2020.102214

Ortega, M. X., Bermudez, J. D., Happ, P. N., Gomes, A., & Feitosa, R. Q. (2019). EVALUATION of DEEP LEARNING TECHNIQUES for DEFORESTATION DETECTION in the AMAZON FOREST. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*(2/W7), 121–128. https://doi.org/10.5194/isprs-annals-IV-2-W7-121-2019

Pacheco-Pascagaza, A. M., Gou, Y., Louis, V., Roberts, J. F., Rodríguez-Veiga, P., Bispo, P. da C., Espírito-Santo, F. D. B., Robb, C., Upton, C., Galindo, G., Cabrera, E., Cendales, I. P. P., Santiago, M. A. C., Negrete, O. C., Meneses, C., Iñiguez, M., & Balzter, H. (2022). Near Real-Time Change Detection System Using Sentinel-2 and Machine Learning: A Test for Mexican and Colombian Forests. *Remote Sensing*, *14*(3), 1–21. https://doi.org/10.3390/rs14030707

Paisitkriangkrai, S., Sherrah, J., Janney, P., & Van-Den Hengel, A. (2015). Effective semantic pixel labelling with convolutional networks and Conditional Random Fields. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2015-Octob*, 36–43. https://doi.org/10.1109/CVPRW.2015.7301381

Phuntsho. (2020). *Differentiating Healthy and Bark Beetle Infected Spruce Trees With Sentinel-1 Sar*. *July*.

PRODES. (2021). *Monitoring of Deforestation of the Brazilian Amazon Forest by Satellite*. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes

R.Pandit, V., & J. Bhiwani, R. (2015). Image Fusion in Remote Sensing Applications: A Review. *International Journal of Computer Applications*, *120*(10), 22–32. https://doi.org/10.5120/21263-3846

Rahimi, Z. (2020). *Combined Sentinel-1 and - 2 soil moisture retrieval for a corn and wheat field in Twente*.

Rashkovetsky, D., Mauracher, F., Langer, M., & Schmitt, M. (2021). Wildfire Detection from Multisensor

Satellite Imagery Using Deep Semantic Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 7001–7016. https://doi.org/10.1109/JSTARS.2021.3093625

Reynolds, A. H. (2022). *Convolutional Neural Networks (CNNs)*. https://anhreynolds.com/blogs/cnn.html

Riggio, C. (2019). *What's the deal with Accuracy, Precision, Recall and F1?* https://towardsdatascience.com/whats-the-deal-with-accuracy-precision-recall-and-f1-f5d8b4db1021

Rizaldy, A. (2018). *Deep Learning-Based Dtm Extraction From Lidar Point Cloud*.

Sentinelhub. (2022). *EO Browser*. https://apps.sentinel-hub.com/eo-browser/?zoom=10&lat=41.9&lng=12.5&themeId=DEFAULT-THEME&toTime=2022-06-05T06%3A44%3A35.165Z

Shrestha, S. (2018). *Detecting rice fields with different water sources using multi-temporal Sentinel-1A imagery in Central Luzon, Philippines*. *March*, 75.

Shumilo, L., Yailymov, B., Lavreniuk, M., & Bilokonska, Y. (2020). Remote Sensing Approaches for Deforestation Identification in Ukraine. *IDAACS-SWS 2020 - 5th IEEE International Symposium on Smart and Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems, Proceedings*, 17–20. https://doi.org/10.1109/IDAACS-SWS50031.2020.9297054

Sinha, S., Jeganathan, C., Sharma, L. K., & Nathawat, M. S. (2015). *A review of radar remote sensing for biomass estimation*. https://doi.org/10.1007/s13762-015-0750-0

Sirro, L., Häme, T., Rauste, Y., Kilpi, J., Hämäläinen, J., Gunia, K., de Jong, B., & Pellat, F. P. (2018). Potential of different optical and SAR data in forest and land cover classification to support REDD+ MRV. *Remote Sensing*, *10*(6). https://doi.org/10.3390/rs10060942

Torres, D. L., Turnes, J. N., Juan, P., Vega, S., Feitosa, R. Q., Silva, D. E., Junior, J. M., & Almeida, C. (2021). *Deforestation Detection with Fully Convolutional Networks in the Amazon Forest from Landsat-8 and Sentinel-2 Images*. 1–20.

Vakalopoulou, M., Karantzalos, K., Komodakis, N., & Paragios, N. (2015). Building detection in very high resolution multispectral data with deep learning features. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2015-Novem*, 1873–1876. https://doi.org/10.1109/IGARSS.2015.7326158

Vinod, N., & Hinton, G. E. (2017). Rectified Linear Units Improve Restricted Boltzmann Machines. *Journal of Applied Biomechanics*, *33*(5), 384–387. https://doi.org/10.1123/jab.2016-0355

Volpi, M., & Tuia, D. (2017). Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(2), 881–893. https://doi.org/10.1109/TGRS.2016.2616585

Wahab, M. A. A., Surin, E. S. M., & Nayan, N. M. (2021). *An Approach to Mapping Deforestation in Permanent Forest Reserve Using the Convolutional Neural Network and Sentinel-1 Synthetic Aperture Radar*. 59–64. https://doi.org/10.1109/camp51653.2021.9498144

Wu, R. (2021). Monitoring coastal wetlands changes using fusion of high-resolution SAR and optical images. *Cehui Xuebao/Acta Geodaetica et Cartographica Sinica*, *50*(2), 280. https://doi.org/10.11947/j.AGCS.2021.20190504

Yan, J. (2019). *Unsupervised Change Detection Technique Based on Fully Convolutional Network Using Rgbd*.

APPENDIX

Annex 1: Cloud Free Sentinel-1 Training and Validation Loss

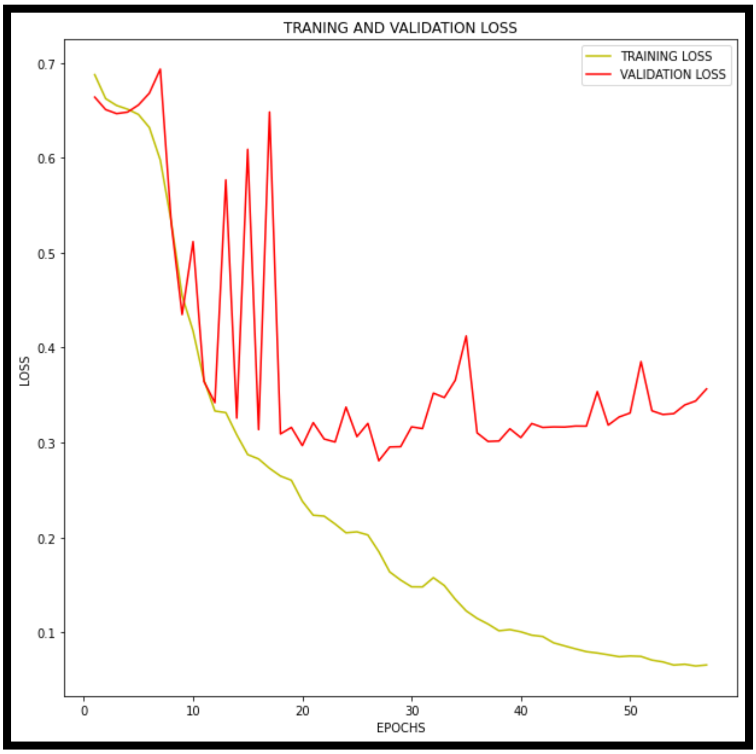


Figure 41: Cloud Free Sentinel-1 Training and Validation Loss

Annex 2: Cloud Free Sentinel-1 Training and Validation Accuracy

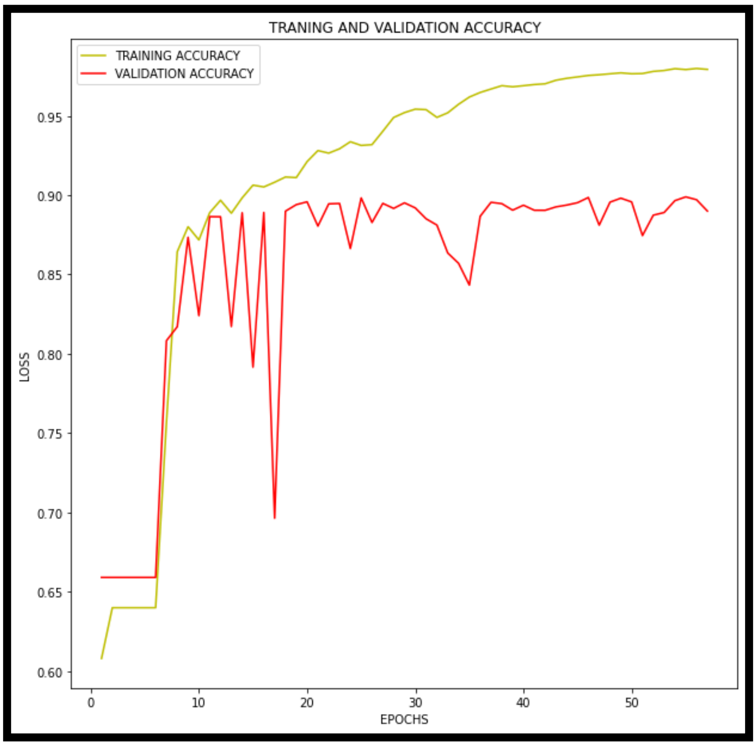


Figure 42: Cloud Free Sentinel-1 Training and Validation Accuracy

Annex 3: Cloud Free Sentinel-2 Training and Validation Loss

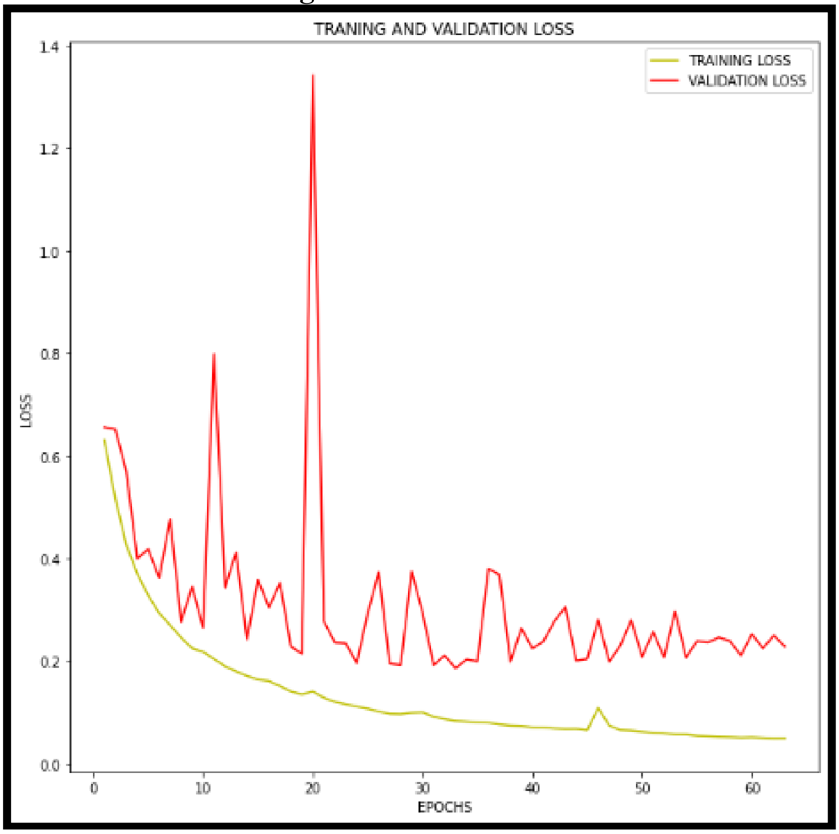


Figure 43: Cloud Free Sentinel-2 Training and Validation Loss

Annex 4: Cloud Free Sentinel-2 Training and Validation Accuracy

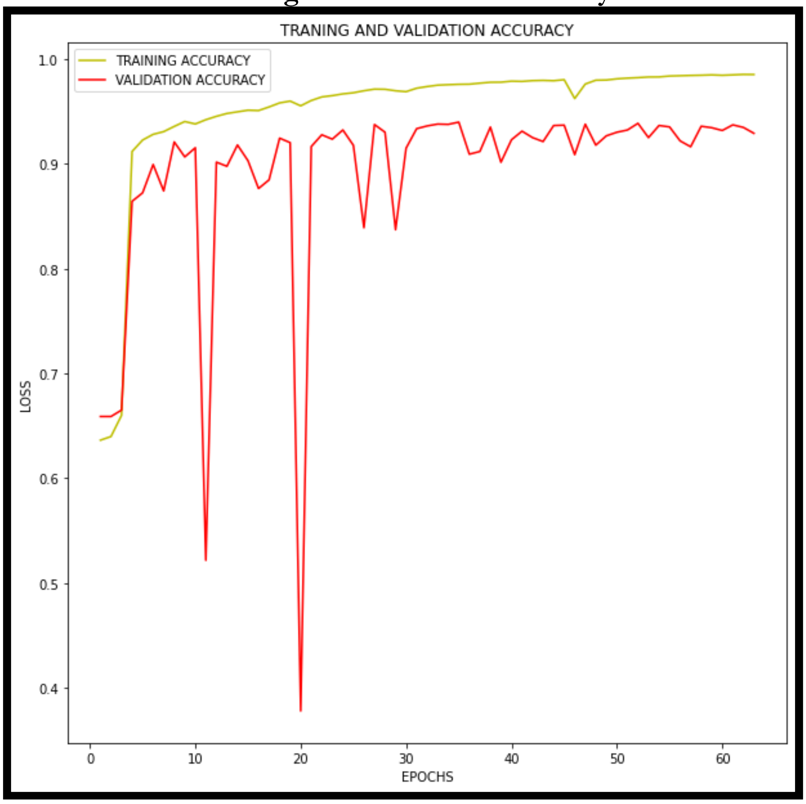


Figure 44: Cloud Free Sentinel-2 Training and Validation Accuracy

Annex 5: Cloud Free Late Fusion Training and Validation Loss

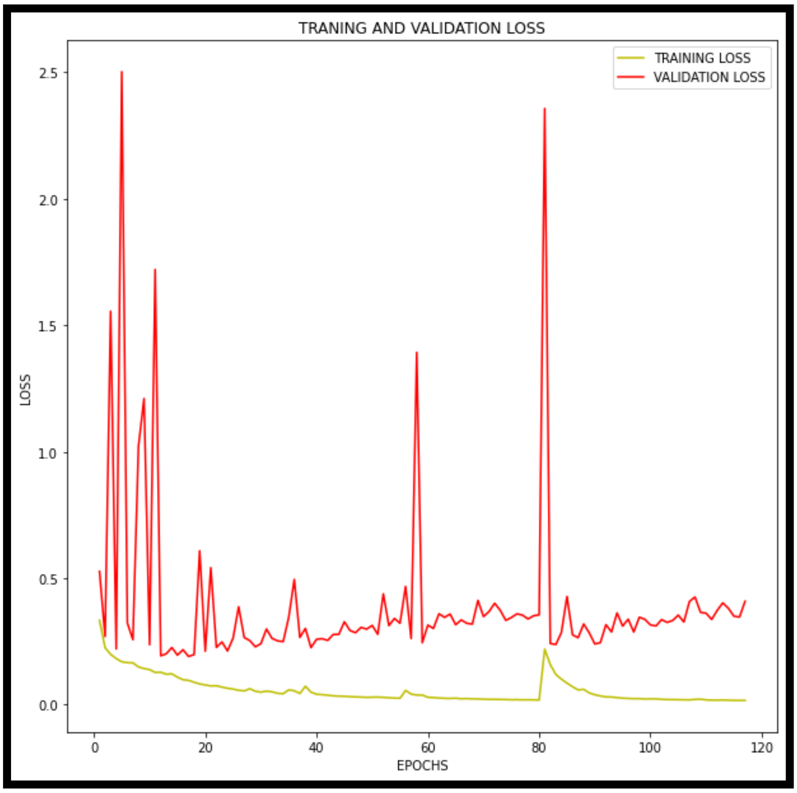


Figure 45: Cloud Free Late Fusion Training and Validation Loss

Annex 6: Cloud Free Late Fusion Training and Validation Accuracy

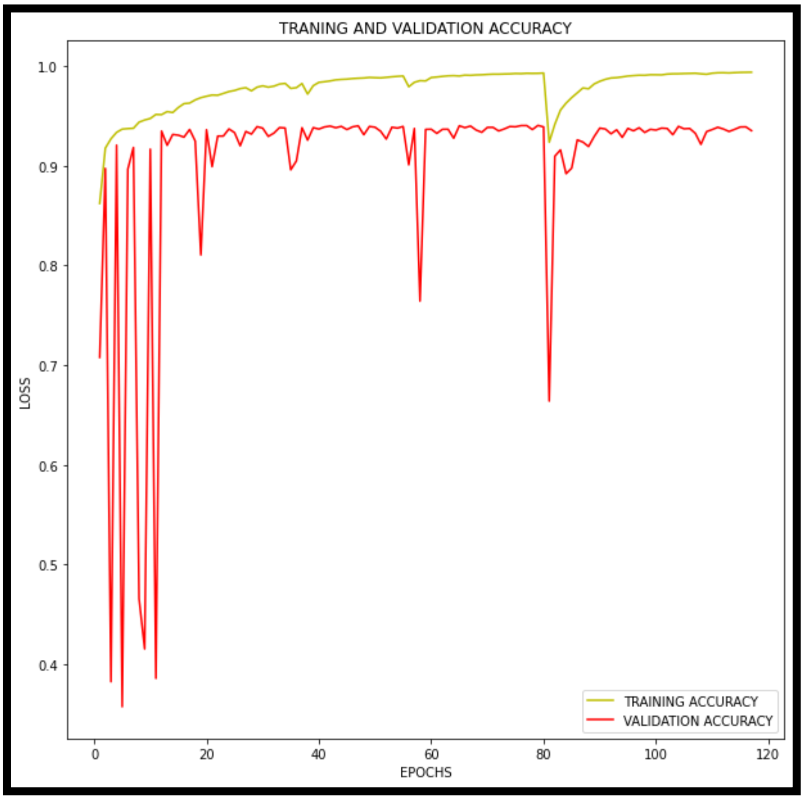


Figure 46: Cloud Free Late Fusion Training and Validation Accuracy

Annex 7: Cloudy Sentinel-1 Training and Validation Loss

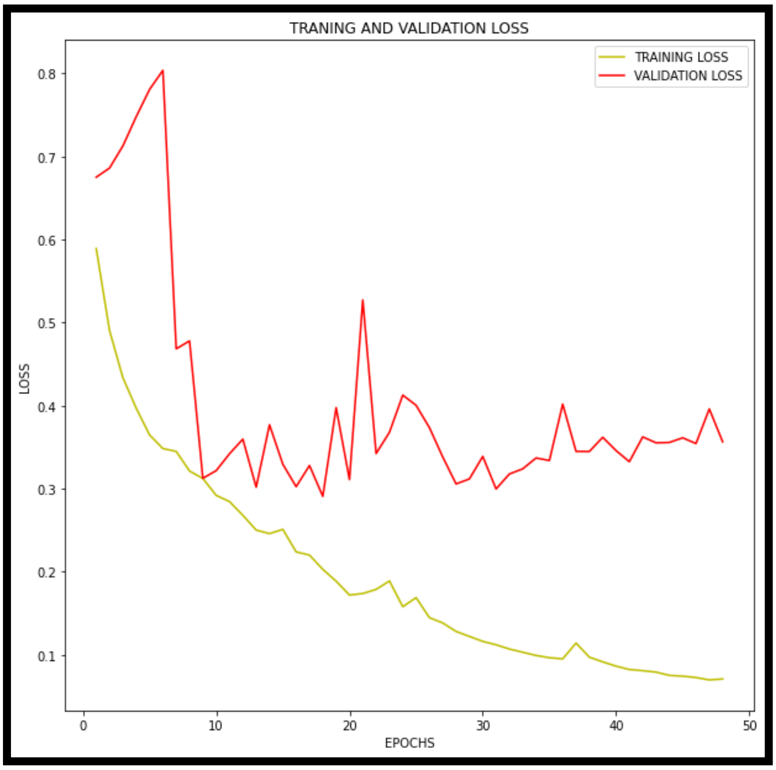


Figure 47: Cloudy Sentinel-1 Training and Validation Loss

Annex 8: Cloudy Sentinel-1 Training and Validation Accuracy

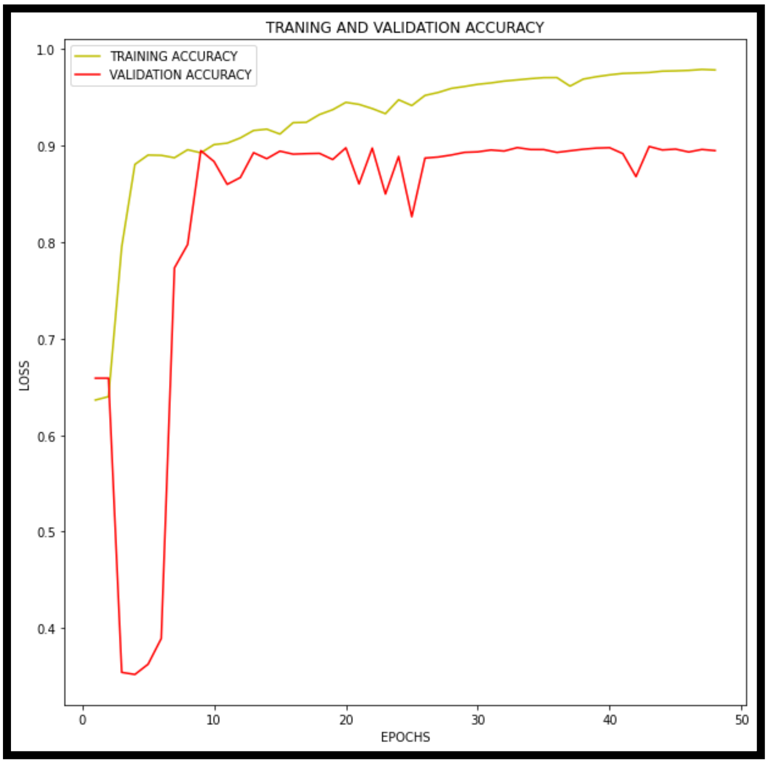


Figure 48: Cloudy Sentinel-1 Training and Validation Accuracy

Annex 9: Cloudy Sentinel-2 Training and Validation Loss

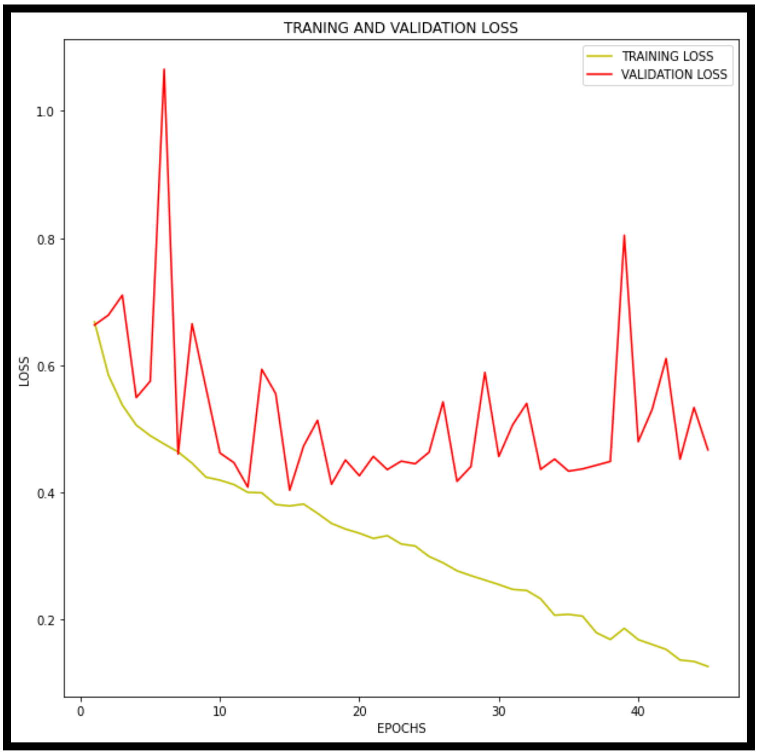


Figure 49: Cloudy Sentinel-2 Training and Validation Loss

Annex 10: Cloudy Sentinel-2 Training and Validation Accuracy

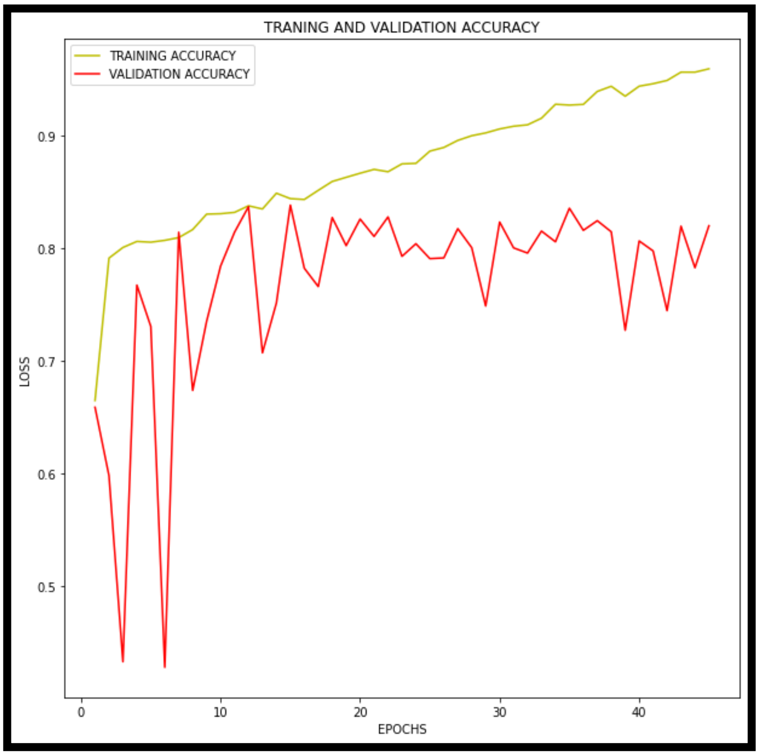


Figure 50: Cloudy Sentinel-2 Training and Validation Accuracy

Annex 11: Cloudy Late Fusion Training and Validation Loss

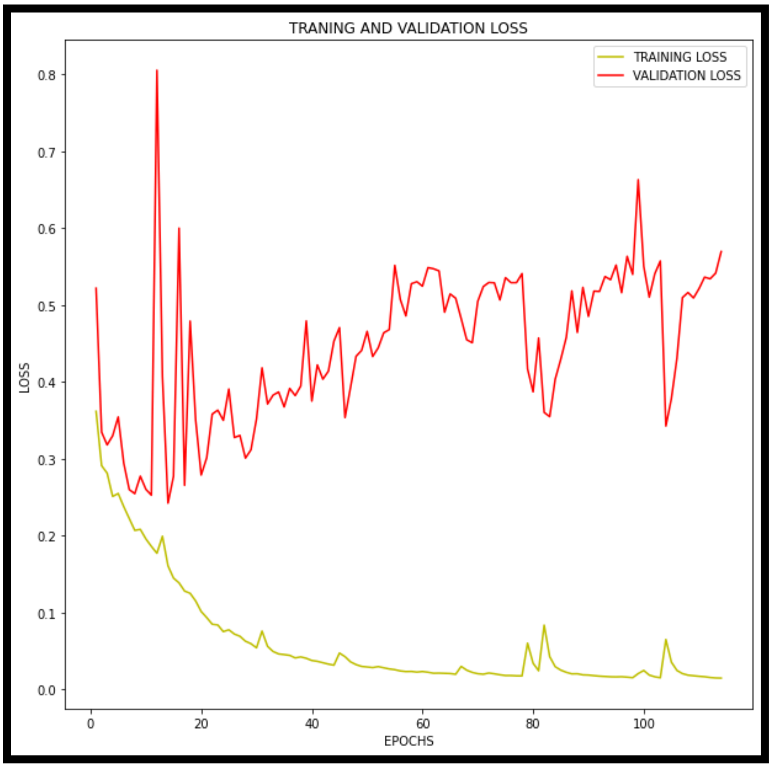


Figure 51: Cloudy Late Fusion Training and Validation Loss

Annex 11: Cloudy Late Fusion Training and Validation Accuracy

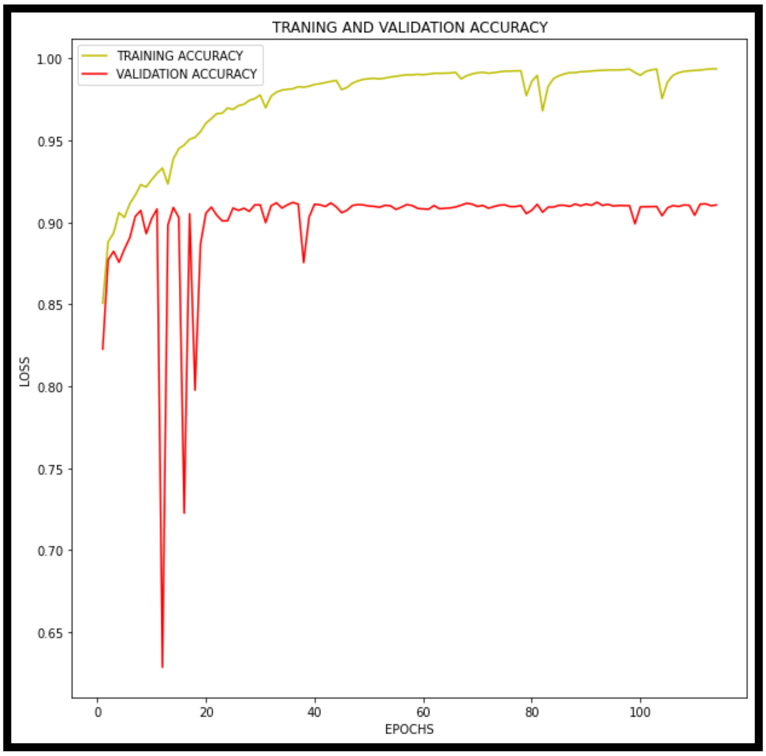


Figure 52: Cloudy Late Fusion Training and Validation Loss