

FOREST PLANTATION DETECTION THROUGH DEEP SEMANTIC SEGMENTATION

F. B. J. R. Dallaqua¹*, R. A. S. Rosa¹, B. Schultz¹, L. R. Faria¹, T. G. Rodrigues¹, C. G. Oliveira¹,
M. E. J. Kieser², V. Malhotra², T. Dwyer², D. S. Wolfe²

¹Visiona Tecnologia Espacial S.A., Estrada. Dr. Altino Bondensan, 500, 12247-016, São José dos Campos-SP, Brazil
(fernanda.dallaqua, rafael.rosa, bruno.schultz, lucas.faria, thiago.rodrigues, cleber.oliveira)@visionaespaical.com.br

²Tesera Systems Inc., PO Box 91073 RPO Royal Oak, Calgary, Alberta, Canada
(michael.kieser, vipul.malhotra, tonja.dwyer, dwight.wolfe)@tesera.com

Commission III, WG III/1

KEY WORDS: Deep learning, Forest plantation, Semantic segmentation, U-net, Sentinel-2, CBERS-4A, Forest inventory.

ABSTRACT:

Forest plantations play an important role ecologically, contribute to carbon sequestration and support billions of dollars of economic activity each year through sustainable forest management and forest sector value chains. As the global demand for forest products and services increases, the marketplace is seeking more reliable data on forest plantations. Remote sensing technologies allied with machine learning, and most recently deep learning techniques, provide valuable data for inventorying forest plantations and related valuation products. In this work, deep semantic segmentation with U-net architecture was used to detect forest plantation areas using Sentinel-2 and CBERS-4A images of different areas of Brazil. First, the U-net models were built from an area of the Centre-East of Paraná State, and then the best models were tested in 3 new areas that present different characteristics. The U-net models built with Sentinel-2 images achieved promising results for areas similar to the ones used in the training set, with F1-score ranging from 0.9171 to 0.9499 and with Kappa values between 0.8712 to 0.9272, demonstrating the feasibility of deep semantic segmentation to detect forest plantations.

1. INTRODUCTION

Forest plantations can be defined as planted forests that are intensively managed, with one or two species, native or exotic, even age class, and regular spacing (Food and Agriculture Organization, 2020). They are of significant economic importance, generating billions of dollars per year (Indústria Brasileira de Árvores, 2021), as a source of a wide variety of products such as wood panels, timber, pulp, paper, biomass, energy, charcoal, and others. According to the Global Forest Resources Assessment 2020 (Food and Agriculture Organization, 2020), the area of forest plantations was 131 million hectares, with the highest share being located in South America and the lowest in Europe. Globally, 44% of forest plantations' total area is composed of introduced species.

Brazil plays an important role in this segment, being the largest exporter of cellulose pulp to the global market and ranking among the 10 largest producers in the world regarding paper, lumber (9th), pulp (2nd), and charcoal (1st). Brazil's forest plantations are composed of the introduced species Eucalyptus, Pine, and Teak, and some native species such as Rubber, Acacia, Araucaria, and Paricá. In 2020, the forest plantations' total area was 9.55 million hectares, with 78% composed of Eucalyptus, 18% of Pine, and the rest of the other species. The states of Minas Gerais, São Paulo, Mato Grosso do Sul, Paraná, Rio Grande do Sul, and Santa Catarina are Brazil's leading producers of forest plantations (Indústria Brasileira de Árvores, 2021).

Remote sensing technologies provide valuable spatial and temporal data to a great variety of applications regarding Earth Observation, which includes the forest plantations sector, whether

in forest mapping, biomass and age estimation, change detection, and others (Trisasonoko and Paull, 2020). Many of these applications rely on machine learning techniques (Dang et al., 2019, Dube et al., 2014, Sibanda et al., 2021, Meng et al., 2022), with its subgroup deep learning drawing attention as it achieves excellent performances and improves information extraction from images (Martins et al., 2021, Cui et al., 2020).

Deep learning is based on neural networks. A neural network is composed of the following layers: the first layer that receives the input data (input layer), the hidden layers, and the output layer. If a neural network contains multiple hidden layers, it is considered a deep neural network, which explains the term deep learning. Some deep learning models are convolutional neural networks, recurrent neural networks, autoencoders, and generative adversarial networks (Ma et al., 2019).

The convolutional neural networks have multiple feature-extraction stages subdivided into three types of layers: (a) convolutional layers, (b) pooling layers, and (c) fully connected layers (Ma et al., 2019). They have been successfully used in various fields, including remote sensing (Zhu et al., 2017, Ma et al., 2019), due to their ability to automatically learn feature representations through training with minimal knowledge of the task (Lecun et al., 1998). They have been applied in a variety of remote sensing image analysis tasks such as image fusion, scene classification, object detection, object-based image analysis (OBIA), land use and land cover (LULC) classification, and semantic segmentation (Ma et al., 2019).

Semantic segmentation aims to generate a pixel-wise classification of images. For remote sensing, the state-of-the-art frameworks are typically composed of encoder and decoder subnetworks (Ma et al., 2019), with U-net (Ronneberger et al., 2015)

* Corresponding author

and Fully Convolutional Network (Long et al., 2014) being examples of architectures for semantic segmentation. U-net summarizes patterns in both the spectral and spatial domain. Besides, it works well with small training datasets, which is advantageous for the remote sensing field, where gathering ground truth annotations can be time-consuming and expensive (Ronneberger et al., 2015, Solórzano et al., 2021, Illarionova et al., 2021).

Deep semantic segmentation with U-net could be used to map forest plantations. In (da Costa et al., 2021) 24 models based on combinations between six architectures, including U-net, and four encoders, were compared regarding Eucalyptus plantations' mapping in Sentinel-2 images with 10 bands. As per (Wagner et al., 2019) very high resolution red-green-blue (RGB) images from the WorldView-3 satellite (0.3 meter of spatial resolution) and U-net were used to segment natural forest and Eucalyptus plantation in the Brazilian Atlantic Forest region.

The goal of this work was to detect forest plantation areas through remote sensing imagery and deep semantic segmentation with U-net architecture. This is the first necessary step in the development of a remote forest plantation inventory and valuation product. The remainder of this work is organized as follows. Section 2 presents the methodology regarding the study areas, the remote sensing images, and the U-net models, Section 3 exhibits the results of the performed experiments, Section 4 concludes the paper and Section 5 highlights some of the next steps of this research.

2. METHODOLOGY

2.1 Study areas

The study areas are divided into two groups: the first was used to train and define the best U-net models, and these best models classified the areas in the second group. Figure 1 presents the study areas, with study area A belonging to the first group and the remaining study areas (B, C, and D) in the second group.

2.1.1 Definition of best models Study area A is shown in more detail in Figure 2. It is in the Centre-East of Paraná State, near Telêmaco Borba and Ibaiti counties. Most of the area is located in the Atlantic Forest biome, which is characterized by ombrophilous (dense, open, and mixed) and seasonal (semideciduous and deciduous) forests. The remaining area is located in the Cerrado biome, which presents forest and countryside formations, with savanna being the most expressive, and its most common physiognomy is sparse trees and shrubs on a grassy carpet (Instituto Brasileiro de Geografia e Estatística, 2019). Its Köppen's climate types are Cfa (Humid subtropical zone with oceanic climate, without dry season, and with hot summer) and Cfb (Humid subtropical zone with oceanic climate, without dry season, and with temperate summer) (Alvares et al., 2013). It presents a flat to slightly undulating topography, and it is mainly composed of forest plantations areas, being the majority of Eucalyptus and the rest of Pine.

Thirteen grids, highlighted in Figure 2, were chosen in which nine are $2 \times 2 km^2$ and four are $2 \times 0.444 km^2$. For the Sentinel-2 dataset, all of these grids were used, being the vast majority of the area (around 95%) composed of the Atlantic Forest biome and the rest being Cerrado. For the CBERS-4A dataset, two grids were used, both located within the Atlantic Forest biome.

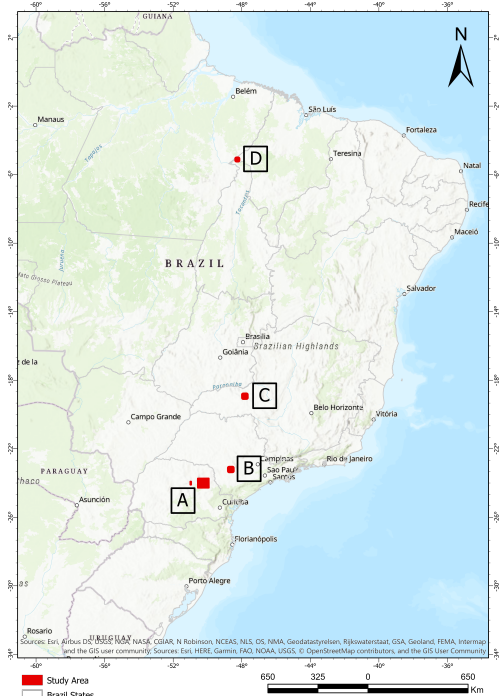


Figure 1. Location map of the study areas.

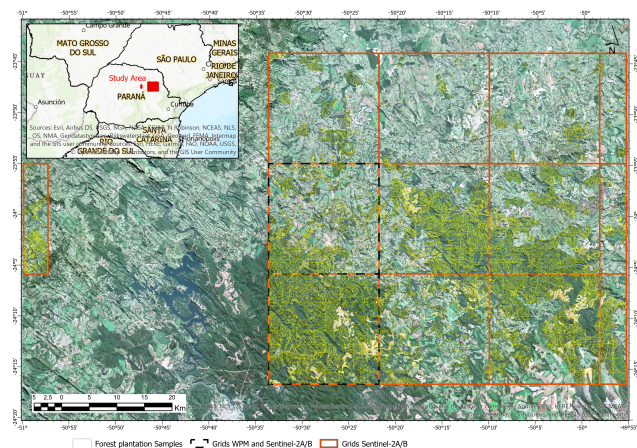


Figure 2. Sentinel-2's RGB composition image of the study area. The grids used to define the best U-net models are highlighted, being 13 for the Sentinel-2 dataset and 2 for CBERS-4A.

2.1.2 Applying best models in new areas Three new areas (B, C, and D), with Eucalyptus plantations, were chosen to attest to the robustness of the generated models, with Table 1 showing their geographical boundaries (coordinates of the upper left and lower right corners). They vary in biome, topography, and climate as can be seen in Table 2. Area B is in the municipality of Itatinga, in the State of São Paulo, which is in Cerrado (majority of the area) and Atlantic Forest biomes. It has a flat to slightly undulating topography, its climate type is Cfa (Alvares et al., 2013), and besides forest plantations, it presents sugar cane, coffee, and orange plantations. Area C is a flat area located in Indianópolis city (State of Minas Gerais), in the Cerrado biome, presenting a Cwb (Humid subtropical zone with dry winter and temperate summer) climate (Alvares et al., 2013) and also having coffee plantations.

Area D is in the municipality of São Pedro da Água Branca, in the State of Maranhão. It is in the Amazon biome, the most extensive biome of Brazil, which is mostly composed of dense ombrophilous forest (Instituto Brasileiro de Geografia e Estatística, 2019). It presents a flat topography, and its climate type is a Tropical zone with dry winter (Aw) (Alvares et al., 2013).

Area	Latitude	Longitude
B	23°07'18.97"S	48°44'12.80"W
	23°17'58.24"S	48°32'18.31"W
C	18°50'42.76"S	47°54'47.94"W
	19°01'20.75"S	47°43'12.86"W
D	05°04'00.53"S	48°18'25.82"W
	05°09'56.47"S	48°12'27.80"W

Table 1. Coordinates of the upper left and lower right corners of the new areas.

2.2 Remote sensing images

The images used in this work are from sensors Multispectral Instrument (MSI) of Sentinel-2A/2B satellites and Multispectral camera and Panchromatic Wide (WPM) of China-Brazil Earth Resources Satellite-4A (CBERS-4A). The MSI sensor has a radiometric resolution of 12 bits and acquires 13 spectral bands from visible and near-infrared (VNIR) to shortwave infrared (SWIR) with different spatial resolutions of 10 (red, green, blue, and near-infrared bands), 20 (red-edge and shortwave infrared bands) and 60 meters (atmospheric correction bands) (European Space Agency, 2015). The WPM sensor acquires 4 VNIR bands (red, green, blue, and near-infrared) with a spatial resolution of 8 meters and a panchromatic band with 2 meters spatial resolution. The radiometric resolution of its images is 10 bits (dos Santos et al., 2022).

Satellite	Area			
	A	B	C	D
Sentinel-2	2020-06-21	2021-07-23	2021-04-24	2021-08-08
CBERS-4A	2020-07-09	2021-07-26	2021-04-24	2021-07-01

Table 3. Acquisition dates for the images used in this work.

Table 3 shows the dates of the acquired images. The images from MSI were at Level 1C, with radiometric processing and geometric correction, and the bands with 20 meters spatial resolution were resampled to 10 meters with the default resampling method (nearest neighbor) in Geospatial Data Abstraction Library (GDAL) (GDAL/OGR contributors, 2022), totaling 10 bands that were used in this work. As the data collected from the WPM sensor still doesn't have radiometric correction (dos Santos et al., 2022), a linear regression was made to standardize WPM data regarding MSI. This was made by collecting MSI and WPM images from the same area and around the same date, sampling approximately 60 pixels, and finding the slope and intercept of the regression line. For WPM's panchromatic band, the corresponding MSI pixels were calculated by the mean of bands blue, green, red, and red-edge 4. The regression lines for the performed radiometric calibration of WPM bands are shown in Equation 1.

$$\begin{aligned}
 blue_{calibrated} &= 5.8432 * blue_{WPM} + (-120.9271) \\
 green_{calibrated} &= 6.4327 * green_{WPM} + (-91.8606) \\
 red_{calibrated} &= 6.2849 * red_{WPM} + (-151.7777) \\
 nir_{calibrated} &= 7.3931 * nir_{WPM} + (-264.7779) \\
 pan_{calibrated} &= 5.6297 * pan_{WPM} + 4.1212
 \end{aligned} \tag{1}$$

The pixels values of each band were transformed to range between 0 to 255. For MSI, two input images with 10 meters spatial resolution were created: one with Landsat-like data where the bands Blue, Green, Red, and Near-Infrared (BGRNir) were chosen, and the other used the 10 available bands (BGRNir plus the four Red-Edge and the two shortwave infrared bands). As for WPM, a BGRNir image was created, and then a fusion (pansharpening operation in GDAL, with nearest as resampling method) between the panchromatic band and this image was performed, resulting in a BGRNir image with 2 meters spatial resolution.

The grids of the study area A (Telêmaco Borba) were divided into smaller images of size 256×256 pixels with minimal overlap, resulting in 640 images for the Sentinel-2 dataset and 3,200 images for CBERS-4A. These images were split into approximately 70% training, 10% validation, and 20% test. To build these sets, the following steps were taken: the proportion of forest plantation pixels was calculated for each 256×256 image; the images were divided into quartiles according to their proportion, and random selection was performed for each quartile to define in which set (training, validation or test) each image would be added.

The training set from Sentinel-2 was augmented by rotation and flip, totaling 3,400 images. As for CBERS-4A, one experiment used its original training set, and the other experiment augmented the training set by rotating the images in 180 degrees, duplicating the number of images (4,478). For all the study areas, the ground truths were built by visual interpretation of the remote sensing images' landcover and manual delineation of forest plantations polygons. Approximately eighty hours of expert work were required to construct the ground truth annotations.

2.3 U-net models

U-net is a Fully Convolutional Neural Network created to segment biomedical images. It has an encoder-decoder architecture that resembles a U shape. In the encoder part, downsampling layers reduce the spatial resolution of the image, and its features are extracted. At the decoder occurs upsampling of the feature map, restoring the image's original dimensions and enabling the pixel-wise classification. This architecture captures the context in the encoder and precise localization in the decoder (Ronneberger et al., 2015).

Two U-net implementations were used to build the models. One is from the Segmentation Models Pytorch repository (Yakubovskiy, 2019) and for its encoder it is used a pre-trained Convolutional Neural Network (backbone) named Efficient-net-b7 (Tan and Le, 2019) (eff7). This backbone was chosen as it achieved the best results to detect eucalyptus plantation areas using Sentinel-2 images in (da Costa et al., 2021). As eff7 has weights trained on the 2012 ILSVRC ImageNet dataset (Deng et al., 2009), it is expected for the model to receive an input image with 3 channels. However, our input images have 4 or 10

Area	City	State	Biome	Size (km^2)	Climate type	Topography
B	Itatinga	São Paulo	Cerrado/Atlantic Forest	2×2	Cfa	Flat/slightly undulating
C	Indianópolis	Minas Gerais	Cerrado	2×2	Cwb	Flat
D	São Pedro da Água Branca	Maranhão	Amazon	1.1×1.1	Aw	Flat

Table 2. Summary of the new areas.

Satellite	Modelname	Accuracy	Precision	Recall	F1-score	Jaccard score	Kappa
Sentinel-2	unetv4_4bands	0.9517	0.8612	0.8564	0.8588	0.7526	0.8297
	smv4_eff7_4bands	0.9363	0.8405	0.7758	0.8069	0.6763	0.7688
	unetv4_10bands	0.9574	0.8775	0.8734	0.8754	0.7785	0.8497
	smv4_eff7_10bands	0.9471	0.8612	0.8241	0.8422	0.7275	0.8105
CBERS-4A	unetv4_4bands	0.9526	0.9364	0.9194	0.9278	0.8654	0.8925
	unetv4_4bands_180degrees	0.9514	0.9270	0.9263	0.9266	0.8633	0.8903

Table 4. Results of the evaluation metrics for the experiments to define the best models.

bands. So, for the models that use this U-net implementation, it was needed to map the bands to 3 channels through an extra convolutional layer. Also, the encoder weights were frozen, so only the decoder weights were trainable.

The other is an implementation of U-net without pre-trained weights. For all models binary cross-entropy was used as loss function and the hyperparameters were: (a) 150 epochs; (b) Adam optimizer; (c) 0.0001 learning rate; (d) sigmoid as activation function, and (e) 1 batch size. To prevent overfitting early stopping was applied, and the models with the best loss in the validation set were saved. The models regarding the repository will be referred to as *smv4_eff7* whereas the others will be *unetv4*.

For the study areas B, C, and D, a mosaicking technique implementation was used to reduce edge errors, a known problem for the U-net architecture (Ronneberger et al., 2015). This implementation has a classification window of size 256×256 pixels that go through the image without overlapping. Six pixels of each window's extremity are disregarded, so only the central 244×244 pixels' classification is taken into account. This procedure occurs five times, each of them starting in a different pixel of the image ($\{x = 0, y = 0\}$, $\{x = 64, y = 64\}$, $\{x = 128, y = 128\}$, $\{x = 192, y = 192\}$, and $\{x = 256, y = 256\}$). Then, the window is applied to missing pixels from the remaining rows and columns, making every pixel of the image be classified five times with probability outputs between 0 and 1. The final pixel value will be the median of the probabilities.

As the U-net models' outputs are values between zero and one, the definition if a pixel is forest plantation or not is given by considering a threshold in which values higher than the threshold is forest plantation and background otherwise. For this work, the conventional threshold of 0.5 was used. The performance of models was assessed through six evaluation metrics: overall accuracy, precision, recall, F1-score, Jaccard similarity coefficient score (Jaccard, 1912), and Kappa (Cohen, 1960).

3. RESULTS

3.1 Definition of best models

Table 4 presents the results for the evaluation metrics. For the Sentinel-2 dataset, all models presented F1-score higher than 0.80 and all except one presented Kappa higher than 0.81 with *unetv4* obtaining the best results for both BGRNir (4 bands)

and 10 bands. For the two U-net implementations (*unetv4* and *smv4_eff7*) the results were improved when 10 bands were used which could be explained by the vegetation characteristics' information of the red-edge bands (Schuster et al., 2012, Immitzer et al., 2016).

Satellite	Modelname	Size of training set	Epoch Time (s)
Sentinel-2	unetv4_4bands	3,400	384
	smv4_eff7_4bands	3,400	560
	unetv4_10bands	3,400	397
	smv4_eff7_10bands	3,400	597
CBERS-4A	unetv4_4bands	2,239	268
	unetv4_4bands_180degrees	4,478	517

Table 6. Mean training time per epoch. The experiments were performed with NVIDIA GPU (Quadro P4000).

Although studies demonstrated that pre-trained Convolutional Neural Networks can perform better than deep learning networks trained from scratch and are well suited for remote sensing image's classification and semantic segmentation (Pan et al., 2019, Cui et al., 2020, Marmanis et al., 2016), in this work these models had the worst results maybe because of the extra convolutional layer that was added to map the bands to 3 channels which results in spectral information loss (Pan et al., 2019). These models also consumed more training time per epoch, as can be seen in Table 6. So, these models weren't considered for CBERS-4A dataset experiments.

F1-score and Kappa were higher than 0.92 and 0.89, respectively, for the CBERS-4A dataset. The model with an augmented training set (*unetv4_4bands_180degrees*) did not improve the results, being very similar to the results of the original training set, and its training time per epoch was almost twice that of the other model.

3.2 Applying best models in new areas

Figure 3 and Table 5 present the results for the classification of the study areas B, C, and D by Sentinel-2's models *unetv4_4bands* and *unetv4_10bands*. For areas B and C, *unetv4_10bands* had the best results for all evaluation metrics except Precision for area B. By analyzing the classified images, it is possible to notice that apparently, *unetv4_4bands* was better to detect the borders between forest plantation polygons for all areas, and

Area	Modelname	Accuracy	Precision	Recall	F1-score	Jaccard score	Kappa
B	unetv4_4bands	0.9409	0.9016	0.9332	0.9171	0.8469	0.8712
	unetv4_10bands	0.9484	0.8992	0.9604	0.9288	0.8671	0.8884
C	unetv4_4bands	0.9561	0.9222	0.9361	0.9291	0.8676	0.8973
	unetv4_10bands	0.9688	0.9369	0.9632	0.9499	0.9045	0.9272
D	unetv4_4bands	0.8043	0.8456	0.5766	0.6857	0.5217	0.5508
	unetv4_10bands	0.8065	0.9741	0.4904	0.6524	0.4841	0.5378

Table 5. Results of the evaluation metrics for the classification of Sentinel-2 images from study areas B, C, and D.

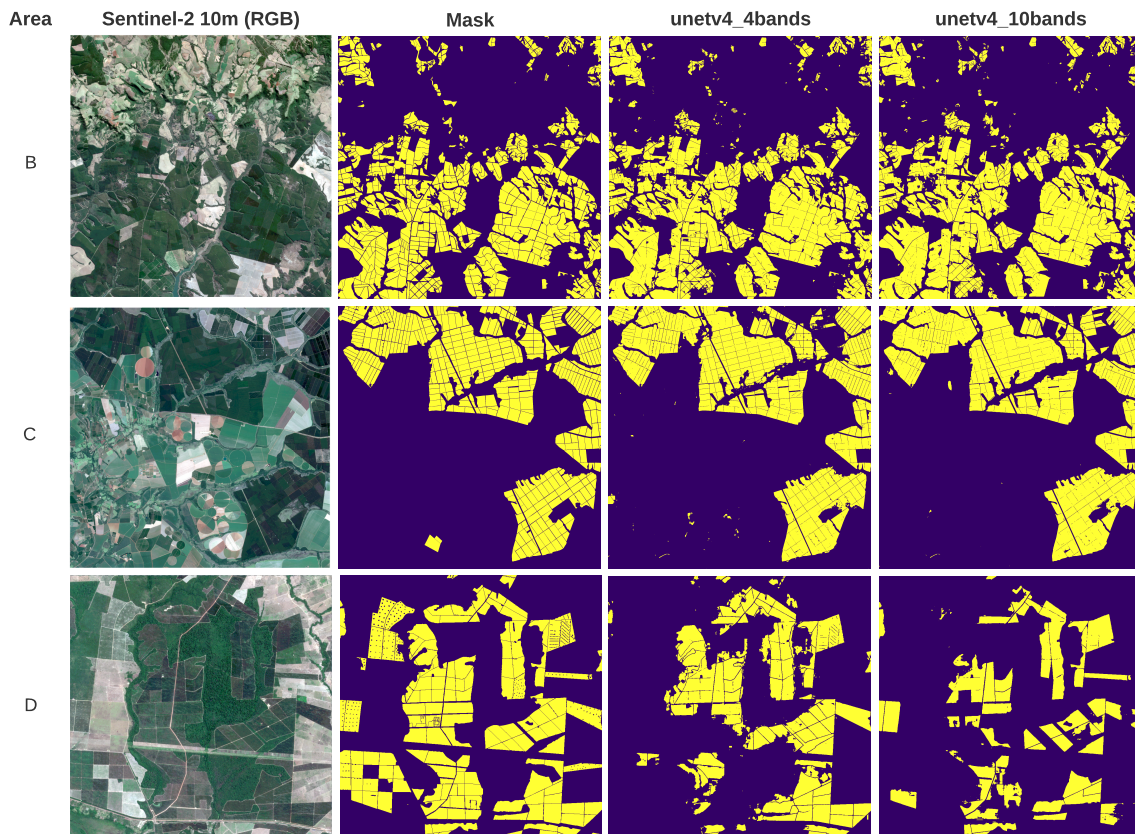


Figure 3. Study areas B, C, and D classified by Sentinel-2’s models unetv4_4bands and unetv4_10bands with a threshold of 0.5.

for areas B and C, unetv4_10bands was able to recognize some forest plantation areas dismissed by the other model.

For area D the Sentinel-2 models had very poor results, with Kappa around 0.55. Both models, especially the one with 10 bands, had good Precision, showing a low rate of false positives but disregarding many forest plantation polygons (very low Recall). Many areas of D are composed of forest plantations with open canopy (background effect) which means that the crowns of the trees do not overlap. As our models were trained with the majority of polygons being with closed canopy, they presented difficulty in classifying forest plantations in this study area. This difficulty also happened with areas B and C but these open canopy plantations were the minority and didn’t impact significantly the results.

CBERS-4A’s model presented the best evaluation metrics regarding study area A (Section 3.1). However, as can be seen in Figure 4 and Table 7, it had unsatisfactory results both visually and quantitatively for the new areas B and C when compared with Sentinel-2’s models. As for area D, it had the best results than the other models.

	Area		
	B	C	D
Accuracy	0.8943	0.8578	0.8636
Precision	0.7904	0.6956	0.7583
Recall	0.9656	0.9803	0.9463
F1-score	0.8692	0.8138	0.8420
Jaccard score	0.7687	0.6860	0.7270
Kappa	0.7820	0.7041	0.7246

Table 7. Results of the evaluation metrics for the classification of CBERS-4A images from study areas B, C, and D by unetv4_4bands model.

The best metric was Recall, with values higher than 0.94, showing that the model is correctly classifying a high rate of forest plantations pixels. However, its low Precision shows that the model labels a considerable rate of non-forest plantations areas as forest plantations. For B, sugar cane and native forest areas were mistaken as forest plantations. For area C, coffee plantations were misclassified, whereas, for D, the wrong classifications happened in native forest areas.

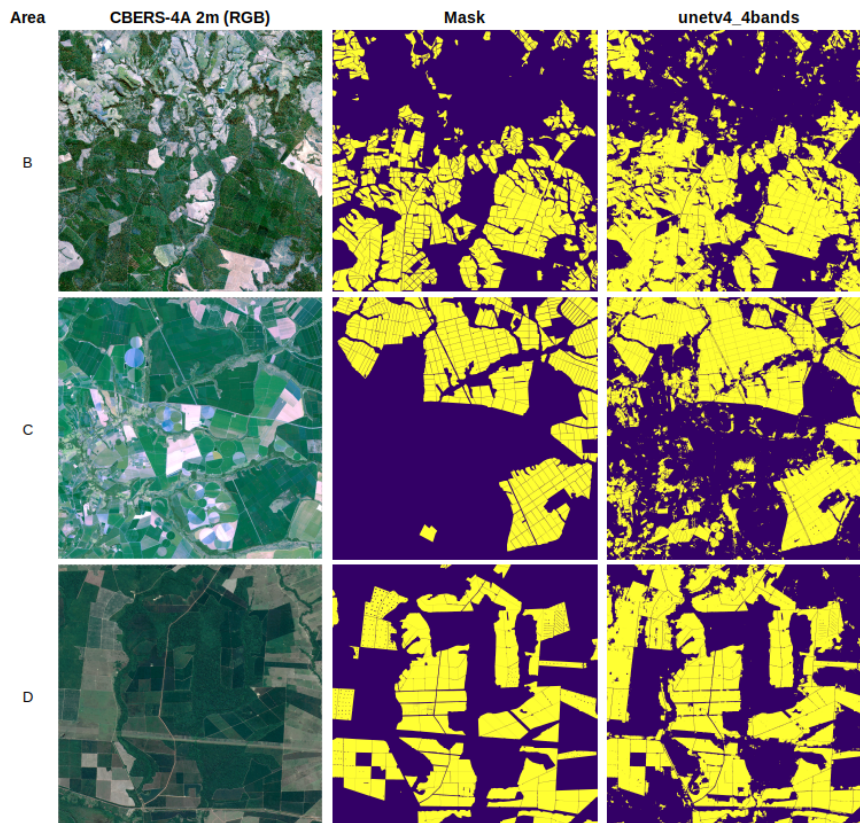


Figure 4. Study areas B, C, and D classified by CBERS-4A's model unetv4.4bands with a threshold of 0.5.

CBERS-4A's model was able to detect the forest plantations areas with open canopy that the Sentinel-2's models had difficulty classifying. A possible explanation for this could be the spatial resolution of the images used for each model. As the images from CBERS-4A have 2 meters spatial resolution, they carry more details, being able to detect background effect in areas that in Sentinel-2's images (10 meters spatial resolution) appear to be closed canopy. With that, CBERS-4A's model was more sensitive to this type of areas, detecting the open canopy forest plantations but also wrongly classifying agriculture areas. As for the native forest areas that were mistaken, it was noticed that these areas have a polygon pattern which could have misled the model. More studies are needed to improve this model.

4. CONCLUSION

This work used remote sensing imagery from Sentinel-2 and CBERS-4A and deep semantic segmentation with U-net architecture to detect forest plantations in different areas of Brazil. These areas varied in biome, topography, and climate. The use of 10 bands from Sentinel-2 improved the evaluation metrics. Models with pre-trained encoders had the worst results as it was needed to map the bands into 3 channels with an extra convolutional layer, resulting in spectral information loss.

Although CBERS-4A achieved good results regarding the evaluation metrics for study area A, it considered many pixels as forest plantations in the other areas, having unsatisfactory results for areas B and C. For area D, although it wrongly classified some native forest pixels, it was able to detect the forest plantations areas with open canopy that Sentinel-2's models had difficulty classifying.

The models built with Sentinel-2 images achieved good evaluation metrics for forest plantations areas similar to the ones used in the training set (closed canopy). With improvements in the training set, open canopy areas may be correctly classified. Generally, the results look promising, demonstrating the feasibility of deep semantic segmentation to detect forest plantations, which could be an important alternative to support forest plantations' management.

5. FUTURE WORK

For future work can be highlighted the improvement of the training set, adding samples from other areas, with different characteristics regarding biome, topography, climate, forest plantations species and ages, and non-forest plantations areas (agriculture, native forest, and others). Studies will be made regarding the classification of forest plantations areas with open canopy. Also, it is planned the classification of forest plantation species and the use of time series for age estimation.

As access to training data, ground truth annotations, is always a challenge, a semi-automated methodology used to successfully build training data on smaller scale projects will be tested: automated landscape polygon segmentation and random selection, feature calculation, k-means clustering and initial labelling, followed by manual human label adjustment.

ACKNOWLEDGEMENTS

The authors would like to thank the São Paulo Research Foundation (FAPESP) and the National Research Council of Canada's Industrial Research Assistance Program (NRC IRAP). This work

is part of the International Consortium Canada/Brazil Collaborative R&D Project of NRC IRAP entitled “Develop forest market intelligence remote sensing solution using Tesera’s High Resolution Inventory Solutions” “Project Number 957110” and FAPESP entitled “Solução para avaliação de florestas por sensoriamento remoto” (Remote sensing forest valuation product) number “2020/01468-0 - NRC - PIPE - Fase 2 Direta”.

REFERENCES

- Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. d. M., Sparovek, G. et al., 2013. Köppen’s climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6), 711–728.
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>.
- Cui, B., Chen, X., Lu, Y., 2020. Semantic Segmentation of Remote Sensing Images Using Transfer Learning and Deep Convolutional Neural Network With Dense Connection. *IEEE Access*, 8, 116744-116755.
- da Costa, L. B., de Carvalho, O. L. F., de Albuquerque, A. O., Gomes, R. A. T., Guimarães, R. F., de Carvalho Júnior, O. A., 2021. Deep semantic segmentation for detecting eucalyptus planted forests in the Brazilian territory using sentinel-2 imagery. *Geocarto International*, 0(0), 1-13. <https://doi.org/10.1080/10106049.2021.1943009>.
- Dang, A. T. N., Nandy, S., Srinet, R., Luong, N. V., Ghosh, S., Senthil Kumar, A., 2019. Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park, Vietnam. *Ecological Informatics*, 50, 24-32.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.
- dos Santos, B. D., de Pinho, C. M. D., Oliveira, G. E. T., Korting, T. S., Escada, M. I. S., Amaral, S., 2022. Identifying Precarious Settlements and Urban Fabric Typologies Based on GEOBIA and Data Mining in Brazilian Amazon Cities. *Remote Sensing*, 14(3). <https://www.mdpi.com/2072-4292/14/3/704>.
- Dube, T., Mutanga, O., Elhadi, A., Ismail, R., 2014. Intra-and-Inter Species Biomass Prediction in a Plantation Forest: Testing the Utility of High Spatial Resolution Spaceborne Multispectral RapidEye Sensor and Advanced Machine Learning Algorithms. *Sensors*, 14(8), 15348–15370. <https://www.mdpi.com/1424-8220/14/8/15348>.
- European Space Agency, 2015. Sentinel-2 User Handbook. *Standard document, Issue 1 Rev 2*.
- Food and Agriculture Organization, 2020. Global Forest Resources Assessment 2020: Main report. <https://doi.org/10.4060/ca9825en>. Accessed 2022-March-22.
- GDAL/OGR contributors, 2022. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation.
- Illarionova, S., Nesteruk, S., Shadrin, D., Ignatiev, V., Pukalchik, M., Oseledets, I., 2021. Object-based augmentation for building semantic segmentation: Ventura and santa rosa case study. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 1659–1668.
- Immitzer, M., Vuolo, F., Atzberger, C., 2016. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing*, 8(3). <https://www.mdpi.com/2072-4292/8/3/166>.
- Indústria Brasileira de Árvores, 2021. 2021 Ibá Annual Report. <https://iba.org/datafiles/publicacoes/relatorios/relatorioiba2021-compactado.pdf>. Accessed 2022-March-23.
- Instituto Brasileiro de Geografia e Estatística, 2019. Biomas e sistema costeiro-marinho do brasil. <https://www.ibge.gov.br/apps/biomas/home>. Accessed 2022-March-24.
- Jaccard, P., 1912. The distribution of the flora in the alpine zone.1. *New Phytologist*, 11(2), 37-50. <https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8137.1912.tb05611.x>.
- Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Long, J., Shelhamer, E., Darrell, T., 2014. Fully Convolutional Networks for Semantic Segmentation. *CoRR*, abs/1411.4038. <http://arxiv.org/abs/1411.4038>.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B. A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166-177. <https://www.sciencedirect.com/science/article/pii/S0924271619301108>.
- Marmanis, D., Datcu, M., Esch, T., Stilla, U., 2016. Deep Learning Earth Observation Classification Using ImageNet Pre-trained Networks. *IEEE Geoscience and Remote Sensing Letters*, 13(1), 105-109.
- Martins, J. A. C., Nogueira, K., Osco, L. P., Gomes, F. D. G., Furuya, D. E. G., Gonçalves, W. N., Sant’Ana, D. A., Ramos, A. P. M., Liesenberg, V., dos Santos, J. A., de Oliveira, P. T. S., Junior, J. M., 2021. Semantic Segmentation of Tree-Canopy in Urban Environment with Pixel-Wise Deep Learning. *Remote Sensing*, 13(16). <https://www.mdpi.com/2072-4292/13/16/3054>.
- Meng, Y., Wei, C., Guo, Y., Tang, Z., 2022. A Planted Forest Mapping Method Based on Long-Term Change Trend Features Derived from Dense Landsat Time Series in an Ecological Restoration Region. *Remote Sensing*, 14(4). <https://www.mdpi.com/2072-4292/14/4/961>.
- Pan, B., Shi, Z., Xu, X., Shi, T., Zhang, N., Zhu, X., 2019. CoinNet: Copy Initialization Network for Multispectral Imagery Semantic Segmentation. *IEEE Geoscience and Remote Sensing Letters*, 16(5), 816-820.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.
- Schuster, C., Förster, M., Kleinschmit, B., 2012. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing*, 33(17), 5583-5599. <https://doi.org/10.1080/01431161.2012.666812>.

Sibanda, M., Buthelezi, S., Ndlovu, H. S., Mothapo, M. C., Mutanga, O., 2021. Mapping the Eucalyptus spp woodlots in communal areas of Southern Africa using Sentinel-2 Multi-Spectral Imager data for hydrological applications. *Physics and Chemistry of the Earth, Parts A/B/C*, 122, 102999.

Solórzano, J. V., Mas, J. F., Gao, Y., Gallardo-Cruz, J. A., 2021. Land Use Land Cover Classification with U-Net: Advantages of Combining Sentinel-1 and Sentinel-2 Imagery. *Remote Sensing*, 13(18). <https://www.mdpi.com/2072-4292/13/18/3600>.

Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*, PMLR, 6105–6114.

Trisasongko, B. H., Paull, D., 2020. A review of remote sensing applications in tropical forestry with a particular emphasis in the plantation sector. *Geocarto International*, 35(3), 317-339. <https://doi.org/10.1080/10106049.2018.1516245>.

Wagner, F. H., Sanchez, A., Tarabalka, Y., Lotte, R. G., Ferreira, M. P., Aidar, M. P. M., Gloor, E., Phillips, O. L., Aragão, L. E. O. C., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation*, 5(4), 360-375.

Yakubovskiy, P., 2019. Segmentation models. https://github.com/qubvel/segmentation_models.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8-36.