TOWARDS A TREE COVER CHANGE EARLY WARNING SYSTEM BASED ON SENTINEL-1 DATA AND A NEURAL NETWORK ARCHITECTURE

A. Fiche [1], C. Hardy[1], B. Mora[1*]

[1] CS GROUP, Toulouse, France

KEY WORDS: Deforestation, Tropical forests, Early warning systems, Sentinel-1, Radar, Time series, Recurrent Neural networks, LSTM.

ABSTRACT:

This study aimed at exploring the potential of neural networks composed of convolutional and Long Short-Term Memory (LSTM) layers to handle dense Sentinel-1 data time series to develop a tree cover loss Early Warning System (EWS). The study area was in the Madre de Dios region in Peru that hosts a humid tropical forest. The second objective of this study was to investigate the potential of large free open-source datasets such as those from the GLAD and Geobosques alerts to calibrate and validate the models. The study demonstrated the capacity of the tested NN models to improve the detection of tree cover loss compared to a classical random forest algorithm thanks to their capacity to handle explicitly both spatial and temporal data. The accuracies of the best model compare reasonably well with those from similar studies. However, the observed overestimation of the Cover-loss class in the output map can be mainly linked to the quality of the alert datasets used as input data. Avenues to overcome the identified limits of this preliminary study are presented. This work provides a solid knowledge basis on the potential of a NN-based EWS and opens potential avenues for further improvements.

1. BACKGROUND AND OBJECTIVES

Deforestation is responsible for about 11% of the global greenhouse gas emissions, as estimated in 2019. The Global Forest Watch (GFW) initiative of the World Resources Institute estimated that deforestation amounted 12.2 million hectares in 2020, of which approximately one third occurred in tropical forests (GFW, 2021). The Reducing Emissions from Deforestation and Forest Degradation (REDD+) framework proposed by the Intergovernmental Panel on Climate Change (IPCC) to address deforestation in tropical regions can be complemented by early warning systems (EWS) that aim to track deforestation in near-real time. In the pan-tropical region, the demand for such EWS has been expressed by several national institutions. In addition to providing alerts that can be used by Legal authorities, EWS also foster transparency on ongoing forestry activities, support land empowerment initiatives, and other international conventions like the one on Biological Diversity (Mora, 2018). For example, the Peruvian Ministry of Environment operates the Geobosques EWS platform (Ministerio del Ambiente, 2022). In Brazil, DETER-B has been developed by the National Institute for Space Research (Diniz et al., 2015). Such national initiatives are complemented by others from Academia, Space Agencies, and non-governmental organizations. For example, the Global Land Analysis & Discovery (GLAD) Forest Alerts has a pan-tropical monitoring coverage (Hansen et al., 2016). More recently, Reiche et al. (2021) proposed the RAdar for Detecting Deforestation (RADD) alert system, and Mermoz et al. (2021) proposed the TropiSCO alert system. The Japanese government proposes the JICA–JAXA Forest EWS in the Tropics (JJ-FAST) system (Watanabe et al., 2021).

Several EWS rely on optical data time series. The Geobosques EWS platform provides alerts based on Landsat-7 and -8 data. Sentinel-2 data is used as a second stage to estimate surface of change (Ministerio del Ambiente, 2022). In Brazil, DETER-B also relies on optical data from the Advanced Wide-Field Sensor (AWIFS) offering a 56m spatial resolution (Diniz et al., 2015). The GLAD alerts are also based on Landsat data. However, the use of optical imagery hampers early cover change detection due to cloud cover that is frequently present over tropical areas, especially during the wet season (Hansen et al., 2016). At best, change detection can be shifted by several weeks. To overcome the limitation of optical imagery, the JJ-FAST, RADD, TropiSCO alert system rely on SAR data time series. While the first system uses L-band data from the ALOS-2 PALSAR-2, the two former systems use C-band Synthetic Aperture RADAR (SAR) Sentinel-1 data. The Sentinel-1A, -1B, provide data since 2014 and 2016, respectively. This current constellation will be soon expanded and continued by Sentinel-1C and -1D satellites, further facilitating the development and use of operational EWS based on long free dense data time series.

In parallel with the sustained availability of free SAR datasets, the interest for deep learning methods has been renewed in the field of Earth observation thanks to more affordable processing resources, and technological developments like those observed for the Graphics Processing Units (Baji, 2018). In the field of remote sensing applications, Ban et al. (2020) demonstrated the added value of Convolutional Neural Networks (CNN) over a the traditional log-ratio operator to map change due to wildfires. The authors used Sentinel-1 data time series. Torres et al. (2021) compared a series of six different Fully-CNN applied to optical data time series (Landsat and Sentinel-2) to map deforestation in the Brazilian Amazon. The best neural network (NN) was based on a Residual U-Net architecture that consistently presented the best trade-off between accuracy and training times. (He et al., 2016). Kislov et al. (2021) demonstrated the potential of deep CNNs compared to classical machine learning methods (e.g., random forest) to map forest disturbances with optical imagery as input data.

The diversity of NN architectures is high as reported by Alzubaidi et al. (2021). Among such architectures, Hamedianfar

* corresponding author: brice.mora@csgroup.eu

et al. (2022) identified the potential of recurrent neural networks (RNN) for forest inventory applications among other architectures. RNNs have the capacity to establish patterns from input datasets in both time and space dimensions (Hochreiter and Schmidhuber, 1997). Chang et al. (2019) showed that RNN-based architecture could outperform classical methods such as random forest and support vector machine algorithms to map forest cover classes. Aerial photography, Landsat imagery and elevation data were used as input data. Ye et al. (2019) used a specific RNN model architecture called Long Short-Term Memory (LSTM) to project Australia's forest cover change. The projection performance of the LSTM model significantly outperformed the one from a spatial econometric model. Time series of biophysical variables were used as input datasets. Parente et al. (2019) obtained better results with U-Net and LSTM NN architectures compared to a random forest model mapping pasture lands in Central Brazil. The authors used optical PlanetScope data to train the models.

We excluded mountainous areas from this preliminary study as such areas are known to pose challenges due to geometric distortions in SAR data to map forest cover (Wu et al., 2021). Our study did not discriminate tree cover change resulting from anthropogenic activities from natural ones such as landslides, windthrows, fire or flooding. To mitigate false alerts, the minimum mapping unit considered for tree cover change was equivalent to 10 Sentinel-1 pixels, i.e., 0.1 ha. Detected small-scale change events represented by a small number of Sentinel-1 pixels can result from local moisture fluctuations or remaining speckle noise (Bouvet et al., 2018; Reiche et al., 2018). The two thematic classes considered in this tree cover change study were: Stable cover and Cover loss.

The objectives of this study were first, to investigate the capacity of a deep learning approach compared to a classical machine learning method, to improve tree cover change detection in tropical regions using dense radar data time series and second, assess the capacity of free open-source alerts to serve as calibration and validation data.

2. MATERIALS AND METHODS

2.1 Study region

The study was performed in the Madre de Dios region located in the southeast of Peru. The region belongs to the Amazon basin and is characterised by a hot and humid climate. The average temperature is 26 °C and annual precipitation can be as high as 3 meters. The region hosts a tropical rain forest. Population density is 1.7 hab./km^2. Gold mining and illegal logging are among the most important ongoing human activities impacting the region. Such activities are the main cause of the deforestation and forest degradation observed in the region. Caballero Espejo et al., (2018) estimated that artisanal-scale gold mining was responsible for nearly 100,000 hectares of forest loss in the Madre de Dios region between 1984 and 2017.

2.2 Data

2.2.1 Sentinel-2 imagery: We used Sentinel-2 L1C images to generate forest masks. Sentinel-2A and -B satellites provide optical data with a spatial resolution of 10m in the visible range and first near-infrared band. For year 2018, the mask was generated with a cloud-free image from Autumn 2017, while for year 2019, the mask was generated with a cloud-free image acquired during Autumn 2018.

2.2.2 Sentinel-1 imagery: We built Sentinel-1 data time series considering dual-polarized (VV and VH) Ground Range Detected (GRD) products acquired in interferometric wide mode. The GRD images have a spatial resolution of 20 by 22 m registered in a 10 by 10 m pixel spacing (ESA, 2020). The period considered in this study spanned from the end of 2017 to year 2019, amounting a total of 170 images corresponding to tile #19LCF within the Sentinel-2 tiling grid.

2.2.3 GLAD and Geobosques datasets: We crossed the GLAD and Geobosques alerts for year 2018 and 2019 to calibrate, validate and test the models. Although GLAD alerts are considered conservative by design (Hansen et al., 2016), we aimed at further strengthening the quality of our datasets crossing these two data sources. We only kept alerts present in both datasets. While Geobosques alerts provided spatial locations of change, GLAD alerts provided both spatial information and dates of detected change.

2.2.4 Dove dataset: We used optical very high spatial resolution (VHSR) imagery from the Dove constellation to visually check samples of the test datasets. The Dove satellites acquire optical imagery at 3-to-5-meter spatial resolution. In this study we used mosaics and daily acquisitions made freely accessible by the Norway's International Climate and Forest Initiative (NICFI) for research purposes.

2.3 Pre-processing

2.3.1 Sentinel-1 imagery: We used the free open-source toolbox S1Tiling (CNES, 2019) to perform a gamma radiometric calibration of the images, as performed by Ballère et al. (2021), followed by a terrain correction and the projection of the images in the Sentinel-2 tiling grid reference. Then we used the free open-source Orfeo Toolbox to apply a multi-temporal speckle filtering Quegan et al. (2000). Four images were used for this purpose, with a 3-by-3-pixel spatial window.

2.3.2 GLAD, Geobosques datasets: The GLAD and Geobosques patch alerts were downloaded from their respective online repositories. Follow up analysis in a geographic information system consisted in selecting the intersection of the spatially agreeing patches between the two datasets for years 2018 and 2019.

2.3.3 Sentinel-2 and Dove datasets: No pre-processing step was performed on the two the Sentinel-2 images used separately to generate the two forest masks, nor on the Dove images used for the photo interpretation of the test samples.

2.4 Processing

2.4.1 Forest masks: The forest masks were generated using images close and prior to the beginning of the periods of interest to set the forest baseline. The first forest mask was made for year 2018 (model calibration and validation period) and year 2019 (model test period). Both forest masks were computed using a random forest algorithm (Breiman, 2001) applied to each Sentinel-2 image using the SNAP software. We considered the following classes: tree-covered areas, grass, urban areas, roads, rivers, bare soil, and clouds, to train and validate the models. Samples were generated via photointerpretation using the Sentinel-2 images.

2.4.2 Mountain mask: We excluded the mountainous areas applying a maximum altitude threshold adapted to the study region. Forested areas below 400m were selected using the Advanced Spaceborne Thermal Emission and Reflection (ASTER) Global Digital Elevation Model (GDEM).

2.4.3 Calibration, validation, and test datasets development: Calibration and validation datasets were built using patches corresponding to year 2018, while the test datasets, used solely for the best NN models, were built with patches from year 2019.

Patches of the validation and test datasets were selected within a series of geographical boxes scattered across the study region. The calibration dataset was generated in the remaining area of the forest mask. We randomly generated 20 boxes using a Poisson-Disc approach (Bridson, 2007) within the Sentinel-2 tile area. The centres of the boxes were spatially away from each other by a minimal distance of 8 km. Then we randomly split the set of boxes in two equal subsets. One subset was used for the selection of the validation patches while the second one was used for the selection of the test patches. Each box had a size of 5 by 5 km and was surrounded by a 1-km buffer to limit spatial autocorrelation between the three datasets. The establishment of the box set followed two additional constraints.: First, mountainous areas should represent less than 40% of each box area and second, the forested area should represent at least 40% of each box area. The values and thresholds for the establishment of the box set were chosen to obtain at least 10,000 patches for the Cover-loss class in both the validation and test datasets.

Stable-cover patches of the calibration and validation datasets were split in two separate sub-classes to facilitate NN learning process. The first sub-class was composed of central Stable-cover pixels only surrounded by Stable-cover pixels while the second class was composed of central Stable-cover pixels surrounded by at least one Cover-loss pixel.

The stratified random patch selection approach within the boxes for the validation and test datasets was performed to obtain 45% of the pixels labelled as Cover-loss, 25% labelled as Stable-cover with at least one Cover-loss neighbour, and 30% labelled as Stable-cover without Cover-loss neighbour. Pixels of the calibration datasets were randomly selected within the forest mask outside the boxes and their buffer zones. Class proportions were identical as those used for the validation and test datasets.

The building of the test datasets was made using the output map of each of the best NN models identified with the validation dataset. Each time, we photo interpreted 400 samples of which, 150 for the Stable-cover class.

2.4.4 Neural network modelling: The tested models combined convolutional and Long Short-Term Memory (LSTM) layers. We chose the LSTM among RNNs approaches for its ability to address the sensitivity decay issue that can appear over time with other RNN-based architectures (Alzubaidi et al., 2021). The free open-source Pytorch library was used for the modelling.

We built separate Sentinel-1 data time series as inputs to the models considering the ascending and descending orbits, the pre- and post-change detection date, and the VV/ VH polarizations. Different input image patch sizes were tested (5 by 5, 7 by 7, 9 by 9 pixels). We chose the widely used Rectified Linear Unit (ReLU) as activation function for the models. We tested different sizes of the pre-change data time series with either 4, 8, 12, or 16 images. For the pre-change Sentinel-1 time series, we also tested the influence of a safety time buffer applied between the date of the first change detection and the date of the latest image of the Sentinel-1 pre-change time series to account for potential time discrepancy between the true date of change and the detection date reported in the calibration and validation datasets. Time buffers of 30 and 60 days were tested. We tested the impact of using one or three post-change images. Furthermore, we tested the inclusion of such images in the NN itself, or as a post-processing approach to evaluate the model output using a simple class-output unanimity rule. In the latter case, the Cover-loss class was chosen if the three model outputs computed with each new observation following a detected change, were agreeing.

2.4.5 Random forest modelling: We used a 16-image pre-change Sentinel-1 time series with a post-change time series either made of one or three images to build the models. Either backscatter values or mean with standard deviation backscatter values of the data time series were provided as input to the models. Samples used to calibrate the models correspond to a subset of those used for the NN models to limit its size, while the validation dataset remained identical. The calibration samples were picked following a stratified random selection approach. We kept the same class proportions as those chosen for the NN modelling. A total of 15,000 patches were used for the calibration dataset.

3. RESULTS

1.1 Calibration and validation datasets

Table 1 provides the sizes in number of patches of the sample datasets used for the NN modelling.

	Cover-loss class	Stable-cover class	
		with Cover-loss neighbour(s)	with Stable neighbours
Calibration dataset	203,978	113,321	135,985
Validation dataset	15,255	8,475	10,170

Table 1. Sample dataset sizes used for the NN modelling.

1.2 Neural network modelling

Tables 2a and 2b provide accuracies obtained with the validation dataset using different patch sizes and different sizes of pre-change Sentinel-1 data time series.

	4 images		8 images	
	Stable cover	Cover loss	Stable cover	Cover loss
5 * 5 pixels	94%	82%	97%	79%
7 * 7 pixels	96%	84%	96%	88%
9 * 9 pixels	96%	88%	96%	88%

Table 2a. Accuracies of the models depending on patch sizes and pre-change data time series sizes.

	12 images		16 images	
	Stable cover	Cover loss	Stable cover	Cover loss
7 * 7 pixels	97%	86%	98%	87%
9 * 9 pixels	97%	89%	98%	88%

Table 2b. Accuracies of the models depending on patch sizes and pre-change data time series sizes.

Table 3 provides accuracies considering one or three post-change image time series sizes integrated in the NN model as input data. These accuracies were obtained using the best NN model reported in the previous table, i.e., with a 9 by 9 input image patch size, and 16 pre-change Sentinel-1 images. Validation dataset was used.

	Stable cover	Cover loss
One post-change image	98%	88%
Three post-change images	99%	94%

Table 3. Accuracies of best model with varying number of post-change input data images.

Tables 4 and 5 provide User accuracies (UA) and Producer accuracies (PA) obtained using the post-change images as input data either for the NN model, or for the post-processing step. The two test datasets designed specifically for each model were used either selecting all patch pixels or core-patch pixels only.

	Stable cover		Cover loss	
	All pixels	Core pixels	All pixels	Core pixels
User's Accuracy	99%	99%	57%	78%
Producer's Accuracy	61%	85%	98%	98%

Table 4. Accuracies of best model with three post-change images integrated as model input data.

	Stable cover		Cover loss	
	All pixels	Core pixels	All pixels	Core pixels
User's Accuracy	99%	99%	84%	94%
Producer's Accuracy	80%	95%	99%	99%

Table 5. Accuracies of best model with three post-change images used post-processing input data.

For conciseness purposes, we do not provide results for the varying time buffers as observed accuracy differences were not significant between the tested cases.

Figure 1 provides the NN architecture that led to the best results provided in Table 5.

1.3 Random forest modelling

Table 6 provides the best results obtained with a random classifier considering mean and standard deviation values of the time series (pre and post change, ascending or descending, VV or VH polarizations), also with either one or three post-change images as input data. Accuracies were obtained with the validation dataset.

	Stable cover	Cover loss
One post-change image	92%	56%
Three post-change images	94%	63%

Table 6. Accuracies of the random forest classifier considering one or three post-change images.

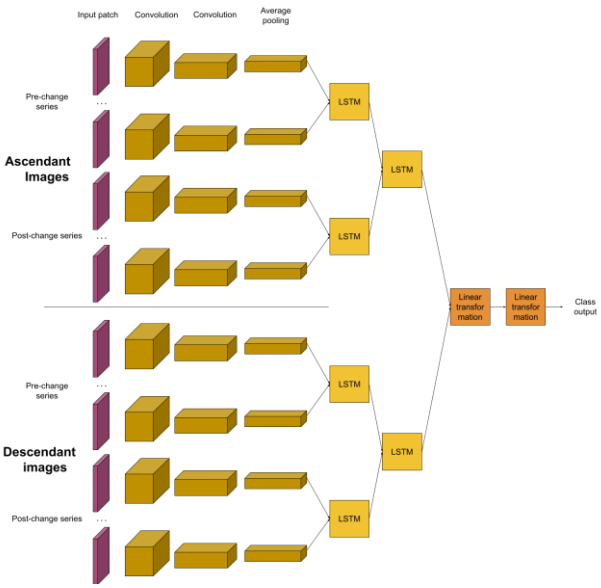

Figure 1. Best NN architecture.

4. DISCUSSION

Deep NN modelling requires large calibration datasets to be trained accurately (Cholet, 2017). We used large data sources openly available to build the calibration and validation datasets to meet this requirement. Following the defined protocol, we were able to obtain calibration and validation dataset sizes of 200K+ and 10K+ samples, respectively (Table 1). Each of the two test datasets were composed of 400 samples only. Though, the photointerpretation procedure allowed us reaching a higher degree of confidence on the quality of these samples.

Tables 2a and 2b demonstrate that NN model accuracies increased with the size of the image patches and the number of pre-change Sentinel-1 time series. The combined effect of the two parameters led to a slow and steady increase of the accuracies for the two classes. The 9-by-9-pixel patch size combined with a 16-image pre-change Sentinel-1 time series provided the best accuracies. The higher impact of the increase of the patch size observed on the Cover-loss class can be explained by the improved ability of larger patches to depict the

limits of the areas of change and their surroundings in the context of this study. This observation is coherent with previous studies that demonstrated for a given application, that larger patch sizes tend to lead to an increase of the accuracies thanks their ability to capture more contextual information (Farabet et al., 2013; Pinheiro, 2014). The positive effect of the increase in the size of the pre-change Sentinel-1 time series can be explained by the higher amount of information provided that can either confirm a positive trend for both classes or decrease the impact of local and temporary noise. For example, heavy rain showers and remaining stems after logging can increase soil backscatter values in C-band radar data (Woodhouse et al., 1999, Reiche et al., 2018).

Table 3 demonstrates the introduction of two additional post-change Sentinel-1 images in the model had a significant positive effect on the Cover-loss class accuracy (+ 6 percentage points). This accuracy improvement can again be linked to the richer amount of information that increases the confidence in the output class selection. However, this increase in the accuracy was made at the expense of the timeliness of a potential action on the ground to investigate on the detected change. The addition of two post-change images delayed the potential confirmation of change by 24 days. Such an approach to confirm changes using post-change images has already been implemented in previous studies like Reiche et al. (2021). In this study performed in the Congo basin with Sentinel-1 data, the authors had a validation database of confirmed changes. A post-change period of 84 days was needed to confirm all changes of their database.

Accuracies obtained with random forest algorithms demonstrate an increase in the accuracies when using three post-change images (Table 6). However, such accuracies are lower than those obtained with NN modelling (Table 3), especially for the Cover-loss class (-31 percentage points). Random forest algorithms have been successfully used in a wide range of Earth observation applications, notably in land cover mapping (Pelletier et al., 2016), or forest mapping (Mora et al., 2013). However, the capacity of random forests algorithms to handle dense, detailed spatial and temporal information as input data is limited compared to NNs designed to perform convolutions (spatial dimension) and handle dense data time series with LSTM layers (temporal dimension).

Accuracies reported in Table 4 were obtained with a test dataset specifically designed for this analysis. As expected, this fully independent dataset provided lower accuracies compared to those reported in Table 3 with identical modelling conditions. More precisely, the results indicate a mapping overestimation of the Cover-loss class (Producer's Accuracy of 98%) compared to a Producer's Accuracy of 61% for the Stable-cover class. These misclassifications can be due to false alerts that remained in the calibration and validation datasets used during the first modelling tests. As indicated, these samples are model outputs. Furthermore, the dates of detected changes provided by the GLAD alerts can be shifted by several days compared to the true date of change. Finally, GLAD and Geobosques alerts are provided at a 30m spatial resolution. Sentinel-1 images have a 20 m by 22 m spatial resolution registered in a 10 by 10 m pixel spacing (ESA, 2020). This discrepancy can further add noise during the training and validation steps.

The comparison of results reported in Tables 4 and 5 demonstrates a significant increase in the accuracies when using the post-change images in a post-processing step, i.e., after the NN modelling. Producer's Accuracy of the Stable-cover class is increased by 19 percentage points. The use of post-change information allowed correcting output model errors. Reiche et al. (2021) achieved a Producer's Accuracy of 95% for the disturbance class for change areas equal or larger than 0.2 ha. However, the inclusion of disturbances smaller than 0.2 ha lowered the accuracy of their model down to 83.5%. In our study, we achieved a Producers' Accuracy of 80% for the Cover-loss class considering change areas equal or greater than 0.1 ha. Mermoz et al. (2021) performed a forest loss study using Sentinel-1 data time series in Southeast Asia. The authors achieved a Producer's Accuracy of 90% for the forest disturbance class with a minimum mapping unit of 0.1 ha. Their method based on shadow detection was adapted to local circumstances, i.e., small, and scattered disturbances. Such shadows appear at the boundary of forest loss patches.

Results reported in Tables 4 and 5 demonstrate higher accuracies when considering only core-patch pixels of the test samples. This result was expected since backscatter values of Sentinel-1 images tend to be less ambiguous for core pixels compared to pixels located at the edge of the Cover-loss areas. We performed this analysis to evaluate the capacity of the model to detect each individual Cover-loss patch. Table 5 demonstrates that when considering core pixels, Producer's Accuracy remained identical for the Cover-loss class (99%) while Producer's Accuracy for Stable-cover class improved significantly to 95% (+15 percentage points). Such results indicate the model keeps overestimating the Cover-loss class in the output map, although the trend has been significantly mitigated thanks to the three post-change images used in a post-processing step. As discussed earlier, the samples used to calibrate and validate the models are outputs from previous models and hence, can include false alerts despite filtering efforts, and generate ambiguous samples. A further analysis of the misclassified Stable-cover class pixels revealed that such pixels tend to be nearer to Cover-loss pixels (avg. distance: 1.3 pixels) than those that were correctly classified (avg. distance: 24 pixels). This significant difference indicates that the spatial resolution combined with the pixel spacing of the Sentinel-1 images limit the capacity of the model to discriminate the two classes at fine scale, and thus contributes to the observed errors.

The comparison of NNs incorporating convolutional and LSTM layers with a random forest algorithm confirmed the superiority of the former thanks to capacity to handle explicitly both spatial and temporal data. The different tests allowed us to optimise both the model architecture and the parameter settings. The analysis of the accuracies obtained with the best architecture are encouraging although revealing the limits of using model outputs as calibration and validation dataset.

The best model of this study was notably based on the use of the three post-changes images, delaying the confirmation of a potential change by 24 days. A boundary-versus-core sample accuracy analysis performed by Reiche et al. (2021) demonstrated that 74% of all core pixels could be confirmed within 24 days and 95% within 48 days. The number of days for an acceptable trade-off between delay and accuracy will vary from one region to another depending on local circumstances such as the drivers of change and the policy enforced to tackle deforestation. Therefore, there is no EWS system architecture that fits all needs and model tuning will be necessary.

Building on the outcomes of this preliminary study, future work may consist in adapting the model to another region where GLAD alerts, but also national alerts could also be available.

Such a study could be an opportunity to test larger input image patch sizes (> 9 by 9 pixels).

Another NN architecture also known as U-Nets could be tested (Ronneberger et al., 2015). Such U-Net architectures have proven their added value for image segmentation in different fields of application such as autonomous cars (Tran and Le, 2019) or precision agriculture (Liu et al., 2020, Zhou et al., 2021). A dedicated cost function to better handle class disproportions like in this study will also be an option. Another avenue may consist in developing calibration and validation datasets based on fieldwork and photo interpretation which limited size could be further augmented using pseudo labelling as proposed by Shi et al., 2022.

Testing the model over mountainous areas is also another step envisioned by the authors of this study. Due to the complexity of such environments, verified cover loss occurrences will be necessary. Modelling could benefit from altitude, slope, and aspect information as input data.

5. CONCLUSION

The study demonstrated the capacity of NN composed of convolutional and LSTM layers to improve the detection of tree cover loss compared to a classical random forest algorithm thanks to their capacity to handle explicitly both spatial and temporal data. The accuracies of the best model compare reasonably well with those from similar studies. However, the analysis of the results indicates an overestimation of the Cover-loss class in the output map. The main explanation lies in the use of information resulting from previous modelling efforts as calibration and validation datasets. We used free and open-source datasets that led to the development of calibration and validation dataset sizes of 200K+ and 10K+ samples, respectively. The quality of such alert datasets has been discussed and demonstrated the need for higher quality calibration and validation data. In the context of NN modelling, the development of large datasets could be facilitated with the use of pseudo labelling approaches. The analysis of the results showed that the limits of the Sentinel-1 spatial resolution also contribute to the overestimation of the Cover-loss class. Other powerful NN approaches like U-Nets are also worth being considered to further improve this EWS prototype. To date, the model developed in this preliminary study cannot be deployed for operational activities. However, this work provides a solid knowledge basis on the potential of a NN-based EWS and opens potential avenues for further improvements.

5.1 References

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 53: (2021).

Baji, T., 2018: Evolution of the GPU Device widely used in AI and Massive Parallel Processing. *IEEE 2nd Electron Devices Technology and Manufacturing Conference (EDTM)*, pp. 7-9.

Ballère, M., Bouvet, A., Mermoz, S., Le Toan, T., Koleck, T., Bedeau, C., André, M., Forestier, E., Frison, P.L., Lardeux, C., 2021: Sar data for tropical forest disturbance alerts in French Guiana: Benefit over optical imagery. *Remote Sensing of Environment*, 252: 112159.

Ban, Y., Zhang, P., Nascetti, A., Bevington, A.R., Wulder. M.A., 2020: Near real-time wildfire progression monitoring with sentinel-1 sar time series and deep learning. *Scientific Reports*, 10(1):1–15.

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., Le Toan T., 2018: Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. *Remote Sensing*, 10:1250.

Breiman, L., 2001: Random forests. *Machine learning*, 45(1):5–32.

Bridson, R., 2007: Fast poisson disk sampling in arbitrary dimensions. ACM SIGGRAPH, 08.

Caballero Espejo, J., Messinger, M., Román-Dañobeytia, F., Ascorra, C., Fernandez, L.E., Silman, M., 2018: Deforestation and Forest Degradation Due to Gold Mining in the Peruvian Amazon: A 34-Year Perspective. *Remote Sens*ing, 10: 1903.

Chang, T., Rasmussen, B.P., Dickson, B.G., Zachmann, L.J., 2019: Chimera:amulti-task recurrent convolutional neural network for forest classification and structural estimation. *Remote Sensing*, 11(7): 768.

Chollet, F., 2017: *Deep learning with python*. Manning Publications Co.

CNES, 2019: S1Tiling: Building Analysis Ready Data of Sentinel-1 time series.https://github.com/CNES/S1Tiling (8 March, 2022).

Diniz, C.G., Souza, A.A.D.A., Santos, D.C., Dias, M.C., Luz, N., de Moraes, D.R.V., Maia, J.S.A., Gomes, A.R., Narvaes, I.D.S., Valeriano, D.M., et al., 2015: DETER-B: The New Amazon Near Real-Time Deforestation Detection System. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7), 3619-3628.

European Space Agency, 2020: Sentinel-1 SAR user guide. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar (7 March, 2022).

Farabet C., Couprie C., Najman L., Lecun Y., 2013: Learning hierarchical features for scene labelling. *IEEE Trans. Pattern Anal. Mach. Intell*. 35(8): 1915-1929.

Hansen, M.C., Krylov1, A., Tyukavina, A., Potapov, P.V., Turubanova1, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016, Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters*, 11(3):034008.

He, K., Zhang, X., Ren, S., Sun, J., 2016: Deep residual learning for image recognition. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

Hochreiter, S., Schmidhuber, J., 1997: Long Short-term Memory. *Neural Computation*, 9(8): 1735–1780.

Kislov, D.E., Korznikov, K.A.., Altman, J., Vozmishcheva, A.S., Krestov, P.V., 2021: Extending deep learning approaches for forest disturbance segmentation on very high-resolution satellite images. *Remote Sensing in Ecology and Conservation*, 7(3): 355-368.

Liu, Y., Zhang, S., Yu, H., Wang, Y., Feng, Y., Sun, J., Zhou, X., 2020: Straw segmentation algorithm based on modified unet in complex farmland environment. *Computers, Materials Continua*, 66: 247-262.

Ministerio del Ambiente, 2015. GEOBOSQUE Plataforma. https://geobosques.minam.gob.pe/ (25 February 2022).

Mora, B., Wulder, M.A., White, J.C. Hobart, G. (2013) Modeling Stand Height, Volume, and Biomass from Very High Spatial Resolution Satellite Imagery and Samples of Airborne LiDAR. *Remote Sensing*, 5(5): 2308-2326.

Mora, B., 2018: User Needs Assessment for Forest Change Early Warning Systems. Global Forest Observations Initiative - Food and Agriculture Organization of the United Nations, Rome, Italy, 45p.

Mermoz, S., Bouvet, A., Koleck, T., Ballère, M., Le Toan, T. 2021: Continuous Detection of Forest Loss in Vietnam, Laos, and Cambodia Using Sentinel-1 Data. *Remote Sensing*, 13, 4877.

Parente, L., Taquary, E., Silva, A.P., Souza Jr., C., Ferreira, L., 2019: Next Generation Mapping: Combining Deep Learning, Cloud Computing, and Big Remote Sensing Data. *Remote Sensing*, 11(23): 2881.

Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016: Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187, 156-168.

Pinheiro, P.H., Collobert, R., 2014: Recurrent convolutional neural networks for scene labelling. Proceedings of the 31st International Conference on Machine Learning, 32(1): 82-90.

Quegan, S., Le Toan, T., Yu, J.J., Ribbes, F., Floury, N., 2000: Multitemporal ers sar analysis applied to forest mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):741–753.

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., 2018: Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sensing of Environment*, 204:147–61.

Reiche, J., Mullissa, A., Slagter, B., Gou, Y., Tsendbazar, N.E., Odongo-Braun, C., Vollrath, A., Weisse, M., Stolle, F., Pickens, A., Donchyts, G., Clinton, N., Gorelick, N., Herold, M., 2021: Forest disturbance alerts for the congo basin using sentinel-1. *Environmental Research Letters*, 16(2):024005.

Ronneberger, O., Fischer, P., Brox, T., 2015: U-net: Convolutional networks for biomedical image segmentation. *Computing Research Repository - arXiv*, abs/1505.04597.

Shi, J., Wu, T., Qin, A.Q., Lei, Y., Jeon, G., 2022: Semisupervised adaptive ladder network for remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, Early Access, 1–1.

Watanabe, M., Koyama, C.N., Hayashi, M., Nagatani, I., Tadono, T., Shimada, M., 2021: Refined algorithm for forest early warning system with ALOS-2/PALSAR-2 ScanSAR data in tropical forest regions. *Remote Sensing of Environment*, 265: 112643.

Tran, L.A., Le, M-H, 2019: Robust u-net-based road lane markings detection for autonomous driving. In 2019 International Conference on System Science and Engineering (ICSSE), 62–66.

Woodhouse, I., van der Sanden, J.J., Hoekman, D.H., 1999: Scatterometer observations of seasonal backscatter variation over tropical rain forest. *IEEE Trans. Geosci. Remote Sens*. 37: 859–861.

Wu, L., Wang, H., Li, Y., Guo, Z., Li, N., 2021: A Novel Method for Layover Detection in Mountainous Areas with SAR Images. *Remote Sensing*, 13(23):4882.

Ye, L., Gao, L., Marcos-Martinez, R., Mallants, D., Bryan, B.A., 2019: Projecting Australia's forest cover dynamics and exploring influential factors using deep learning. *Environmental Modelling & Software*, 119, 407-417.

Zou, K., Chen, X., Wang, Y., Zhang, C., Zhang, F., 2021: A modified u-net with a specific data argumentation method for semantic segmentation of weed images in the field. *Computers and Electronics in Agriculture*, 187:106242.