
A Novel Integrated ML Approach Utilizing Radar & Satellite Imagery for Selective Logging Detection

Saraswathy Amjith

Massachusetts Institute of Technology
Cambridge, MA 02139
swathy@mit.edu

Joshua Fan

Cornell University
Ithaca, NY 14850
jyf6@cornell.edu

Abstract

Illegal selective logging is devastating to the world’s biodiversity and forests, and many countries face challenges in monitoring these forests. In the Amazon, 94% of deforestation is attributed to illegal activities, primarily conducted through selective logging—a practice that has received far less research attention than clear-cut deforestation. This study introduces a novel machine learning approach that enhances the detection of selective logging by integrating Sentinel-1 Synthetic Aperture Radar (SAR) data with Sentinel-2 optical imagery, moving beyond conventional methods that rely on a single data source. Requisite data was collected through Google Earth Engine for Sentinel-1 and Sentinel-2 and used to create three new datasets of 134,606 samples (5x other studies). These datasets were then used to train and evaluate three different classification architectures: Convolutional Neural Networks (CNN), Random Forest, and Gradient Boosted Trees. The CNN model trained on both modalities outperformed CNNs trained on individual Sentinel datasets, achieving an accuracy of 95.08%, compared to 57.51% for Sentinel-1 and 91.95% for Sentinel-2. The Gradient Boosted Trees model, while slightly less accurate with a 94.39% accuracy and a 94.03 F1 score, required significantly less computational power. Our findings demonstrate that the integration of optical and radar data substantially improves the prediction of selective logging activities, achieving up to a 7.08% improvement over previous selective logging detection research and offering a scalable solution for forestry monitoring in resource-limited settings.

1 Introduction

Illegal logging is estimated to be a \$150 billion industry, eventually contributing towards the 1.5 gigatons of carbon resulting annually from deforestation (World Wildlife Fund, n.d.). In the Brazilian Amazon, selective logging is particularly common; rather than clearing the entire forest and leaving barren land, most loggers opt for more subtle methods of harvesting only a small amount of highly profitable trees. In doing so, measures such as canopy cover, biodiversity, or carbon content are loosely preserved. Despite its importance, selective logging is notoriously difficult to detect and quantify via remote sensing due to the subtle and diffuse nature of canopy damage.

While developing countries often lack resources to monitor protected forests, machine learning now enables satellite-based disturbance alerts to support land management and combat illegal logging (Hansen et al., 2016). Satellite monitoring systems focus on detecting full deforestation (complete canopy removal), mainly using medium-resolution optical imagery such as Landsat ((Kuck et al., 2021); (Vargas Zeppetello et al., 2020) ; (Hansen et al., 2016). However, Landsat’s 30m resolution is often too coarse to detect individual tree fall gaps (10-20m across). Furthermore, the areas most prone to deforestation are in the humid tropics, experiencing persistent cloudiness (especially in the wet

season) which creates temporal gaps in optical monitoring and makes real-time detection of logging challenging (Reiche et al., 2021). Consequently, many selective logging occurrences go undetected, often disappearing within weeks due to the fast regeneration of vegetation. Additionally, although methods for clear-cut deforestation detection are well developed and achieve accuracies over 90% in many alert systems, selective logging has been less intensely studied, with the most recent study only obtaining an accuracy of 88% (Kuck et al., 2021).

Synthetic Aperture Radar offers a complementary solution that can penetrate clouds. When a large tree is removed, the radar backscatter from that area often changes in characteristic ways: loss of canopy causes reduced volume scattering from leaves/branches and can increase ground returns or double-bounce reflections from trunks and soil, especially at certain wavelengths and polarizations (Kuck et al., 2021). Indeed, a growing body of work has validated SAR’s effectiveness for logging detection. Rauste et al. developed a method to map selective logging in the Republic of Congo using L-band ALOS PALSAR imagery acquired before and after logging, achieving high overall accuracy in identifying logged areas (Rauste et al., 2013). Reiche et al., using only Sentinel-1 data, produced an accuracy of 83.5% (Reiche et al., 2021). More recently, (Kuck et al., 2021). demonstrated that X-band radar data (TerraSAR-X/TanDEM-X) processed with machine learning could detect selective logging in the Brazilian Amazon with 88% classification accuracy .

Given that optical and SAR data provide complementary information, with optical images sensing canopy spectral changes and SAR sensing structural changes, integrating the two could yield a more robust selective logging detection approach than either alone. Since their launches in 2014 (Sentinel-1 radar) and 2016 (Sentinel-2 optical), both platforms provide 10 m resolution imagery every six days, perfectly co-registered across the tropics. Yet, to date, no study has fully fused Sentinel-1 and Sentinel-2 data for selective-logging detection, relying instead on either optical or radar observations alone. While Hethcoat et al. (2021) explored Sentinel-1 time-series break-point detection, their work did not integrate optical information nor employ deep-learning models on fused data. We address this gap by developing a novel approach that fuses features from radar and optical data. By leveraging a feature-fusion strategy, our approach overcomes some limitations of single-sensor methods (e.g. optical-only methods failing under cloud cover, or SAR-only methods missing spectral cues of vegetation stress).

2 Materials and Methods

2.1 Data Set Creation

The data set’s location was selected to be Jamari National Park, an area managed by the Brazilian forest ministry using sustainable selective logging. We utilized multi-source satellite data, primarily Sentinel-1 SAR imagery (C-band, 10 m, dual-polarization) and Sentinel-2 optical imagery (10–20 m, multi-spectral), acquired contemporaneously during logging periods in 2016. Sentinel-1 GRD images in Google Earth Engine are preprocessed for orbit, thermal-noise and border-noise removal, radiometrically calibrated to σ^0 , and terrain-corrected. Sentinel-2 Level-1C images are atmospherically corrected to surface reflectance and cloud-masked to include only cloud-free pixels. Model inputs consist of concatenated Sentinel-1 radar and Sentinel-2 optical bands over a two-week window, capturing imagery both before and after logging events. Using a time series instead of a single snapshot reduces false positives from static forest or soil variations by detecting change relative to an initial baseline. Additionally, NDVI (Normalized Difference Vegetation Index) was computed for the Sentinel-2 imagery, and RVI (Radar Vegetation Index) was calculated for the Sentinel-1 imagery. NDVI quantifies vegetation health by comparing near-infrared light, which vegetation strongly reflects, with red light, which it absorbs, while RVI uses radar data to assess the scattering variability of vegetation. Class imbalance – only a small fraction of the area is logged in any given year – was addressed by up-sampling the logging class in training and using a stratified random sampling of negatives to have nearly equivalent class sizes.

Three distinct datasets were generated: one using only Sentinel-1 data, one with only Sentinel-2 data, and one that combined both. For the combined dataset, geospatially aligned satellite and radar bands from each image and location were concatenated. In total, the datasets comprised 134,606 samples.

2.2 Machine Learning Models

We evaluated three supervised methods: Random Forest (RF), XGBoost, and a U-Net CNN. The U-Net (Milesial, n.d.) is a fully convolutional network—initially developed for biomedical segmentation on limited data—that uses repeated blocks of unpadded 3×3 convolutions (doubling channels), batch-norm, ReLU, and 2×2 max-pooling (halving spatial dimensions). The original version was designed for larger images. To address this, one down-sampling layer and one up-sampling layer were removed from the U-Net architecture. RF and XGBoost are nonparametric, tree-based classifiers trained on the same feature set; their probabilistic outputs are averaged for the final prediction. The classifiers were tuned via cross-validation. We trained the models to classify each data point as either “Selective Logging” or “No Logging”.

3 Results

Standard accuracy metrics were calculated: overall accuracy, recall, F1 score, and AUC. Additionally, to quantify the benefit of the multi-sensor approach, we trained and tested the model using only optical features, and only SAR features, to see how the fused approach compared.

The data is analyzed using each type of ML model (random forest, CNN, gradient boosted tree) and then calculated on accuracy, precision, recall, and F_1 score for each type of model.

Models (Sentinel 1)	Accuracy (%)	Precision (%)	Recall (%)	F_1 (%)
CNN	78.2	79.5	76.8	78.1
Random Forest	57.5	58.3	55.6	56.9
Gradient Boost	58.1	59.2	56.7	57.9

Table 1: Performance metrics using solely Sentinel 1 data.

Models (Sentinel 2)	Accuracy (%)	Precision (%)	Recall (%)	F_1 (%)
CNN	91.95	92.3	91.5	91.9
Random Forest	86.77	87.5	85.9	86.7
Gradient Boost	90.40	90.8	89.9	90.3

Table 2: Performance metrics using solely Sentinel 2 data.

Models Sentinel-1 & 2	Accuracy	Precision	Recall	F_1
CNN	95.08	93.32	96.28	94.73
Random Forest	92.46	92.42	91.09	92.64
Gradient Boost	94.39	93.58	94.49	94.03

Table 3: Performance Metrics Using Fused Sentinel-1 and Sentinel-2 Data.

When both radar and optical data are integrated, overall accuracy increases for all models, with the CNN maintaining the top performance. Specifically, the CNN achieves 95.08% accuracy, 93.52% precision, 94.28% recall, and a 94.73 F_1 score. The improved results confirm that the fusion of optical and radar observations provides complementary information—radar’s structural insights augmenting optical spectral signals—thereby enhancing the detection of smaller or more subtle logging disturbances.

Figure 1 compares the accuracy of three machine learning model on different input configurations (Sentinel-1 only, Sentinel-2 only, and integrated radar + optical). Each model has lowest accuracy using radar-only data and improves significantly when switched to optical data. The greatest accuracy gains occur when radar and optical imagery are fused.

These logging maps in Figure 2 are generated after merging Sentinel-1 (radar) and Sentinel-2 (optical) satellite data, distinguishing between logged (red) and stable (green) forest areas.

4 Discussion

Our integrated dataset reached 95.08% detection accuracy—surpassing the ~88% previously reported in the region (Jackson & Adam, 2020) and the 50–75% rates at $<15\text{m}^3\text{ha}^{-1}$ logging intensity documented by Hethcoat et al. (2021), underscoring the novelty and significance of combining optical (Sentinel-2) and radar (Sentinel-1) satellite data. By leveraging the complementary strengths

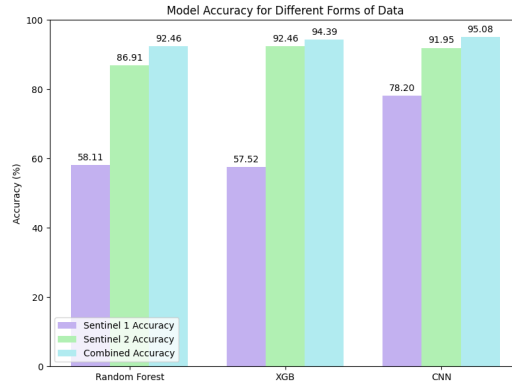


Figure 1: Model Accuracy For the Different Datasets Created on the Various Models Tested.

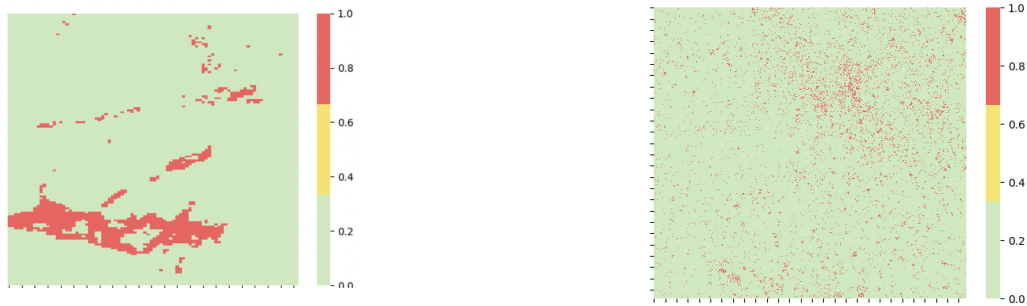


Figure 2: Logging Map for Small Test Regions with Red Indicating Logged & Green Indicating Stable Forest.

of spectral (optical) and structural (radar) information, the new approach consistently outperforms current models that rely solely on one data source, improving accuracy by at least 3%, by leveraging the complementary strengths of SAR and optical bands to better distinguish between logged and stable forests.

From a methodological standpoint, the integration of both sensors addresses key limitations of prior models that rely on a single data source. Radar data, which can penetrate cloud cover and detect subtle structural changes in canopy layers, is particularly advantageous in persistently overcast regions or during rainy seasons. Meanwhile, optical imagery captures nuanced spectral signals linked to vegetation health and canopy stress. The fused dataset therefore offers a more comprehensive representation of forest conditions, allowing the detection of smaller-scale or transient logging disturbances that might remain undetected in single-sensor approaches.

Especially as illegal logging activities show an upward trend of decreasing in size but increasing in frequency, the enhanced detail and performance gained by integrating Sentinel-1 and Sentinel-2 data enable more effective detection of selective logging, offering a promising path forward in combating deforestation.

References

- [1] U.S. Agency for International Development. (2022, February 8). Illegal logging and deforestation. Retrieved from <https://www.usaid.gov/biodiversity/illegal-logging-and-deforestation>
- [2] World Wildlife Fund. (n.d.). Deforestation and forest degradation. Retrieved from <https://www.worldwildlife.org/threats/deforestation-and-forest-degradation>
- [3] MAAP. (2021, August 30). MAAP #144: The amazon & climate change: Carbon sink vs carbon source. Retrieved from <https://www.maaproject.org/2021/amazon-carbon-flux/>

- [4] Lynch, J., Maslin, M., Balzter, H., et al. (2013). Choose satellites to monitor deforestation. *Nature*, 496, 293–294. <https://doi.org/10.1038/496293a>
- [5] Vargas Zeppetello, L. R., et al. (2020). [Title]. *Environmental Research Letters*, 15, 084012.
- [6] Hansen, M. C., et al. (2016). [Title]. *Environmental Research Letters*, 11, 034008.
- [7] Reiche, J., Mullissa, A., Slagter, B., Gou, Y., Tsendbazar, N.-E., Odongo-Braun, C., ... Herold, M. (2021). Forest disturbance alerts for the Congo Basin using Sentinel-1. *Environmental Research Letters*, 16(2), 024005. <https://doi.org/10.1088/1748-9326/abd0a8>
- [8] Hethcoat, M. G., Edwards, D. P., Carreiras, J. M. B., Bryant, R. G., França, F., & Quegan, S. (2019, February 1). A machine learning approach to map tropical selective logging. *Remote Sensing of Environment*. Retrieved November 7, 2023, from <https://doi.org/10.1016/j.rse.2018.11.044>
- Hethcoat, M. G., Carreiras, J. M. B., Edwards, D. P., Bryant, R. G., Quegan, S. (2021). [9] Detecting tropical selective logging with C-band SAR data may require a time-series approach. *Remote Sensing of Environment*, 260, 112411. <https://doi.org/10.1016/j.rse.2021.112411>
- [10] Jackson, C. M., & Adam, E. (2020). Remote sensing of selective logging in tropical forests: Current state and future directions. *iForest - Biogeosciences and Forestry*, 13(4), 286–300. <https://doi.org/10.3832/ifor3301-013>
- [11] Kuck, T. N., Sano, E. E., Bispo, P. d. C., Shiguemori, E. H., Silva Filho, P. F. F., & Matricardi, E. A. T. (2021). A comparative assessment of machine-learning techniques for forest degradation caused by selective logging in an Amazon region using multitemporal X-Band SAR images. *Remote Sensing*, 13(17), 3341. <https://doi.org/10.3390/rs13173341>
- [12] Gis and Science. (2012, July 25). Detection and spatial analysis of selective logging with geometrically corrected Landsat images. Retrieved from <https://gisandscience.wordpress.com/2012/07/25/detection-and-spatial-analysis-of-selective-logging-with-geometrically-corrected-landsat-images/>
- [13] Forest logging alerts from Radarsat-2 SAR data. (n.d.). Retrieved from https://www.researchgate.net/publication/333324232_FOREST_LOGGING_ALERTS_FROM_RADARSAT-2_SAR_DATA
- [14] Milesial. (n.d.). [GitHub repository]. Retrieved from <https://github.com/milesial/Pytorch-UNet>
- [15] Asner, G. P.; Knapp, D. E.; Broadbent, E. N.; Oliveira, P. J.; Keller, M.; Silva, J. N. 2005. Selective logging in the Brazilian Amazon. *Science* 310(5747), 480–482. <https://doi.org/10.1126/science.1118051>.
- [16] Rauste, Y.; Antropov, O.; Hame, T.; Ramminger, G.; Gomez, S.; Seifert, F. M. 2013, December. Mapping selective logging in tropical forest with space-borne SAR data. In *ESA Living Planet Symposium*, Vol. 722, p. 168.
- [17] Ortega Adarme, M., Queiroz Feitosa, R., Nigri Happ, P., Aparecido De Almeida, C., & Rodrigues Gomes, A. (2020). Evaluation of deep learning techniques for deforestation detection in the Brazilian Amazon and Cerrado biomes from remote sensing imagery. *Remote Sensing*, 12(6), 910. <https://doi.org/10.3390/rs12060910>

A Feature Selection

Table 4: Sentinel-2 Features

Band	Resolution	Wavelength	Description
B2	10 m	443.9 nm (S2A) / 442.3 nm (S2B)	Blue
B3	10 m	496.6 nm (S2A) / 492.1 nm (S2B)	Green
B4	10 m	560 nm (S2A) / 559 nm (S2B)	Red
B5	20 m	664.5 nm (S2A) / 665 nm (S2B)	Red Edge 1
B6	20 m	703.9 nm (S2A) / 703.8 nm (S2B)	Red Edge 2
B7	20 m	740.2 nm (S2A) / 739.1 nm (S2B)	Red Edge 3
B8	10 m	782.5 nm (S2A) / 779.7 nm (S2B)	NIR (Near Infrared)
B8A	20 m	835.1 nm (S2A) / 833 nm (S2B)	Narrow NIR
B11	20 m	864.8 nm (S2A) / 864.2 nm (S2B)	SWIR 1
B12	20 m	945 nm (S2A) / 943 nm (S2B)	SWIR 2

The Sentinel-1 dataset included 4 band values (VV December, VH December, VV January, VH January), the Sentinel-2 dataset contained 20 band values (B2 December, B2 January, B3 December,

B3 January, B4 December, B4 January, B5 December, B5 January, B6 December, B6 January, B7 December, B7 January, B8 December, B8 January, B8A December, B8A January, B11 December, B11 January, B12 December, and B12 January), and the combined dataset featured 24 band values.

B Model Info

For Random Forest, we optimized the number of trees, tree depth, and learning rate; for XGBoost, we tuned the learning rate, max depth, and number of boosting rounds.

A PyTorch implementation of U-Net for semantic segmentation of high-quality images was adapted for our purposes (Milesial, n.d.). The primary hyperparameters tuned were the learning rate and batch size—with higher learning rates favoring smaller batch sizes, and vice versa. The model achieved its best performance in terms of loss and accuracy with a batch size of 64 and a learning rate of $1.1e-4$. Moreover, using the Adam optimizer resulted in superior performance compared to gradient descent, with Cross Entropy Loss serving as the loss function.

The Random Forest model was developed using the scikit-learn library. This ensemble method determines whether a pixel has been logged based on its surface reflectance values. Essentially, the model constructs multiple decision trees, each of which independently votes on the pixel’s classification, with the final label determined by the majority vote.

Gradient Boosted Trees are ensemble methods that, like Random Forests, combine multiple decision trees. However, while Random Forests build independent trees and aggregate their predictions at the end, Gradient Boosted Trees employ a sequential boosting approach, where each tree is trained to correct the errors of its predecessors. For this study, we utilized XGBoost—a popular implementation renowned for its rapid training times, high performance, and advanced regularization features that enhance the model’s generalization capabilities.

The dataset was divided into training and testing sets using a 90:10 split. Additionally, select small regions were designated as testing sites to generate prediction maps for logged and stable forests. Pixel-level predictions were then used to create detailed logging maps with MATLAB and the Seaborn library.

C Results & Analysis

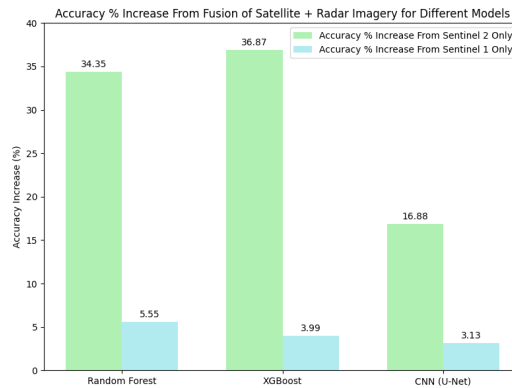


Figure 3: Accuracy Increase From the Fusion of Satellite + Radar Imagery for Various Models Tested.

(Figure 3 compares the accuracy gains achieved by fusing radar (Sentinel-1) and optical (Sentinel-2) imagery, relative to using either sensor alone, across three different models: Random Forest, XGBoost, and CNN (U-Net). The left bar in each pair (green) shows the accuracy increase from Sentinel-1-only to the fused dataset, whereas the right bar (blue) represents the accuracy increase from Sentinel-2-only to the fused dataset. All three models exhibit much larger accuracy improvements when transitioning from Sentinel-1-only data to the integrated dataset. For instance, Random Forest shows an increase of 34.35% in accuracy, while XGBoost and the CNN show increases of 36.87% and

16.88%, respectively. This indicates that radar data alone captures less of the selective logging signal, and that adding optical features provides significant complementary information. The accuracy gains from Sentinel-2-only to the fused dataset are comparatively smaller, ranging from about 3% to 5%. Random Forest experiences the largest overall improvement (34.35% from Sentinel-1 to fused), suggesting that it benefits most from the added optical information. XGBoost shows a similar pattern but with an even greater jump of nearly 37%. The CNN also improves notably, though its jump from Sentinel-1 to fused (17%) is smaller than those of the ensemble methods, likely reflecting the CNN’s stronger baseline performance on Sentinel-1 alone.

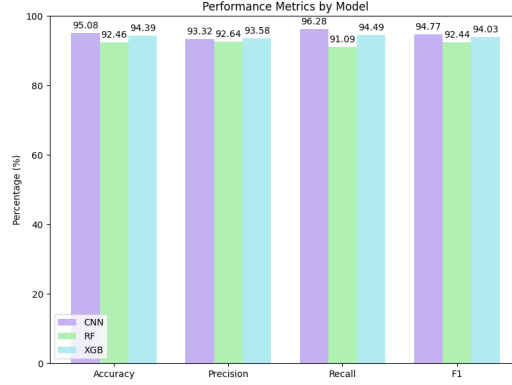


Figure 4: Performance Metrics For the Various Models Tested on the Integrated Radar & Satellite Dataset.

(Figure 4 focuses exclusively on the integrated dataset, showing four key metrics—accuracy, precision, recall, and F_1 score—for each model. All three methods achieve high accuracy, exceeding 90%. However, the breakdown reveals small differences in precision and recall, indicating trade-offs in how each model handles false positives versus false negatives. The CNN tends to balance these metrics well, leading to a slightly higher F_1 score overall. XGBoost, on the other hand, often shows comparable or marginally lower precision/recall, suggesting it may prioritize capturing true positives at the expense of a few more false alarms or vice versa. Random Forest still performs well across all metrics but remains slightly behind the other two methods, reflecting the consistent trend seen in the overall accuracy measurements from (Figure ??).

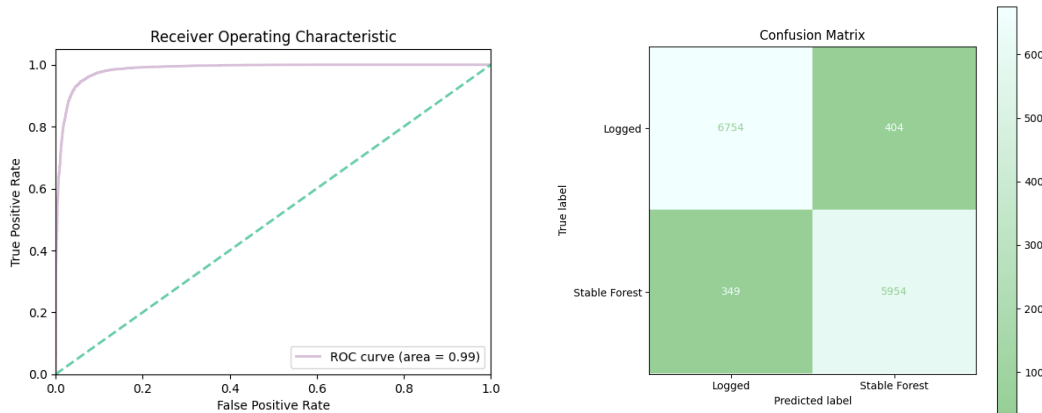


Figure 5: (Left) Area Under Curve for the CNN U-Net implementation. (Right) Confusion matrix for the CNN U-Net implementation.

Figure 5 depicts the trade-off between the true positive rate (TPR) and the false positive rate (FPR) across different classification thresholds. The area under the curve (AUC) is 0.99, indicating that the CNN U-Net model can effectively separate logged and stable forest classes. An AUC close to 1.0 suggests that the model consistently achieves high TPR while maintaining a low FPR, reflecting

strong overall discrimination between the two classes. The confusion matrix provides a more detailed view of model predictions compared to ground truth. Most samples lie along the diagonal (top-left and bottom-right cells), signifying correct predictions of “Logged” as logged and “Stable Forest” as stable. The off-diagonal entries are relatively small, indicating few false positives (predicting logged when stable) and few false negatives (predicting stable when logged). This balance aligns with the high AUC observed in Figure 4(a) and confirms that the CNN U-Net performs well in minimizing both missed detections and false alarms.

The comparative analysis of machine learning algorithms further supports the efficacy of the integrated data. The CNN (U-Net) emerged as the top-performing model, consistently achieving the highest accuracy and F_1 scores, likely because it learns spatial features directly from the imagery. Nevertheless, XGBoost demonstrated performance that was only marginally lower, while being less computationally demanding. Random Forest, although slightly behind in accuracy, remains appealing for its interpretability and minimal tuning requirements.