Evaluation of a U-Net Based Method for Segmenting Unauthorized Airstrips in the Amazon Rainforest using SAR Imagery

Leandro da Silva Gomes[1], Elcio Hideiti Shiguemori[2], Tahisa Neitzel Kuck[2], Dimas Irion Alves[1]
[1]Technological Institute of Aeronautics (ITA), São José dos Campos/SP - Brasil
[2]Institute for Advanced Studies (IEAv), São José dos Campos/SP - Brasil

***Abstract*— In recent times, the Amazon Rainforest in Brazil has faced significant environmental damage as a result of illegal human activities often facilitated by unlawful landing strips. These activities include mining and drug trafficking. This research presents a method for detecting these airstrips in 16-bit per pixel synthetic aperture radar (SAR) images using a U-Net model. The rainforest is extensive and known for its dense vegetation. The initial phase involved creating a database of SAR images from the Sentinel-1 satellite in their original TIF format. The database included 646 images for training, 277 for validation, and 117 for generalization testing. The findings suggest that the U-Net network can effectively detect features corresponding to the airstrips and accurately segment the targets in synthetic aperture radar images with 16 bits.**

***Keywords*— SAR Images; Unregistered Airstrips; Illegal Airstrips; Sentinel-1; U-NET**

I. INTRODUCTION

The Brazilian Amazon, spanning over 5 million square kilometers, is an expansive and diverse region with biodiversity and natural resources [1]. Nevertheless, its area also has challenges associated with illegal practices. A significant issue is the existence of unauthorized airstrips, often employed for illicit activities such as mining and drug trafficking [2], [3]. Their detection and monitoring are frequently impeded when attempting to use conventional optical satellite images due to the dense cloud cover characteristic of the Amazon, particularly during certain seasons. This frequent cloud cover hinders the detection of illegal activities and the assessment of environmental impacts [4]. In response, Synthetic Aperture Radar (SAR) has become a tool that provides high-resolution images of terrestrial and marine surfaces, regardless of weather conditions or daylight. The capability of SAR radars to identify targets and track environmental changes across large areas under diverse atmospheric conditions is well-documented and acknowledged in both scientific and practical applications [5], [6]. Figure 1 presents a SAR image obtained even under adverse weather conditions.

Nevertheless, the complex nature of these images can pose significant challenges for human interpretation [7]. Given this complexity, integrating computational technologies, such as Artificial Intelligence (AI), particularly Convolutional Neural Networks (CNNs), has proven effective in extracting meaningful information from SAR images. With their deep learning capabilities and capacity to process large datasets, CNNs aid in accelerating decision-making utilizing these images [4], [7]–[11]. Research utilizing CNNs for the automatic detection of features in radar images suggests a promising future for the application of AI in this domain [7], [8], [10], [11].

The dissertation presented in [7], for instance, explores the deployment of a Convolutional Neural Network (CNN) for the automated detection of small maritime vessels within SAR imagery of the ICEYE constellation. The aim is to improve detection methods in contexts where conventional visual approaches fall short, utilizing the YOLOv4 network customized for spotting vessels in SAR imagery with a spatial resolution of 2.5 meters per pixel. Key outcomes of the central experiment include a Mean Average Precision (mAP) of 64% and a global accuracy of 82.75% in the generalization test. Nevertheless, the algorithm demonstrated a notable rate of false alarms in oceanic regions not covered in the training data. This suggests more comprehensive data is needed to enhance the model's generalization abilities.

Pang et al. [8] propose utilizing a more streamlined YOLO model for object detection, specifically the YOLOv5-MNE, tailored for ship detection in SAR imagery. This model addresses the challenges of unclear target outlines and complex backgrounds. The approach optimizes training speed and memory usage by reducing the number of model parameters while maintaining credible accuracy across extensive datasets. The study introduces the MNEBlock, a redesigned lightweight block based on the MobileNetV3 architecture. This block utilizes the principles of efficient channel attention (ECA) to enhance the model's feature extraction capabilities without significantly increasing computational overhead. The architecture also substitutes the SiLU activation function with the ReLU function at the network's output to minimize parameters and integrate the Coordinated Attention (CA) mechanism to boost detection capabilities. The findings indicate that YOLOv5-MNE attains a precision of up to 94.8%. The research further evaluates the impact of various modules and settings on the model's performance, underscoring the effectiveness of the MNEBlock and CA mechanism in enhancing ship detection precision in SAR images.

In another study, Suresh et al. [12] introduce one of the initial methods for automatically detecting oil spills in SAR images from the ENVISAT satellite. Observations noted that these patches appear dark in images due to the diffraction effects in backscatter, and the detection methods successfully differentiated oil spills from other dark structures, achieving an accuracy of up to 76% in pinpointing the desired targets.

When training some AI models, sometimes there is an intrinsic challenge when the object of interest occupies sig-

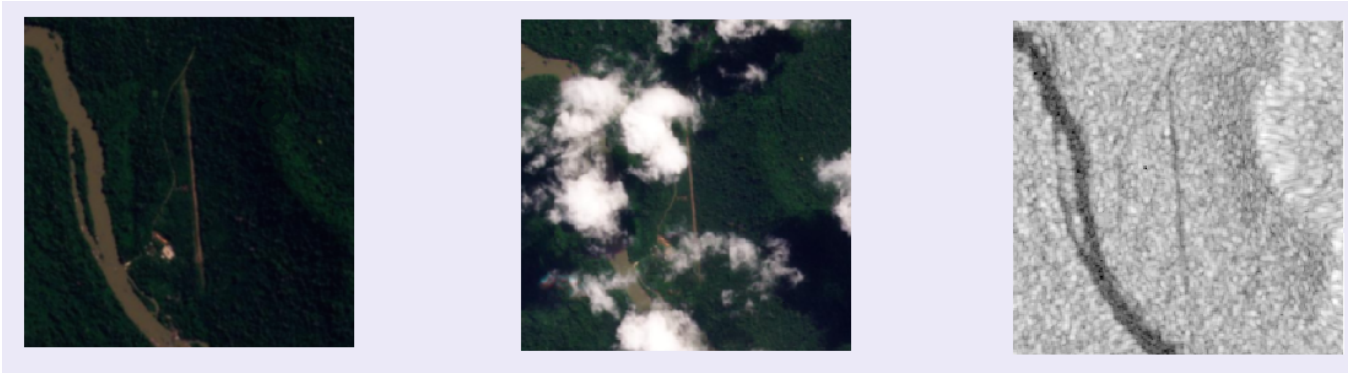

Fig. 1. Airstrip near Bacajá river, Brazil. a) PLANET image with clouds. b) PLANET image without clouds. c) Sentinel-1 SAR image

nificantly fewer pixels than the background. Considering this, [13] uses a specific loss function. The study is about developing a fully convolutional neural network, termed V-Net, architected explicitly for the segmentation of Magnetic Resonance Imaging (MRI) prostate volumes. This network uses volumetric convolutions suitable for 3D data. It employs a specific loss function (based on the Dice coefficient) that mitigates the bias towards large background areas typical in medical imaging. The team augmented the data by applying random non-linear transformations and histogram matching to deal with the limited number of annotated volumes available for training. The approach achieved good performances on challenging test data while requiring only a fraction of the processing time needed by other previous methods. In the central experiment, the V-Net achieved the better result with a Dice coefficient of 0.869 ± 0.033 and a Hausdorff distance of 5.71 ± 1.20 mm.

Considering the detrimental effects of illicit airstrips in the Amazon, the promising results from AI studies with the CNN models, alongside the benefits of utilizing SAR imagery and the existing gap in research for detecting such airstrips with SAR technology, this study is designed to assess how effectively the CNN model U-Net can identify these targets.

The organization of this paper is as follows: Section 2 describes the technical aspects, features, and capabilities of the U-Net architecture. Section 3, titled Materials and Methods, describes the methodological approaches used, including how data was gathered, the neural network configuration, the training processes, and the evaluation metrics employed. Section 4 details the study's results, evaluating how well the U-Net model segments unauthorized airstrips. The paper concludes with a summary of the main findings and recommendations for further research in the Final Remarks section.

II. The U-Net Model

The segmentation of objects is a critical task in computer vision, and it is effectively carried out by the U-Net, a convolutional neural network (CNN) designed for this purpose [14]. Developed by Olaf Ronneberger, Philipp Fischer, and Thomas Brox at the University of Freiburg, U-Net is distinguished for its architecture, which integrates a contracting path to capture context and a symmetrically expanding path that ensures precise localization of segmentation. This structure has set new benchmarks, particularly in challenges like the International Symposium on Biomedical Imaging (ISBI) competition to segment neuronal structures in electron microscopic stacks.

The U-Net architecture is described as a U-shaped map, facilitating a balance between image context capture and localization. This balance is achieved through an expansive path that features upsampling operations, which enhance the output's resolution and allow for detailed segmentation at a pixel level. The Figure 2 is an adaption from [14] and shows the architecture that employs a series of convolutional operations. Each layer in the expansive path corresponds to a layer in the contracting path, maintaining a symmetry that is key for the network's performance.

This network has been applied to various biomedical segmentation tasks, demonstrating its efficiency. For instance, in the ISBI cell tracking challenge 2015, U-Net's application to different types of light microscopy images resulted in superior segmentation accuracy, winning the challenge by a significant margin. The network's ability to segment complex images often in less than a second for a 512x512 image on modern GPUs, underscores its practical utility in clinical and research settings [15]. In conclusion, the U-Net architecture has impacted the field of image segmentation, combining innovative architectural features with training strategies to set new standards in accuracy and efficiency [16].

III. Materials and Methods

To evaluate the U-Net model for segmenting illegal landing strips, we had to create a database of SAR images since there is no existing dataset for this purpose. For each image, we needed to generate its segmentation mask. To do this, the initial step involved confirming the locations of airstrips through MapBiomas, a collaborative initiative in Brazil that includes academic institutions, technology companies, and non-governmental organizations [17]. This project is dedicated to mapping and monitoring land use and cover changes nationwide through remote sensing techniques and geospatial analytics. It offers geographical data encompassing sanctioned and unsanctioned airstrips, which aids in logistical coordination and field research within the Amazon region [18]. Figure 3 depicts the airstrips identified by the project.

A compilation of geographic coordinates, presented in decimal degrees for longitude and latitude, was utilized to pinpoint areas with unauthorized airstrips. This compilation was integrated into a Python script run within the Google Colab platform. The script's initial steps included installing and setting up the Google Earth Engine (GEE) API, enabling access to its data handling and geospatial analysis features.

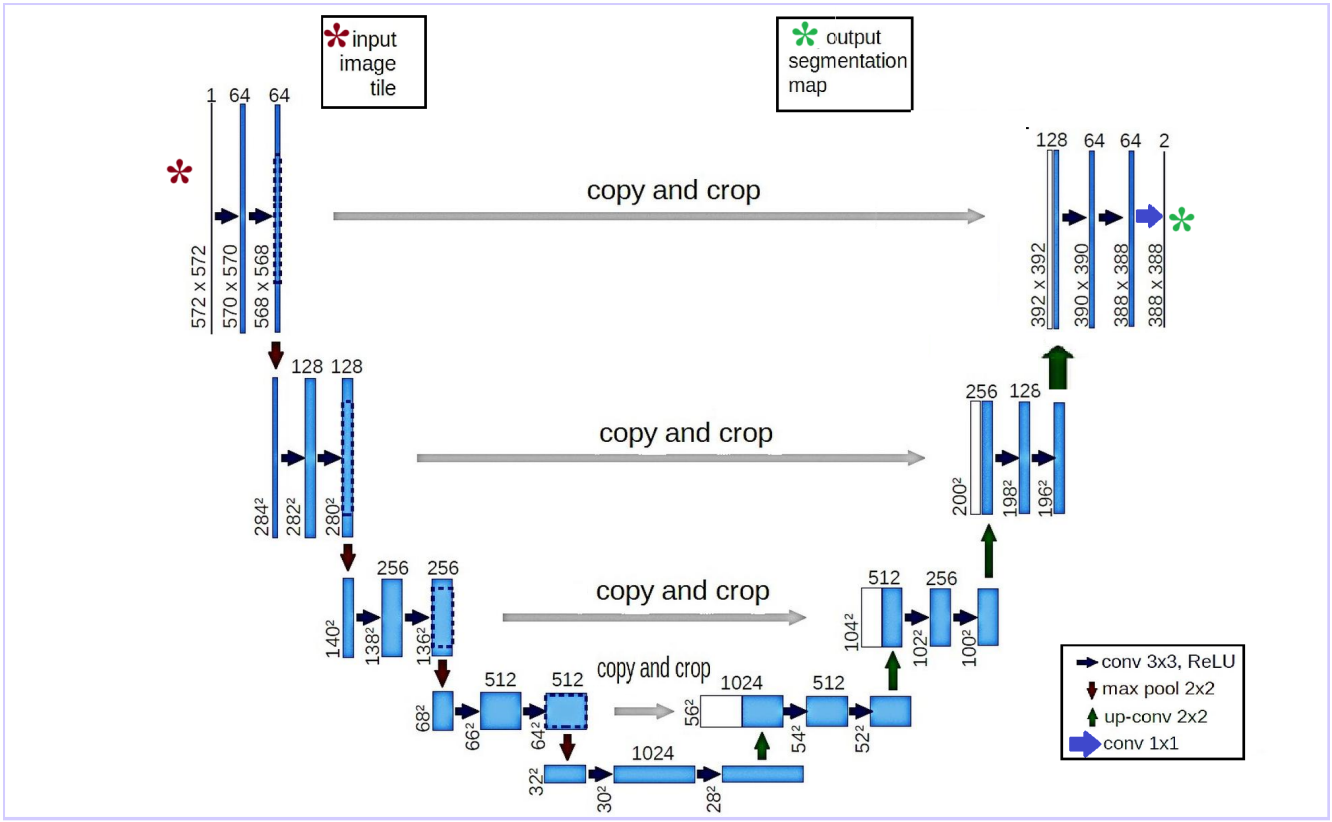

Fig. 2. Architecture of U-Net.

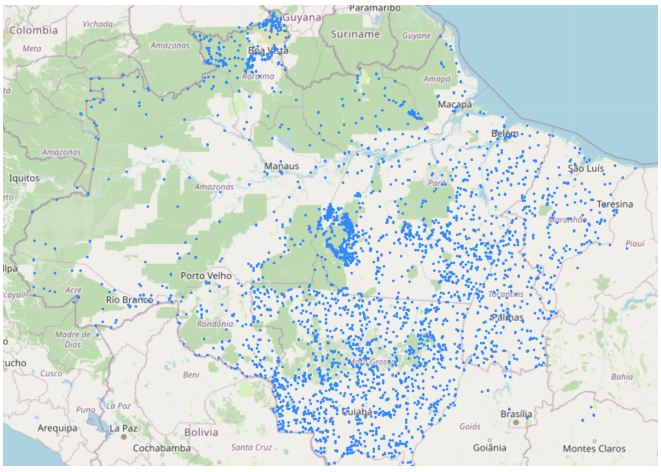

Fig. 3. Airstrips in the Amazon Rainforest recorded by MapBiomas.

The script generated a satellite image covering a 4 km square area for each identified airstrip location, represented in 512 × 512 pixels images. These images, sourced from the Sentinel-1 satellite for the year 2023—the most recent update year for the MapBiomas airstrip database—were captured in Interferometric Wide (IW) mode and utilized vertical-vertical (VV) polarization. The targets were segmented, generating images where each pixel that makes up the target of interest receives a value of 255, and pixels that make up the background receive a value of 0. This process was aided by using high-resolution images from the PLANET constellation, corresponding to the exact dates as the SAR images. This approach was adopted due to the need to correctly identify the tracks since the visual characteristics of these structures are often poorly discernible in radar images. Furthermore, targets are volatile: tracks may be temporary, with vegetation quickly covering previously deforested areas, or appear suddenly due to illicit activities, such as mining, requiring verification of visible spectrum images regarding its existence for correct segmentation. Figure 4 exemplifies two images of illegal airstrips located in the Amazon generated for the dataset with their respective segmentation masks.

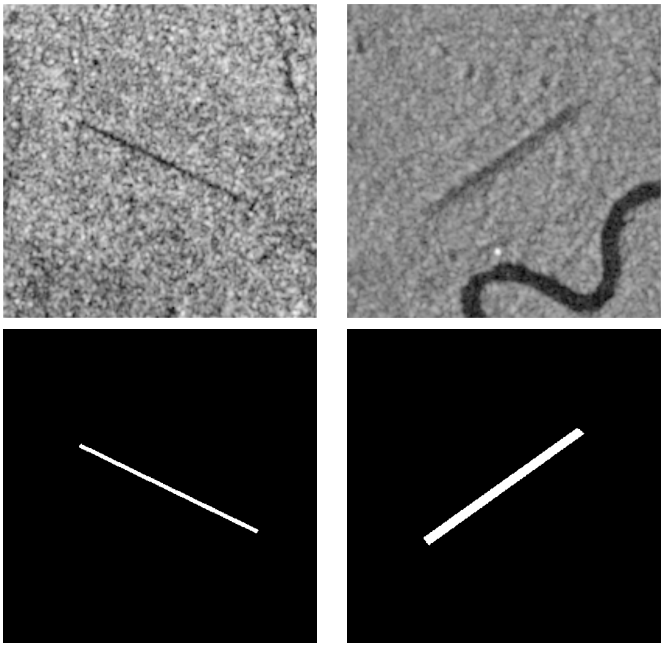

Fig. 4. Examples of images generated for the dataset.

After completing the image acquisition process, 1040 ima-

ges were obtained. To ensure statistical significance and reduce model bias, we balanced the database by generating an image without an airstrip for each sample containing an airstrip. Balancing the data helps to prevent the model from overfitting to the more frequent class and improves its ability to generalize to new data. Of the 1040 airstrips, 117 were exclusively reserved for the generalization test. This set aside of images for generalization testing helps evaluate how well the model performs on unseen data, thus measuring its ability to generalize beyond the training and validation datasets. A Numpy array containing float32 data was used to feed the UNet model, which was implemented using the Keras TensorFlow library. Initially, an empty array with the shape (0, 512, 512, 1) was created to store the images, and another array with the shape (0, 512, 512, 2) was created for the categorical masks. The choice of 512x512 image size, as supported by [15], facilitates efficient computational processing while retaining sufficient detail for effective learning and segmentation. The images were loaded one by one and added to the image array. Simultaneously, the corresponding masks were loaded and added to the mask array. After data preparation, the image and mask arrays were split into training and validation sets using 30% of the data allocated for validation. Considering the results obtained by [13], the loss function based on the Dice Coefficient shown in equation (1) was used during the training:

$$D = \frac{2\sum_{i=1}^{N} p_i g_i}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2}, \tag{1}$$

where p_i represents the predicted binary segmentation for the i-th pixel and g_i represents the ground truth binary segmentation for the i-th pixel. The sums run over all N pixels in the images. This coefficient ranges from 0 to 1, with 1 indicating perfect overlap. The Dice loss, therefore, is calculated as $1 - D$, which is minimized during training to improve segmentation accuracy. The gradient of the Dice coefficient concerning each pixel in the prediction is computed and used to optimize the model, allowing it to effectively handle the imbalance between foreground and background pixels, leading to improved segmentation performance.

IV. Evaluation of Results

The U-Net model was trained using the dice loss function proposed by [13] and the Adam optimizer in the experiment. Around epoch 40, the training loss continued to fall, but validation loss did not improve significantly. The training was stopped because there were no performance improvements on the validation data until epoch 100. The network's best weights were selected and used in the generalization test images. Figure 5 shows the train and validation loss curves.

The accuracy presented in the validation data exceeded 98%. However, it is acknowledged that this metric may not be the most suitable for evaluating the model's performance in this context. The primary reason is that accuracy is calculated on a per-pixel basis, and the objects of interest in our dataset occupy significantly fewer pixels than the background. Consequently, the model may achieve high accuracy by predominantly predicting the background class and skewing the metric. As an experiment, a refined evaluation was adopted for our segmentation model to effectively address the limitations of traditional accuracy metrics in this scenario where the

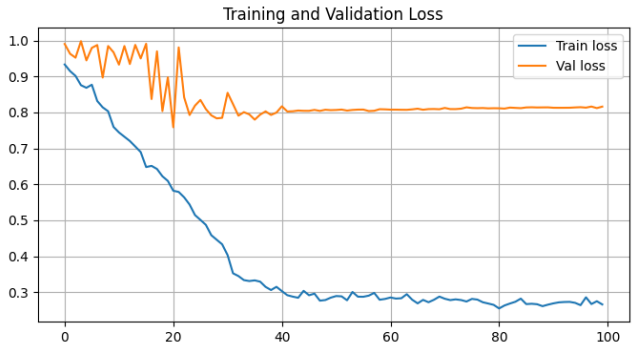

Fig. 5. The validation and train loss curves

object of interest occupies a significantly smaller area than the background. We implemented a 50% binarization threshold, where a pixel's probability above this threshold denotes it as part of the object class. Furthermore, for a predicted mask to be deemed successful in detecting an object, it required a minimum of 50% overlap with the true mask. This method was applied to the generalization test images. For each image, we calculated the intersection between the binarized predicted mask and the true mask, defining a detection as successful if at least half of the true mask pixels were covered by the predicted mask. This comprehensive evaluation approach, with this level of stringent criteria, led to the successful detection of 27 airstrips across the test dataset. Figure 6 presents an example of a successful detection.

This result suggests that, with less rigorous criteria, the model could identify even more features, demonstrating its utility in accurately detecting relevant features in complex backgrounds. However, this potential increase in detections could come at the cost of a higher rate of false positives, requiring careful consideration of the trade-off between detection sensitivity and precision. Figure 7 presents 1 of 21 examples of cases that were below the 50% overlap rate between real and predicted masks but were able to segment part of the targets without confusion with background component structures.

Considering the total of images, there were seven cases in which the presence of roads led the algorithm to segment these structures incorrectly. Figure 8 presents one of these cases.

The rest of the non-segmented landing strips generally correspond to targets with few features visible in the images. They are landing strips with dimensions of an average of 20 meters wide, which generates a few digitized pixels in Sentinel-1 images with a spatial resolution of 10 meters per pixel. Figure 9 exemplifies one of these cases.

The results reported herein reflect the inherent complexities associated with SAR images, which are known for their capabilities in penetrating weather conditions such as cloud cover, which is a significant advantage over optical imagery. However, they also present substantial challenges due to their high sensitivity to surface roughness and moisture content, leading to considerable noise and speckle effects. These factors inherently complicate the task of accurate feature detection. Furthermore, SAR imagery's dependency on the angle of incidence and complex interactions with different materials can result in ambiguous and non-intuitive image features, adding another layer of difficulty for segmentation models. Despite

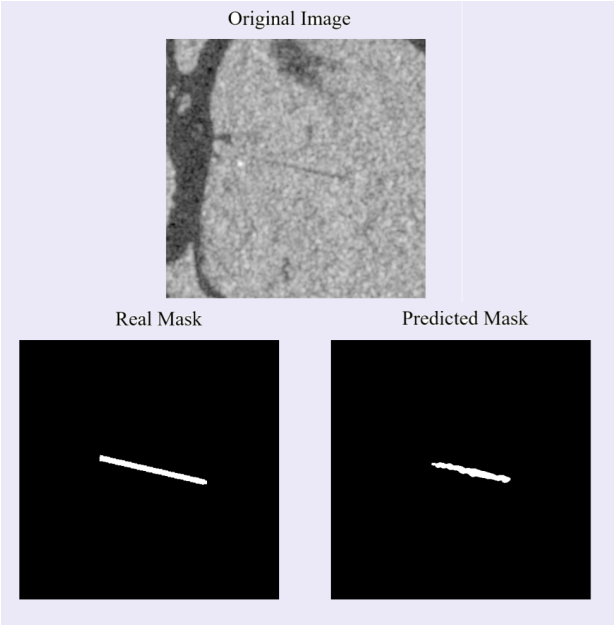

Fig. 6. Example of a correctly detected track according to the established criteria

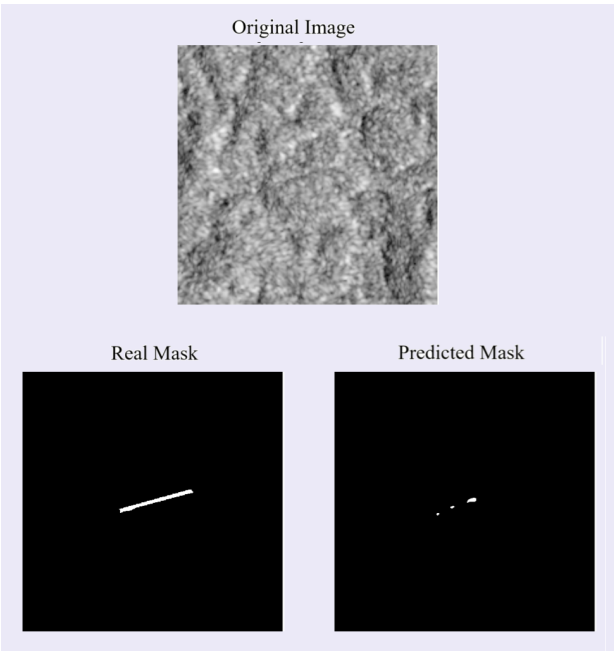

Fig. 7. Example of a correctly detected track below to the established criteria

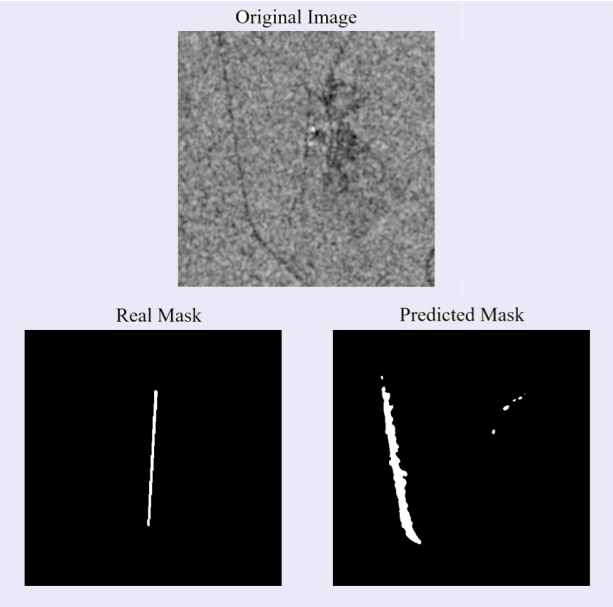

Fig. 8. Road segmented as airstrip

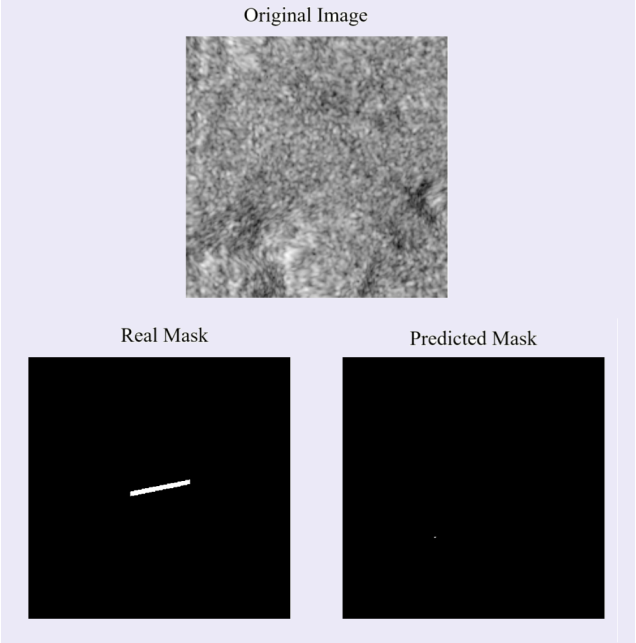

Fig. 9. Road segmented as airstrip

these challenges, this study represents a preliminary model that successfully demonstrates the ability to detect airstrips under conditions where visible imagery would fail, such as under cloud cover. This capability showcases the potential of SAR imagery for reliable earth observation regardless of weather conditions. It sets a promising foundation for future refinements and applications of deep learning models in interpreting SAR data.

V. Concluding Remarks

This study evaluates the use of the U-Net model in segmenting airstrips from SAR images, an endeavor characterized by significant challenges due to the complex nature of the dataset and inherent properties of SAR imagery. Despite these complexities, the model demonstrated a capability to detect airstrips under conditions where traditional optical sensors would be ineffective, such as in cloudy or obscured weather conditions. The U-Net model's performance, marked by detecting 27 airstrips from the generalization test dataset, underscores its potential in earth observation and remote sensing. The experiment's outcomes reinforce the need for a nuanced approach to evaluating model performance, especially in highly imbalanced classes where the object of interest, such as airstrips, occupies a relatively small portion of the image data. Implementing a 50% binarization threshold and the requirement for a 50% overlap between the predicted and true masks indicated effectiveness in emphasizing the practical utility of the model over mere statistical accuracy. Furthermore, the model's ability to detect even partial airstrips

using less stringent criteria suggests the possibility of its application in monitoring and surveillance tasks in remote and less accessible regions. However, false positives, primarily due to the misidentification of roads as airstrips, highlight areas for future improvement. These issues highlight the importance of refining the model to enhance its precision and reduce the likelihood of false detections. A suggestion for a future study would be to add images with roads and highways to the dataset to help the algorithm better identify these structures and distinguish them from landing strips. The search for better neural network hyperparameters can also help.

The continued development and refinement of deep learning techniques for interpreting SAR data are essential. We can improve these technologies' accuracy and utility, facilitating their broader application in challenging and diverse environments by advancing these technologies. The insights gained from this study lay the foundation for future research, aiming to harness the full potential of SAR imagery for reliable and effective monitoring of geographic features and human activities, even under adverse conditions. This study demonstrates the capability of deep learning models like U-Net in a novel application. It highlights that it opens the door for more sophisticated approaches to segmenting and analyzing SAR data for environmental and security monitoring.

VI. Acknowledgments

The authors of this article would like to thank the Brazilian Air Force (FAB), the Postgraduate Program in Operational Applications (PPGAO), and the Electronic Warfare Laboratory (LabGE-ITA) for their support during this work.

Referências

[1] IBGE. (2024) Amazônia legal. Accessed : 14 april 2024. [Online]. Available: https://www.ibge.gov.br/geociencias/cartas-e-mapas/mapas-regionais/15819-amazonia-legal.html

[2] I. S. (ISA), "Yanomami under attack: Illegal mining on yanomami indigenous land," Instituto Socioambiental, Technical Report, 2022, available online: https://acervo.socioambiental.org/acervo/documentos/yanomami-under-attack-illegal-mining-yanomami-indigenous-land (Accessed: 14 April 2024).

[3] A. Y. Hutukara and A. W. Ye'kwana, "Scars in the forest: evolution of illegal mining in the yanomami indigenous land in 2020," in *Monitoring System of Illegal Mining in the Yanomami Indigenous Land*, I. S. (ISA), Ed. Instituto Socioambiental (ISA), 2021, pp. 1–50.

[4] T. Kuck, P. Silva Filho, E. Sano, P. Bispo, E. Shiguemori, and R. Dalagnol, "Change detection of selective logging in the brazilian amazon using x-band SAR data and pre-trained convolutional neural networks," *Remote Sens.*, vol. 13, p. 4944, 2021, accessed: 12 april 2024. [Online]. Available: https://doi.org/10.3390/rs13234944

[5] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. March 2013, pp. 6–43, accessed: 20 maio 2024. [Online]. Available: https://doi.org/10.1109/MGRS.2013.2248301

[6] F. Meyer, "Spaceborne synthetic aperture radar: Principles, data access, and basic processing techniques," in *The SAR Handbook*. Fairbanks, AK, USA: University of Alaska Fairbanks, 2021, pp. 1–53.

[7] H. C. T. De Oliveira, "Detecção automática de pequenas embarcações em imagens SAR da constelação iceye com uso de uma rede neural convolucional," São José dos Campos, p. 84, 2021.

[8] L. Pang, B. Li, F. Zhang, X. Meng, and L. Zhang, "A lightweight yolov5-mne algorithm for SAR ship detection," *Sensors*, vol. 22, p. 7088, 2022.

[9] T. Kuck, E. Sano, P. Bispo, E. Shiguemori, P. Silva Filho, and E. Matricardi, "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 Sens.*, vol. 13, p. 3341, 2021, accessed: 12 abril 2024. [Online]. Available: https://doi.org/10.3390/rs13173341

[10] S. Lei *et al.*, "Srsdd-v1.0: A high-resolution SAR rotation ship detection dataset," *Remote Sens.*, vol. 13, p. 5104, 2021.

[11] T. Zhang *et al.*, "SAR ship detection dataset (SSDD): Official release and comprehensive data analysis," *Remote Sens.*, vol. 13, p. 3690, 2021.

[12] G. Suresh, C. Melsheimer, J.-H. Körber, and G. Bohrmann, "Automatic estimation of oil seep locations in synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 4218–4229, 2015.

[13] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," *ArXiv*, vol. 1606.04797, 2016. [Online]. Available: https://arxiv.org/abs/1606.04797

[14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, pp. 234–241, 2015. [Online]. Available: http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/

[15] A. Mali. (2018) Ultrasound image segmentation using u-net. Accessed:11 de abril de 2024. [Online]. Available: https://ashm8206.github.io/2018/03/07/Ultrasound-Image-Segementation-Using-Unet.html

[16] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "Unet 3+: A full-scale connected unet for medical image segmentation," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1055–1059.

[17] MapBiomas. (2024) O projeto mapbiomas. Accessed: 12 abril 2024. [Online]. Available: https://brasil.mapbiomas.org/o-projeto/

[18] M. P. Ilegais. (2024) Pistas de pouso mapeadas pelo porjeto mapbiomas. Accessed: 12 abril 2024. [Online]. Available: https://plataforma.brasil.mapbiomas.org/pistas-de-pouso