GEOPROCESSING AND CONVOLUTIONAL NEURAL NETWORKS: ANALYSIS OF LAND COVER AND LAND USE IN THE ALMADA RIVER BASIN (STATE OF BAHIA, BRAZIL)

Hercules da Silva Carvalho
Universidade Federal do Sul da Bahia,
Centro de Formação em Tecnociências e Inovação, Ilhéus, BA, Brasil
herculesi1343@gmail.com

Vinícius de Amorim Silva
Universidade Federal do Sul da Bahia,
Centro de Formação em Tecnociências e Inovação, Ilhéus, BA, Brasil
vinicius@ufsb.edu.br

Paulo Sérgio Vila Nova Souza
Universidade Federal do Sul da Bahia,
Centro de Formação em Ciências Agroflorestais, Ilhéus, BA, Brasil
paulosvn@gfe.ufsb.edu.br

ABSTRACT

Geoprocessing techniques associated with convolutional neural network models (CNN) emerge as a viable alternative for obtaining data to subsidize decision-making. In this context, this work aimed to evaluate the application of CNN algorithms for the classification and detection of land cover and use classes in satellite images of the Almada River Basin (ARB). To achieve the goal, logical steps were structured: (i) information collection; (ii) processing of Dataset Eurosat; and (III) the evaluation of the models. Classification results demonstrated more than 90% precisions in class recognition. As for the detection model, a 70% accuracy was identified for the “Forest” and “Pasture” classes, which have large extensions within ARB. Both models showed their versatility in application and viability as tools for monitoring the physical and environmental conditions of ARB. In this sense, the effectiveness of the models is emphasized in the identification and location of land cover and use classes, emphasizing the importance of building a dataset that highlights the characteristics of the study area. This contributes to obtaining reliable results and improving the practical use of CNN models.

Keywords: Deep learning. Computer vision. Geospatial analysis. Geographic information system. Remote sensing.

GEOPROCESSAMENTO E REDES NEURAIS CONVOLUCIONAIS: ANÁLISE DA COBERTURA E USO DA TERRA NA BACIA HIDROGRÁFICA DO RIO ALMADA (BAHIA – BRASIL)

RESUMO

As técnicas de geoprocessamento associadas aos modelos de Redes Neurais Convolucionais (CNN) emergem como uma alternativa viável para a obtenção de dados que possam subsidiar a tomada de decisões. Neste contexto, este trabalho tem como objetivo avaliar a aplicação de algoritmos de CNN para a classificação e detecção das classes de cobertura e uso da terra nas imagens de satélite da Bacia Hidrográfica do Rio Almada (BHRA). Para alcançar o objetivo, são estruturadas etapas lógicas: (i) coleta de informações; (ii) processamento do *dataset* EuroSat; e, (iii) a avaliação dos modelos. Os resultados da classificação demonstram precisões superiores a 90% no reconhecimento das classes. Quanto ao modelo de detecção, identifica-se uma precisão de 70% para as classes “Forest” e “Pasture”, que têm grandes extensões dentro da BHRA. Ambos os modelos evidenciaram sua versatilidade na aplicação e na viabilidade como ferramentas de monitoramento das condições físicas e ambientais da BHRA. Nesse sentido, ressalta-se a eficácia dos modelos na identificação e localização das classes de cobertura e uso da terra, enfatizando a importância da construção de um *dataset* que evidencia as características da área de estudo. Isso contribui para a obtenção de resultados confiáveis, aprimorando a utilidade prática dos modelos de CNN.

Palavras-chave: Aprendizagem profunda. Visão computacional. Análise geoespacial. Sistema da informação geográfica. Sensoriamento remoto.

INTRODUCTION

Geoprocessing and Convolutional Neural Networks (CNN) techniques have emerged as effective tools for analyzing land cover and land use. They play a crucial role in investigating measures that strengthen resilience and mitigate adverse impacts on agriculture. Furthermore, they address climate change, which, in turn, is also influenced by land cover and land use. These aspects are highlighted in Sustainable Development Goals (SDGs), especially in SDGs 2, 12 and 13. In this context, geoprocessing and CNN provide valuable information for identifying and classifying areas with the potential to contribute to carbon credit initiatives, aimed at reforestation or recovery of the degraded regions (HEGAZY et al., 2015; SILVA, 2023).

The integration of geospatial information on land cover and land use with CNN allows, after model training, the precise assignment of categories to sets of pixels in satellite images. This approach not only finds but also identifies different classes of land cover and use with considerable accuracy. The results obtained become valuable for government entities and interested parties, facilitating the rapid and effective identification of areas of deforestation and human activities that degrade environmental conditions (REX et al., 2018; MAGIRI, 2023).

Furthermore, CNN is evident as a powerful approach, capable of integrating with other technologies, as demonstrated in the methodological proposal with CNN and decision trees, used to map rice plantations, where the results obtained revealed remarkable global accuracy of 93.56%, surpassing other previously investigated methods (ZHAO et al., 2019). The justification also stands out, in which, due to limitations in the acquisition of high-resolution remote sensing images, a new fusion model is proposed that makes full use of CNN to derive high-resolution images. This method not only demonstrates effectiveness but also presents a more robust resolution compared to conventional algorithms (TAN et al., 2018).

In this way, with the advancement of technologies and access to high-resolution satellite images, new opportunities for exploring information have emerged, enabling the construction of analysis scenarios applicable in different areas of knowledge (REX et al., 2018; LISBOA et al., 2023; MAGIRI, 2023). Among these advances, the possibility of using CNN to analyze fragility in the Almada River Basin (ARB) stands out. The most fragile areas include urban areas, sandy sediments exposed in beach areas, mangroves, and degraded pockets in the coastal plains of the eastern portion and areas with exposed soil in the western region of the basin (SEI,1999; FREIRE, 2006; NASCIMENTO et al., 2007; FRANCO et al., 2011; GOMES et al., 2012; LAVENÈRE-WANDERLEY, 2018).

Furthermore, the use of CNN in the study of environmental vulnerability in ARB is an approach capable of providing insights into engineering works, such as the Porto Sul Intermodal Logistics Complex. Through analyses conducted with CNN and geoprocessing techniques, it is possible to contribute to the understanding of potential risks to environmental degradation, with the identification of pressure caused by anthropogenic actions integrating natural factors, due to human settlements and changes in environmental dynamics (SENA; VEIGA; SILVA, 2023), consolidating itself as a possible monitoring tool with results that guide investments in both the public and private sectors.

Another important aspect is the use of CNN architectures to estimate the classification probabilities of the presence of land abandonment in specific grid elements, based on satellite images, where the validation results generated by the model reinforced the perception of CNNs as tools effective for the automated detection of land abandonment (KRYSIAK et al., 2020). Thus, the ability of these auxiliary tools to recognize these areas, as well as areas of deforestation, to register them for the future implementation of Degraded Area Recovery Plans (PRAD) becomes evident.

However, for initiatives like this to come to fruition, it is essential to carry out mapping of the ARB with the identification of points that contribute to the assessment of the basin as a whole, including the classification of land cover and land use, identification of deforested or affected areas by flooding. In this context, this work aimed to evaluate the application of CNN models for the classification and detection of land cover and land use classes in ARB satellite images.

METHODOLOGY

To develop CNN models, an approach structured in logical steps is adopted. Initially, (a) information is collected, with the help of the Google Earth Engine Platform, a free platform for educational and research use, and QGIS (version 3.22.16), distributed under the GNU General Public License, which contributes to the manipulation of satellite images. Subsequently, (b) data processing proceeds, which includes normalization of input information by adjusting the pixel values for an interval between 0 and 1, and dividing

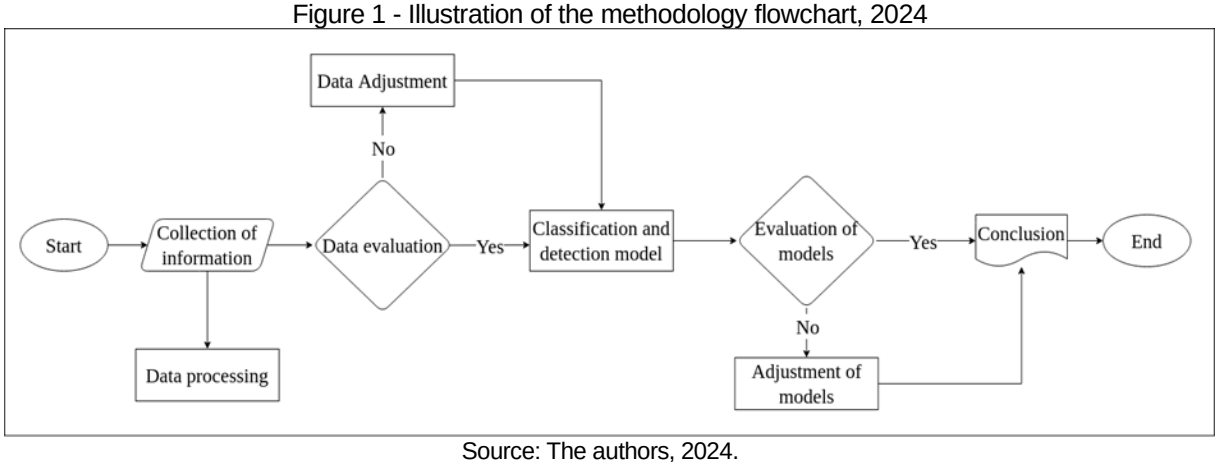
the pixels by the maximum value of the sensor, essential for training and testing classification and feature detection models. Finally, (c) data evaluation is carried out, playing a crucial role in the subsequent stages of the process.

To determine the visualization scale, the dimensions of the ARB image, with 2.400 pixels in width and 1.300 pixels in height, also from Sentinel-2, are used as a reference. Considering the scale of 1:10.000, get is resulting scale after resizing to 64x64 pixels approximately 1:400.000. This scale is adopted as the standard for other digital product cartography.

In the data collection phase, to obtain insights, the string “ALL=(land coverage and land use) AND ALL=(Convolutional Neural Network)” was searched in the scientific database of Web of Science, and for vegetation conditions in ARB, analyzing the Normalized Difference Vegetation Index (NDVI), based on the guidance available in the Google Earth Engine documentation.

In the second phase, within the Google Colaboratory environment, a free platform, with the configuration optimized for the use of Graphics Processing Units (GPU), the classification and detection models are trained and applied. The evaluation results are compiled in reports, which highlight the conclusions related to the accuracy, precision and quality of the information (Figure 1).

Figure 1 - Illustration of the methodology flowchart, 2024

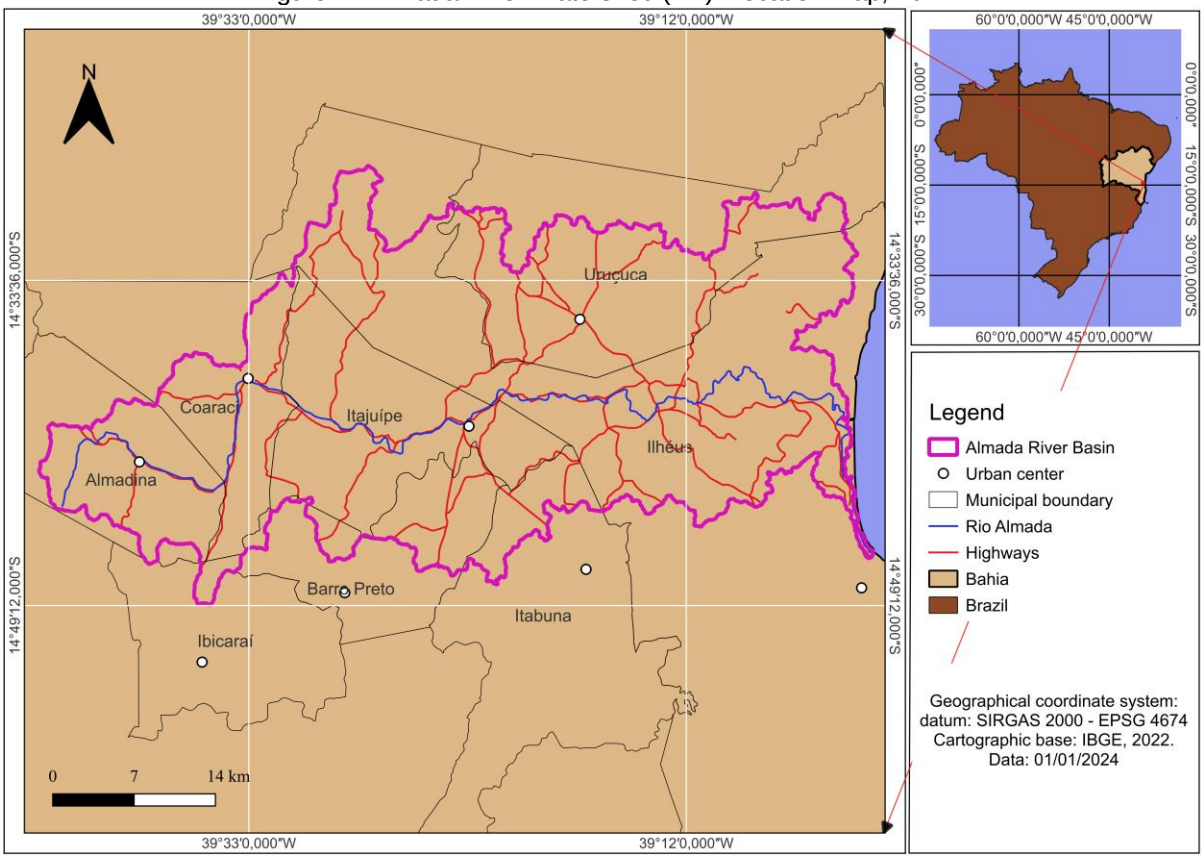


Source: The authors, 2024.

Study Area

This work was carried out on samples from ARB, a biophysical unit located in the Atlantic Forest biome, in the south of the state of Bahia. The study area covers the municipalities of Almadina, Barro Preto, Coaraci, Ibicaraí, Itajuípe, Itabuna, Ilhéus, and Uruçuca (Figure 2).

Figure 2 - Almada River Watershed (BA): Location map, 2024

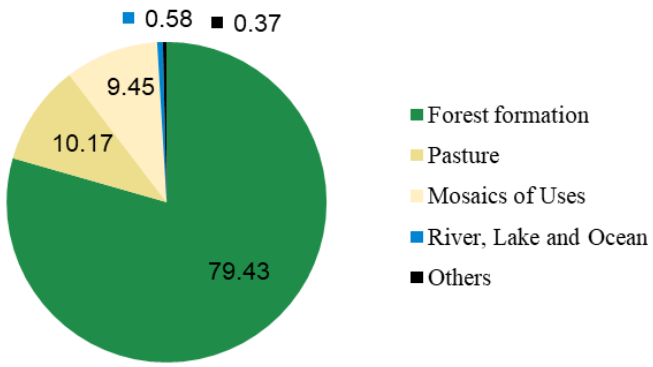


Source: IBGE, 2022. Elaboration: The authors, 2024.

The ARB covers an area of approximately 1.573 km², limited to the North and West by the Rio de Contas Hydrographic Basin and to the South by the Rio Cachoeira Hydrographic Basin. Its upstream is located in Serra do Chuchu, municipality of Almadina, at approximately 600 m altitude. The river runs for approximately 188.000 m, flowing north of the city of Ilhéus (GOMES et al., 2012).

Furthermore, it presents distinct patterns of land cover and land use. This characterization is based on analysis related to information obtained from collection 8 of the MapBiomas Project, CC-BY-SA license. Predominantly, the areas in the basin include forest formations, pastures, mosaics of uses, and bodies of water, such as rivers, lakes, and the ocean (Figure 3).

Figure 3 - Almada River Watershed (BA): Illustration of the main land coverage and land use at the ARB, 2024

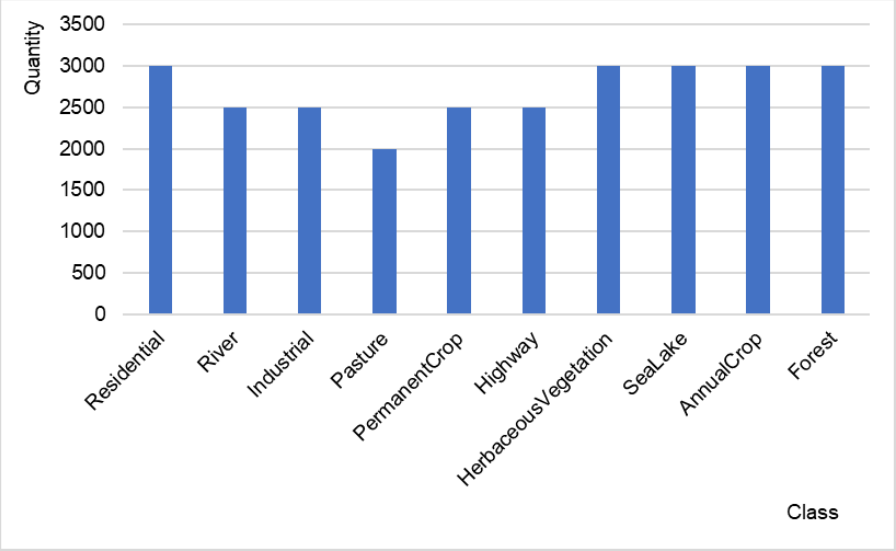


Source: MapBiomas Project. Elaboration: The authors, 2024.

Optical Data

EuroSat is a dataset developed with the specific purpose of contributing to deep learning in the area of land cover and land use classification, available for free on the official TensorFlow website. Composed of 27.000 images from the Sentinel-2 satellite, duly labeled and georeferenced, this set offers a wide range of information (Figure 4).

Figure 4 - Distribution of images from the EuroSat set in the 10 classes, 2024



Source: HELBER et al., 2018. Elaboration: The authors, 2024.

EuroSat presents two distinct data models: the first, called “rgb”, exclusively encompasses the optical bands R, G, and B, in the “jpg” file format, classified as uint8, indicating integer data without a positive sign, with storage of 8 bits. Meanwhile, the second, titled “all”, incorporates data in “tif” format, covering all 13 original Sentinel-2 bands, with 64-bit storage.

In Figure 5, the classifications related to land cover and land use are presented, accompanied by illustrations of the classes. Each image has dimensions of 64x64, with a format of (64,64,3) for “rgb” mode, indicating three color channels, and (64,64,13) for “all” mode.

Figure 5 - Illustration of the classes of land coverage and use in the EuroSat, 2024

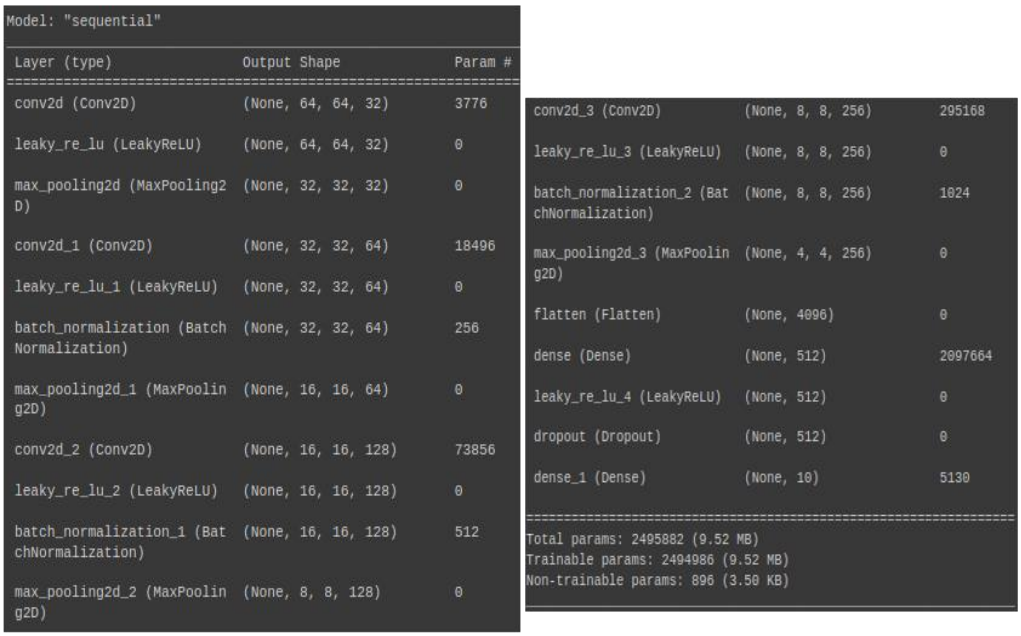
Source: HELBER et al., 2018. Elaboration: The authors, 2024.

It is worth highlighting the definition of the Sentinel-2 bands to understand how the information is distributed in the data set. This set comprises four visible and infrared bands, six shortwave infrared bands, and three atmospheric bands.

Image Classification Model

The chosen CNN model adopts a sequential structure, where each layer is organized with a predetermined number of neurons. Developed with the TensorFlow library (version 2.15.0), this model incorporates the Leaky Rectified Linear Unit (LeakyReLU) activation function and uses the Adaptive Moment Estimation (Adam) optimizer. Furthermore, the “categorical_crossentropy” loss function widely recognized in multiclass classification problems is adopted (Figure 6).

Figure 6 - Illustration of the sequential type classification model summary, 2024



```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 64, 64, 32)        3776

 leaky_re_lu (LeakyReLU)     (None, 64, 64, 32)        0

 max_pooling2d (MaxPooling2  (None, 32, 32, 32)        0
 D)

 conv2d_1 (Conv2D)           (None, 32, 32, 64)        18496

 leaky_re_lu_1 (LeakyReLU)   (None, 32, 32, 64)        0

 batch_normalization (Batch  (None, 32, 32, 64)        256
 Normalization)

 max_pooling2d_1 (MaxPoolin  (None, 16, 16, 64)        0
 g2D)

 conv2d_2 (Conv2D)           (None, 16, 16, 128)       73856

 leaky_re_lu_2 (LeakyReLU)   (None, 16, 16, 128)       0

 batch_normalization_1 (Bat  (None, 16, 16, 128)       512
 chNormalization)

 max_pooling2d_2 (MaxPoolin  (None, 8, 8, 128)         0
 g2D)

 conv2d_3 (Conv2D)           (None, 8, 8, 256)         295168

 leaky_re_lu_3 (LeakyReLU)   (None, 8, 8, 256)         0

 batch_normalization_2 (Bat  (None, 8, 8, 256)         1024
 chNormalization)

 max_pooling2d_3 (MaxPoolin  (None, 4, 4, 256)         0
 g2D)

 flatten (Flatten)           (None, 4096)              0

 dense (Dense)               (None, 512)               2097664

 leaky_re_lu_4 (LeakyReLU)   (None, 512)               0

 dropout (Dropout)           (None, 512)               0

 dense_1 (Dense)             (None, 10)                5130

=================================================================
Total params: 2495882 (9.52 MB)
Trainable params: 2494986 (9.52 MB)
Non-trainable params: 896 (3.50 KB)
_________________________________________________________________
```

Source: The authors, 2024.

In the model training and testing process, 50% of the data from the EuroSat database is utilized, which corresponds to a standard of 1.350 images for each class. It is important to highlight that, to explore the feasibility of assessments with different band configurations, data belonging to the “all” category are used, indicating the presence of all 13 Sentinel-2 bands. Geospatial data manipulation in Python is performed with the Geospatial Data Abstraction Library (GDAL) (version 3.6.4).

Consequently, in order to avoid the need for data reprocessing, and provide a more efficient later use in terms of speed and computational cost, the normalized training and testing data are stored in databases using the Pickle library (version 4.0).

Model of image detection

For the detection of objects within an image, two critical points should be highlighted: the choice of CNN architecture and labeling, made through bounding boxes. For the architecture, we opted for You Only Look Once (YOLO), version four, with Darknet framework, both open-source applications, which present a regression approach and simultaneous classification.

In addition, the model is executed in the Google Colaboratory environment, after adjustments in the files "Makefile", geo.cfg, geo.names, and geo.data, which, according to the model source, play the following roles:

- Makefile: this file contains essential rules and definitions for compiling the Darknet source code.

- *.cfg: this file is responsible for defining the specific architecture of the neural network, including the number of layers, training hyperparameters, and other fundamental configurations.
- *.names: in this file, classes of the objects of study are stored, essential for proper recognition and classification.
- *.data: This file houses the crucial settings for model training, including the paths to the training set, and test, as well as other pertinent configuration files.

The definition of the bounding boxes is conducted manually with the labeling application (version 1.8.6), an open-source application. In this procedure, the "rgb" category of EuroSat is selected. From the previous analysis of the land cover and land use, are chosen as labeling classes "Forest", "Pasture" and "River", choosing to use 500 images of each category.

Evaluation of the CNN models

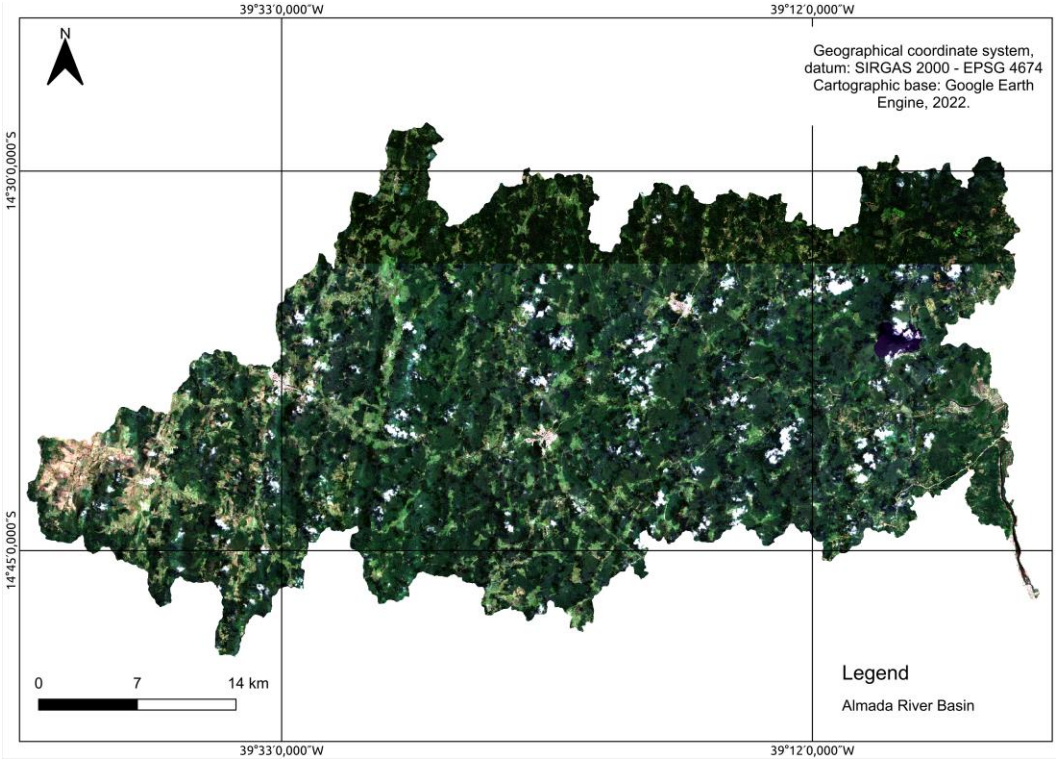
In the evaluation of the classification model are employed the accuracy metrics, which measure the proportion of correct predictions in relation to the total evaluations; The precision, which evaluates the accuracy of the predictions as for the total classifications; and the recall, which highlights the false negatives (KIM et al., 2023). In addition to the techniques mentioned, the evaluation of the confusion is used for this work.

Within the detection model, since the YOLO model provides the probability of confidence associated with each detection, the correct detections are counted and the number of false positives is recorded, considering the percentages presented by the model. In addition, the performance of the model is evaluated by detecting classes coming from another base, such as 'Planet: Understanding The Amazon from Space' by Goldenberg et al. (2017), used by Magiri (2023), with dimensions of 256x256.

Application of the CNN models

After completing the evaluations on the quality of the information obtained in the models, these are used in the interpretation of the land cover and land use. The models aims, for the classification model, to identify the specific class for the regions of pixels in the image, and to locate the classes present in the image, in the case of the detection model. For this purpose, ARB images captured by the Sentinel-2 satellite and obtained from Google Earth Engine, with tolerance settings of 20% for cloudy pixels, are used to prevent clouds obscuring crucial information in the image, acquired on 2022-06-09, with the wavelength of 492 nm for band B2, used for detecting water bodies, 559 nm for band B3, employed for vegetation differentiation analysis, and 665 nm for band B4, used for vegetation detection (Figure 7).

Figure 7 - Almada River Watershed (BA): Illustration of the Sentinel-2 image of bands B4, B3 and B2, 2024



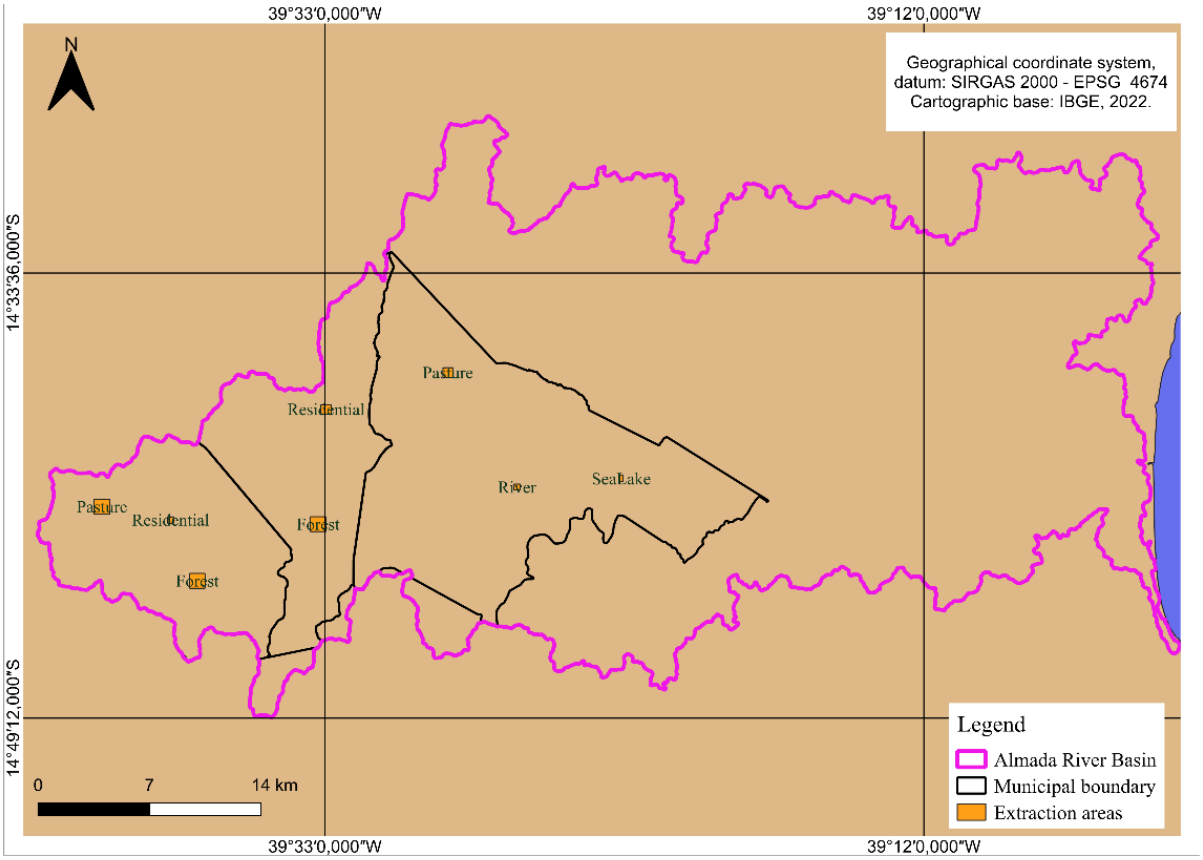
Source: Google Earth Engine, 2022. Elaboration: The authors, 2024.

Implementing the classification model involves the selection of images cut with the standard dimensions of the EuroSat data set (64x64), previously classified by the authors. This approach enables the evaluation of the effectiveness of the model. In addition, the choice of images follows established criteria, such as (a) the specific purpose of image recognition; (b) the presence of only one class for the classification of the image; and, (c) the level of image identification for the human eye.

In the context of detection, the images are used in the same format used in the training, where the criteria: (a) the specific purpose of the identification of the classes in the image; and (b) the presence of multiple classes in a single image are established. These criteria are carefully defined to highlight the advantages derived from the use of models for the analysis of land cover and land use.

By ensuring that images meet these criteria, it is also possible to verify the ability of the model to meet the given requirements. In this context, we chose to select images originating from the territory of the cities located within the perimeter of the basin, especially Almadina, Coaraci, and Itajuípe. These chosen areas are regions of recognition of the authors and contain representative characteristics, such as forest formations, pastures, and water bodies (Figure 8).

Figure 8 - Almada River Watershed (BA): Illustration of the locations of the selected areas based on the criteria.

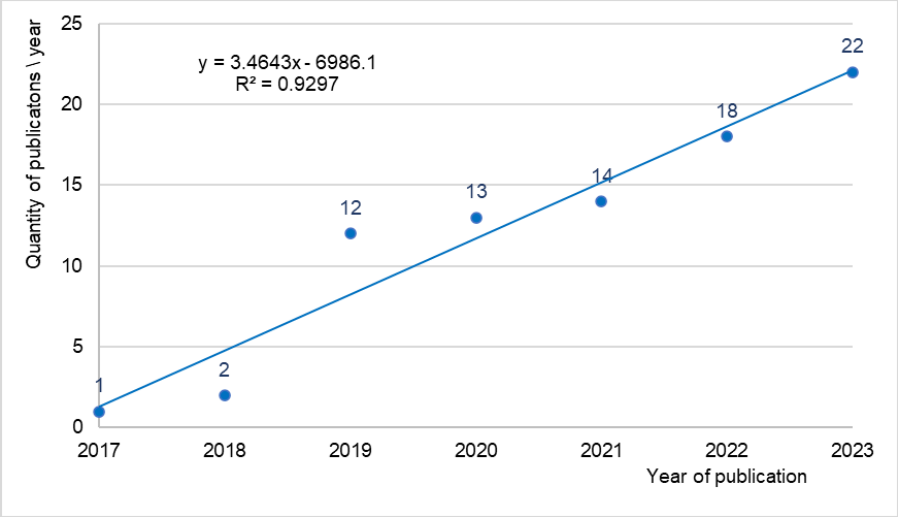


Source: IBGE, 2022. Elaboration: The authors, 2024.

RESULTS AND DISCUSSIONS

The results of the Web of Science indicate that the first bibliographic record of the relationship between coverage and land use with CNN dates back to 2017. Additionally, a growing trend of publications over subsequent years was identified. There was a significant 61% increase in the number of publications between the years 2020 and 2023 (Figure 9).

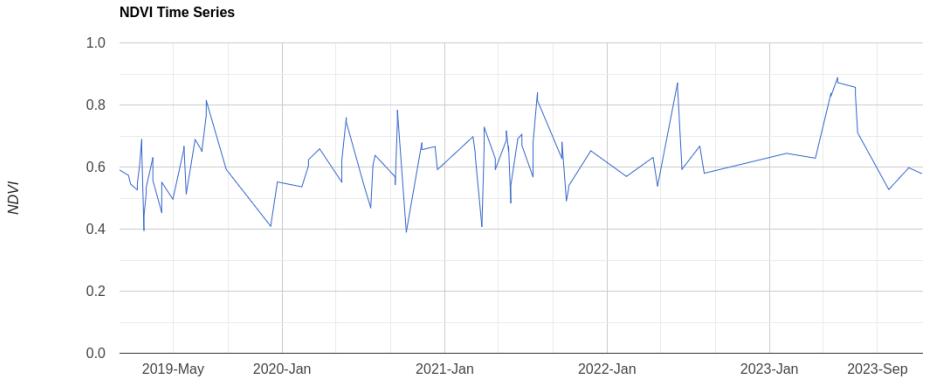
Figure 9 - Quantity of publications on “Land coverage and land use AND Convolutional Neural Network”, 2024



Source: The authors, 2024.

Figure 9 reveals that the use of CNN for the analysis of satellite images was an effective and expanding tool. This approach reveals substantial results in the classification of land cover and land use, which have strategic potential decision-making. Likewise, to improve the analyses of the basin characteristics, the NDVI identification provided a comprehensive evaluation of the variable conditions for the years 2019 to 2023, being in the range between 0.4 and 0.8, which is evaluated as robust indicative of the presence of vegetation and biodiversity of the Atlantic Forest biome (Figure 10).

Figure 10 - Almada River Watershed (BA): Illustration of the NDVI time series, 2024



Source: The authors, 2024.

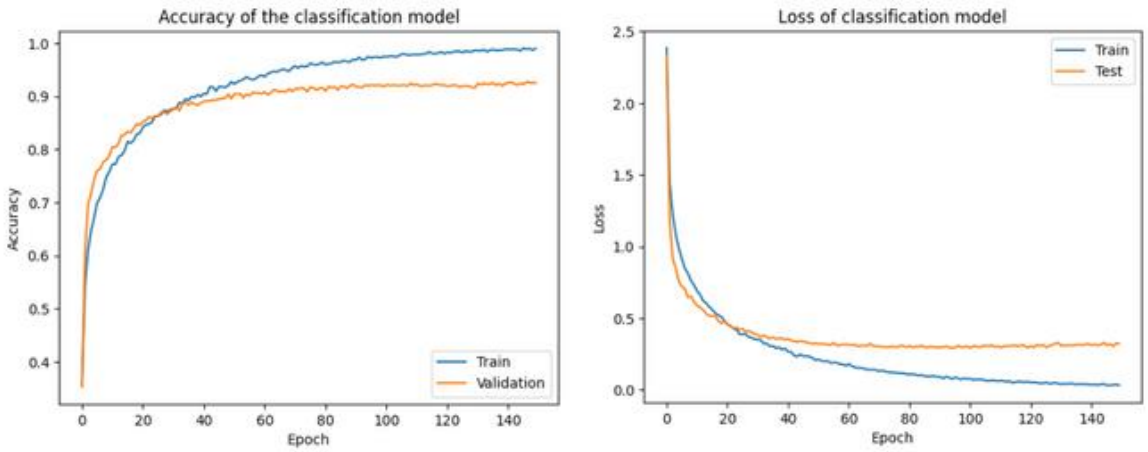
Additionally, regarding the training and application of the models, the critical points of computational consumption, for performance and execution time, were concentrated in the data processing and labeling phase. For this reason, the use of Google Colaboratory enabled the stages of processing and training of the models to occur with greater efficiency and shorter execution times.

It is important to note that before the application of the models, the accuracy of the classifications and detections of the characteristics are evaluated. Subsequently, the results obtained from the models, including the data sets processed for classification and the YOLO configuration files, are properly stored and are available on Zenodo’s free platform (CARVALHO; SILVA; SOUZA, 2024).

Evaluation of the image classification and detection model

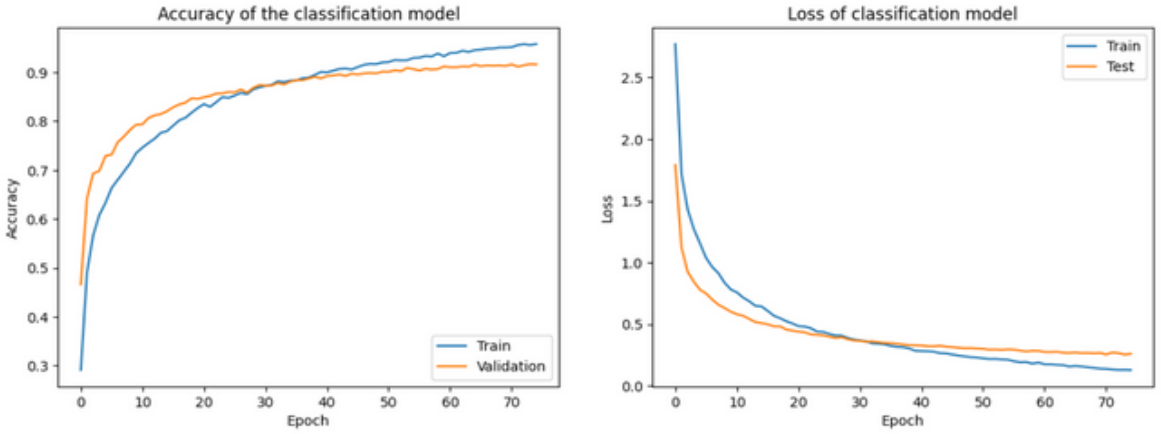
In Figure 11, as the model converges during training, the loss decreases, while the accuracy approaches the validation data, indicating an effective optimization. However, throughout the training process, the model tended to stabilize, maintaining a consistent performance, achieving an accuracy of 92%. Therefore, it was decided to end the training earlier, after 75 times, when analyzing the interval between 60 and 80 times, recognizing the stability of the model (Figure 12).

Figure 11 - Accuracy and loss chart for the classification model with 150 times, 2024



Source: The authors, 2024.

Figure 12 - Accuracy and loss graph for the classification model with 75 epochs, 2024



Source: The authors, 2024.

Moreover, the classification results showed that the model achieves a better performance after adjustments in the activation and loss functions. In addition, when using the EarlyStopping function, the model maintained stability, stopping after 68 times, close to the manually determined number. This indicates that the training occurred satisfactorily, showing no signs of overfitting or underfitting.

The model demonstrated proficiency by assigning specific categories with precise pixel regions in satellite imagery; this capability encompasses the ability to distinguish between classes of land cover and use. Evidenced in the precision and recall metrics, where they present the ability of the model to evaluate the class of a region and the probability of classification is correct (Table 1).

Table 1 - Results of the classification model evaluation metrics, 2024

Class	Precision	Recall	F1-Score
Residential	0.95	0.94	0.95
River	0.97	0.95	0.96
Industrial	0.91	0.97	0.94
Pasture	0.91	0.93	0.92
PermanentCrop	0.89	0.90	0.90
Highway	0.82	0.75	0.78
HerbaceousVegetation	0.91	0.88	0.89
SeaLake	0.99	0.99	0.99
AnnualCrop	0.91	0.92	0.91
Forest	0.94	0.98	0.96

Source: The authors, 2024.

Additionally, the standardization of the number of images for each class mitigated the risk of the model becoming biased toward a specific type of information. Identified, (Table 1) where the model accuracy remained predominantly above 90%, indicating a robust performance in the test dataset. In other studies, CNN models with accuracy above 90% are considered a powerful tool for the classification of supervised images, assisting in making geospatial decisions with a higher level of confidence (KAIDI et al., 2022; BRITZ; DUKER, 2023; WESSON; BRITZ; DUKER, 2023).

As for recall, most classes exhibited values above 90%, whereas only two classes were below this threshold. This performance highlights the model's ability to correctly identify all positive examples. The minimization of false negatives suggests that the model identifies most land cover and land use classes, with a probability above 90% of being correct.

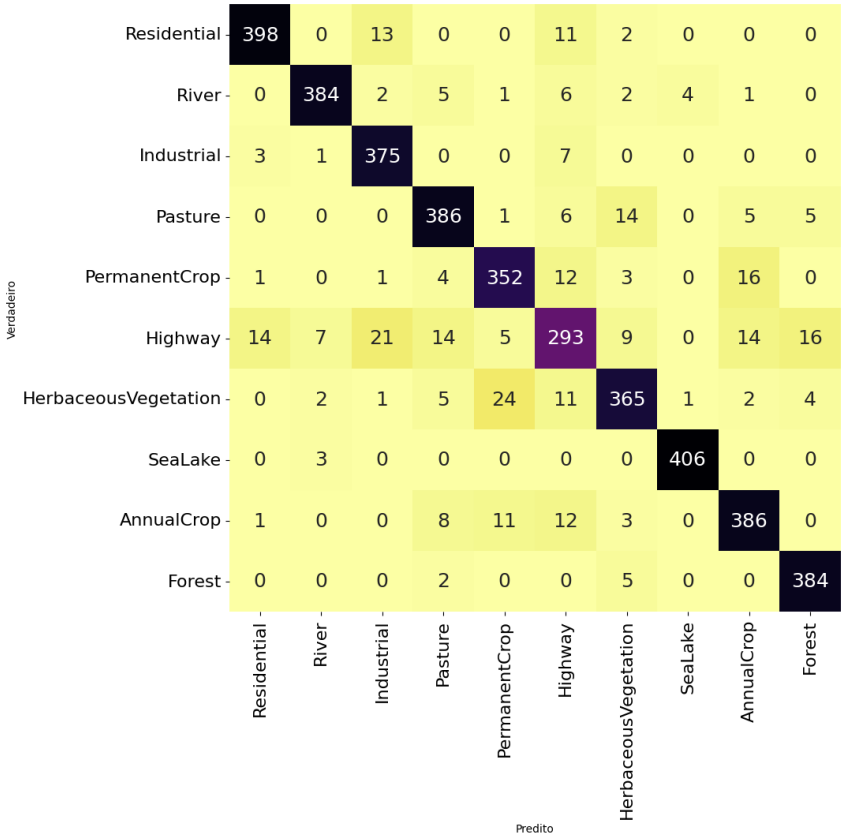
In addition, the F1-Score metric represents the balance between accuracy and recall as a single measure of performance, where it is revealed that only the class "Highway" and "Herbaceous Vegetation" were below 90%. Notably, these classes have color patterns and features close to the "Industrial" and "PermanentCrop" classes, respectively. In a deeper analysis, it is considered the possibility that these classes reflect information of multiple characteristics. This is because there are roads that may be connected to industrial areas, in the same way that permanent harvests may be associated with certain characteristics of herbaceous vegetation.

Finally, when exploring the confusion matrix, we obtained a detailed view of the model's predictions in relation to the different classes. In this matrix, true and false positives are clearly delineated, with correct representations occupying the main diagonal and mistaken classifications positioned outside this diagonal (Figure 13).

The relationships between classes, such as "Forest" and "SeaLake" stood out in the confusion matrix, which has distinct characteristics in terms of texture, shapes, and spatial contexts, exhibiting high precision. However, the class that shares close structures makes accurate label assignment more challenging for the classification model, increasing the number of false positives. This similarity of characteristics between classes can negatively impact a model's ability to distinguish them with high precision.

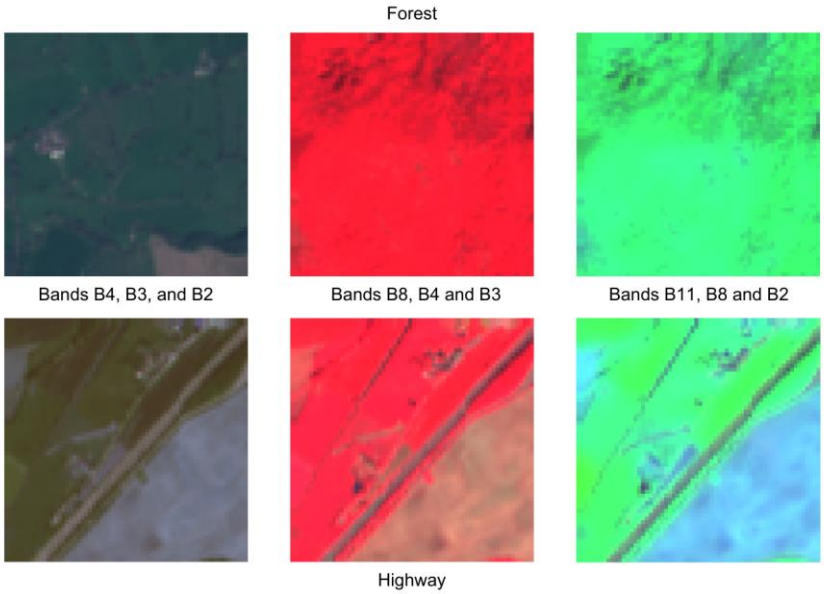
Therefore, an alternative approach is the enhancement of colors or distinct characteristics, based on the optical frequencies resulting from the configurations of the satellite bands. As a result, the images can highlight areas with vegetation, highlighting shades of green more precisely through the simulation of true colors, with the Sentinel-2 Bands B4, B3, and B2, or bands B8 with a wavelength of 833 nm, used for detecting changes in vegetation cover, and B11 with 1.610 nm, with applications for detecting materials on land surfaces. This strategy aims to improve the visualization of specific characteristics, providing a more detailed and interpretative analysis of satellite images (Figure 14).

Figure 13 - Illustration of the confusion matrix for the classification model, 2024



Source: The authors, 2024.

Figure 14 - Band configuration in Sentinel-2 images highlighting model training characteristics, 2024



Source: HELBER et al., 2018. Elaboration: The authors, 2024.

In addition, another point in the processing involves the submission of the images to atmospheric correction and, subsequently, the resampling of the bands, highlighting the importance of adjusting the resolutions

according to the specific objectives of the study. These band configurations can be applied to other satellites, such as Landsat-8 (REX et al., 2018).

Importantly, in sensitive applications such as road safety or industrial detection where reliability is essential, continuous evaluation of model performance and frequent adjustments are indispensable to ensure reliable and ethical results. Classification errors can have serious implications, justifying the need for cautious approaches in critical contexts. Therefore, the data processing should be guided by the objective of the work to be performed (REX et al., 2018; ZHAO et al., 2019).

Regarding the detection model, the results show a clear understanding of the performance concerning object detection, considering the trade-off between precision and recall. The specific details of the detection offer valuable insights into the model's behavior in terms of successes and errors about the objects present in the analyzed images. We highlight an accuracy rate above 70% for the "Forest" and "Pasture" classes, values close to those found by (KRYSIAK et al., 2020), which showed effectiveness in the automatic detection of land abandonment.

In a comprehensive analysis of the model's detections, 5.870 detections were identified, covering a total of 783 unique truths. The detections were divided into different classes, and the detection accuracy for each class was evaluated (Table 2).

Table 2 - Results of the evaluation of the detection model with YOLO, 2024

Class	Precision / %	True positives	False positives
Forest	72.09	142	8
Pasture	70.23	143	11
River	8.62	41	198

Source: The authors, 2024.

In an analysis of the performance of the model with a confidence threshold set at 0.25, the results indicated an accuracy of 60%, a recall of 40%, and an F1-score of 50%. These metrics are essential to assess the model's ability to make accurate and comprehensive detections.

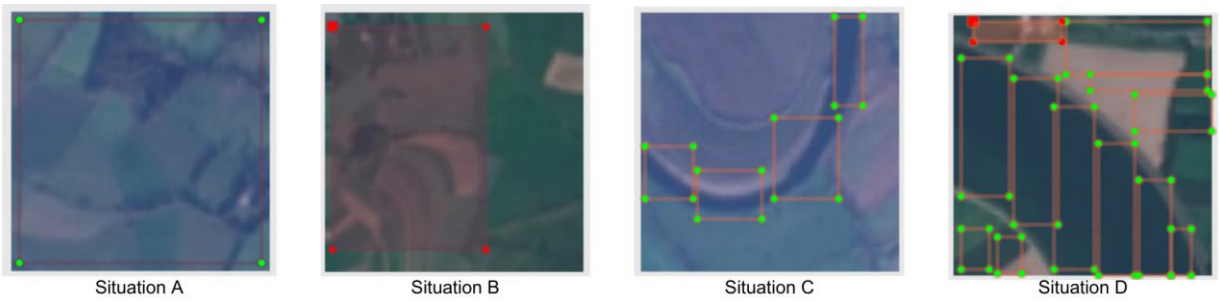
Additionally, evaluating the performance in terms of intersection over union, with a threshold of 50%, the model presents an average of 49.48%. This indicates the effectiveness of the model in terms of overlap between the predicted areas and the real areas. For the same threshold of 50%, the metric of average accuracy in all classes is calculated, resulting in a value of 50.32%. This metric provides an overview of the model's ability to make accurate detections in different classes.

However, the "River" class stands out with an accuracy of 8% and a considerable number of false positives. This result highlights the importance not only of the detailed construction of the bounding boxes but also understanding of the dynamics of land cover and use, accompanied by their distinct characteristics.

Unlike the "Forest" classes, which exhibit more static patterns without relevant changes from one portion to another, the "River" class presents characteristic sinuosities, which can be attributed to processes such as erosion, sedimentation, and human occupation, such as changes in the course of rivers due to pasture areas and mosaic use.

In the context of bounding box construction, it is possible to highlight four distinct scenarios. In situation "A", the presence of only one class was identified, delimited in the total extension of the image. In contrast, in situation "B", we identified two classes, as well as in situation "D", which comprises the three classes of the training set. In these cases, it becomes necessary to create a bounding box for each type of class to mitigate the risk of false positives. Finally, situation "C" presents a limitation because the bounding boxes present difficulty following sinuous scenarios, resulting in distinct cutouts to compose the entire area, Figure 15.

Figure 15 - Scenarios identified during the construction of the bounding boxes, 2024



Source: GOLDENBERG et al., 2017. Elaboration: The authors, 2024.

Finally, in the validation of the model's performance on external data, using the dataset 'Planet Understanding The Amazon from Space', results similar to those obtained in EuroSat are found. In particular, the "Forest" class stood out, with a confidence average above 90%, while "Pasture" and "River" registered 50% and 35%, respectively. Figure 16 illustrates the bounding boxes, highlighting the confidence associated with each detection, as well as the classifications. It is relevant to note the ability of the model to identify more than one class in a single image, providing valuable data that can enrich decision-making strategies.

Figure 16 - Detection of characteristics for each class with bounding boxes, 2024

Source: GOLDENBERG et al., 2017. Elaboration: The authors, 2024.

It should be noted that this analysis opens possibilities for the recognition of the distinct species that make up the agroforestry system, especially the cabruca system. This, however, requires the construction of a robust database with information that helps the model in identification, allowing the accounting and determination of forest conditions more accurately and comprehensively.

Application of the classification and detection model in ARB

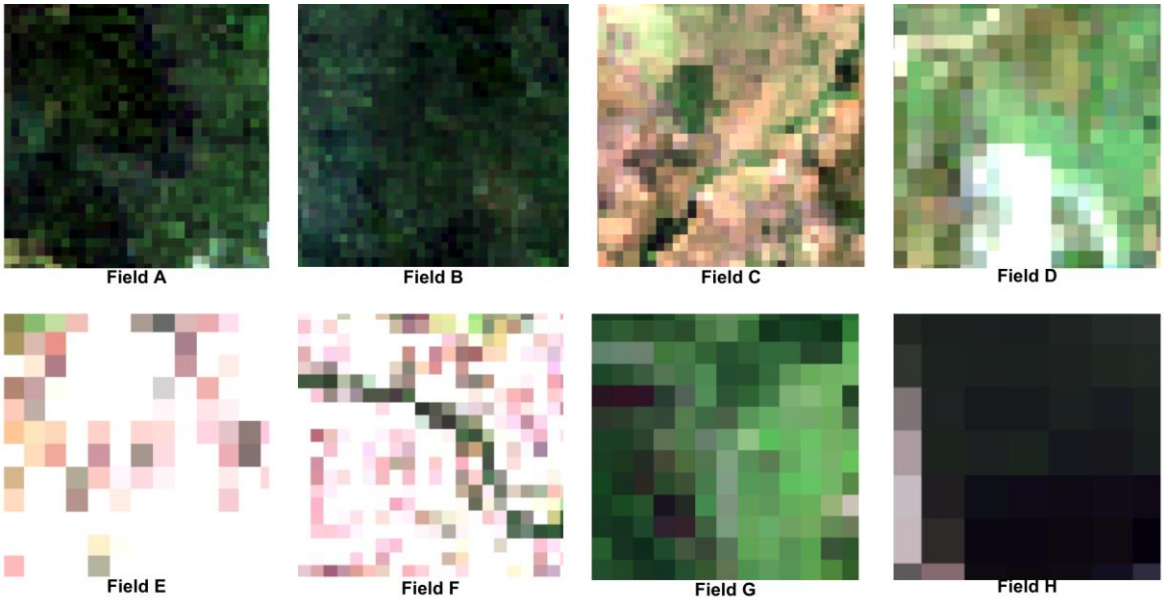
The discussion is enriched by the practical applicability of these models in diverse contexts such as environmental monitoring, urban planning, locating deforestation areas, and responding to environmental disasters.

In this context, when applying the model, it is crucial to consider its versatility, which allows the analysis of combinations between optical bands. This characteristic expands the possibilities of identification, especially in obtaining data for mapping land cover and use.

This mapping provides valuable subsidies to understand how the study area has been historically used and occupied, evidencing the distribution of vegetation cover and anthropogenic where it is emphasized that zoning techniques by use and occupation classes are essential for environmental analysis and territorial planning (VIEGA; SILVA, 2018; LISBOA et al., 2023; TOLENTINO et al., 2023).

Thus, in the identification of the land cover and land use classes, the model's practical ability to classify the selected areas is verified, which are defined and labeled based on the criteria established in the methodology, where eight areas are chosen (Figure 17). From these images, the classification algorithm produced the results compiled in Table 3.

Figure 17 - Almada River Watershed (BA): Section of study areas selected for classification using the CNN model, 2024



Source: Google Earth Engine, 2022. Elaboration: The authors, 2024.

Table 3 - CNN model classification result for the selected areas, 2024

Field	Classified by the author	Classification by the model	Probability / %
A	Forest	Forest	95.90
B	Forest	Forest	97.97
C	Pasture	Pasture	95.36
D	Pasture	Pasture	96.84
E	Residential	Residential	93.96
F	Residential	Residential	94.24
G	River	River	95.96
H	SeaLake	SeaLake	98.98

Source: The authors, 2024.

In Table 3, the model's efficacy in accurately classifying satellite images was identified, maintaining an accuracy exceeding 90% in practical analyses. It is noteworthy that the ARB encompasses various forested and vegetated regions, as evidenced by NDVI analysis. These areas were identified by the model with an accuracy above 95%, suggesting its capability to produce reliable results across extensive areas of the watershed.

In addition, one of the criteria for sample selection is the difficulty of human eye recognition of land cover and land use, particularly evident in fields "E" and "F." The samples are cutouts measuring 64x64 pixels, which may result in a loss of visual quality; nevertheless, they preserve the essential features of relief, texture, and spatial context. These attributes are crucial for enabling the algorithm to accurately distinguish between different land cover and use classes. Finally, it is highlighted that the "SeaLake" class exhibited an accuracy close to that found in the model validation, reinforcing reliability and practical utility in the classification.

Methods like these are essential for assisting in watershed monitoring. This is due to the difficulty of staying up to date with the complexity of monitoring and evaluating the dynamics of urban construction, such as the Porto Sul Intermodal Logistics Complex. This project has impacts on the three pillars of sustainability (economic, social, and environmental) and is located near the mouth of the ARB. Furthermore, the ARB is one of the main natural systems in the cocoa region, encompassing areas of the Atlantic Forest biome, including secondary forests, sandbanks, and mangroves (GOMES et al., 2012). In this context, it is crucial to highlight that these elements are not labeled in EuroSat, hindering the classification of these areas.

Given this scenario, it becomes relevant to construct a custom dataset containing essential information, such as fragility points, exposed sandy sediments, and degraded pockets in the plains, as well as data for recognizing the cocoa region. This aims to identify the phenomena in the ARB area and support decision-making.

Thus, an evident solution is the combination of CNN and machine learning algorithms, aligning them with the features of relief, texture, shape, and spatial context in satellite images. This aims to refine information, providing more concise and accurate data. Such information can assist in identifying potential deforestation or environmental risk areas at the ARB boundary.

In the context of land cover and land use class detection, exporting converted satellite images to the "jpg" format, compatible with model training data, results in a reduction in detection quality, leading to a decrease in confidence levels. This phenomenon may be attributed to the loss of essential features of relief and texture, which can make the interpretation by the model challenging.

However, the algorithm still manages to distinguish various land cover and land use classes for which it was trained. This is especially notable when identifying features in the study area, including forest formations and water bodies (Figure 18).

Figure 18 - Municipality of Itajuípe (BA): the result of applying the model to a Sentinel-2 image, 2024

Source: Google Earth Engine, 2022. Elaboration: The authors, 2024.

Figure 18 follows the established methodology criteria, presenting multiple classes of land cover and land use. The detection of multiple classes allows for the identification and interpretation of types of land cover and land use, aiding in the analysis of their surroundings. Additionally, the specific aim of detection is highlighted, which enabled the verifying of the model's capability in bounding box building to highlight areas of riverbeds near civil constructions, which is common in the ARB, where river sections surround urban areas.

As established by Law nº 12,651/2012, in Article 4, paragraph I, the marginal strips of any perennial and intermittent natural watercourse must be preserved from the edge of the regular bed channel, with a minimum width of 30 m, especially in urban areas. Additionally, the law also mandates the obligation of rural property registration. Therefore, the detection model stands out as an alternative to locating points of interest according to the criteria established by law.

This capability is evidenced in the delineation of the boxes, which can be made using Python algorithms to determine, for example, the distance between the riverbed and nearby civil constructions. It is worth noting the possibility of this tool being framed as a monitoring mechanism for identifying flood-prone areas, which are common in urban areas within the watershed, similar to what occurs in the Hydrographic Basin of the Cachoeira River (TOLENTINO et al., 2023).

Additionally, this class detection process is exemplified in an image from Google Earth Pro with coordinates: 14°40'44.4"S, 39°28'33.8"W. In this representation, the bounding boxes encompass part of the river stretch with nearby civil constructions, Figure 19.

Figure 19 - Municipality of Itajuípe (BA): the result of applying the model to a Google Earth Pro, 2024

Source: Google Earth Pro, 2024. Elaboration: The authors, 2024.

Furthermore, the use of images obtained from Google Earth Pro is highlighted, where, in all tests conducted, the model demonstrated consistency by presenting confidence levels above 50% for each class detected. This result emphasizes the relevance of using higher-resolution images. Alternatively, the use of satellites with higher resolutions, such as CBERS-4A, which offers a 5-meter resolution, is recommended to enhance both the quality and accuracy of detections.

Another strategy is to increase the number of samples for model training, which requires time and attention. However, since EuroSat is a dataset with its own labeling standardization, automating the generation of bounding boxes for the same index becomes feasible. This standardization can cover the entire image extent through the use of algorithms that generate files with the same coordinates for each point. Although this approach disregards situations where there is more than one class in the same image, it still contributes to expediting the construction of bounding boxes and increasing the number of samples available for model training (Figure 20).

Figure 20 - Illustration of the organization of bounding box files, 2024

Class index	0
Coordinate of point X	0.507812
Coordinate of point Y	0.507812
Bounding box width	0.984375
Bounding box height	0.984375

Source: GOLDENBERG et al., 2017. Elaboration: The authors, 2024.

Finally, it is crucial to consider the versatility and capability of CNN models to adapt to the goals of the work, provided they are trained with a dataset containing specific characteristics for the desired analyses. However, it is relevant to highlight the limitations of CNN models regarding the dimensions and format of the data used

in training. The 64x64 dimension presents a restriction, especially when compared to the usual size of satellite images or study areas.

A solution to overcome this limitation was proposed by Silva (2022), involving the partitioning of the original image into several mosaics with dimensions matching the training data. Subsequently, the reconstruction of the original dimension is done by concatenating the binary prediction images corresponding to the various mosaics, ensuring the best possible result.

FINAL CONSIDERATIONS

The CNN models proved to be effective for the classification and detection of land cover and land use classes in ARB satellite images. Analyzing the classification results, the efficiency of recognizing land cover and land use classes is evident, achieving precision above 90%. As an alternative to improving class classifications, specific combinations of satellite bands can emphasize characteristics that provide a more detailed and interpretative analysis of images.

For the detection model, not only the precise construction of bounding boxes is emphasized but also the need to understand the complex dynamics of land cover and land use, considering their distinct characteristics. Furthermore, the detection model demonstrated remarkable effectiveness in identifying multiple classes, allowing the interpretation of various types of land cover and land use and aiding in the analysis of their surroundings. This highlights its ability to overcome challenges and maintain consistent performance, with corresponding confidence values.

In the context of the model application, the versatility of CNN models along with geoprocessing techniques stands out, enabling the use of Python algorithms for unsupervised tasks, capable of generating information based on criteria established by law, such as defining the distance between riverbeds and civil constructions, as well as detecting rural properties. For this purpose, the use of higher-resolution images or the processing of images with atmospheric correction and band resampling is recommended, giving preference to band combinations that enhance features and facilitate model training.

Additionally, the possibility of using them as a tool for monitoring watershed conditions is identified, serving as a mechanism for the identification and location of areas of interest, such as those prone to flooding or deforestation, common in the ARB area. Moreover, the critical point is emphasized regarding the need for a dataset with the relief, texture, and spatial context characteristics of specific areas in the watershed, such as sandbanks and mangroves, in addition to points that aid in detecting vulnerabilities. Thus, besides the construction of the custom dataset, a strategy adopted is to prioritize factors that allow analyzing the environmental dynamics of the ARB, contributing to robust assessments and guiding assertive decisions.

Therefore, the results of CNN models, both for classification and detection, are expected to open new research possibilities, focusing on the integration of geoprocessing and CNN for satellite image analysis to improve decision-making with greater precision. As a perspective for future work, the investigation of image segmentation techniques is highlighted, aiming to identify or isolate specific regions of interest using CNN for the creation of thematic maps of the ARB, highlighting changes in land cover and land use.

ACKNOWLEDGEMENTS

To the National Council for Scientific and Technological Development (CNPQ) for funding through the Universal Project, Process nº 409304/2021-2. To the Federal University of Southern Bahia for providing a CNPQ Scientific Initiation scholarship to one of the authors. To the Coordination for the Improvement of Higher Education Personnel (CAPES) for granting a graduate scholarship to one of the authors.

REFERENCES

BRASIL. **Lei no 12.651, de 25 de maio de 2012.** Dispõe sobre a proteção da vegetação nativa; altera as Leis no 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as Leis no 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória no 2.166-67, de 24 de agosto de 2001; e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/L12651compilado.htm. Acesso em: 29 de dezembro de 2023.

CARVALHO, H. S.; SILVA, V. A.; SOUZA, P. S. V. N. Data from convolutionary neural network models for land coverage and use analysis in the Almada River Basin (Bahia - Brazil) [Data set]. 2024. Zenodo. https://doi.org/10.5281/zenodo.10650112.

EUROSAT. TENSORFLOW DATASETS. 2022. Disponível em: https://www.tensorflow.org/datasets/catalog/eurosat?hl=pt-br. Acesso em: 11 fev. 2024.

FRANCO, G. B.; MARQUES, E. A. G.; GOMES, R. L.; CHAGAS, C. DA S.; SOUZA, C. M. P. DE; BETIM, L. S. Fragilidade Ambiental da Bacia Hidrográfica do Rio Almanda – BAHIA. **Revista de Geografia**, v. 28. 2011.

FREIRE, A.F.M., 2006. A Sequência Holocênica na Plataforma Continental Central do Estado da Bahia – Costa do Cacau. 172f. **Dissertação** (Mestrado em Geologia). Universidade Federal da Bahia. 2006.

GOLDENBERG, B.; UZKENT, B.; CLOUGH, C.; FUNKE, D.; DESAI, D.; JESUSMARTINEZMANSO, G.; SCOTT, K.; RISDAL, M.; PETE, M. R.; HOLM, R.; NAIR, R.; HERRON, S.; STAFFORD, T.; KAN, W. Planet: Understanding the Amazon from Space. Kaggle. 2017. Disponível em:https://kaggle.com/competitions/planet-understanding-the-amazon-from-space.

GOMES, R. L.; MORAES, M. E. B. de; MOREAU, A. M. D. S.; MOREAU, M. S.; FRANCO, G. B.; MARQUES, E. A. G. Aspectos físico-ambientais e de uso e ocupação do solo da bacia hidrográfica do rio Almada-BA. **Boletim de Geografia**, v. 30, n. 2, 2012. https://doi.org/10.4025/bolgeogr.v30i2.16423

GOOGLE COLABORATORY. Google Colaboratory. Disponível em: https://colab.research.google.com

GOOGLE EARTH ENGINE. Google Earth Engine. Disponível em: https://earthengine.google.com

GOOGLE LLC. Google Earth Pro (versão 7.3.6.9345) [software]. Disponível em: https://www.google.com/earth/versions

HEGAZY, I. R.; KALOOP, M. R. Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt. International Journal of Sustainable Built Environment. 2015. https://doi.org/10.1016/j.ijsbe.2015.02.005

HELBER, P.; BISCHKE, B.; DENGEL, A.; BORTH, D. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. 2018. [Data set]. In EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification, v. 12, n. 7, p. 2217–2226. Zenodo. https://doi.org/10.1109/JSTARS.2019.2918242

KAIDI, P.; HUAN, X.; QI, X.; PEIQI, H.; ZIYI, L. A Physics-Assisted Convolutional Neural Network for Bathymetric Mapping Using ICESat-2 and Sentinel-2 Data, **IEEE Journals & Magazine**, v. 60, p. 1-13, 2022. Disponível em: https://ieeexplore-ieee-org.ez10.periodicos.capes.gov.br/document/9915464. Acesso em: 13 jan. 2024.

KRYSIAK, S.; PAPIŃSKA, E.; MAJCHROWSKA, A.; ADAMAIAK, M.; KOZIARKIEWICZ, M. Detecting Land Abandonment in Łódź Voivodeship Using Convolutional Neural Networks. **Land**, v. 9, n. 3, p. 82, 2020. https://doi.org/10.3390/land9030082

LAVENÈRE-WANDERLEY, A. A. O. Mobilidade Sedimentar na Plataforma Leste Brasileira entre o Rio de Contas (Ba) eo Rio Doce (ES): Controle Morfológico e do Clima de Ondas. 2018. **Tese** (Doutorado em Oceanografia Geológica) – Instituto Oceanográfico, University of São Paulo, 2018.

LISBOA, G. P.; MOREAU, M. S.; GOMES, R. L.; LISBOA, G. dos S.; SILVA, V. de A.; MAGALHÃES, R. da S.; FRANÇA, L. C. de J.; SANTOS, A. M. dos; HADDAD, H. C. Environmental fragility of land systems in a hydrographic basin located in the South region of the state of Bahia, Brazil. Caminhos de Geografia, v. 24, n. 91, p. 189–207, 2023. https://doi.org/10.14393/RCG249162194

MAGIRI, D. S. Classificação de imagens de satélite com redes neurais convolucionais. 2023. Disponível em: https://repositorio.ufscar.br/handle/ufscar/17726. Acesso em: 01 de dezembro de 2023.

NASCIMENTO, L., BITTENCOURT, A.C.S.P., SANTOS, A., DOMINGUEZ, J.M.L. Deriva Litorânea ao Longo da Costa do Cacau, Bahia: Repercussões na Geomorfologia Costeira. **Revista Pesquisas em Geociências**, v. 34, n. 2, p. 45–56, 2007. https://doi.org/10.22456/1807-9806.19471

PROJETO MAPBIOMAS – Coleção 8 da Série Anual de Mapas de Cobertura e Uso da Terra do Brasil.

QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project. Disponível em: https://www.qgis.org/

REX, F. E.; KAFER, P. S.; DEBASTIANI, A. B.; KAZAMA, V. S. POTENCIAL DE IMAGENS MSI (SENTINEL-2) PARA CLASSIFICAÇÃO DO USO E COBERTURA DA TERRA. ENCICLOPÉDIA BIOSFERA, **Centro Científico Conhecer**, v. 15 n. 27, p. 219, 2018. https://doi.org/10.18677/EnciBio_2018A6

SEI, Superintendência de Estudos Econômicos e Sociais da Bahia. Balanço hídrico do Estado da Bahia, p. 250. 1999.

SENA, S. R. de; VEIGA, R. de S.; SILVA, V. D. A. Análise de Áreas de Risco à Degradação Ambiental no Município de Porto Seguro, Bahia, Brasil, **Revista Brasileira de Geografia Física**, v. 16, n. 6, p. 3059–3072, 2023. https://doi.org/10.26848/rbgf.v16.6.p3059-3072

SILVA, C. C. L. Definição das Áreas de Preservação Permanente (APPs) de cursos de rio e análise de necessidade de restauração florestal no município de Boituva, SP. 2023. Disponível em: https://repositorio.unesp.br/items/51b5f984-54fc-4d88-8b43-bf3c7320e0db. Acesso em: 21 de dezembro de 2023.

SILVA, H. C. V. Redes neurais de convolução na classificação de edifícios em imagens de alta resolução espacial. 2022. Disponível em: https://repositorio.ul.pt/handle/10451/55447. Acesso em: 02 fev. 2024.

TAN, Z.; YUE, P.; DI, L.; TANG, J. Deriving High Spatiotemporal Remote Sensing Images Using Deep Convolutional Network, **Remote Sensing**, v. 10, n. 7, p. 1066, 2018. https://doi.org/10.26848/rbgf.v16.6.p3059-3072

TOLENTINO, G. M.; SILVA, V. de A.; LISBOA, G. dos S.; SANTOS, M. S. T.; SENA, S. R. de. Áreas de risco a alagamento e inundação na foz do rio Cachoeira, em Ilhéus (Bahia – Brasil): Flooding and inundation risk areas at the mouth of the Cachoeira River in Ilhéus (Bahia – Brazil). **Revista de Geociências do Nordeste**, v. 9, n. 2, p. 83–95, 2023. https://doi.org/10.21680/2447-3359.2023v9n2ID32143

TZUTALIN. LabelImg. GitHub repository, versão 1.8.6, 2020. Disponível em: https://github.com/tzutalin/labelImg

VEIGA, R. de S.; SILVA, V. de A. USO, COBERTURA E OCUPAÇÃO DA TERRA NO MUNICÍPIO DE PORTO SEGURO, BA: UMA ANÁLISE ESPAÇO TEMPORAL (1985-2016). **Caminhos de Geografia**, v. 19, n. 65, p. 232–244, 2018. https://doi.org/10.14393/RCG196517

WESSON, C.; BRITZ; W.; DUKER, R. Investigating the efficiency and capabilities of UAVs and Convolutional Neural Networks in the field of remote sensing as a land classification tool. South African **Journal of Geomatics**, v. 12, n. 2, p. 172-189, 2023. https://doi.org/10.4314/sajg.v12i.2.5

ZHAO, S.; LIU, X.; DING, C.; LIU, S.; WU, C.; WU, L. Mapping Rice Paddies in Complex Landscapes with Convolutional Neural Networks and Phenological Metrics. **GIScience & Remote Sensing**, v. 57, n. 1, p. 37–48, 2019. https://doi.org/10.1080/15481603.2019.1658960

Recebido em: 28/03/2024
Aceito para publicação em: 24/06/2024