Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul
Programa de Pós-Graduação em Recursos Naturais

LUCIENE SALES DAGHER ARCE

IDENTIFICATION OF A FOREST SPECIES OF SOCIO-ENVIRONMENTAL INTEREST BASED ON MACHINE LEARNING

Campo Grande, MS
Julho/2021

Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul
Programa de Pós-Graduação em Recursos Naturais

LUCIENE SALES DAGHER ARCE

IDENTIFICATION OF A FOREST SPECIES OF SOCIO-ENVIRONMENTAL INTEREST BASED ON MACHINE LEARNING

Dissertação apresentada à Secretaria do Curso de Pós-Graduação em Recursos Naturais como pré-requisito para obtenção do título de mestre em Recursos Naturais.

Orientadora: Prof. Drª. Camila Aoki
Co-orientador: Prof. Dr. José Marcato Junior

Campo Grande, MS
Julho/2021

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus, único e soberano.

Aos meus orientadores Camila Aoki e José Marcato Junior pelas valiosas contribuições dadas durante todo o processo, se dedicando, mesmo distante, durante esse difícil período que enfrentamos.

Aos colegas do curso e do Laboratório de Geomática que compartilharam dos inúmeros desafios que enfrentamos, sempre com o espírito colaborativo.

Também quero agradecer à Universidade Federal de Mato Grosso do Sul e o seu corpo docente que demonstrou estar comprometido com a qualidade e excelência do ensino.

Por último, mas não menos importante à minha família e amigos pelo apoio que sempre me deram durante toda essa jornada.

SUMMARY

LIST OF FIGURES

LIST OF TABLES

RESUMO

Um novo método de aprendizagem profunda capaz de mapear espécies única de palmeira *Mauricia flexuosa*, conhecida como Buriti em imagem aérea RGB foi proposto. Buriti é uma palmeira importante para comunidades e para fauna, além de ser um indicador de áreas de veredas, sendo importante seu mapeamento. A primeira sessão dessa dissertação apresenta um breve relato sobre a flora mundial e a legislação pertinente a espécie estudada. A segunda sessão apresenta um novo método baseado em Rede Neural Convulacional (CNN) que possibilita identificar e geolocalizar a espécie proposta além de comparar o desempenho com outras redes de detecção de objetos. Foram rotuladas manualmente um total de 5.334 palmeiras em um conjunto de 1.394 recortes de ortoimagens, retornando um erro absoluto médio (MAE) de 0,75 árvores e uma medida F de 86,9%. Os resultados apresentaram melhores que os métodos Faster-RCNN e RetinaNet. Conclui-se que o método proposto é eficiente para lidar em ambiente de alta densidade e complexidade florestal, podendo mapear e localizar as espécies de *Mauricia flexuosa* com alta acurácia.

Palavras-chave: *Mauricia flexuosa*, detecção de objeto; rede neural convulacional; aprendizado de máquina, monitoramento ambiental.

ABSTRACT

A new deep learning method capable of mapping unique Mauricia flexuosa palm species, known as Buriti in aerial RGB imagery was proposed. Buriti is an important palm tree for communities and fauna, in addition to being an indicator of path areas, and its mapping is important. The first session of this dissertation presents a brief report on the world flora and legislation pertaining to the studied species. A second session presents a new method based on the Convulational Neural Network (CNN) that makes it possible to identify and geolocate a species, in addition to comparing its performance with other object detection networks. They were manually labeled a total of 5.334 palm trees in a set of 1,394 ortoimage oatches, returning a mean absolute error (MAE) of 0.75 trees and F score of 86.9%. The results are better than Faster-RCNN and RetinaNet methods. Concluding that the proposed method is efficient to deal with high density and forest complexity environment, being able to map and locate as species of *Mauricia flexuosa* with hight accuracy.

Keywords: *Mauricia flexuosa;* object detection; convolutional neural network; deep learning; environmental monitoring.

1 INTRODUCTION

The rational use of natural resources is essential for nature conservation. Its sustainable use, the conservation of ecosystems, and the regeneration of degraded environments are important for life on Earth. Forest conservation and regeneration are promising strategies for improving water, energy, and food security (MELO, 2020). Forests are important as they provide food and habitat sources for fauna (LORENZI, 2002).

In the world, the total forest area has about 4.06 billion hectares, which is 31% of the total land area (FAO, 2020). Despite intensive studies and laws to reduce environmental degradation, around 12.2 million hectares of tree cover in the tropics were lost in 2020, which is 12% higher than the previous year. From this total, 4.2 million hectares occurred in areas of tropical rainforests, which are important for carbon storage and biodiversity (WORLD RESOURCES INSTITUTE, 2021).

Based on data from 2001 to 2020 available on the Global Forest Watch map (2021), Brazil lost 26.7 million hectares of humid primary forest, representing 45% of its total loss of tree cover (non-tropical and dry primary forests, secondary forests, and tree plantations, and rainforests) in the same period. The total area of primary rainforest in Brazil decreased by 7.7% in this period.

Although the changes in the legal frameworks governing Brazilian forests (Forest Code) are well documented, it has not yet been possible to reverse the short-term decline in deforestation (BRASIL, 2015).

With a wide variety of natural characteristics, Brazil has ecosystems with very distinct characteristics in terms of vegetation and fauna (MMA, 2020), called biomes: Amazon, Caatinga, Cerrado, Atlantic Forest, Pampa, and Pantanal. Amazon, Cerrado, Atlantic Forest, and Pantanal are present in the Brazilian Midwest. Cerrado and Atlantic Forest biomes are considered global *hotspots* in biodiversity (MYERS *et al*., 2000).

In Mato Grosso do Sul, there are three biomes: the Atlantic Forest, the Cerrado, and the Pantanal. The Cerrado, which comprises most of the state, is recognized as the richest and most diverse tropical savanna globally, home of around 11,627 cataloged plant species (MMA, 2019).

The Cerrado, in recent decades, has been degraded by the expansion of the agricultural frontier, with growing pressure to increase the production of meat and grains, with progressive depletion of the natural resources (INPE, 2021). Furthermore, there is large exploitation of woody material to produce charcoal in the Cerrado biome (MMA, 2019), thus reducing species of socio-environmental interest.

The local population's sustainable use of native flora and the valorization of natural resources are alternatives for preserving over 200 species that have some economic value, whether food, wood, or medicinal (EMBRAPA, 2004 BORTOLOTTO, *et al*. 2018). In addition to its environmental importance, the Cerrado provides a wide variety of fruits containing considerable amounts of protein, fiber, phenolic compounds, vitamins, calcium, phosphorus, and fatty acids (VIEIRA, et al., 2011); contributing to the sustainable development and quality of life of the population (AVIDOS and FERREIRA, 2000).

The fruits with the greatest potential for exploitation that are consumed by the local population (indigenous, riverside, settlers) and that are sold in urban centers are: Marolo (*Annona crassifolia*), Cagaita (*Eugenia dysenterica*), Cajuzinho do cerrado (*Anacardium humile*), Bacupari (*Salacia crassifolia*), seeds of Barú (Dipteryx alata), Buriti (*Mauritia flexuosa*), Jatobás (*Hymenaea courbaril, Hymenaea stigonocarpa*), Mangaba (*Hancornia speciosa*) and Pequi (*Caryocar brasiliense*) and others (BRASIL, 2015).

Mauritia flexuosa, known as Buriti, grows spontaneously in floodplains, veredas, gallery forests saturated with water near-permanent watercourses, igapós, iguarapés and springs, forming dense populations. It is a palm that defines the environments of veredas, both for its ecological importance and for its high density compared to the few tree species that occur there (RESENDE, 2012).

According to the Brazilian Forest Code, the palm tree is one of the indicators of the vereda area. In article 3, item XII, the vereda is classified as savanna phytophysiognomy, found in hydromorphic soils, usually with the emerging arboreal palm *M. flexuosa*, without forming a canopy, amid clusters of

herbaceous and shrub species, protected by permanent preservation (BRASIL, 2012).

Also, in the Forest Code, article 4, item XI, the delimitation of the permanent preservation area in a 50-meter strip from the swampy and waterlogged area (BRASIL, 2012) is fundamental for the preservation of this phytophysiognomy, the conservation of the local fauna and flora and the existing water resource.

The palm tree serves as a source of food and a place of shelter and reproduction for various elements of fauna (RESENDE, 2012). Besides, it has been consumed in cuisine, and Buriti nuts can also be used to produce beauty products and activated charcoal to remove heavy metals in a viable process to control water pollution (PINTO, 2013).

With the exploitation of forest species of great socio-environmental interest, multidisciplinary studies are essential, which is important for the management and conservation of vulnerable ecosystems. The use of technology, such as aerial vehicles, has shown great potential for classifying and mapping rural and urban areas. They can be used to identify forest species when equipped with sensors, such as RGB, multispectral, hyperspectral, LiDAR (Light Detection And Ranging), and radar (MORALES *et al*., 2018). Remote sensing enables the acquisition of images from areas that are often inaccessible, and when combined with machine learning, can contribute to the knowledge of the local flora (BREIMAN, 2001).

Our objective is to investigate machine learning methods to detect trees of environmental and social interest through RGB aerial images to identify and geolocate a species in a complex forest environment. Specifically, we focused on Buriti species in the urban area of Campo Grande, Mato Grosso do Sul, Brazil. We employed supervised learning methods, and for this, we generated a large dataset, with more than 5000 labeled trees.

1.2 Organization of the dissertation

The dissertation is organized into two sections. Section 1 presents a general context about the importance of conservation and laws. We also present a general discussion on the identification/mapping of tree species based on remote sensing and machine learning. Section 2 presents an article where we automatically mapped multipurpose palm trees on aerial images using deep learning-based approaches.

REFERENCES

AVIDOS, M. F. D.; FERREIRA, L. T. Frutos dos cerrados: Preservação gera muitos frutos. **Biotecnologia Ciência & Desenvolvimento**, 2000, *3*(15): 36-41. Disponível em: <http://www.almanaquedocampo.com.br/imagens/files/frutos%20do%20Cerrado.pdf>. Acessado em: 10 de janeiro de 2021.

BORTOLOTTO, I. M., DAMASCENO-JUNIOR, G. A. POTT, A., 2018. Lista preliminar das plantas alimentícias nativas de Mato Grosso do Sul, Brasil. **Iheringia, Série Botânica**, 73, 101-116. https://doi.org/10.21826/2446-8231201873s101.

BRASIL. Lei 12.651, de 12 de maio de 2012. **Dispõe sobre a proteção da vegetação nativa;** altera as Leis n. 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 n. 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória n. 2.166-67, de 24 de agosto de 2001; e dá outras providências. Diário Oficial da União, 28 de maio de 2012.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. **Alimentos regionais brasileiros** / Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica. – 2. ed. – Brasília: Ministério da Saúde, 2015. 484 p. Disponível em: <https://bvsms.saude.gov.br/bvs/publicacoes/alimentos_regionais_brasileiros_2ed.pdf>. Acessado em: 08 de março de 2020.

BREIMAN, L. **Random Forests. Machine Learning,** v. 45, p. 5-32, 2001. Available at: <https://link.springer.com/content/pdf/10.1023/A:1010933404324.pdf >. Accessed on: 08 de março de 2020.

EMBRAPA. **A EMBRAPA nos biomas brasileiros.** 16p. 2004. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/82598/1/a-embrapa-nos-biomas-brasileiros.pdf>. Acessado em: 10 de janeiro de 2021.

FAO, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. GLOBAL FOREST RESOURCES ASSESSMENT, 2020. **Key findings.** Available at: <http://www.fao.org/3/ca8753en/CA8753EN.pdf>. Accessed on**:** 19 de june de 2021.

GLOBAL FOREST WATCH. **Tree cover loss in Brazil.** Available at: <www.globalforestwatch.org>. Accessed on: 19 de june de 2021.

INPE. INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. COORDENAÇÃO GERAL DE OBSERVAÇÃO DA TERRA. PRODES – **Incremento anual de área desmatada no Cerrado Brasileiro.** Disponível em: <http://www.obt.inpe.br/cerrado>. Acessado em: 08 de março de 2021.

LORENZI, H. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas do brasil, vol. 1. **Instituto Plantarum,** Nova Odessa, SP, 2002, 4a. edição.

MELO, F. P. L., *et al.* Adding forests to the water–energy–food nexus. **Nat Sustain 4**, 85–92 (2021). https://doi.org/10.1038/s41893-020-00608-z.
MMA, Ministério do Meio ambiente. **O Bioma Cerrado.** 2019. Disponível em: <http://www.mma.gov.br/biomas/cerrado>. Acessado em: 09 de dezembro de 2019.

MMA, Ministério do Meio ambiente. **O Bioma.** 2020. Disponível em: <https://www.mma.gov.br/biomas.html>. Acessado em: 10 de maio de 2020.

MORALES, G. *et al*. **Automatic segmentation of *Mauritia flexuosa* in unmanned aerial vehicle (UAV) imagery using deep learning.** 2018. ISBN 12. https://res.mdpi.com/d_attachment/forests/forests-09-00736/article_deploy/forests-09-00736-v2.pdf.

MYERS, N. *et al*. Biodiversity hotspots for conservation priorities. **Nature** 403, 853–858 (2000). https://doi.org/10.1038/35002501.

PINTO, M. V. de S.; SILVA, D. L. da; SARAIVA, A. C. F. Production and characterization of the activated carbon from buriti stone (*Mauritia flexuosa* L. f.) to evaluate the adsorption's process of copper (II). **Acta Amaz.**, Manaus, v. 43, n. 1, p. 73-80, mar. 2013. https://doi.org/10.1590/S0044-59672013000100009.

RESENDE, I. L. de M., *et al*. Estrutura etária de populações de *Mauritia flexuosa* L. F. (Arecaceae) de veredas da região central de Goiás, Brasil. **Revista Árvore,** 2012, v. 36, n. 1 pp. 103-112. Disponível em: <https://doi.org/10.1590/S0100-67622012000100012>. Acessado 20 Junho 2021.

VIEIRA, L. M. et al. Fenólicos totais e capacidade antioxidante *in vitro* de polpas de frutos tropicais. Revista Brasileira de Fruticultura, 2011; *33*(3): 888-897. http://dx.doi.org/10.1590/S0100-29452011005000099.

WORLD RESOURCES INSTITUTE. **Global Forest Review.** Primary Rainforest Destruction Increased 12% from 2019 to 2020. Available at: <https://research.wri.org/gfr/forest-pulse>. Accessed on: 19 de june de 2021.

2 MAPPING SINGLE-SPECIE PALM-TREES IN FOREST ENVIRONMENTS USING CONVOLUTIONAL NEURAL NETWORK

ABSTRACT

Accurately mapping individual tree-species in densely forested environments, using only RGB images, is a challenging task. Remote sensing research has focused more on multispectral, hyperspectral, and LiDAR data to perform this despite the comparatively higher cost. The main reason for that is the spectral similarity between species in RGB scenes, which can be a hindrance for most automatic methods. State-of-the-art deep learning methods could be capable of identifying tree-species with an attractive cost, accuracy, and computational load in RGB images. This paper presents a deep learning approach to detect an important multi-use species of palm trees (Mauritia *flexuosa*; i.e., Buriti) in Brazil with aerial RGB imagery. In South-America, this palm-tree is essential for many indigenous and local communities because of its characteristics. The species is also a valuable indicator of water resources, which comes as a benefit for mapping its location. The method is based on a Convolutional Neural Network (CNN) to identify and geolocate singular tree-species in a high-complexity forest environment, and considers the likelihood of every pixel in the image to be recognized as a possible tree. This study also compares the performance of the proposed method against state-of-the-art object detection networks. For this, images from a dataset composed of 1,394 scenes, where 5,334 palm-trees were manually labeled, were used. The results returned a mean absolute error (MAE) of 0.75 trees and F-measure of 86.9%, respectively. These results are better than both Faster R-CNN (MAE of 0.984 trees and F-measure of 81.7%) and RetinaNet (MAE of 3.761 trees and F-measure of 69.8%) considering equal experiment conditions. The proposed method provided fast solutions to detect the palm-trees, with a delivering image detection of 0.073 seconds and a standard deviation of 0.002 using the GPU. In conclusion, the method presented is efficient to deal with a high-density forest scenario, and can accurately map the location of single-species like the M *flexuosa* palm-tree.

Keywords**:** object detection; convolutional neural network; deep learning; environmental monitoring.

2.1 **INTRODUCTION**

The unplanned development and land occupation in both urban and rural areas are the main reasons behind deforestation, contributing to environmental degradation in riparian zones and modify the natural landscape. Multidisciplinary research is necessary to ascertain the population of vegetative species, identifying their locations and distribution patterns. Such information is essential for the management and conservation of vulnerable ecosystems, and mapping these environments may help governmental entities to control or mitigate environmental damage properly. In the last decade, remote sensing data have been widely applied for monitoring vegetation health (Näsi et al., 2015), biomass (Navarro et al., 2020), forest management (Reis et al., 2019), biodiversity (Saarinen et al., 2018), among others (Feng et al., 2015; Ferentinos et al., 2018; Casapia et al., 2019; Li et al., 2020; Santos et al., 2020; Miyoshi et al., 2020). Thus, remote sensing approaches have been used to investigate areas with difficult terrain access, demonstrating great potential for the classification and detection of forest vegetation.

Remote sensing platforms can be embedded with different sensors such as RGB (Red-Green-Blue), multispectral and hyperspectral, LiDAR (Light Detection and Ranging), and others (Morales and Kemper et al., 2018). The identification of arboreous vegetation with remote sensing data depends on the spatial and spectral resolutions (Voss and Sugumaran, 2008). LiDAR sensors can produce accurate data on the height of the trees, which is an excellent variable to be adopted by automatic extraction methods in forest environments (Andersen, Reutebuch, et al., 2006; Ganz et al., 2019). Multispectral and hyperspectral sensors have the advantage of recording the spectral divergence of the vegetation, which is important for enhancing differences between species configurations, health status, etc. (Ozdarici et al., 2015; Csillik et al. 2018; Mioshi et al., 2020). Still, in recent years, high spatial resolution images acquired by RGB sensors have been used in many studies for vegetation identification (Berveglieri et al., 2018; Cao et al., 2018; Lobo Torres et al., 2020; Santos et al., 2020; Weinstein et al., 2020). These sensors have a relatively low cost in comparison with others but offer limited information regarding the spectral range.

For single-tree species mapping, the literature already investigated different methods by evaluating multispectral and hyperspectral data (Clark et al., 2005, Jones et al., 2010; Colgan et al., 2012; Liu et al., 2017; Maschler et al., 2018; Hennessy et al., 2020;), airborne LiDAR point clouds (Dalponte, Ørka, et al., 2014), and multi-sensory data fusion (Dalponte, Bruzzone, et al., 2012; Cho et al., 2012; Apostol et al., 2020). Atzberger et al. (2012) were able to classify tree-species in a temperate forest using satellite multispectral imagery. Franklin and Ahmed (2018) evaluated UAV (Unmanned Aerial Vehicle)-based multispectral image to map deciduous vegetation. Dalponte, Ørka, et al. (2013) used hyperspectral data to detect boreal tree-species at pixel-level, achieving high accuracies forest mapping. Most of these studies were conducted with hyperspectral sensors and LiDAR sensors. However, both hyperspectral and LiDAR data cost and process-demand are non-attractive for rapid decision models. This is different from RGB sensors, which have a lower cost and are highly available. Guimarães et al. (2020) demonstrated that the majority of recent applications are implementing RGB imagery data in the vegetation detection scenario.

The visual inspection of remote sensing imagery is a time-consuming, labor-intensive, and biased task. Therefore, various studies have developed multiple methods regarding the automated extraction of the vegetation features (Jones et al., 2010; Özcan, et al., 2017; Miyoshi et al., 2020). Accurately mapping individual tree-species in densely forested environments is still a challenging task, even for more robust methods. The increase in quality and quantity in remote sensing data, alongside the rapid improvement of technological resources, allowed for the development of intelligent methods in the computational vision community. By combining remote sensing data with artificial intelligence techniques, it is possible to properly map tree-species and improve accuracy in applications regarding vegetation monitoring. In recent years, multiple frameworks have been implemented to assess the performance of such algorithms to accomplish this task (Feng et al., 2015; Belgiu et al., 2016; Näsi et al., 2018; Özcan et al. 2017; Nezami et al., 2020; Navarro et al., 2020).

During the past years, the detection and extraction of trees in high-resolution imagery were performed with more traditional machine learning

algorithms, like support vector machine (SVM), random forest (RF), artificial neural networks (ANN), and others (Nevalainen et al., 2017; Tuominen et al., 2018; Raczko, Zagajewski, 2017; Xie et al., 2019). They returned interesting outcomes in plenty of studies regarding vegetation analysis (Maxwell et al., 2018; Osco et al., 2019; Pham et al., 2019; Miyoshi et al., 2020; Marrs and Meister, 2019; Imangholiloo et al., 2019). However, these learners (known as shallow learners) are limited due to data complexity and may return lower accuracies in comparison against deep learning methods. When considering adverse conditions in a given forest dataset, deeper methods are required.

Identifying individual species in a scene can be a challenging task since (Colgan et al., 2012). However, state-of-the-art deep learning-based methods should be capable of identifying single tree-species with an attractive accuracy and computational load even in RGB images. Recently, deep learning-based methods have been implemented in multiple remote sensing, specifically for image segmentation, classification, and object detection approaches (Nezami et al., 2020; Safonova et al., 2019; Li et al., 2017; Ma et al., 2019). Deep learning techniques are among the most recent adopted approaches to process remote sensing data (Schmidhuber, 2015; Kemper et al., 2018; Kamilaris and Prenafeta-Boldú, 2018). In a general sense, deep learning can return better performance than shallow learners, especially in the presence of large quantities of data or if the input data is highly complex (Sothe et al., 2020; Khamparia and Singh, 2019).

In heavy-dense forested environments, the identification of single-tree species can become a challenge even for robust methods like current state-of-the-art deep networks. This motivated several studies to investigate the performance of deep learning methods to evaluate their performance on this task. A recently published research tested the performance of object detection deep networks like YOLOv3 (Redmon and Farhadi, 2018), RetinaNet (Lin et al., 2017), and Faster-RCNN (Ren et al., 2015) to detect tree-canopy in RGB imagery covering an urban area (Santos et al., 2019). Another study modified the VGG16 (Simonyan and Zisserman, 2015) to monitor the health conditions of vegetation (Sylvain et al., 2019). A combination of LiDAR and RGB images was also used with the RetinaNet to identify tree-crowns in UAV images (Weinstein et al., 2019).

The DenseNet (Hartling et al., 2019) was also implemented multispectral data to classify tree species.

The spatial and spectral divergences between the tree and non-tree are essential for automatic methods (Özcan et al., 2017; Csillik et al., 2018). In highly-dense scenarios like heavily forested areas, the individual detection of trees could be difficult. RGB sensors are not capable of providing the same amounts of spectral data as multispectral or hyperspectral sensors, which offers a potential hindrance for automatic extraction methods. Nonetheless, state-of-the-art deep learning methods based on confidence maps, instead of object detection approaches, could be capable of identifying single tree-species in highly dense areas using RGB images. Methods that could accurately identify a species, among others, may help optimize several applications in environmental planning and forest management.

In the presented context, this paper presents a deep learning approach to detect individual fruit species of palm-trees (Mauritia *flexuosa*; L.f. Buriti) in aerial RGB orthoimage. As the contribution of this approach, a framework to identify and geolocate a single species in a high-complexity forested environment is discussed. The study also compares the performance of the proposed method with other state-of-the-art object detection deep neural networks to test its robustness. The palm-tree M. *flexuosa* is valuable as a source of food, remedy, fiber, and light wood for both indigenous communities and local populations (Almeida et al. 2008, Bortolotto et al. 2018, Gilmore et al. 2013; Hoek et al., 2019). It is also considered a native species of the Brazilian flora with high current and potential economic value (Martins, et al. 2016). Besides, this species has high ecological importance, constituting a food source, nest sites, and habitat to a wide variety of species and provides multiple ecosystem services (Delgado et al. 2007, Khorsand Rosa & Koptur 2013; Hoek et al. 2019; Goulding & Smith 2007, Resende et al. 2012).

2.2 **MATERIALS AND METHOD**

2.2.1 General Description of our method

The approach proposed in this paper is composed of three main phases (see Figure 1): (1) The dataset was composed of aerial RGB orthoimages obtained from a riparian zone of a well-known populated region of M. *flexuosa* palm-trees. With specialist assistance, the palm-trees in the RGB images were visually identified and labeled in a Geographical Information System (GIS). The image and labeled data were split into groups of training, validation, and testing subsets; (2) The object detection method was applied in a computational environment; (3) The performance of the proposed method was compared with other networks.

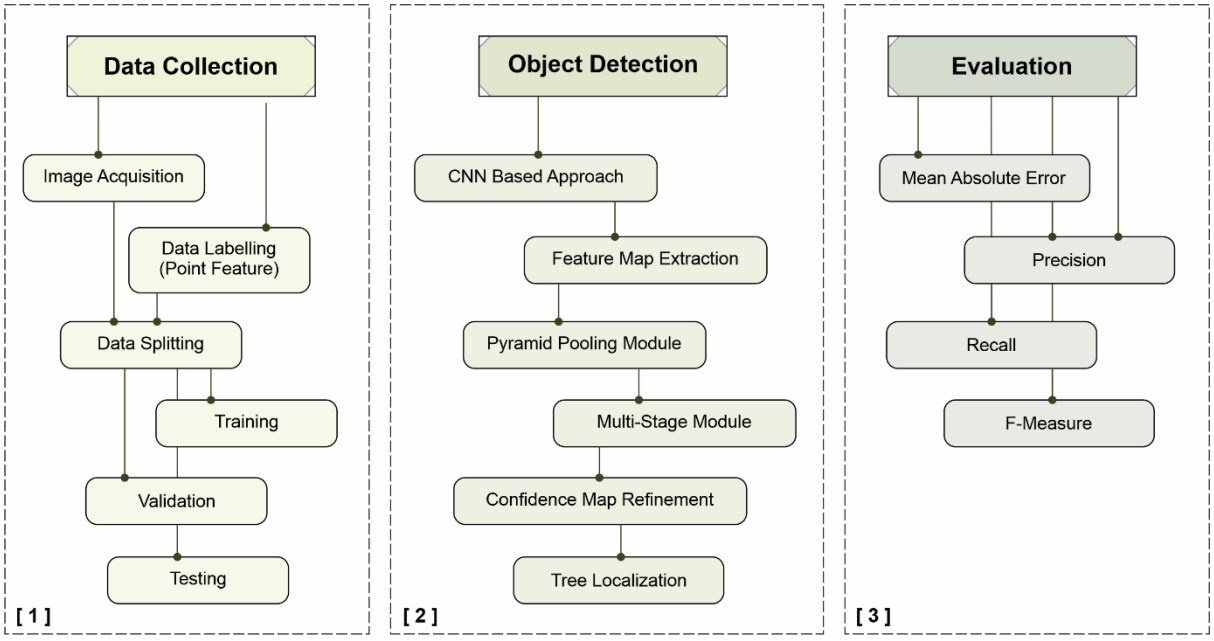

Figure 1. Summarized phases of the proposed approach.

2.2.2 Study Area and Mapped Species

The riparian zone of the upper-stream of the Imbiruçu brook, located near the city of Campo Grande, in the state of Mato Grosso do Sul, Brazil was selected for the study (Figure 2). This stream, formed by a dendritic drainage system, is inserted in the hydrographical basin of the Paraguay River and covered by the

Cerrado (Brazilian Savanna) biome. This area is composed of a highly complex forest patch containing a widespread of palm-tree species Mauritia flexuosa (popular name Buriti). The Arecaceae is a dioecious (Holm, Miller, et al., 2008) long-living species and grows naturally in flooded areas, providing water balance for rivers and other water bodies (Ramos, Curi, et al., 2006). In highly dense, monodominant stands in flooded areas, mature M. flexuosa palm-trees reach 20 m high (Holm, Miller, et al., 2008). The evaluated site in our experiment, in specific, is one of the well-known locations where a large number of samples of this species is sufficient to train a deep neural network.

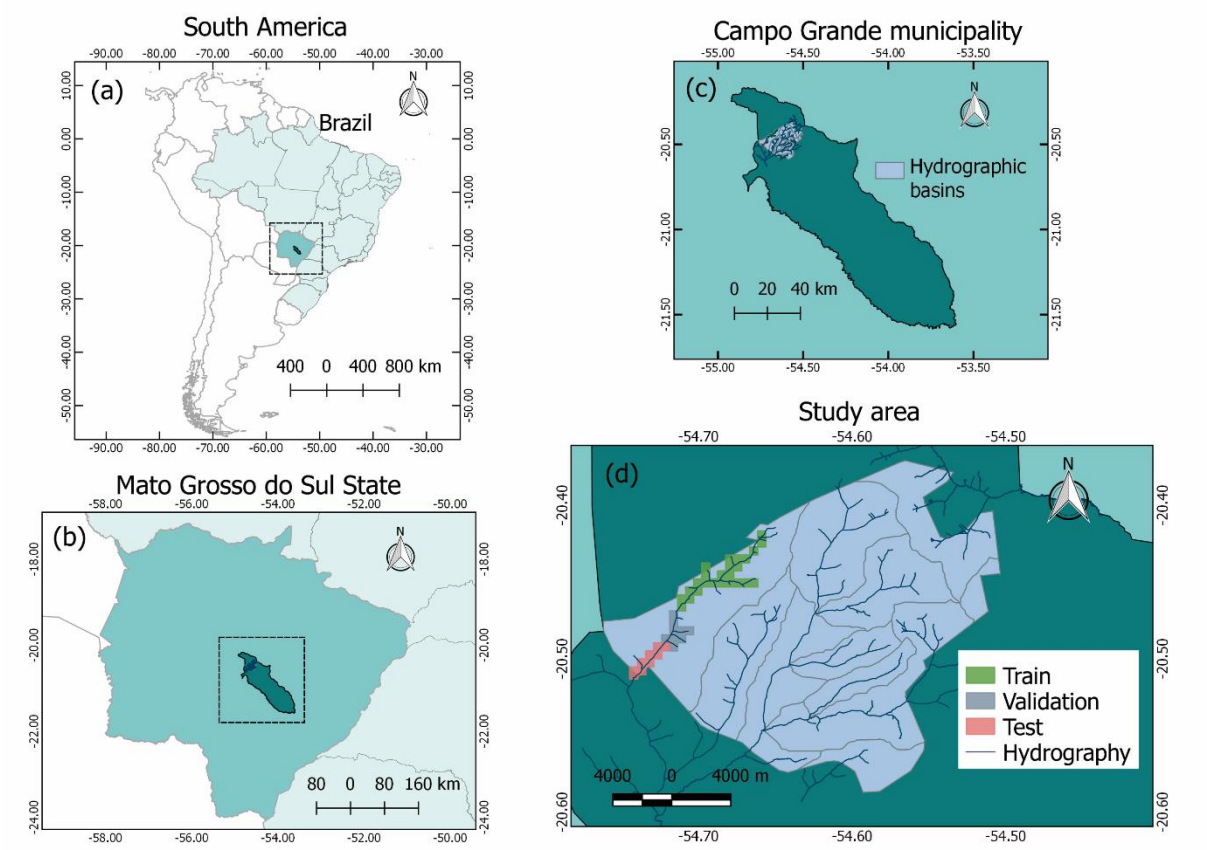

Figure 2. Location map of the study area in (a) South America and Brazil, (b) Mato Grosso do Sul, (c) Campo Grande, and (d) Study area.

The aerial RGB orthoimages were provided by the city of Campo Grande, State of Mato Grosso do Sul, Brazil. The ground sample distance (GSD) of the orthoimages is 10 cm. A total of 43 orthoimages with dimensions 5619 x 5946 pixels were used in the study. This aerial image dataset was composed of 1,394

scenes, where 5,334 palm-trees were manually labeled and used as ground-truth (Figure 3).

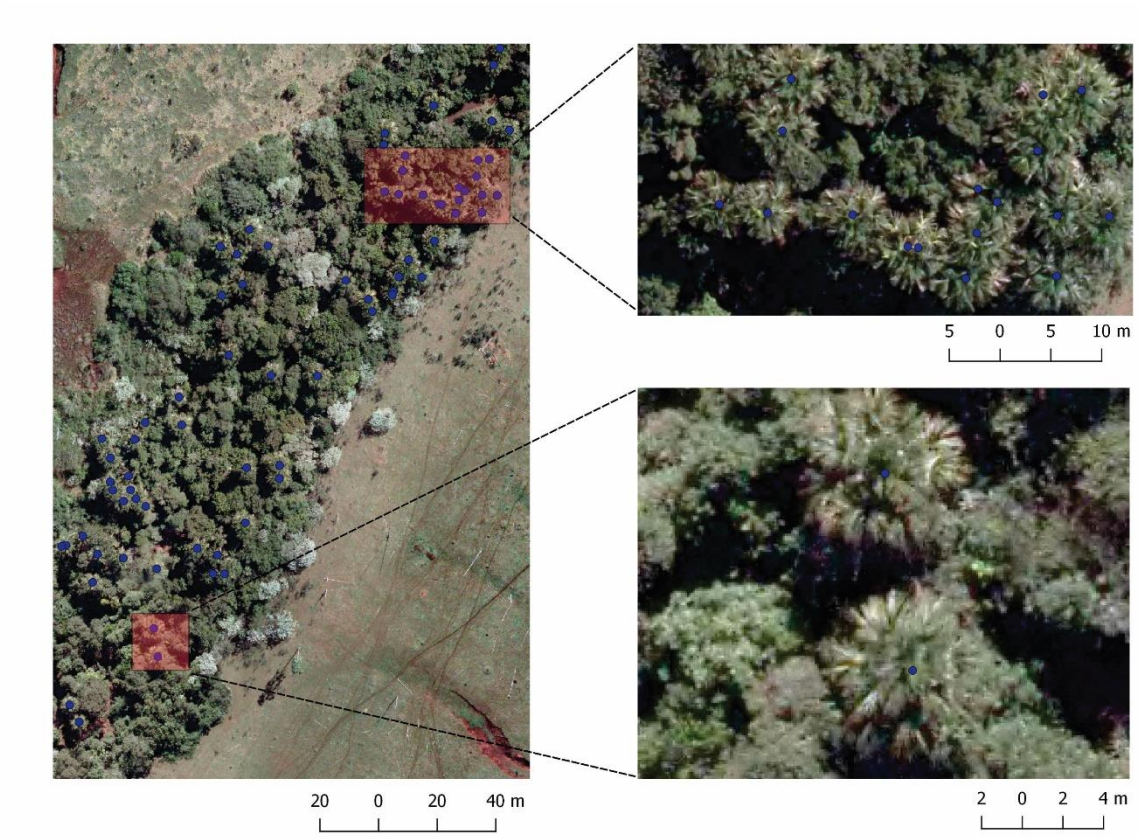

Figure 3. Examples of the labeled dataset. M. flexuosa palm-trees are represented with blue dots.

2.2.3 Proposed Method

This study proposes a CNN method that uses the RGB image as an input and, throughout a confidence map refinement, returns a prediction map with the tree locations (Figure 4). The objects' position is calculated after a 2D confidence map estimation, based on previous works (Aich and Stavness, 2018; Osco et al., 2020). The first step of the architecture extracts the feature map. In a sequential step, the feature map goes through the Pyramid Pooling Module (PPM) (Zhao et al., 2017). The last step of the architecture produces a confidence map in a Multi-Stage Module (MSM) that enhances the position of the tree by adjusting the prediction to its center.

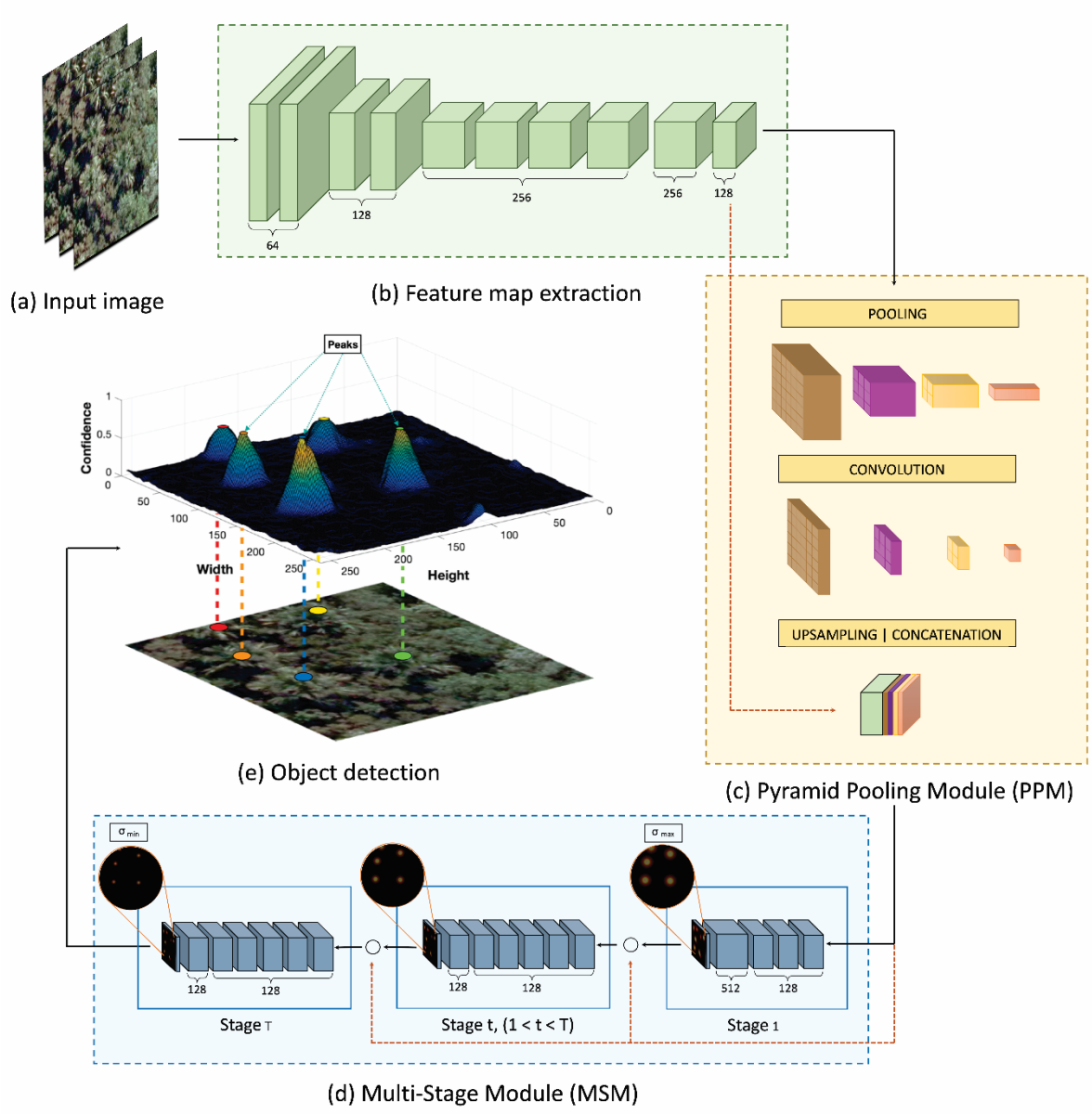

Figure 4. Proposed CNN. The feature map (b) is extracted from the input image (a) and improved by the PPM (c). The result is used as input at the MSM step (d), where T stages enhance the prediction positions (e).

2.2.3.1 Feature Map Extraction

For the feature map extraction (Figure 4(b)), the proposed CNN is based on the VGG19 (Simonyan and Zisserman, 2015). Here, the CNN is composed of 8 convolutional layers with 64, 128, and 256 filters with a 3 x 3 size window, with Rectified Linear Units (ReLU) functions in all layers. The spatial volume size was reduced in half after the second and fourth layers by a max-pooling layer (2 x 2

window). The PPM (Zhao et al., 2017) was used (Figure 4(c)) to extract global and local information, which helps the CNN to be less sensitive to tree scale changes. The feature maps extracted by PPM are up sampled to sizes equivalent of the input feature map and concatenated with it to create an enhanced characterization of the image.

2.2.3.2 Tree Localization

The MSM step (Figure 4(d)) estimates the confidence map from the feature map extracted in the previous module. The MSM is composed of *T* refinement stages, where the first stage contains 3 layers of 128 filters with *3 x 3* size, 1 layer with 512 filters of *1 x 1* size, and one final layer with 1 filter that generates the confidence map C_1 from the first stage. The position of the trees predicted in the first stage is refined in the *T - 1* stages. In each stage *t ∈ [2, 3, ..., T]*, the prediction *C_{t-1}* returned from a previous stage t-1 and the feature map from the PPM module are concatenated. The concatenation process allows for both global and local context information to be incorporated in it . The final layer in this step has a sigmoid activation function since the method considers the probability of occurrence of a tree to exist or not [0,1]. At the end of each stage, a loss function Equation 1 is adopted to avoid the vanishing gradient problem. The general loss function is calculated by the following Equation 2.

$$f_t = \sum_p \| \hat{C}_t(p) - C_t(p) \|_2^2, \quad (1)$$

where $\hat{C}_t(p)$ is the ground-truth confidence map with location (*p*) in the stage (t).

$$f = \sum_{t=1}^{T} f_t \quad (2)$$

The confidence map is generated by a 2D Gaussian kernel at the center of the labeled tree. A standard deviation σ_t controls the spread of the peak for each Gaussian kernel (Figure 5). Different values of σ_t were used to refine the predictions. The value of σ_1 in the MSM is set to maximum (σ_{max}) while the $\boldsymbol{\sigma}\_T$

in the final stage is set to minimum (σ_{min}). In the early phases of the experiment, different values for $\boldsymbol{\sigma}$ were adopted to improve its robustness. Finally, the tree location is estimated by the peaks of the confidence map (Figure 5). These peaks are considered the local maximum of the confidence map, representing a high probability of a tree occurrence . The *p = (xp, yp)* is considered as a local maximum if $C_T(p)>C_T(v)$ for all the neighbors *v*. Here, v is given by *(xp1, yp)* or *(xp, yp1)*.

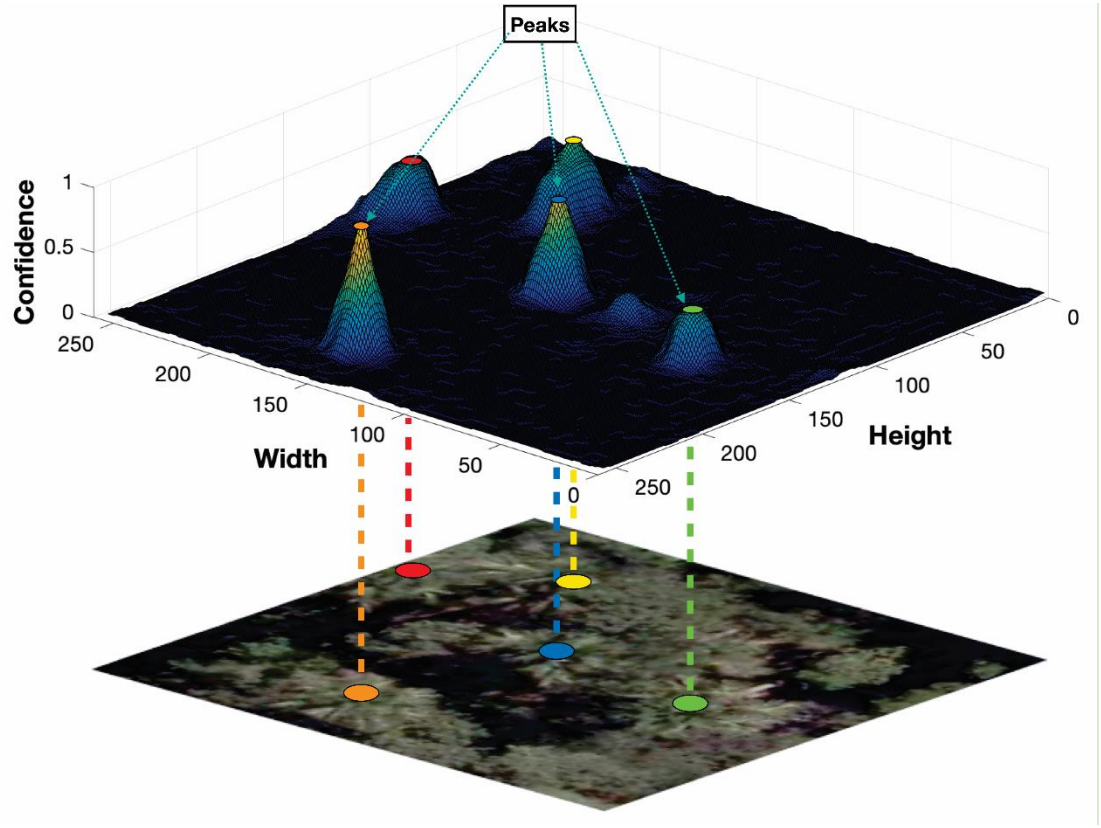

Figure 5. Tree localization example from a refined confidence map.

A peak in the confidence map is defined as a tree if $C_T(p)> \tau$. To prevent the network from confusing trees in a nearby range from each other, a distance of δ is defined. For this study, δ equal to 1 pixel and τ equal to 0.35 were defined as valid metrics. These values were defined during a previous experimental phase.

2.2.4 Experimental Setup

For the experimental setup, the RGB ortohomosaics were separated into 25 (60%), 9 (20%), and 9 (20%) for training, validation, and testing, respectively (Figure 6). Later, they were split into non-overlapping patches of 256 x 256 pixels (25.6 m x 25.6 m). The patches were then divided into training (42.3%), validation (34.5%), and testing (23.2%) sets. Table 1 lists the number of samples (trees) and image patches, and Figure 6 displays examples of the orthomosaics used to extract the datasets. For the training process, the CNN was initialized with pre-trained weights from ImageNet (< http://www.image-net.org/ >) and a Stochastic Gradient Descent optimizer was applied with a moment equal to 0.9. For this, the validation set was used to adjust the learning rate and the number of epochs, which were set to 0.001 and 100, respectively.

Table 1. *Description of the training, validation, and testing sets of the M. flexuosa palms-trees dataset.*

Dataset	Number of Patches	Number of Samples
Training	590	1,784
Validation	481	2,296
Testing	323	1,254

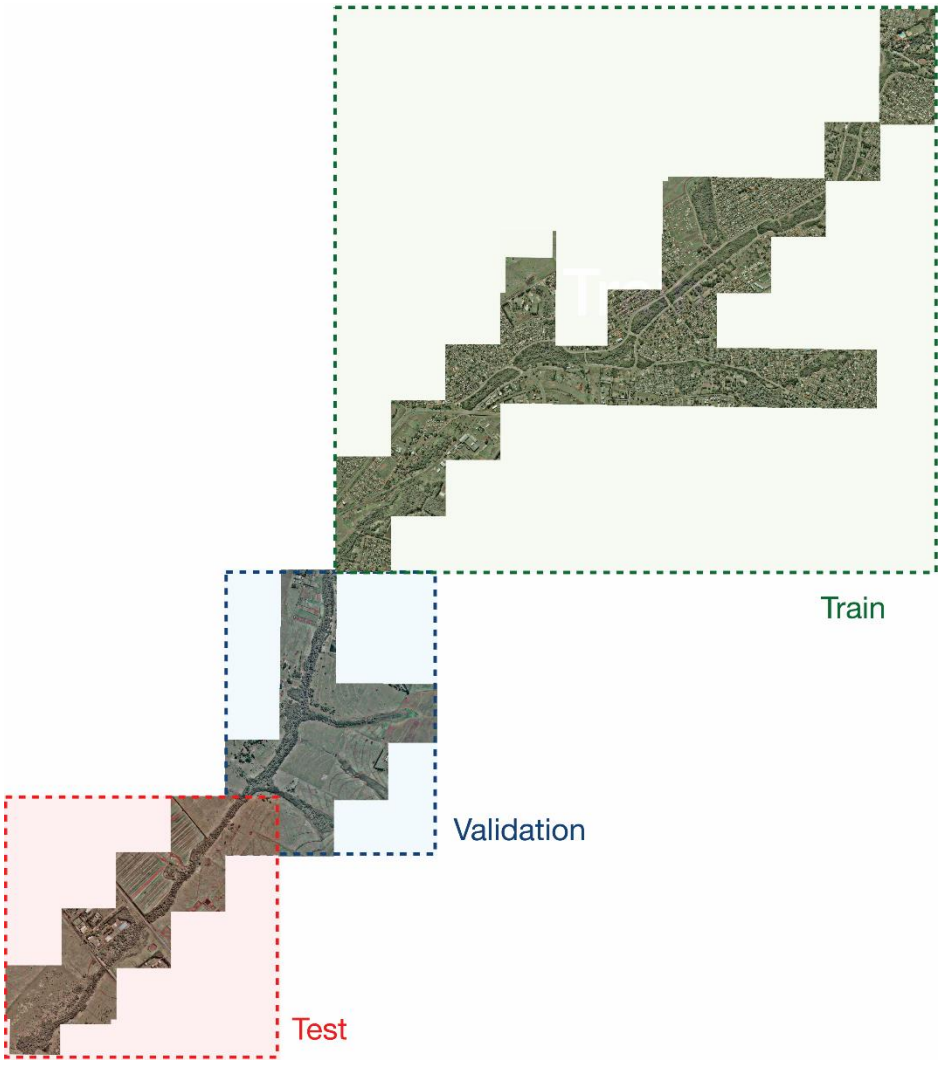

Figure 6. Training, validation, and testing datasets separated per region.

The performance of the proposed network was assessed with the following metrics: mean absolute error (MAE); precision (P); recall (R), and; F-measure (F_1). The results were compared with Faster R-CNN and RetinaNet methods. Since these methods are based on bounding-boxes, the plant-position (x, y) from the labeled ground-truth was used as a center of the box. The correct size of the box corresponds with the size occupied by the tree-canopy. To perform this comparison, the same conjunct of training, validation, and testing datasets were adopted for the three methods. Likely, an inverse process was applied during the test phase, where the position of the tree was obtained by the center of the point inside the predicted bounding-box of the RetinaNet and Faster R-CNN methods. This allowed applying the same metrics (MAE, P, R, and F_1) for comparing the performances of each neural network.

2.3 **RESULTS**

2.3.1 Validation of the parameters

The proposed approach parameters σ_{min}, σ_{max} , and the number of stages T, are responsible for refining the prediction map. Initially, the influence of these parameters was evaluated on the M. *flexuosa* palms-trees validation set. Table 2 shows the evaluation of the number of stages T used in the MSM refinement phase. In this experiment, the values of σ_{min} = 1, σ_{max} = 4 and ranged T from 1 to 5 were set, and it was discovered that T = 4 achieved the best performance among the numbers of analyzed stages, reaching an MAE of 0.852 trees and an F-measure of 87.1%.

Table 2. *Influence of the number of stages (T) in counting and detection of M. flexuosa palms-trees (σ_{min} = 1 and σ_{max} = 4 were adopted).*

Stages (T)	MAE	Precision (%)	Recall (%)	F-measure (%)
1	0.933	85.1	86.4	83.8
2	0.943	93.5	83.6	86.9
4	**0.852**	**91.5**	**85.5**	**87.1**
5	0.966	93.9	83.1	86.6

The values of σ_{min} and σ_{max} applied in the refinement stage were also evaluated. For this, the number of stages T = 4 was adopted in subsequent experiments since it obtained the best results (see Table 2). Since the σ_{min}values represent the dispersion of the density maps around the center of the trees, it was found that smaller values do not correctly cover the trees and, therefore, can impair the detection. On the other hand, higher σ_{min} values are also harmful as they cover more than one tree per area. Thus, the best results were obtained with σ_{max} = 4, indicating that it fits better with the M. *flexuosa* palms-trees characteristics, and generates a more accurate refinement map.

Table 3. Influence of the σ_{max} in counting and detection of M. flexuosa palms-trees (σ_{min} = 1 and stages T = 4 were adopted).

σ_{max}	**MAE**	**Precision (%)**	**Recall (%)**	**F-measure (%)**
3	0.931	86.7	88.7	85.8
4	**0.852**	**91.5**	**85.5**	**87.1**
5	1.611	91.4	69.6	76.8

Table 4 presents the evaluation of different values of σ_{min} responsible for the last stage of the MSM. For this, σ_{max} = 4 and T = 4 were adopted since they obtained better results in the previous experiments (Tables 2 and 3). When σ_{min}= 1, the proposed approach returned the best performances among the analyzed values. Therefore, the refinement step implemented with values of σ_{min} = 1, σ_{max} = 4, and T = 4 generated a more accurate refinement on the validation set.

Table 4. Influence of the σ_{min} in counting and detection of M. flexuosa palms-trees (σ_{max} = 4 and stages T = 4 were used).

σmin	**MAE**	**Precision (%)**	**Recall (%)**	**F-measure (%)**
0.75	1.879	90.7	64.9	73.3
1	**0.852**	**91.5**	**85.5**	**87.1**
1.5	1.012	83.9	88.9	84.3
2	1.671	72.4	91.0	77.6

2.3.2 Comparative results between object detection methods

The proposed method returned better performances when compared against different object detection methods like Faster R-CNN and RetinaNet. The MAE, precision, recall, and F-measure metrics were calculated for each of them, and results are displayed in Table 5. The proposed approach achieved high precision and good F-measure values but returned a slight-lower recall value

when confronted against them. Nonetheless, it is essential to consider the tradeoff in recall difference (-6.58% from the Faster R-CNN and -12.35% from the RetinaNet) with the precision difference (+14.52 from the Faster R-CNN and +35.49% from the RetinaNet).

Since the F-measure uses both the precision and the recall values to compute the test results, it can be assumed that the proposed approach performs better and returns a better balance between true-positive predicted and true-positive rates concerning the identification of palm-trees. Nonetheless, the results are consistent with recent literature where object detection applications were applied for the identification of single tree-species (Li et al., 2019; Djerriri et al., 2018; Osco et al., 2020; Scillik et al., 2018; Santos et al., 2019); but performed in the non-RGB image domain. The low precision values for the bounding-boxes methods may be explained by a high-density of objects (i.e., M. *flexuosa* palm-trees). This condition is described as problematic for deep networks based on these characteristics, especially when the boxes have high-intersections with similar objects (Goldman et al., 2019; Hsieh et al., 2017).

Table 5. Comparative results between our method and Faster R-CNN and RetinaNet.

Method	MAE	Precision (%)	Recall (%)	F-measure (%)
Faster R-CNN	0.984	79.0	90.8	81.7
RetinaNet	3.761	58.0	**96.6**	69.8
Proposed Method	**0.758**	**93.5**	84.2	**86.9**

To verify the potential of the proposed approach in real-time processing, a comparison of its performance with the other state-of-the-art methods was conducted. Table 6 shows the average processing time and standard deviation for 100 images of the test set. The values of σ_{min} = 1, σ_{max} = 4 and T = 4 were used to obtain the best performance in previous experiments. The results show that the approach can achieve real-time processing, delivering image detection in 0.073 seconds with a standard deviation of 0.002 using a GPU. Similarly,

RetinaNet and Faster R-CNN methods obtained an average detection time and standard deviation of 0.057, 0.046, and 0.002, 0.001, respectively.

Table 6. Processing time evaluation of the analyzed approaches.

Method	CPU		GPU	
	Average Time (sec)	**Standard deviation**	**Average Time (sec)**	**Standard deviation**
Faster R-CNN	1.57	0.031	0.05	0.001
RetinaNet	1.93	0.028	0.06	0.002
Proposed Method	1.26	0.051	0.07	0.002

Figure 7 presents the qualitative results of the proposed method where the annotations of M. *flexuosa* palm-trees are marked with yellow circles, and the blue dots indicate the correctly detected positions. This approach correctly detects the M. *flexuosa* palm-trees in different capture conditions, such as overlapping trees (Figure 7(a)), partial occlusion of the treetops (Figure 7(b)), and highly dense vegetation areas (Figure 7(c)), highlighted by orange circles. Moreover, the predicted positions have a satisfactory level of accuracy, generating detections (blue dots) close to the annotations (centers of the yellow circles).

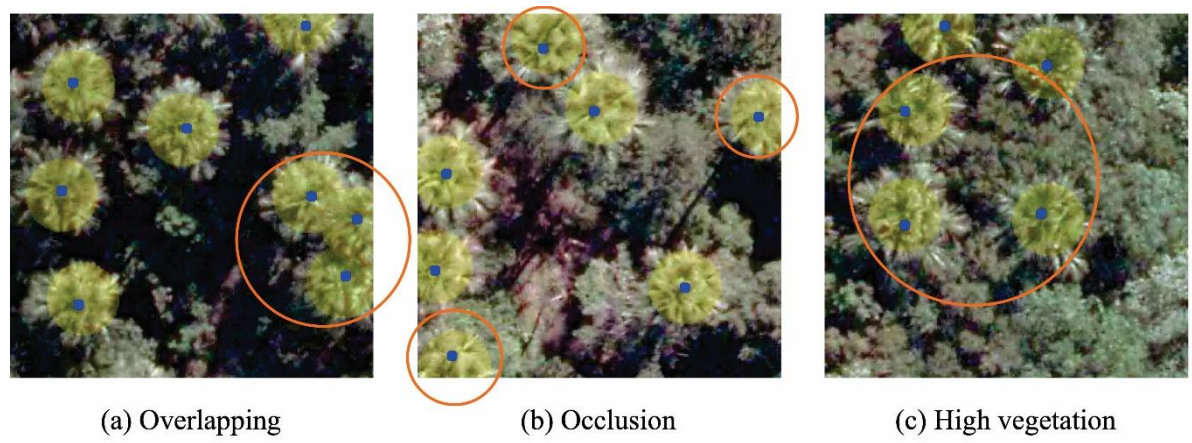

Figure 7. Qualitative results of the proposed method in three scenes: (a) an example of the detected nearby trees with overlapping, (b) detected trees with parts of the canopy occluded at the edge of the image, and (c) demons detected trees in areas of high vegetation. The orange circles highlight the detections.

Although the method obtained good results in the detection of M. *flexuosa* palm-trees, it faces some challenges. Figure 8 presents areas where the incorrect detections are shown by the red circles. The main challenge is the detection of trees with a high level of occlusion at the image boundary or by overlapping of trees (highlighted by the orange circles). However, even in these few cases, the proposed approach can correctly detect most of the palm-trees.

Figure 8. Examples of the challenges faced by our method in the M. flexuosa palm-tree detection task. The orange circles indicate challenging detections.

The visual comparison of the palm-tree detection approaches is shown in Figure 9. Column (a) displays the detections obtained by the proposed method, while Columns (b) and (c) are related to the compared methods: Faster R-CNN and RetinaNet, respectively. The approach that obtained the worst result was RetinaNet (Figure 9(c)), generating many false-positives (red dots) close to the M. *flexuosa* palm-trees detections. On the other hand, Faster R-CNN (Figure 9(b)), despite having fewer false-positives, did not properly learn the characteristics of the palm-trees and incorrectly detected other tree species among them. Following the quantitative results shown in Table 5, the proposed approach has the greater precision in detecting M. *flexuosa* palm-trees (Figure 9(a)), while having the least number of incorrect detections (false-positives).

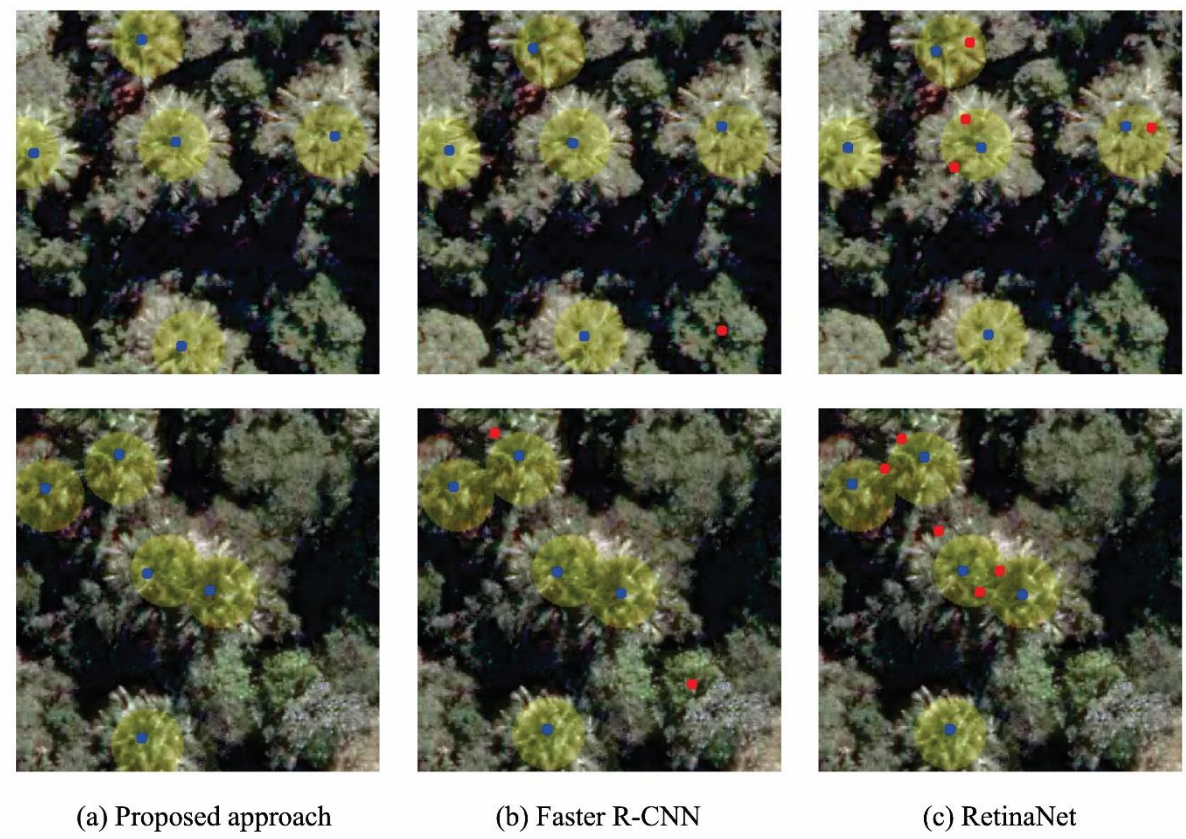

Figure 9. Visual comparison of the analyzed methods. (a) shows the detections obtained by the proposed approach; (b) indicates the detections of the Faster R-CNN and; (c) demonstrates the detections of the RetinaNet. Blue and red points correspond to correct and incorrect detection positions, respectively, and the yellow circle to M. flexuosa palms-trees annotation.

2.4 **DISCUSSION**

This study demonstrated a feasible method to automatically map single palm-tree species of the M. *flexuosa* plant genera using an RGB imagery dataset. *Mauritia flexuosa* frequently occurs at low elevations, with high density on river banks and lake margins around water sources, and in inundated or humid areas (Martins et al. 2012a). This is one of the most widely distributed palm trees in South America, Brazil. This species occurs in the Amazon region, Caatinga, Cerrado, and Pantanal (Lorenzi et al. 2004), and is one of the palm trees mostly used by humans, being an important item in the diet of many indigenous groups and rural communities (Martins et al. 2012a, b).

Mapping the M. *flexuosa* palm-trees is an important practice for multiple regions of South America, like Brazil, where this plant is viewed as a valuable resource. This palm is widely used for several purposes, is considered a species of multiple-use (Santos and Coelho-Ferreira 2012, Bortolotto et al. 2018), occurs in areas of “Veredas”, considered protected by the Brazilian forest code, but there is still a great lack of characterization of the habitats of this species in this country. Mapping and identifying populations of the palm *M.flexuosa* is relevant because it is a reliable indicator of water resources, such as streams inside dense gallery forests, slow-flowing swamp surface water, and shallow groundwater in the Cerrado region, vital for humans and wildlife, besides being a valuable source of several non-timber forest products. The approach provides useful information for sustainable economic use and conservation.

As described, single tree-species identification is a challenging task even for state-of-the-art deep neural networks when considering only RGB imagery. Mainly because forest environments are constituted by multiple spectral-spatial information, overlapping canopies, leaves and branches, sizes, growth-stages, and densities, among others. In this manner, studies considered different data to help solve this issue like density-points information, canopy high, digital terrain and surface models, spectral divergence, etc. (Özcan et al., 2017; Franklin et al., 2018; Saarinen et al., 2018; Sothe et al., 2020; Tuominen et al., 2018). Regardless, in this paper, it is proposed a simplification of this process by adopting little input information (i.e., label feature as points and RGB imagery) and a robust method that once trained, can rapidly perform and resolve the said task even in a real-time context.

The results of the present approach achieved satisfactory precision (93.5%), recall (84.2%), and F-measure (86.9%) values, respectively), and a small MAE (0.758 trees). Studies that applied deep neural networks for detecting other types of arboreal vegetation returned approximated metrics. For the identification of citrus-tree, a CNN method was able to provide 96.2% accuracy (Csillik et al., 2018), and in oil palm-tree detections, a deep neural network implementation returned an accuracy of 96.0% (Li et al., 2019). One different kind of palm-trees than the ones evaluated in our dataset was investigated with a modification of the AlexNet CNN architecture and returned high prediction values

(R^2 = 0.99, with the relative error between 2.6% to 9.2%) (Djerriri et al., 2018). A study (Santos et al., 2019) achieved accuracy higher than 90% to detect single tree-species using the RetinaNet and RGB images. However, in all aforementioned papers, the tree density patterns differentiate from ours, and the evaluated individual trees are more spaced from each other, which makes a simpler object detection problem.

In the described manner, the proposed method may help in mapping the M. *flexuosa* palm-tree with little computational load and high accuracy. Since this approach can compute point features as labeled objects, it reduces the amount of label-work required from the human counterpart. Additionally, the method provided a fast solution to detect the palm-tree's locations with a delivering image detection of 0.073 seconds and a standard deviation of 0.002 using a GPU. This information is essential for properly calculating the cost and effectiveness of the method. The presented approach may help new research while providing primary information for exploring environmental management practices in the experiment context (i.e., evaluating a keystone tree-species). The proposed method could also be incorporated into areas and regions to help detect the M. *flexuosa* palm-tree and contribute to decision-making conservation measures of said species.

2.5 **CONCLUSION**

This paper presents an approach based on deep networks to map single species of fruit palm-trees (Mauritia *flexuosa*) in aerial RGB imagery. According to the performance assessment, the method returned an MAE of 0.75 trees and F-measure of 86.9%. A comparative study also shows that the proposed method returned better accuracy than state-of-the-art methods like Faster R-CNN and RetinaNet under the same experimental conditions. Besides, this approach took a shorter time to detect the palm-trees with 0.073 seconds for delivering image detection and achieved a standard deviation of 0.002 using the GPU. In future implementations, it should be possible to add new strategies in this CNN architecture to overcome challenges regarding other tree patterns. Still, the identification of individual species can help to assist in both monitoring and mapping important singular species. As such, the proposed method may assist

in new research for the forest remote sensing community that includes data obtained with RGB sensors.

Acknowledgments: This research was partially funded by CNPq (p: 433783/2018-4, 303559/2019-5 and 304052/2019-1) and CAPES Print (p: 88881.311850/2018-01). The authors acknowledge the support of the UFMS (Federal Universityof Mato Grosso do Sul) and CAPES (Finance Code 001), and NVIDIA© for the donation of the Titan X graphics card used in the experiments.

AUTHORS

Luciene Sales Dagher Arce [a], Mauro dos Santos de Arruda [a], Lucas Prado Osco [a,b,*], Ana Paula Marques Ramos [c], Camila Aoki [a], Arnildo Pott [a], Sarah N. Fatholahi [d], Jonathan Li [d], Wesley Nunes Gonçalves [a,e], José Marcato Junior [a]

[a] *Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul (UFMS), Campo Grande, Mato Grosso do Sul, Brazil; E-mail: lucienearcepos@gmail.com, mauro.arruda@ufms.br, camila.aoki@ufms.br, arnildo.pott@gmail.com, jose.marcato@ufms.br*

[b] *Faculty of Engineering and Architecture and Urbanism, University of Western São Paulo (UNOESTE), Presidente Prudente, São Paulo, Brazil; E-mail: lucasosco@unoeste.br*

[c] *Graduate Program in Environment and Regional Development, University of Western São Paulo (UNOESTE), Presidente Prudente, São Paulo, Brazil; E-mail: anaramos@unoeste.br*

[d] *Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada; nfatholahi7@gmail.com, junli@uwaterloo.ca,*

[e] *Faculty of Computer Science, Federal University of Mato Grosso do Sul (UFMS), Campo Grande, Mato Grosso do Sul, Brazil; E-mail: wesley.goncalves@ufms.br*

** Corresponding author: E-mail address: lucasosco@unoeste.br (L. P. Osco)*

REFERENCES

Almeida, S. P., Costa, T. S. A., Silva, J. A., 2008. Frutas nativas do Cerrado caracterização físico-química e fonte potencial de nutrientes. **In**: SANO, S. M.; ALMEIDA, S. P.; RIBEIRO, J. F. (Eds.). Cerrado: ecologia e flora. Brasília: Embrapa Cerrados/Embrapa Informação Tecnológica, p.351-381.

Ampatzidis, Y., Partel, V., 2019. UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence. **Remote Sens**. 11(4), 410. https://doi.org/10.3390/rs11040410

Andersen, H.E., Reutebuch, S.E., McGaughey, R.J., 2006. A rigorous assessment of tree height measurements obtained using airborne LiDAR and conventional field methods. **Can. J. Remote Sens**. 32, 355–366. https://doi.org/10.5589/m06-030

Apostol, B., Petrila, M., Lorenţ, A., Ciceu, A., Gancz, V., Badea, O., 2020. Species discrimination and individual tree detection for predicting main dendrometric characteristics in mixed temperate forests by use of airborne laser scanning and ultra-high-resolution imagery. **Sci. Total Environ**. 698. https://doi.org/10.1016/j.scitotenv.2019.134074

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. **Remote Sens**. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Berveglieri, A., Imai, N.N., Tommaselli, A.M.G., Casagrande, B., Honkavaara, E., 2018. Successional stages and their evolution in tropical forests using multi-temporal photogrammetric surface models and super pixels. ISPRS J. Photogramm. **Remote Sens**. 146, 548–558. https://doi.org/10.1016/j.isprsjprs.2018.11.002

Bortolotto, I. M., Damasceno-Junior, G. A. Pott, A., 2018. **Lista preliminar das plantas alimentícias nativas de Mato Grosso do Sul, Brasil**. Iheringia, Série Botânica, 73(supl.):101-116.

Cao, J., Leng, W., Liu, K., Liu, L., He, Z., Zhu, Y., 2018. Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. **Remote Sens**. 10(1), 89. https://doi.org/10.3390/rs10010089

Casapia, X.T., Falen, L., Bartholomeus, H., Cárdenas, R., Flores, G., Herold, M., Honorio Coronado, E.N., Baker, T.R., 2019. Identifying and quantifying the abundance of economically important palms in tropical moist forest using UAV imagery. **Remote Sens**. 12(1), 9. https://doi.org/10.3390/rs12010009

Cho, M.A., Mathieu, R., Asner, G.P., Naidoo, L., van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I.P.J., Erasmus, B., 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. **Remote Sens. Environ**. 125, 214–226. https://doi.org/10.1016/j.rse.2012.07.010

Clark, M.L., Roberts, D.A., Clark, D.B., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. **Remote Sens. Environ**. 96, 375–398. https://doi.org/10.1016/j.rse.2005.03.009

Colgan, M.S., Baldeck, C.A., Féret, J.-B., Asner, G.P., 2012. Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. **Remote Sens**. 4(11), 3462–3480. https://doi.org/10.3390/rs4113462

Csillik, O., Cherbini, J., Johnson, R., Lyons, A., Kelly, M., 2018. Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks. **Drones**. 2(4), 39. https://doi.org/10.3390/drones2040039

Dalponte, M., Bruzzone, L., Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. **Remote Sens**. Environ. 123, 258–270. https://doi.org/10.1016/j.rse.2012.03.013

Dalponte, M., Bruzzone, L., Vescovo, L., Gianelle, D., 2009. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. **Remote Sens. Environ**. 113, 2345–2355. https://doi.org/10.1016/j.rse.2009.06.013

Dalponte, M., Ørka, H.O., Ene, L.T., Gobakken, T., Næsset, E., 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. **Remote Sens. Environ**. 140, 306–317. https://doi.org/10.1016/j.rse.2013.09.006

Dalponte, M., Ørka, H.O., Gobakken, T., Gianelle, D., Næsset, E., 2013. Tree species classification in boreal forests with hyperspectral data. **IEEE Trans. Geosci. Remote Sens**. 51, 2632–2645. https://doi.org/10.1109/TGRS.2012.2216272

Delgado, C., Couturier, G., Mejia, K., 2007. Mauritia *flexuosa* (Arecaceae: Calamoideae), an Amazonian palm with cultivation purposes in Peru. **Fruits**, 62: 157-169.

Feng, Q., Liu, J., Gong, J., 2015. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. **Remote Sens**. 7, 1074–1094. https://doi.org/10.3390/rs70101074

Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. **Comput. Electron. Agric**. 145, 311–318. https://doi.org/10.1016/j.compag.2018.01.009

Franklin, S.E., Ahmed, O.S., 2018. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. **Int. J. Remote Sens**. 39, 5236–5245. https://doi.org/10.1080/01431161.2017.1363442

Ganz, S., Kaber, Y., Adler, P., 2019. Measuring Tree Height with Remote Sensing-A Comparison of Photogrammetric and LiDAR Data with Different Field Measurements. **Forests** 10, 694. https://doi.org/10.3390/f10080694

Gilmore MP, Endress BA, Horn CM., 2013. The socio-cultural importance of Mauritia *flexuosa* palm swamps (aguajales) and implications for multi-use management in two Maijuna communities of the Peruvian Amazon. **J Ethnobiol Ethnomed** 9:29

Guimarães, N., Pádua, L., Marques, P., Silva, N., Peres, E., Sousa, J.J., 2020. Forestry remote sensing from unmanned aerial vehicles: A review focusing on the data, processing and potentialities. **Remote Sens**. 12(6), 1046. https://doi.org/10.3390/rs12061046

Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., Carron, J., 2019. Urban tree species classification using a WorldView-2/3 and LiDAR data fusion approach and deep learning. **Sensors**. 19(6), 1284. https://doi.org/10.3390/s19061284

Hennessy, A., Clarke, K., Lewis, M., 2020. Hyperspectral classification of plants: A review of waveband selection generalisability. **Remote Sens**. 12(1), 113. https://doi.org/10.3390/rs12010113

Hoek, Y., Solas, S.Á., Peñuela, M.C., 2019. The palm Mauritia *flexuosa*, a keystone plant resource on multiple fronts. **Biodivers. Conserv**. 28, 539–551. https://doi.org/10.1007/s10531-018-01686-4

Holm, J.A., Miller, C.J., Cropper, W.P., 2008. Population dynamics of the dioecious Amazonian palm Mauritia *flexuosa*: Simulation analysis of sustainable harvesting. **Biotropica** 40, 550–558. https://doi.org/10.1111/j.1744-7429.2008.00412.x

Imangholiloo, M., Saarinen, N., Markelin, L., Rosnell, T., Näsi, R., Hakala, T., Honkavaara, E., Holopainen, M., Hyyppä, J., Vastaranta, M., 2019. Characterizing seedling stands using leaf-off and leaf-on photogrammetric point clouds and hyperspectral imagery acquired from unmanned aerial vehicle. **Forests**. 10, 415. https://doi.org/10.3390/f10050415

Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree species classification with Random forest using very high spatial resolution 8-band worldView-2 satellite data. **Remote Sens**. 4, 2661–2693. https://doi.org/10.3390/rs4092661

Jones, T.G., Coops, N.C., Sharma, T., 2010. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. **Remote Sens. Environ**. 114, 2841–2852. https://doi.org/10.1016/j.rse.2010.07.002

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. **Comput. Electron. Agric**. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Khamparia, A., Singh, K., 2019. A systematic review on deep learning architectures and applications. **Expert Syst**. 36, e12400. https://doi.org/10.1111/exsy.12400

Khorsand Rosa R, Koptur S., 2013. New findings on the pollination biology of Mauritia *flexuosa* (Arecaceae) in Roraima, Brazil: linking dioecy, wind, and habitat. **American Journal of Botany**. Mar;100(3):613-621.

Li, L., Chen, J., Mu, X., Li, W., Yan, G., Xie, D., Zhang, W., 2020. Quantifying understory and overstory vegetation cover using UAV-based RGB Imagery in forest plantation. **Remote Sens**. 12(2), 298. https://doi.org/10.3390/rs12020298

Li, W., Fu, H., Yu, L., Cracknell, A., 2017. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. **Remote Sens**. 9. https://doi.org/10.3390/rs9010022

Lin, T.-Y., Goyal, P., Girshick, R.B., He, K., Dollár, P., 2017. Focal loss for dense object detection. **CoRR**, abs/1708.02002.

Liu, L., Song, B., Zhang, S., Liu, X., 2017. A novel principal component analysis method for the reconstruction of leaf reflectance spectra and retrieval of leaf biochemical contents. **Remote Sens**. 9(11), 1113. https://doi.org/10.3390/rs9111113

Lobo Torres, D., Feitosa, R., Nigri Happ, P., Cue La Rosa, L., Junior, J., Martins, J., Bressan, P., Gonçalves, W., Liesenberg, V., 2020. Applying fully convolutional architectures for semantic segmentation of a single tree species in urban environment on high resolution UAV optical imagery. **Sensors**. 20(2), 563. https://doi.org/10.3390/s20020563

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. **ISPRS J. Photogramm. Remote Sens**. 152, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

Marrs, J., Ni-Meister, W., 2019. Machine learning techniques for tree species classification using co-registered LiDAR and hyperspectral data. **Remote Sens.** 11(7), 819. https://doi.org/10.3390/rs11070819

Martins, R.C., Agostini-Costa T.S., Santelli P., Filgueiras T.S., 2016. Mauritia *flexuosa* Buriti. **In**: Espécies nativas da flora brasileira de valor econômico atual ou potencial: Plantas para o Futuro: Região Centro-Oeste / Ministério do Meio Ambiente. Secretaria de Biodiversidade; Roberto Fontes Vieira (Ed.). Julcéia Camillo (Ed.). Lidio Coradin (Ed.). – Brasília, DF: MMA.

Maschler, J., Atzberger, C., Immitzer, M., 2018. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. **Remote Sens**. 10(8), 1218. https://doi.org/10.3390/rs10081218

Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int. J. Remote Sens. 39, 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Miyoshi, G.T., Imai, N.N., Garcia Tommaselli, A.M., Antunes de Moraes, M.V., Honkavaara, E., 2020. Evaluation of hyperspectral multitemporal information to improve tree species identification in the highly diverse Atlantic forest. **Remote Sens.** 12(2), 244. https://doi.org/10.3390/rs12020244

Morales, G., Kemper, G., Sevillano, G., Arteaga, D., Ortega, I., Telles, J., 2018. Automatic segmentation of Mauritia *flexuosa* in unmanned aerial vehicle (UAV) imagery using deep learning. **Forests** 9. https://doi.org/10.3390/f9120736

Näsi, R., Honkavaara, E., Blomqvist, M., Lyytikäinen-Saarenmaa, P., Hakala, T., Viljanen, N., Kantola, T., Holopainen, M., 2018. Remote sensing of bark beetle damage in urban forests at individual tree level using a novel hyperspectral camera from UAV and aircraft. Urban Forest. **Urban Green**. 30, 72–83. https://doi.org/10.1016/j.ufug.2018.01.010

Näsi, R., Honkavaara, E., Lyytikäinen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., Viljanen, N., Kantola, T., Tanhuanpää, T., Holopainen, M., 2015. Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. **Remote Sens**. 7(11), 15467–15493. https://doi.org/10.3390/rs71115467

Navarro, A., Young, M., Allan, B., Carnell, P., Macreadie, P., Ierodiaconou, D., 2020. The application of Unmanned Aerial Vehicles (UAVs) to estimate above-ground biomass of mangrove ecosystems. **Remote Sens**. Environ. 242, 111747. https://doi.org/10.1016/j.rse.2020.111747

Nevalainen, O., Honkavaara, E., Tuominen, S., Viljanen, N., Hakala, T., Yu, X., Hyyppä, J., Saari, H., Pölönen, I., Imai, N.N., Tommaselli, A.M.G., 2017. Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging. **Remote Sens**. 9(3), 185. https://doi.org/10.3390/rs9030185

Nezami, S., Khoramshahi, E., Nevalainen, O., Pölönen, I., Honkavaara, E., 2020. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks. **Remote Sens**. 12(7), 1070. https://doi.org/10.3390/rs12071070

Osco, L.P., de Arruda, M. dos S., Marcato Junior, J., da Silva, N.B., Ramos, A.P.M., Moryia, É.A.S., Imai, N.N., Pereira, D.R., Creste, J.E., Matsubara, E.T., Li, J., Gonçalves, W.N., 2020. A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery. I**SPRS J. Photogramm. Remote Sens**. 160, 97–106. https://doi.org/10.1016/j.isprsjprs.2019.12.010

Osco, L.P., Marques Ramos, A.P., Roberto Pereira, D., Akemi Saito Moriya, É., Nobuhiro Imai, N., Takashi Matsubara, E., Estrabis, N., de Souza, M., Marcato Junior, J., Gonçalves, W.N., Li, J., Liesenberg, V., Creste, J.E., 2019. Predicting canopy nitrogen content in citrus-trees using random forest algorithm associated to spectral vegetation indices from UAV-imagery. **Remote Sens**. 11(24), 2925. https://doi.org/10.3390/rs11242925

Özcan, A.H., Hisar, D., Sayar, Y., Ünsalan, C., 2017. Tree crown detection and delineation in satellite images using probabilistic voting. Remote Sens. Lett. 8, 761–770. https://doi.org/10.1080/2150704X.2017.1322733Ozdarici-Ok, A., 2015. Automatic detection and delineation of citrus trees from VHR satellite

imagery. **Int. J. Remote Sens**. 36(17), 4275–4296. https://doi.org/10.1080/01431161.2015.1079663

Pacheco Santos, L.M., 2005. Nutritional and ecological aspects of M. *flexuosa* or aguaje (Mauritia *flexuosa* Linnaeus filius): A carotene-rich palm fruit from Latin America. **Ecol. Food Nutr**. 44, 345–358. https://doi.org/10.1080/03670240500253369

Pham, T.D., Yokoya, N., Bui, D.T., Yoshino, K., Friess, D.A., 2019. Remote sensing approaches for monitoring mangrove species, structure, and biomass: opportunities and challenges. **Remote Sens**. 11(3), 230. https://doi.org/10.3390/rs11030230

Raczko, E., Zagajewski, B., 2017. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. **Europ. J. Remote Sens**. 50, 144–154. https://doi.org/10.1080/22797254.2017.1299557

Redmon, J., Farhadi, A., 2018. YOLOv3: An incremental improvement. **arXiv**:1804.02767

Reis, B.P., Martins, S.V., Fernandes Filho, E.I., Sarcinelli, T.S., Gleriani, J.M., Marcatti, G.E., Leite, H.G., Halassy, M., 2019. Management recommendation generation for areas under forest restoration process through images obtained by UAV and LiDAR. **Remote Sens**. 11(13), 1508. https://doi.org/10.3390/rs11131508

Ren, S., He, K., Girshick, R.B., Sun, J., 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. **CoRR**, abs/1506.01497.

Saarinen, N., Vastaranta, M., Näsi, R., Rosnell, T., Hakala, T., Honkavaara, E., Wulder, M.A., Luoma, V., Tommaselli, A.M.G., Imai, N.N., Ribeiro, E.A.W., Guimaraes, R.B., Holopainen, M., Hyyppa, J., 2018. Assessing biodiversity in boreal forests with UAV-based photogrammetric point clouds and hyperspectral imaging. **Remote Sens**. 10(2), 338. https://doi.org/10.3390/rs10020338

Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y., Herrera, F., 2019. Detection of Fir Trees (Abies sibirica) Damaged by the Bark Beetle in Unmanned Aerial Vehicle Images with Deep Learning. **Remote Sens**. 11(6), 643. https://doi.org/10.3390/rs11060643

Santos, A.A., Marcato Junior, J., Araújo, M.S., Di Martini, D.R., Tetila, E.C., Siqueira, H.L., Aoki, C., Eltner, A., Matsubara, E.T., Pistori, H., Feitosa, R.Q., Liesenberg, V., Gonçalves, W.N., 2019. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs. **Sensors** (Switzerland) 19, 1–11. https://doi.org/10.3390/s19163595

Schmidhuber, J., 2015. Deep learning in neural networks: An overview. **Neural Networks** 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. **arXiv** 1409.1556.

Sothe, C., Almeida, C.M.D., Schimalski, M.B., Rosa, L.E.C.L., Castro, J.D.B., Feitosa, R.Q., Dalponte, M., Lima, C.L., Liesenberg, V., Miyoshi, G.T., Tommaselli, A.M.G., 2020. Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. **GISci. Remote Sens**. 57, 369–394. https://doi.org/10.1080/15481603.2020.1712102

Sylvain, J.-D., Drolet, G., Brown, N., 2019. Mapping dead forest cover using a deep convolutional neural network and digital aerial photography. **ISPRS J. Photogramm. Remote Sens**. 156, 14–26. https://doi.org/10.1016/j.isprsjprs.2019.07.010

Tuominen, S., Näsi, R., Honkavaara, E., Balazs, A., Hakala, T., Viljanen, N., Pölönen, I., Saari, H., Ojanen, H., 2018. Assessment of classifiers and remote sensing features of hyperspectral imagery and stereo-photogrammetric point clouds for recognition of tree species in a forest area of high species diversity. **Remote Sens.** 10(5), 714. https://doi.org/10.3390/rs10050714

Voss, M., Sugumaran, R., 2008. Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. **Sensors** 8, 3020–3036. https://doi.org/10.3390/s8053020

Weinstein, Ben.G., Marconi, S., Bohlman, S., Zare, A., White, E., 2019. Individual tree-crown detection in RGB imagery using self-supervised deep learning neural networks. **bioRxiv** 2019. https://doi.org/10.1101/532952

Xie, Z., Chen, Y., Lu, D., Li, G., Chen, E., 2019. Classification of land cover, forest, and tree species classes with ZiYuan-3 multispectral and stereo data. **Remote Sens.** 11(2), 164. https://doi.org/10.3390/rs11020164