PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
DOUTORADO EM MEIO AMBIENTE E
DESENVOLVIMENTO REGIONAL

DANIELLE ELIS GARCIA FURUYA

IDENTIFICAÇÃO DE ESPÉCIE ARBÓREA INVASORA EM IMAGENS PANORÂMICAS URBANAS POR APRENDIZAGEM PROFUNDA

Presidente Prudente – SP
2024

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
DOUTORADO EM MEIO AMBIENTE E
DESENVOLVIMENTO REGIONAL

DANIELLE ELIS GARCIA FURUYA

IDENTIFICAÇÃO DE ESPÉCIE ARBÓREA INVASORA EM IMAGENS PANORÂMICAS URBANAS POR APRENDIZAGEM PROFUNDA

Tese apresentada à Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade do Oeste Paulista, como parte dos requisitos para obtenção do título de Doutora em Meio Ambiente e Desenvolvimento Regional.
Área de Concentração: Ciências Ambientais.

Orientadora: Ana Paula Marques Ramos
Co-orientador: Lucas Prado Osco
Paulo Antônio da Silva (UNOESTE).

Colaboradores Externos: David L. Des Marais (Massachusetts Institute of Technology – MIT – Estados Unidos)
Wesley Nunes Gonçalves (UFMS).

Presidente Prudente - SP
2024

581.7 Furuya, Danielle Elis Garcia.
F992i Identificação de espécie arbórea invasora em imagens panorâmicas urbanas por aprendizagem profunda. / Danielle Elis Garcia Furuya. – Presidente Prudente, 2024.
116 f.: il.

Tese (Doutorado em Meio Ambiente e Desenvolvimento Regional) - Universidade do Oeste Paulista – Unoeste, Presidente Prudente, SP, 2024.
Bibliografia.
Orientador: Ana Paula Marques Ramos.

1. Leucaena leucocephala. 2. SAM. 3. YOLOv8. 4. DETR. 5. Imagens panorâmicas. I. Título.

Catalogação na Fonte: Maria Letícia Silva Vila Real – CRB 8/10699

DANIELLE ELIS GARCIA FURUYA

IDENTIFICAÇÃO DE ESPÉCIE ARBÓREA INVASORA EM IMAGENS PANORÂMICAS URBANAS POR APRENDIZAGEM PROFUNDA

Tese apresentada à Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade do Oeste Paulista, como parte dos requisitos para obtenção do título de Doutora em Meio Ambiente e Desenvolvimento Regional – Área de concentração: Ciências Ambientais.

Presidente Prudente, 21 de março de 2024.

BANCA EXAMINADORA

____
Prof. Dra. Ana Paula Marques Ramos
Universidade Estadual Paulista –
UnespPresidente Prudente - SP

____
Prof. Dr. Paulo Antonio da Silva
Universidade do Oeste Paulista –
UnoestePresidente Prudente - SP

____
Prof. Dr. José Marcato Júnior
Universidade Federal de Mato
Grosso do SulCampo Grande - MS

____
Prof. Dr. Jonathan de Andrade Silva
Universidade Federal de Mato
Grosso do SulCampo Grande - MS

____
Prof. Dr. Renan Furlan de Oliveira
Universidade Federal de Santa
CatarinaSanta Catarina - SC

RESUMO

Identificação de espécie arbórea invasora em imagens panorâmicas urbanas por aprendizagem profunda

O mapeamento da vegetação arbórea contribui com o planejamento urbano e facilita a gestão das áreas verdes nas cidades. Identificar espécies arbóreas auxilia no diagnóstico ambiental e permite verificar se as espécies presentes em áreas urbanas contribuem com o equilíbrio e fornecem os serviços ecossistêmicos ideais. As áreas verdes nem sempre são compostas com a quantidade de espécies nativas recomendadas, e podem apresentar espécies arbóreas exóticas capazes de prejudicar o ambiente urbano, pois certas espécies passam a se tornar invasoras. Nesse contexto, a espécie arbórea leucena (*Leucaena leucocephala*) está presente com frequência em diversas cidades e corresponde a uma espécie exótica invasora em diversos países, incluindo o Brasil. Os dados de sensoriamento remoto processados por algoritmos de inteligência artificial (especialmente de aprendizagem profunda), podem consistir em uma estratégia eficiente e eficaz para identificar Leucenas em ambientes urbanos. No entanto, essa é uma lacuna persistente na literatura. Neste trabalho, busca-se explorar o potencial de algortimos de deep learning (detecção e segmentação semântica) para identificar a espécie arbórea Leucena em regiões urbanas a partir de imagens panorâmicas do Google Street View e de celular. Essas imagens estão sendo exploradas para a identificação das Leucenas tanto na perspectiva de segmentação semântica, quanto de deteção de objetos. O modelo SAM (Segment Anything Model), lançado em abril de 2023, foi testado nessas imagens para, inicialmente, diferenciar vegetação arbórea sem levar em consideração a espécie. Os resultados parciais indicam capacidade do SAM nessa tarefa, alcançando um F1-score de 99%. Nos testes com o SAM para e espécie leucena, o modelo mostrou ser mais vantajoso em imagens do Street View quando comparado a imagens de celular, com uma média de acurácia de 92% e 63% respectivamente. Na perspectiva de deteção de objetos, foram realizados testes com as redes YOLOv8 e DETR. Lançada em janeiro de 2023, a YOLOv8 é uma das opções mais recentes para a tarefa de detecção de objetos em imagens, já a DETR corresponde a uma rede baseada em transformers e tem potencial para diversas tarefas. Os resultados indicam uma capacidade promissora da YOLOv8 na tarefa proposta de detecção da espécie leucena em imagens panorâmicas. Além dos testes, uma revisão cienciométrica foi construída para averiguar como a problemática de mapeamento arbóreo geral, e por espécie, tem sido tratada a partir de dados de sensoriamento remoto e aprendizagem de máquina (rasa e profunda). Os testes realizados reforçam a originalidade do presente trabalho, mostrando uma carência de estudos que explorem esse tipo de dados com técnicas avançadas de deep learning, para identificar nas cidades a presença de uma das 100 piores espécies invasoras do mundo, a Leucena.

Palavras-chave: Leucaena leucocephala; SAM; YOLOv8; DETR; Imagens panorâmicas.

ABSTRACT

Identification of invasive tree species in urban panoramic images using deep learning

Mapping tree vegetation contributes to urban planning and facilitates the management of green areas in cities. Identifying tree species helps in environmental diagnosis and allows us to verify whether the species present in urban areas contribute to the balance and provide ideal ecosystem services. Green areas are not always composed of the recommended number of native species, and may contain exotic tree species capable of harming the urban environment, as certain species become invasive. In this context, the leucaena tree species (Leucaena leucocephala) is frequently present in several cities and corresponds to an invasive exotic species in several countries, including Brazil. Remote sensing data processed by artificial intelligence algorithms (especially deep learning) can be an efficient and effective strategy for identifying Leucaena in urban environments. However, this is a persistent gap in the literature. In this work, we seek to explore the potential of deep learning algorithms (detection and semantic segmentation) to identify the Leucaena tree species in urban regions using panoramic images from Google Street View and cell phones. These images are being explored to identify Leucenes both from the perspective of semantic segmentation and object detection. The SAM model (Segment Anything Model), launched in April 2023, was tested on these images to initially differentiate tree vegetation without taking the species into account. The partial results indicate SAM's ability in this task, reaching an F1-score of 99%. In tests with SAM for the Leucaena species, the model proved to be more advantageous in Street View images when compared to cell phone images, with an average accuracy of 92% and 63% respectively. From an object detection perspective, tests were carried out with the YOLOv8 and DeTR networks. Launched in January 2023, YOLOv8 is one of the most recent options for the task of object detection in images, while DeTR corresponds to a transformer-based network and has the potential for several tasks. The results indicate a promising capacity of YOLOv8 in the proposed task of detecting the Leucaena species in panoramic images. In addition to the tests, a scientometric review was constructed to investigate how the problem of tree mapping in general, and by species, has been addressed using remote sensing data and machine learning (shallow and deep). The tests carried out reinforce the originality of this work, showing a lack of studies that explore this type of data with advanced deep learning techniques, to identify the presence of one of the 100 worst invasive species in the world, the Leucaena.

Keywords: Leucaena leucocephala; SAM; YOLOv8; DETR; Panoramic images.

LISTA DE TABELAS

LISTA DE FIGURAS

SUMÁRIO

1 CONSIDERAÇÕES INICIAIS

Graduada em Engenharia Ambiental e Sanitária pela Universidade do Oeste Paulista (2018). Especialista em Ciência de Dados e Big Data pela Pontifícia Universidade Católica de Minas Gerais - PUC Minas (2023). Mestre em Meio Ambiente e Desenvolvimento Regional pelo Programa de Pós Graduação da Universidade do Oeste Paulista - PPGMADRE - UNOESTE (2021). Durante o mestrado fui bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e desenvolvi pesquisas na área de ciências ambientais, a partir do uso de sensoriamento remoto e inteligência artificial. Na dissertação de mestrado, atuei na avaliação de algoritmos de machine learning para identificar vegetação em zonas ripárias com imagens do satélite Sentinel-2. Fui integrante dos seguintes grupos de pesquisa durante o doutorado: Núcleo de Estudos Ambientais – NEAmb (antigo NEAGEO) desde 2019 e Geomática e Inteligência Artificial Aplicada a Análise Ambiental (Gi3A) desde 2019. Entrei no doutorado em março de 2021, também com bolsa CAPES, e tenho trabalhado com métodos de deep learning para identificar a espécie arbórea *Leucaena leucocephala* (Leucena) em áreas urbanas com imagens panorâmicas do Street View e de celular. Dentro do PPGMADRE, me mantive na linha de pesquisa 2 - Planejamento Ambiental e Desenvolvimento Regional. Realizei o período de Doutorado Sanduíche com bolsa CAPES (novembro de 2022 a abril de 2023) no Massachusetts Institute of Technology (MIT) em Cambridge nos Estados Unidos. Durante esse período tive a supervisão do Professor David L. Des Marais no departamento de Engenharia Civil e Ambiental.

Este documento está organizado em quatro capítulos principais. Primeiro é apresentado a introdução geral, com a problemática e objetivos da presente pesquisa. O primeiro capítulo é composto por um levantamento científico sobre a literatura que compõe a pesquisa em formato de manuscrito. Neste capítulo, o levantamento é feito por uma análise cienciométrica e revisão de estudos que usaram machine learning/deep learning para mapear vegetação em área urbana com utilização de sensoriamento remoto. No segundo capítulo são apresentados os testes com o Segment Anything Model (SAM) para segmentar vegetação arbórea em geral com imagens panorâmicas. O terceiro capítulo traz os resultados do SAM para a espécie leucena com imagens do Street View e de celular. Em seguida, no quarto capítulo é apresentado uma abordagem para detectar a espécie leucena em área urbana em imagens do Street View e de celular com as redes YOLOv8 e DETR. Por fim, é apresentado uma conclusão geral dos resultados obtidos na tese.

2 INTRODUÇÃO E JUSTIFICATIVA

A vegetação urbana fornece serviços ecossistêmicos que influenciam a saúde e o bem-estar das pessoas. As sociedades devem considerar a melhor forma de garantir que todos os residentes urbanos possam se beneficiar desses serviços ecossistêmicos (Nesbitt *et al.,* 2019). O monitoramento da vegetação arbórea urbana é fundamental para auxiliar no planejamento das cidades e contribuir com um ambiente equilibrado diante dos impactos ambientais que ocorrem no meio urbano. A espécie arbórea Leucena (*Leucaena leucocephala*) é uma espécie originária do México e da América Central que tem sido identificada como uma espécie exótica e invasora principalmente em áreas degradadas nas regiões tropicais e subtropicais. Essa espécie se tornou um obstáculo para a gestão de florestas locais, pois inibe o desenvolvimento de plantas nativas (Bartholomeu; de Sousa; Brazolin, 2020).

A Leucena foi listada como uma das 100 piores espécies arbóreas exóticas e invasoras do mundo (Lowe *et al.*, 2000; Bodevan; Coelho; Marques, 2016). Realizar a identificação dessa espécie invasora em áreas urbanas é necessário para uma melhor gestão de áreas verdes, pois auxilia no equilíbrio entre o ambiente natural e artificial. Uma estratégia para identificar vegetação arbórea em amplas áreas geográficas é a utilização de dados de sensoriamento remoto (Demattê *et al.*, 2017).

O sensoriamento remoto tem sido aplicado em estudos que buscam identificar características de vegetação. A extração da vegetação a partir de imagens de sensoriamento remoto é o processo de extração de características com base em elementos de interpretação, como a cor da imagem, textura, tom, padrão e informações de associação (Xie *et al.*, 2008). Diversos métodos foram desenvolvidos para permitir a extração destas informações como a utilização de bandas espectrais, classificação supervisionada e não-supervisionada.

A escolha das imagens panorâmicas do Google Street View e de celular para o mapeamento arbóreo urbano das Leucenas, em detrimento das imagens aéreas ou orbitais, baseia-se na necessidade de identificar detalhes específicos que são fundamentais para distinguir essa espécie. As Leucenas, que possuem características físicas semelhantes a outras espécies de sua família, apresentam vagens distintas, sendo este um dos aspectos visuais mais marcantes para a sua diferenciação, inclusive em relação a espécies correlatas. No entanto, tais características são perceptíveis apenas em imagens de alta resolução e vistas panorâmicas, como as fornecidas pelo Google Street View. A vantagem dessas imagens consiste na capacidade de capturar detalhes finos a nível do solo, tais como a textura das folhas, os padrões de ramificação e, crucialmente, as vagens características das Leucenas. Além disso, as imagens do Street View oferecem uma perspectiva próxima à visão humana, ideal para analisar ambientes urbanos onde as árvores podem estar inseridas. A facilidade de acesso, a abrangente cobertura e as atualizações regulares do Google Street View facilitam a coleta de dados em diversas áreas urbanas (Biljecki; Ito, 2021), enquanto a integração com outras fontes de dados geoespaciais possibilita uma análise mais abrangente e multidimensional do papel das Leucenas no contexto urbano.

A área de Inteligência Artificial, mais especificamente de aprendizagem de máquina (Machine Learning – ML), a qual inclui a aprendizagem profunda (Deep Learning - DL), tem sido fortemente integrada à área de Geomática nos últimos anos para o desenvolvimento de abordagens robustas para realizar a análise digital de imagens. As análises incluem tarefas de classificação, detecção de objetos, segmentação semântica e, mais recentemente, segmentação de instâncias em imagens orbitais, aéreas ou terrestres (Minaee *et al.,* 2020).

A aprendizagem profunda corresponde a uma subclasse da aprendizagem de máquina e possui redes neurais com mais camadas que permitem maiores capacidades de aprendizagem (Kamilaris; Prenafeta-Boldú, 2018). Esses métodos podem apoiar estudos de planejamento e monitoramento ambiental e têm potencial para auxiliar na identificação de espécies arbóreas.

O objetivo deste projeto é explorar o potencial de redes de aprendizagem profunda (detecção e segmentação semântica) para identificar a espécie arbórea Leucena em regiões urbanas a partir de imagens panorâmicas do Google Street View e de celular. Como objetivos específicos, tem-se: (1) Identificar as abordagens de mapeamento de vegetação em áreas urbanas já realizadas, com ênfase no uso de sensoriamento remoto e técnicas de inteligência artificial; (2) Verificar a eficácia do Segment Anything Model (SAM) na segmentação de árvores em geral e da espécie Leucena em imagens panorâmicas disponíveis no Google Street View e de celular; (3) Avaliar a capacidade das redes YOLOv8 e DETR para detectar especificamente a espécie arbórea leucena em imagens panorâmicas do Google Street View e de celular.

CAPÍTULO 1 - RECENT PROGRESS IN MACHINE LEARNING AND DEEP LEARNING WITH REMOTE SENSING FOR MAPPING VEGETATION AND TREE SPECIES IN URBAN AREAS.

Abstract: Understanding urban vegetation is fundamental to creating healthier and more sustainable environments that benefit the environment and society's quality of life. Remote sensing and artificial intelligence are being used in environmental tasks to optimize procedures and decision-making. The combination of different techniques can bring more accurate results and facilitate environmental diagnosis and monitoring. This work presents a comprehensive analysis on the topic of urban vegetation, integrating concepts from remote sensing, machine learning and deep learning. 221 articles available on the Web of Science were analyzed, highlighting emerging trends in this area. We present a list of 149 tree species already mapped in urban environments, in addition to identifying the types of remote sensing data used in these studies. We highlight other important points such as the use of different neural networks, evaluation metrics, countries with the highest contribution and other relevant information. By analyzing all this information, we aim to offer a comprehensive view of the current state of knowledge in this dynamic field, providing valuable insights for researchers, urban planners and professionals involved in the sustainable management of urban areas.

Keywords: Remote sensing; Machine Learning; Deep Learning; Urban Vegetation Mapping; Tree Species.

Resumo: Compreender a vegetação urbana é fundamental para criar ambientes mais saudáveis e sustentáveis que beneficiem o meio ambiente e a qualidade de vida da sociedade. O sensoriamento remoto e a inteligência artificial são utilizados em tarefas ambientais para otimizar procedimentos e tomadas de decisão. A combinação de diferentes técnicas pode trazer resultados mais precisos e facilitar o diagnóstico e monitoramento ambiental. Este trabalho apresenta uma análise abrangente sobre o tema vegetação urbana, integrando conceitos de sensoriamento remoto, aprendizado de máquina e aprendizado profundo. Foram analisados 221 artigos disponíveis na Web of Science, destacando tendências emergentes nesta área. Apresentamos uma lista de 149 espécies arbóreas já mapeadas em ambientes urbanos, além de identificar os tipos de dados de sensoriamento remoto utilizados nesses estudos. Destacamos outros pontos importantes como a utilização de diferentes redes neurais, métricas de avaliação, países com maior contribuição e outras informações relevantes. Ao analisar toda esta informação, pretendemos oferecer uma visão abrangente do estado atual do conhecimento neste campo dinâmico, fornecendo informações valiosas para investigadores, urbanistas e profissionais envolvidos na gestão sustentável de áreas urbanas.

Palavras-chave: Sensoriamento remoto; Aprendizado de máquina, Aprendizado profundo, Mapeamento de vegetação urbana, Espécies arbóreas.

1 Introduction

Studies and analysis in urban areas need robust and efficient techniques to perform accurate environmental mapping and monitoring. The vegetation in urban areas is extremely important because green areas generate many benefits for society, such as providing ecosystem services and reducing pollution and temperature (Yang *et al.*, 2022b). The use of remote sensing increases every year due to the ease of obtaining free-available images and orbital data from different locations and analyzing large areas. Furthermore, it is possible to capture spatial data with higher temporal, spectral and spatial resolution compared to decades ago (Neyns; Canters, 2022).

Studies have shown an increase in the application of remote sensing in urban areas with different satellites and techniques (Shahtahmassebi *et al.*, 2021; Yin *et al.*, 2021; García-Pardo *et al.*, 2022). Applications in the field of urban vegetation mapping are expected to grow rapidly in the coming years due to the increasing number of data sources and awareness of the role of urban vegetation as a provider of multiple ecosystem services (Neyns; Canters, 2022).

The simultaneous application of remote sensing and artificial intelligence has stood out as an innovative and promising area of research in environmental studies. Studies have shown that the use of machine learning algorithms like Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), XGBoost and K-nearest neighbors (KNN) in vegetation mapping has good accuracy (Gašparović; Dobrinić, 2020; Hartling; Sagan; Maimaitijiang, 2021; Marques Ramos *et al.*, 2020; Nguyen *et al.*, 2019; Li; Yuan; Dong, 2021; Furuya *et al.*, 2020; Silva, 2021; Osco *et al.*, 2022;

Furuya *et al.*, 2023).

The literature has shown that Deep Learning (DL) outperforms shallow machine learning methods and demonstrated promising results (Paoletti *et al.*, 2019; Yuan *et al.*, 2020; Kattenborn *et al.*, 2021; Osco *et al.*, 2021). Deep learning is a potential approach that focuses on large-scale and deep artificial neural networks, being responsible for optimizing and returning better learning patterns (Yuan *et al.*, 2020; Osco *et al.*, 2021). Moreover, DL is characterized by a significantly increased number of successively connected neural layers (Kattenborn *et al.*, 2021).

Deep Learning has offered robust and intelligent methods to improve the mapping of the Earth's surface and has been introduced into environmental remote sensing and applied in many aspects, including land cover mapping, environmental parameter retrieval, data fusion and downscaling, tree mapping, and information construction and prediction (Osco *et al.*, 2021; Yuan *et al.*, 2020; de Araújo Carvalho *et al.*, 2022; Arce *et al.*, 2021).

Some studies have analyzed the performance of different algorithms in environmental studies. Yuan *et al*. (2020) made a review of Deep learning in environmental remote sensing. The authors also analyzed the progress on DL in remote sensing of ten more environmental parameters (atmosphere, vegetation, hydrology, air and land surface temperature, evapotranspiration, solar radiation, and ocean color) and discussed new insights on combining DL and physical/geographical laws.

Another study from Kattenborn *et al.* (2021) introduced the principles of Convolution Neural Networks (CNN) and explained why it is suitable for vegetation remote sensing. The authors showed the current trends and developments, including considerations about spectral resolution, spatial resolution, different sensor types, modes of reference data generation, sources of existing reference data, as well as CNN approaches and architectures. This literature review showed that CNN could be applied to various problems, including the detection of individual plants or the pixel-wise segmentation of vegetation classes. All advances in the area of remote sensing and artificial intelligence contribute to the optimization of environmental projects, including urban areas.

In this paper, the main objective is to perform a comprehensive analysis that addresses the intersection of topics related to urban tree species, urban vegetation, machine learning, deep learning, and remote sensing. The research involves an extensive review of the literature, specifically examining studies that focus on urban vegetation mapping including tree species. Additionally, the analysis includes a detailed assessment of the employment of machine learning, deep learning, or a combination of both techniques in the reviewed articles. We provide a comprehensive and up-to-date look at how these fields intertwine, highlighting the critical role of remote sensing technology in analyzing and understanding urban vegetation, especially in the context of tree species. This paper brings the following contributions:

1-Scientometric analysis, showing the characteristics of studies published on the topic of urban vegetation with remote sensing and artificial intelligence;

2- List of 149 tree species that have already been mapped using images and ML and/or DL algorithms;

3- Deep learning networks and satellites that have already been used;

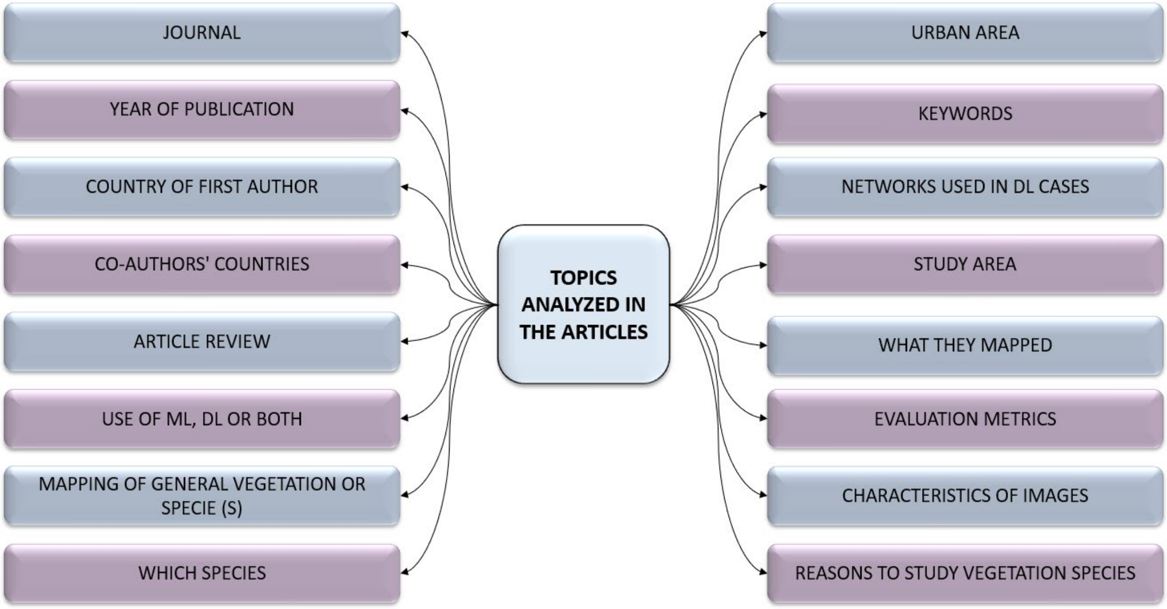
4- Other data that provide information on the literature already published on combining urban vegetation mapping with images and algorithms.

2 Method

The analysis of this study was conducted with data collected from studies available on the Web of Science. We combined the topics of interest (urban vegetation, urban tree species, remote sensing, machine learning and deep learning) to analyze the available literature and check progress in research in this area. The search formula of the advanced method selected was: TS = ((“machine learning” OR “deep learning” OR “convolution neural network” OR recurrent neural network OR vision transformers) AND (urban forest OR urban trees mapping OR urban arboreous vegetation OR urban tree species OR Leucaena leucocephala OR leucaena) AND (remote sensing OR multispectral image OR aerial image OR orbital image OR hyperspectral imagery OR “Google Street View imagery” OR panoramic image)).

Web of Science filtered 515 studies based on the defined search formula. Of this total, only 221 studies correspond to articles on the topic of interest published in scientific journals until February 2024. From these articles, information of interest for this study was collected and analyzed as shown in the Figure 1. These 221 articles specifically deal with mapping urban vegetation, including trees, using remote sensing data and artificial intelligence (machine and deep learning).

Figure 1- Topics of interest analyzed in the 221 articles found on the Web of Science to verify what has been studied and the types of techniques applied in the area in order to map urban vegetation with remote sensing and artificial intelligence.



Fonte: Author (2024).

The topics (Figure 1) were selected to understand the dynamics and trends of the area. We conducted quantitative analyses to examine the annual publication trends, author distribution by country, and identification of journals with the highest publication frequency. This allowed us to assess scientific productivity within the topics of urban vegetation, remote sensing, and artificial intelligence. We also analyzed information related to environmental issues, showing a comprehensive list of tree species already mapped in urban areas using images and ML/DL techniques. Additionally, we identified the primary objectives of the analyzed studies, which can include Land Use and Land Cover (LULC), object detection, vegetation in general, or the mapping of tree species.

Throughout this study, we underscore the justifications provided in the articles for researching tree species, showing the various advantages and contributions that optimize and assist urban planning. From a technical perspective, we also analyzed the types of satellites used, the characteristics of the images, the evaluation metrics found in the studies and the frequency of the neural networks used, observing the comparison of the use of machine learning and deep learning. Of the total of 221 articles, ten correspond to review articles. For this reason, they were not included in some statistics.

We created a word cloud with the keywords of all 221 articles analyzed to check the terms most used in studies and the trend in searches (Figure 2). A total of 1057 keywords were found in the articles of interest.

Figure 2- Word Cloud created with all the 1057 keywords found in 221 studies showing the terms that had the most highlights.

Fonte: Author (2024).

In the word cloud created with the keywords, it is possible to observe that some examples of the most used words in the studies are: machine learning, deep learning, remote sensing, image, urban and vegetation. Another observation is the following words: classification, detection, random forest, support vector machine and neural network, which frequently appear in the keywords of the articles and refer to the techniques and algorithms used in the methodology.

Considering all the articles analyzed, only ten correspond to review articles. Table 1 shows information regarding the ten articles found.

Table 1- Information from the ten review articles found in Web of Science considering the filter of interest for the present study.

	Title	Authors/Year	What they analyzed
1	Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review	Kuras *et al.* (2021)	Features of airborne hyperspectral images, lidar and their fusion for urban land cover classification. Machine learning/deep learning algorithms for classifying individual urban classes were also reviewed.
2	Remotely Sensed Tree Characteriza in Urban Areas: A Review	Velasquez-Camacho *et al.* (2021)	A systematic review of 48 articles published between 2016 and 2020 on remote sensing techniques, data processing and analysis methods for tree classification and segmentation, with a focus on urban vegetation.
3	A Survey of Mobile Laser Scanning Applications and Key Techniques over Urban Areas	Wang *et al.* (2019)	A survey on urban applications and key techniques based on mobile laser scanning (MLS) point clouds. The authors showed the applications of MLS in urban areas, including the use of deep learning and vegetation mapping and inventory.
4	The Digital Forest: Mapping a Decade of Knowledge on Technological Applications for Forest Ecosystems	Nitoslawski *et al.* (2021)	Identification of hardware and software technologies and techniques, methodologies used, as well as challenges and limitations in forest management applications.
5	Mapping of Urban Vegetation with High-Resolution Remote Sensing: A Review	Neyns and Canters (2022)	An overview of the main approaches to mapping urban vegetation from high-resolution remote sensing data. The authors analyze vegetation typology, remote sensing data and mapping approaches.
6	A Review of General Methods for Quantifying and Estimating Urban	Yang *et al.* (2022a)	Assessment of recent research methods in the urban forests, with a

	Trees and Biomass		focus on estimating urban forest biomass and detecting individual tree characteristics.
7	Remote sensing for the assessment of ecosystem services provided by urban vegetation: A review of the methods applied	García-Pardo *et al.* (2022)	A review of methods and applications in remote sensing for the assessment of ecosystem services provided by vegetation in cities.
8	Change detection of urban objects using 3D point clouds: A review	Stilla and Xu (2023)	A review of point-cloud-based 3D change detection for urban objects. In the study, the authors identified techniques in change detection and contextualized the up-to-date uses of point clouds, including for vegetation surveys.
9	A systematic review of urban green space research over the last 30 years: A bibliometric analysis	Farkas *et al.* (2023)	A review of urban green spaces considering the methods, countries and other statistics.
10	Assessing urban greenery by harvesting street view data: A review	Lu *et al.* (2023)	A review of studies that used Street View images to assess urban greenery. The authors analyzed studies between 2010 and 2022.

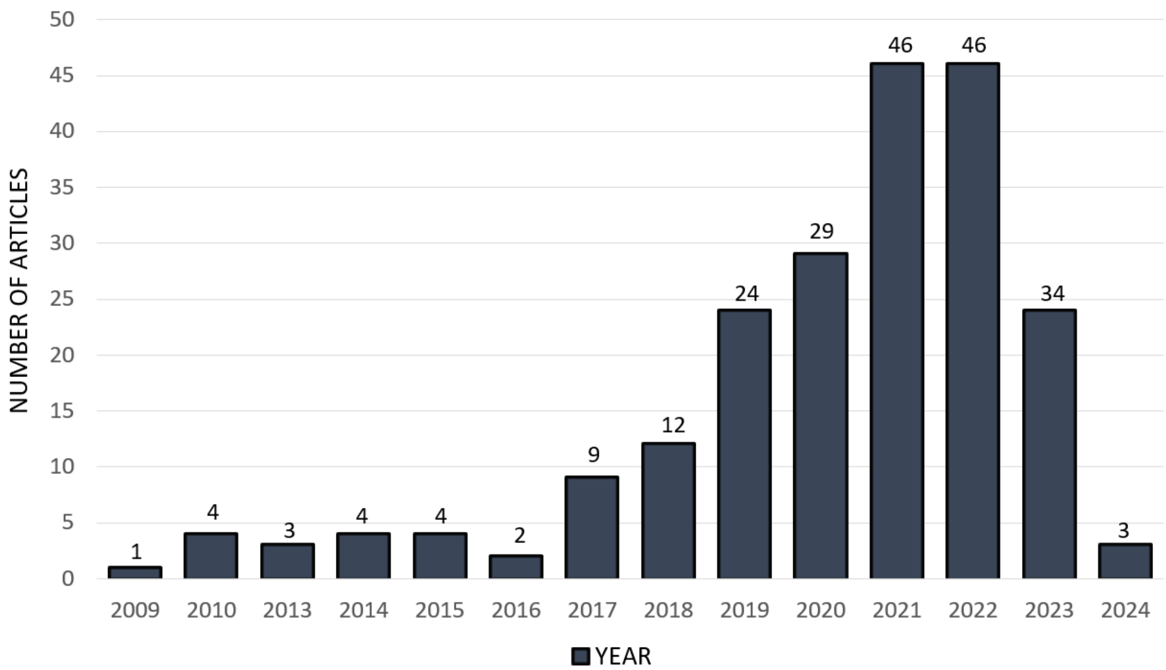
Although review articles show data relating to vegetation in urban areas, no studies were found that mention which tree species are already mapped with remote sensing and artificial intelligence. Furthermore, the last review article corresponds to May, 2023 and as of the date of analysis, more than thirty new articles have already been found in the interest filter. In this context, the studies filtered in Web of Science were analyzed based on the topics in Figure 1.

3 Results and Discussion

3.1. Temporal Analysis and Influential Journals of Publications

Figure 3 shows the distribution of articles by year of publication. Of the 221 articles analyzed, the first record on the topic of urban vegetation using artificial intelligence and remote sensing data is from 2009 titled "Prediction of street tree morphological parameters using artificial neural networks" (Jutras; Prasher; Mehuys, 2009). This study addressed the importance of urban tree management and proposed a method to predict diameter at breast height (DBH), annual DBH increment and crown volume with measurements extracted from aerial LIDAR data. The authors tested seven models of artificial neural networks (ANN) on different tree species, achieving robust results with prediction coefficients above 70%. They concluded that it is feasible to predict features with limited aerial LIDAR information, allowing for the reduction of traditional fieldwork and costs. This study also suggested optimized options for urban tree inventories, using traditional methods or aerial data acquisition, with an average prediction performance of 91%.

Figure 3- Temporal distribution of articles from 2009 to February 2024.



Fonte: Author (2024).

With the quantitative data of articles published per year, we can see that the most recent years stand out with a greater number of studies. Five years ago, in 2018, 12 articles were published, and in the following years, the quantity already exceeded 20 per year. This increase can be explained by the development and application of Artificial Intelligence techniques. Machine learning (ML) and deep learning (DL) tools gained significant popularity and attention in the last decade due to the advances in computational systems in terms of power and performance (Talib *et al.*, 2021; Wang *et al.*, 2020).

In relation to scientific journals, we found studies published in 66 journals. Table 2 shows the journals with more articles according to the interest filter. Most of the analyzed journals have one, two or three articles.

Table 2- Journals with the highest number of articles in the filter of interest (urban vegetation, urban tree species, remote sensing, machine learning and deep learning).

Journal	Number of Articles
Remote Sensing	70
Sensors	12
Remote Sensing of Environment	9
ISPRS International Journal of Geo-information	8
ISPRS Journal of Photogrammetry and Remote Sensing	7
International Journal of Remote Sensing	7
Urban Forestry & Urban Greening	6
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing	5
Journal of Applied Remote Sensing	4
IEEE Transactions on Geoscience and Remote Sensing	4
Computers and Electronics in Agriculture	4
IEEE Access	4

Remote Sensing journal has the most articles published on urban vegetation and the use of ML/DL, with a total of 70 articles out of the 221 analyzed. This journal has a current Impact Factor of 5.349 and a 5-year Impact Factor of 5.786.

3.2. Analysis of Countries in relation to Authors, Co-authors and Study Areas

The selected articles have authors from different countries. We can find authors from two or more countries in the same article. For better data analysis, we separated the countries of the first authors and the countries with the highest number of co-authors (Table 3). Of the 221 articles in this study, we have the first authors from

43 different countries. Most countries have few articles with first authors ranging from one to four articles.

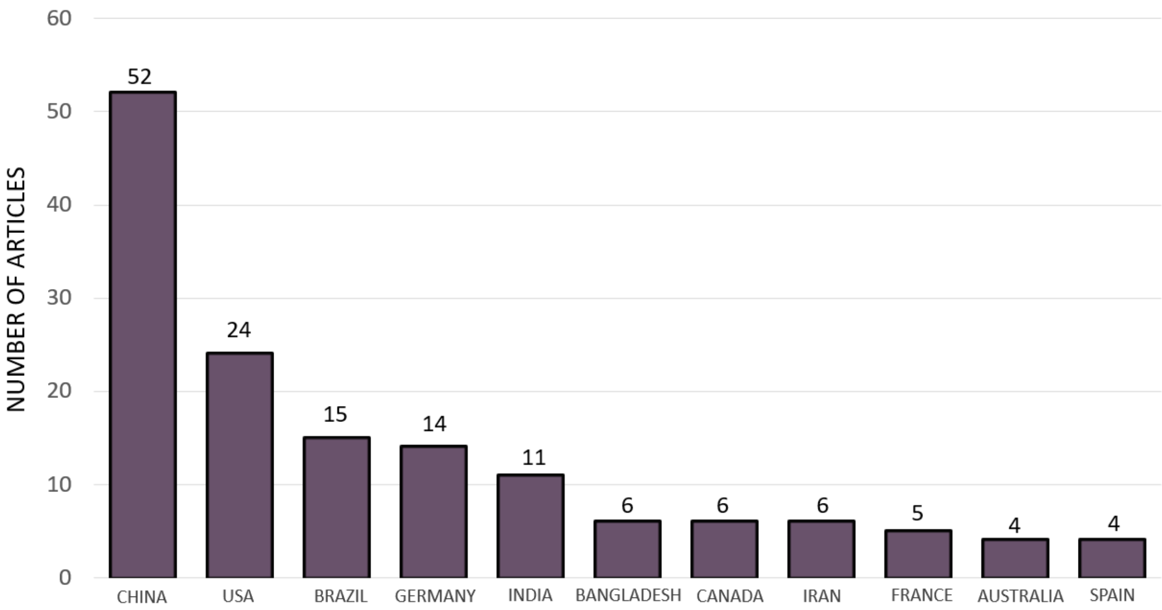
Regarding the countries with the most participation of co-authors, we can see in Table 3 the five countries with the highest number of articles. We found a total of 56 countries in total with the participation of co-authors. It is important to highlight that the presence of the country in the article was considered in this evaluation and not the number of co-authors by nationality.

Table 3- Countries with the highest participation of first authors and co-authors of the 221 articles analyzed.

Country with First Author	**Number of Articles**
China	62
USA	25
Germany	17
Brazil	16
India	11
Country with Co-Authors	**Number of Articles**
China	71
USA	38
Germany	24
Brazil	18
India	16
Canada	16

Regarding the areas of study chosen in the articles, we found 49 countries that already had some area as a location for tests conducted by the authors. Figure 4 shows the countries with the highest frequency of study areas. China has the highest quantity of articles with a total of 52 articles.

Figure 4- Countries with the highest occurrence of Study Areas.



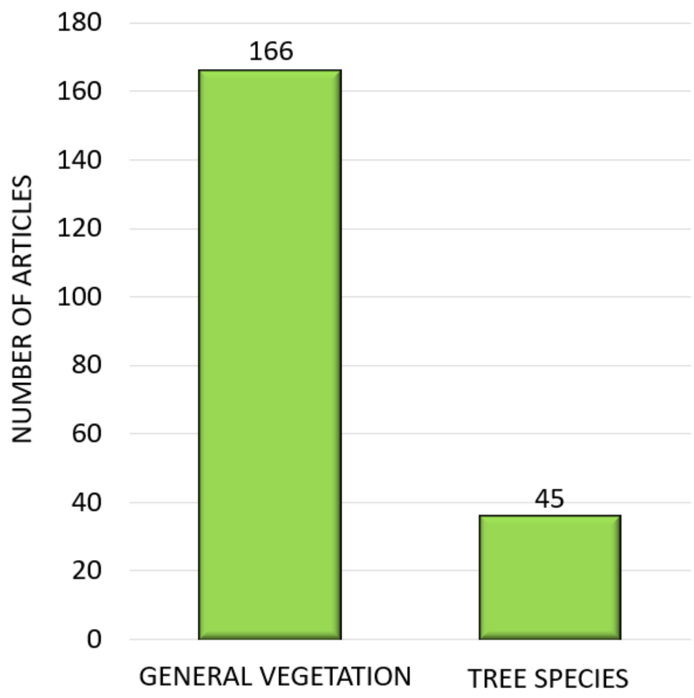
Fonte: Author (2024).

3.3. Urban Vegetation Analysis

The next topic of interest was to analyze whether the articles mapped vegetation in general or tree species. The ten review articles were not considered in this part of the analysis.

Several articles have mapped different levels of vegetation, for example, grasses, sparse vegetation, crops, forest vegetation and others. Even though the authors mapped different levels of vegetation, at this stage we consider it as vegetation in general, as the interest in this section is to analyze which tree species have already been mapped and studied. Figure 5 shows, of the total of 211 articles, how many articles analyzed vegetation in general and how many analyzed one or more tree species.

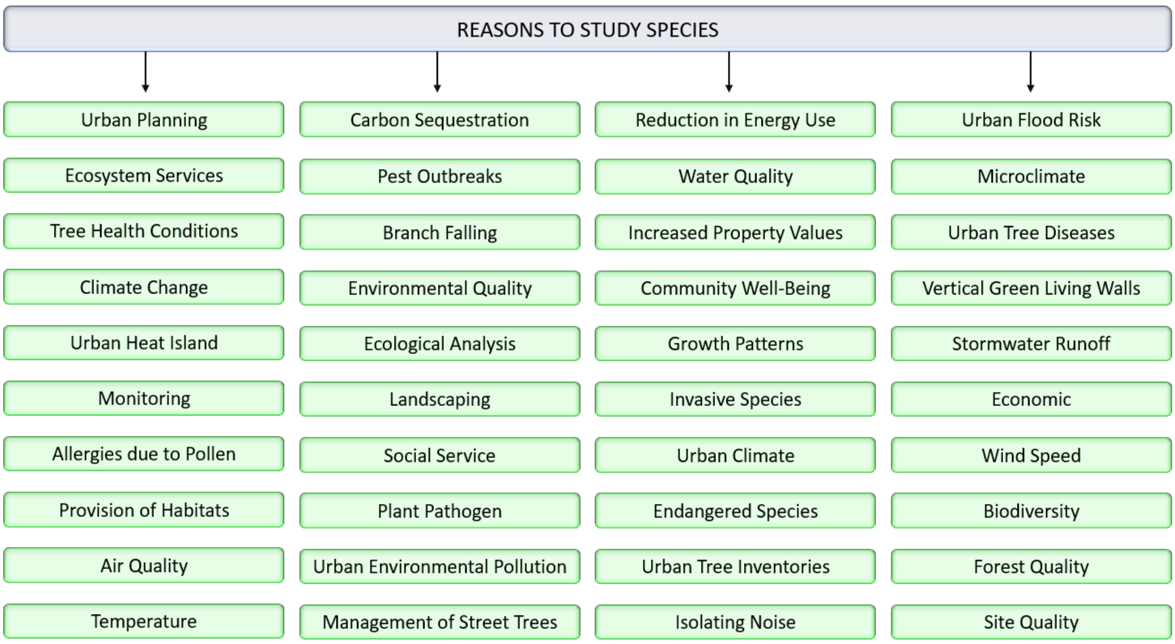
Figure 5- Quantitative analysis of articles that mapped tree species and vegetation in general.



Fonte: Author (2024).

The next step was to identify the reasons for studying tree species mentioned in the studies, as shown in Figure 6. We identified 40 different reasons according to the authors.

Figure 6- Justifications for studying tree species according to the authors.



Fonte: Author (2024).

Studies Studies that aimed to analyze some type of vegetation without considering species showed advantages and good accuracy of algorithms combined with images (Gašparović; Dobrinić, 2020; Xu *et al.*, 2020; Wang; Shao; Kennedy, 2014; Ayhan *et al.*, 2020; Hoang Nhat-Duc, 2021; Sun *et al.*, 2021; Timilsina; Aryal; Kirkpatrick, 2020; Shi *et al.*, 2022; Huerta *et al.*, 2021; Wang *et al.*, 2022; Wang, 2022; Schmohl; Narváez Vallejo; Soergel, 2022; Zhang *et al.*, 2022).

Gašparović e Dobrinić (2020) performed a multitemporal analysis of vegetation using Sentinel 1 images in an urban area and compared six algorithms for this task. The authors tested Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGB), Multi-Layer Perceptron (MLP), AdaBoost.M1 (AB) and Extreme Learning Machine (ELM). They showed that SVM had the best performance in the single-date image analysis, but the MLP classifier yielded the highest overall accuracy in the multitemporal classification scenario. The authors also showed that urban vegetation mapping using single-date imagery is inefficient and it is possible to capture the phenological stages and discriminate various land-cover classes with the dense time series of Synthetic Aperture Radar (SAR) imagery. They verified that the F1 measure exceeded 90% for forest and low vegetation using three and five Sentinel 1 imagery for classification.

Xu *et al.* (2020) proposed a deep learning classification method based on phenological feature constraints for urban green spaces provided by high-resolution GaoFen-2 images using the HRNet (High-Resolution Network). The authors showed that it is possible to solve the problem of misclassification of evergreen and deciduous trees with the introduction of phenological features into HRNet model training because can improve classification accuracy. The authors also proved that combining high-resolution remote sensing images and vegetation phenology can improve deep learning urban green space classification.

Ayhan *et al.* (2020) investigated the performance of three methods for vegetation detection. Two of these methods are based on deep learning and another one is an object-based method that utilizes NDVI (Normalized Difference Vegetation Index), computer vision, and machine learning techniques. They used two deep learning methods, the DeepLabV3+ and a customized convolutional neural network (CNN). Both were evaluated concerning their detection performance when training and testing datasets originated from different geographical sites with different image resolutions. The authors (Ayhan *et al.*, 2020) also proposed a novel object-based vegetation detection approach, which utilizes NDVI, computer vision, and machine learning (ML) techniques. They applied high-resolution airborne color images which consist of RGB and near-infrared (NIR) bands. Still, RGB color images alone were also used with the two deep learning methods to examine their detection performances without the NIR band. The results showed that the NDVI-ML approach performs better than both two deep learning models. According to the authors, the reason is that the training data and testing data are different in appearance making it challenging for deep learning methods. They highlighted that NDVI-ML does not require any training data and may be more practical in real-world applications. They also showed that NDVI-ML does not apply to situations where the NIR band is unavailable.

Huerta *et al.* (2021), evaluated two deep learning model techniques for semantic segmentation of Urban green spaces polygons using different CNN encoders on the U-Net architecture with very high-resolution satellite imagery (WorldView-2). The results showed that the methodology is an accurate digital tool for updating geometrical Urban green spaces databases at the metropolitan level and also extracting these spaces (Dice coefficient of 0.57, recall of 0.8, IoU of 0.75, and kappa coefficient of 0.94). The authors highlighted that as cities grow or change, implementing these models can update Urban green space inventories for better management. In addition, the models can make maps of these spaces more accessible and bring benefits to the management of the conservation of natural resources and environmental services.

Schmohl, Narváez Vallejo e Soergel (2022), exploited the full 3D potential of airborne laser scanning (ALS) point clouds by using a 3D neural network for individual tree detection. For 3D feature extraction, a sparse convolutional network was used. This network fed both semantic segmentation and circular object detection outputs, which were combined for further increased accuracy. The method achieved an average precision of 83% regarding the common 0.5 intersections over the union criterion. They also showed that 85% percent of the stems were found correctly with a precision of 88%, while the tree area was covered by the individual tree detections with an F1 accuracy of 92%. The authors outperformed recent detection networks and traditional delineation baselines.

Although many studies show the advantages of the techniques studied for urban vegetation mapping, some methods did not show good results, proving to be an unfeasible option. However, according to the authors, it is possible to investigate other methods to obtain considerable accuracy based on this analysis. Degerickx, Hermy e

Somers (2020) explored the potential of both passive (2 m-hyperspectral and 0.5 m-multispectral optical imagery) and active (airborne LiDAR) remote sensing technology and proposed a functional urban green typology for mapping the proposed types using object-based image analysis and machine learning. The results demonstrated the potential of the method analyzed, but also showed the limitations of detailed urban green mapping with remote sensing data. The authors mentioned that for classification, the airborne LiDAR data was found to be the most important data source, but in this case required complementary spectral data, for preference of hyperspectral when urban green space has high detail. According to them, the main sources of error resulted in poor classification accuracies (<50%), mostly for shrub and herbaceous vegetation classes. This error is due to the high spectral similarity between types of urban green spaces and close interactions of different objects.

Although only 45 articles mapped tree species, we can see in Table 4 that several species have already been mapped using ML and DL techniques. Most of the 45 articles analyzed the performance of algorithms to map more than one specie, so in this study, we can see a variety of 149 species, families or genera of trees and shrubs found. This table shows the species studied using remote sensing and artificial intelligence (ML/DL).

Table 4- Taxa of species already studied by the authors in the 211 articles.

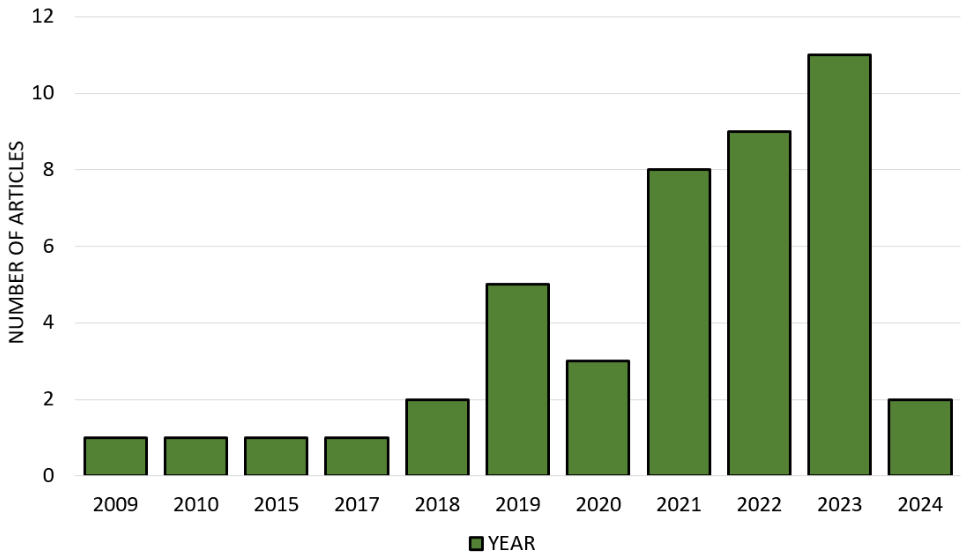
Species		
Abies (firs)	*Dipteryx odorata (tonka beans)*	*Phragmites australis (common reed)*
	Elaeagnus umbellate (autumn olive)	*Pinus massoniana (masson pine)*
Acacia auriculiformis (earleaf acácia)	*Elaeocarpus obtusus (elaeocarpus apiculatus mast)*	*Pinus nigra (austrian pine)*
Acacia confusa (formosan koa)	*Euonymus japonicus (japanese euonymus)*	*Platanus (plane trees)*
Acer negundo (box elder)	*Fagus sylvatica (beech)*	*Platanus spp.*
Acer platanoides (norway maple)	*Fallopia japonica (japanese knotweed)*	*Platanus occidentalis (american sycamore)*
Acer pseudoplatanus	*Ficus benjamina (weeping fig)*	*Platanus x acerifolia (london planetree)*
Acer saccharinum (silver maple)		
Acer saccharum (sugar maple)	*Ficus hispida (hairy fig)*	*Platanus x hispanica (london plane)*
Acerifolia münchh (platanus)	*Ficus macrophylla (australian banyan)*	*Plumeria rubra (frangipani)*
Aeschynanthus radicans (monalisa or lipstick plant)	*Ficus microcarpa (indian laurel fig)*	*Poplar (cottonwood)*
Aesculus x carnea hayne (red horse-	*Ficus nítida (indian laurel columns)*	*Populus balsamifera (balsam poplar)*

chestnut)		
Aesculus hippocastanum (horse chestnut)	Ficus variegata (common red-stem fig)	Populus nigra (black poplar)
Afrocarpus gracilior (fern pine)	Ficus virens (white fig)	Populus tomentosa (chinese white poplar)
Ailanthus altissima	Fraxinus americana (white ash)	Prunus armeniaca (apricot tree)
Albizia (silk tree)	Fraxinus angustifolia (desert Ash)	Prunus avium (cherry)
Albizia lebbeck (indian siris)	Fraxinus excelsior (european ash)	Prunus spp. (PR)
Aleurites moluccana (candlenut tree)	Fraxinus ornus (manna ash)	Pterocarya fraxinifolia (caucasian wingnut)
Alnus glutinosa (black alder)	Fraxinus pennsylvanica (green ash)	Quercus alba
Arecaceae cedar	Fraxinus uhdei (shamel ash)	Quercus ilex L. (holm oak)
Atlas cedar (cedrus atlantica)	Ginkgo biloba (ginkgo)	Quercus palustres (pin oak)
		Quercus robur
Bauhinia variegata (mountain ebony)	Ginko biloba (GI)	Quercus rubra (northern red oak)
Betula papyrifera (paper birch)	Gleditsia triacanthos L. (honey locust)	Quercus virginiana (live oak)
Betula pendula (european white birch – silver birch)	Heracleum mantegazzianum (giant hogweed)	Robinia pseudoacacia (black locust)
Betula pubescens (downy birch)	Jacaranda mimosifolia (jacaranda)	Roystonea regia (florida royal palm)
Betulaceae (birch family)	Juglandaceae (walnut family)	Sapindus saponaria (soapberry tree)
Bischofia polycarpa	Juglans nigra (black walnut)	Senna siamea (kassod tree)
		Sophora japonica
Brachychiton rupestris (bottle tree)	Koelreuteria paniculata (goldenrain tree)	Sorbus aucuparia (rowan)
Broussonetia papyrifera (paper mulberry)	Lantana camara	Spathiphyllum wallisii (regel peace lily)
	Larix decidua (common/european larch—two groups)	
Caesalpinia pulcherrima (peacock flower)	Leucaena leucocephala	Syzygium australe (brush cherry)
	Licania tomentosa (oiti)	
Carpinus betulus (hornbeam)	Ligustrum lucidum (glossy privet)	Tamarindus indica L. (tamarind)
Casuarina equisetifolia (casuarina)	Ligustrum vulgare (common privet)	Taxodium distichum (bald cypress)
Cedrus deodara (deodar cedar)	Liquidambar styraciflua (sweetgum)	Taxus baccata (english yew)
Celtis australis (mediterranean hackberry)	Liriodendron chinense	
	Liriodendron tulipifera (tulip poplar)	Terminalia catappa L. (sea almond)
Celtis occidentalis (hackberry)	Litsea monopetala (many-flowered litsea)	Thuja occidentalis
Celtis sinensis (chinese hackberry)	Lophostemon confertus (brisbane box)	Tilia cordata (small-leaved lime)
Ceratonia siliqua (carob)	Macaranga tanarius (common macaranga)	Tilia platyphyllos (largeleaf linden)
Cercis occidentalis (western redbud)	Machilus chekiangensis (chekiang machilus)	Tilia tomentosa (white and silver linden)
Chionanthus retusus (CH)	Magnolia grandiflora (southern magnolia)	Tipuana tipu
		Tradescantia spathacea (moses in the cradle or wondering jew)
Chlorophytum comosum (spider plant)	Mallotus paniculatus (turn-in-the-wind)	Ulmus glabra (wych elm)
Cinnamomum camphora (camphor tree)	Mangifera indica (mango)	Ulmus parvifolia (chinese elm)
Cinnamomum parthenoxylon (selasian wood)	Metasequoia glyptostroboide (dawn redwood)	Ulmus pumila (siberian elm)
Corylus colurna L (turkish hazelnut)	Mimosaceae (acacia family)	Washingtonia filifera (california fan palm)

Cupaniopsis anacardioides (carrotwood)	*Ocotea odorifera (sassafras)*	*Washingtonia robusta (mexican fan palm)*
Cupressaceae pine (cypress pine)	*Pachira aquática (guiana chestnut)*	*Yulania denudate (magnolia)*
Cupressus sempervirens (italian cypress)	*Peperomia obtusifolia (baby rubber plant)*	*Zelkova carpinifolia (caucasian zelkova)*
Cycas revoluta (sago palm)	*Philodendron hederaceum (philodendron)*	*Zelkova serrata (ZE)*
Delonix regia (royal poinciana)	*Phoenix dactylifera (date palm)*	*Ziziphus (lotus plant)*
Dipteryx alata vogel (baru tree)	*Phoenix roebelenii (dwarf palm)*	

Figure 7 shows the distribution of the 45 articles that studied species according to the year of publication, showing that the increase in studies is recent.

Figure 7- Number of articles published per year for tree species.



Fonte: Author (2024).

In terms of tree species classification, the results are promising and demonstrate the potential of the method for creating urban street tree profile inventory (Choi *et al.*, 2022). The study of Cetin e Yastikli (2022), showed that the classification results serve as the basis for a pilot study in future deciduous and coniferous tree species classification studies using machine learning algorithms in large-scale urban applications. Their results are encouraging with the study area, which included different urban structures with dense trees of different species, sizes and ages. According to the authors, compared with traditional field surveys, this method has benefits, such as automation, especially in terms of study requirements and reducing manpower.

Rodríguez-Puerta *et al.* (2022), compared and combined Airborne Laser Scanning (ALS) point clouds combined with aerial orthophotographs, and street-level imagery from Google Street View (GSV). The authors obtained accurate and robust results with both data sources. According to the authors, the use of ALS and aerial orthophotographs has benefits because are freely available in many countries. In this context, the ALS allows the location of trees anywhere in the city, including public and private parks and gardens. On the other side, the use of street-level images only allows the detection of trees growing in trafficable streets. For future studies, the authors suggested the use of semantic segmentation techniques and mentioned that the next challenge is to analyze how street-level images could allow them to identify some quantitative (height, crown width, etc.) and qualitative (species, genus, health status) characteristics of trees.

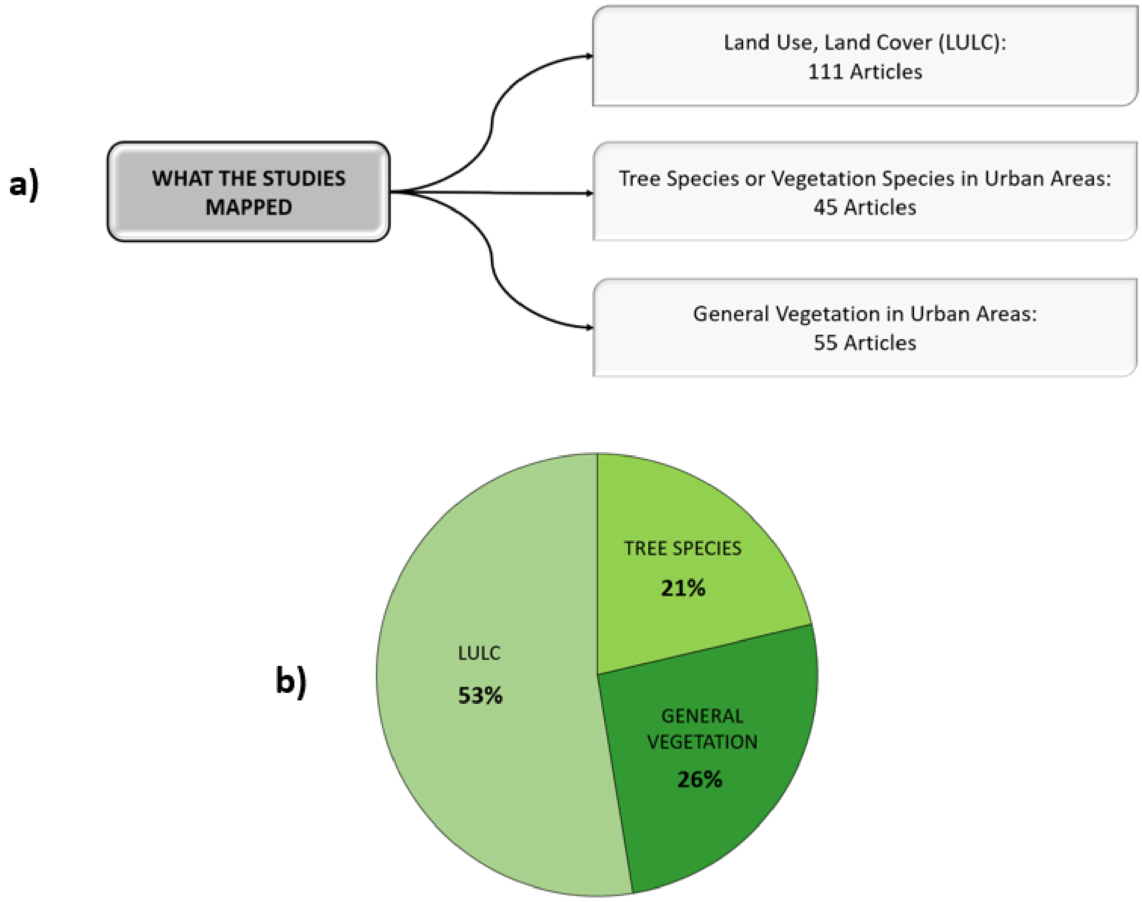
Yang *et al.*, (2022b), integrated the Mask Region-based Convolutional Neural Network (Mask R-CNN) model, an advanced deep learning technology with Google Earth images. The authors wanted to detect tree crown cover in New York's Central Park. The authors obtained 81.8% for the crown area detection rate and the crown detection model was 82.8%. The results indicated the potential of the model and that is possible to identify tree crowns in complex environments. The authors mentioned that additional model testing will be needed using various datasets for natural and urban forests in the future and they want to focus on the CNN model improvement.

Martins *et al.* (2021), showed the potential of CNNs and aerial photographs to map tree species in highly diverse tropical urban settings, providing valuable insights for urban forest management and green space planning. Santos *et al.* (2019) evaluated Convolutional Neural Network (CNN)-based methods combined with Unmanned Aerial Vehicle (UAV) high spatial resolution RGB imagery. The authors wanted to study an important tree species threatened by extinction. They evaluated three methods: Faster Region-based Convolutional Neural Network (Faster R-CNN), YOLOv3 and RetinaNet. The authors obtained an average precision of around 92% with an associated processing time below 30 milliseconds. Despite the low number of articles that analyzed species, the studies have shown good accuracy and promising results using different types of images and ML or DL algorithms.

Regarding what was mapped in the analyzed studies, we could observe that most worked with Land Use Land Cover (LULC). It is important to note that within these LULC articles, we can identify different classes (soil, water, vegetation...). In some articles, there is the occurrence of identification of different species of vegetation, even though the main focus is LULC. Figure 8 shows the number of articles that mapped the options found.

It is important to highlight that in Figure 8 most articles classified as LULC analyzed a certain element/class such as built areas, impermeable areas, water, soil or other, but it was necessary to apply vegetation indices and separate them into a class, explaining why they entered the statistics. The part of the figure "vegetation in general" in an urban area refers to articles whose main objective was to identify vegetation and not a specie.

Figure 8- a) Number of Articles that mapped LULC, species or general vegetation. b) Approximate percentage of articles according to what they mapped.



Fonte: Author (2024).

Studies that performed Land Use Land Cover (LULC) with images, algorithms or neural networks had good results (Camargo *et al.*, 2019; Moayedi *et al.* (2020); Ali *et al.*, 2022; Hossain *et al.*, 2022). Even with good classification and mapping, the studies always present obstacles and difficulties encountered by the algorithms.

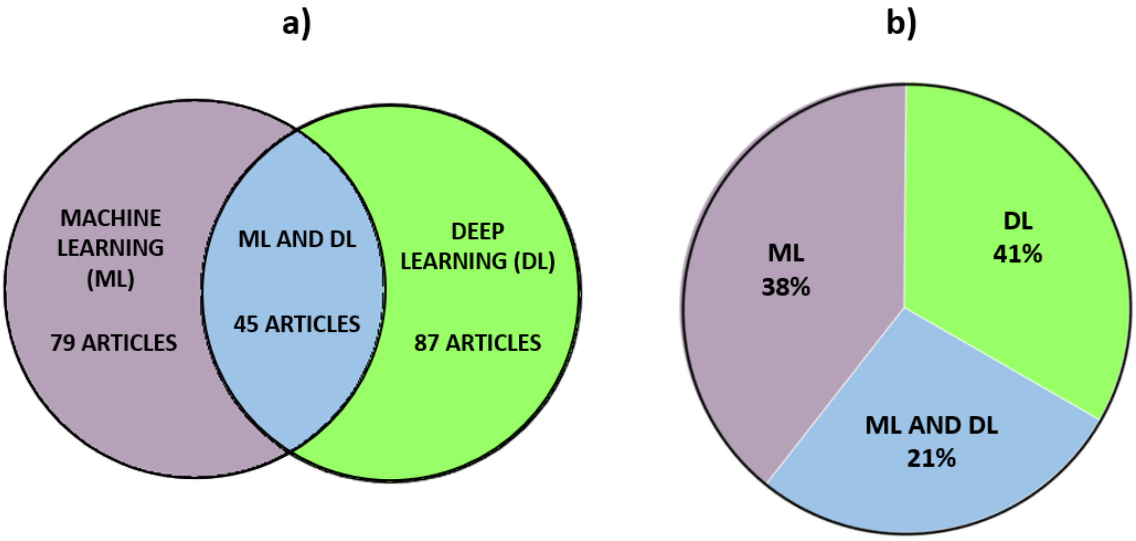
Rousset *et al.* (2021), achieved overall accuracy higher than 73% using XGBoost and neural networks such as ResNet, DenseNet, SegNet and DeepLab for the Land Cover (LC) detection task with RGBNIR as input. In this study, DeepLab achieved an overall accuracy of 81.41%. The authors mentioned that confusion between forest and low-density vegetation was noted in the study, but they explained that distinguishing these classes is difficult even for a human operator. For this reason, the training and testing datasets for LULC classification are never error-free, making it difficult to know whether classification errors are due to a lack of model performance or mislabeling.

Agapiou (2021), mapped LULC and obtained accuracy in the vegetation class of 80%. The author also highlighted the shortcomings found. Three false positive classification results have been observed including the classification of shallow waters as vegetation instead of water bodies. Okujeni *et al.* (2014), mentioned that observed deficiencies in LC mainly relate to known problems arising from spectral similarities or shadowing. Shi e Yang (2017), evaluated LULC in an urban area with algorithms such as RF, SVM and MLP. The authors found that the RF had better overall accuracy than the MLP. These examples show that the application of algorithms with remote sensing has shown good and promising results, but they are not free of presenting errors and obstacles. With the advancement of techniques, it is expected that the errors found can be corrected in order to improve accuracy in the classification, identification and monitoring of vegetation.

3.4 Analysis of the use of Machine Learning and Deep Learning

One of the topics analyzed in this study was whether the articles used machine learning (ML), deep learning (DL) techniques, or both (some studies compared one or more algorithms and neural networks). Figure 9 shows the quantity and approximate percentage of articles that used ML and DL.

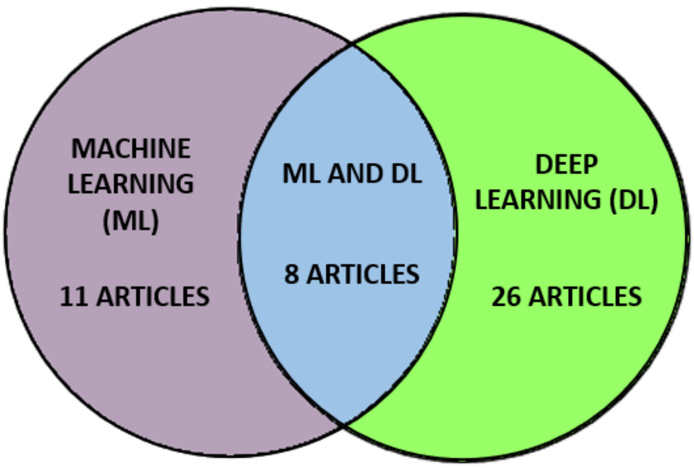
Figure 9- a) Number and b) Approximate percentage of articles that used ML, DL or both in 211 studies (without the 10 review articles).



Fonte: Author (2024).

Regarding tree species mapping (45 articles that studied one or more species), Figure 10 shows the number of articles that used machine learning, deep learning or both, considering only the 45 articles that studied vegetation species in urban areas. Identification and mapping of tree species require more complex techniques and algorithms due to the difficulty of separating trees in images depending on the spatial and spectral resolution. It is possible to observe that most articles that aimed to study one or more species used deep learning techniques.

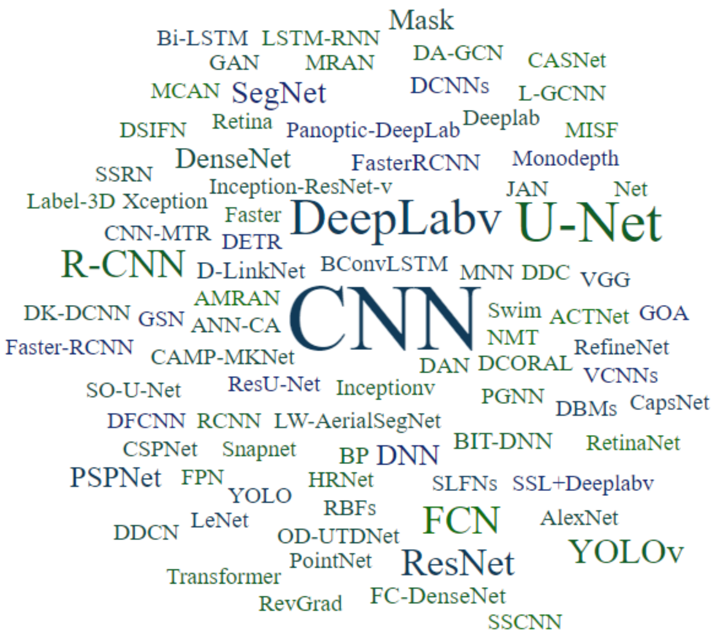
Figure 10- Number of articles that used ML, DL or both for tree species.



Fonte: Author (2024).

The Random Forest (RF) and Support Vector Machine (SVM) algorithms were the most frequent algorithms in the analyzed studies. Considering the most recent techniques, we surveyed the most used neural networks to understand which techniques and algorithms showed good performance for classifying tree species. Figure 11 shows a word cloud with different DL networks found in the 132 articles (45 articles that used ML and DL, and 87 articles that used DL).

Figure 11- Word Cloud with the frequency of Neural Networks used in the articles.



Fonte: Author (2024).

From the data obtained from the Web of Science, it was possible to identify 92 different DL networks used in the studies. It is important to highlight that of the 92 networks identified for the topic of interest, more than 40 networks are present in only one article. Table 5 shows the most used neural networks in the articles. The other neural networks found are present in only one or two articles.

Table 5- Neural Networks with more occurrence in the articles

Neural Networks	Occurrence
Convolutional Neural Network (CNN)	32
U-Net	17
DeepLabv3+	14

Fully Convolutional Networks (FCNs)	8
ResNet	7
SegNet	5
DenseNet	4

From the total of 132 articles that used at least one neural network, we can see that the network used most frequently was the Convolutional Neural Network (CNN) with 32 articles.

Articles that compared the efficiency of deep learning networks with machine learning algorithms for tree species identified an improvement in accuracy. Hartling *et al.* (2019), evaluated the performance of a novel deep learning method, Dense Convolutional Network (DenseNet) to identify tree species in an urban environment within a fused image of WorldView-2 VNIR, Worldview-3 SWIR and LiDAR datasets. They also compared the results of DenseNet against Random Forest (RF) and Support Vector Machine (SVM), revealing that DenseNet is more effective for eight urban tree species classification, presenting an accuracy of 82.6%, compared to 51.8% and 52% for SVM and RF classifiers, respectively. The authors also increased the training sample size and all classifiers improved, but DenseNet in all training sample trials outperformed SVM and RF classifiers. They verified that is possible to achieve more accurate classification of individual tree species by increasing training samples improved DenseNet.

Pearse *et al.* (2021), compared the ResNet-50 architecture, which comprises 49 convolution layers, with the XGBoost algorithm to classify tree species using 10 cm resolution imagery. The authors also tested whether the phenology of the specie could be used to enhance classification by using multitemporal aerial imagery that showed the same trees with and without widespread flowering. The accuracy of XGBoost and deep learning models were 83.2% and 95.2%, respectively. The results of the dataset with strong phenology (flowering) presented higher metrics for both classifiers.

Wang *et al.* (2021), compared the classification of tree species using an unmanned aerial vehicle (UAV) tilt photogrammetry and four algorithms. The authors identified that random forest (RF), support vector machine (SVM), and backpropagation (BP) neural network provide better classification results when using combined parameters compared with those using individual tree attributes alone. The

BP neural network using combined attributes was the best combination found, with an F-measure of 0.872 and a classification precision of 89.1%. The authors showed a low-cost, high-efficiency, and high-precision method for tree species classification, corresponding that optical UAV tilt photogrammetry combined with a machine learning classification algorithm. These studies showed the efficiency of using neural networks compared to machine learning algorithms for urban tree species.

3.5. Analysis of Remote Sensing Data

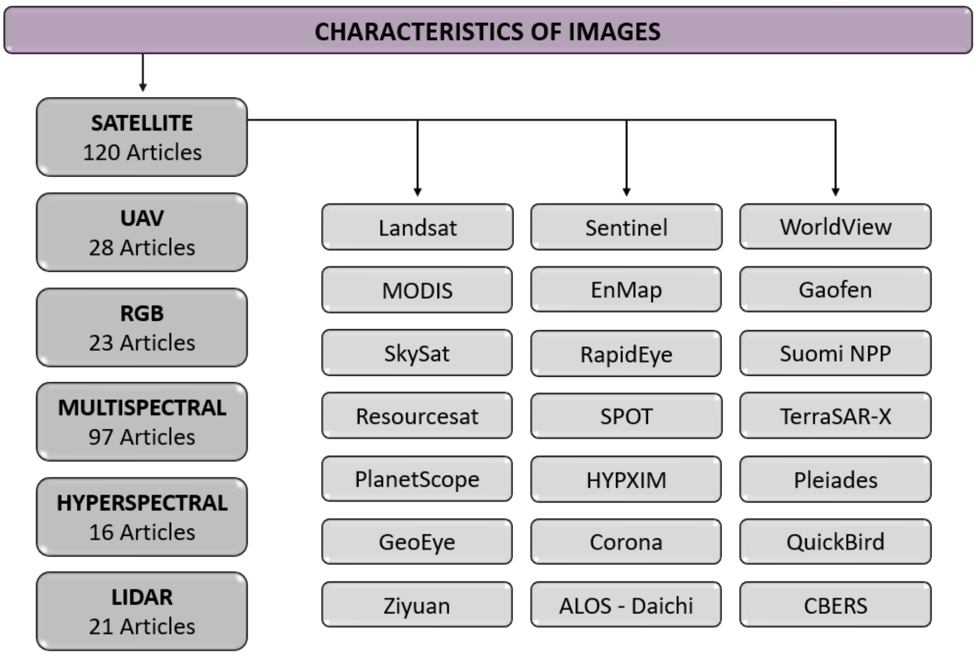
The following step was to analyze the remote sensing used in the studies. The articles present images with different characteristics. Most of them collected data from satellites or Unmanned Aerial Vehicle (UAV).

Most articles used multispectral images. Figure 12 shows the number of articles that used the types of images shown. We found more than 20 satellites used, the figure shows which satellites were found, but the articles may have used different versions of each satellite; for example, the figure shows Landsat, but in the studies, Landsat 5, Landsat 7, and Landsat 8 were found.

Studies have shown the efficient use of hyperspectral images with algorithms to map tree species. Hyperspectral imaging refers to the dense sampling of spectral features with many narrow bands (Zhang; Lin; Wang, 2022). Mozgeris *et al.* (2018) investigated the potential of hyperspectral imaging using ultra-light aircraft for vegetation species mapping in an urban environment. They presented an imaging system based on the simultaneous use of Rikola frame format hyperspectral and Nikon D800E that adopted colour infrared cameras installed onboard a Bekas X32 manned ultra-light aircraft is introduced. The best-performing classification algorithm was multilayer perceptron (MLP) with an overall classification accuracy of 63%. The authors observed that hyperspectral images resulted in significantly better tree species classification than the colour infrared images. They also indicated that ultra-light aircraft-based hyperspectral and colour-infrared imaging was considered to be a technically and economically sound solution. These images can benefit urban green space inventories and ease tree mapping, characterization, and monitoring.

Brabant *et al.* (2019), studied the capability of Very High Spatial Resolution (VHSR) hyperspectral satellite data for discriminating urban tree diversity. They acquired a multispectral Sentinel2 image at 10 m and an airborne HySpex image (408 bands/2 m) from which prototypal spaceborne hyperspectral images (named HYPXIM) at 4 m and 8 m have been simulated. The authors compared and assessed hyperspectral (HySpex and HYPXIM) and Sentinel2 datasets with RF and SVM. Their results show that HYPXIM 4 m and HySpex 2 m reduced by Minimum Noise Fraction (MNF) provide the greatest classification of 14 species using the SVM with an overall accuracy of 78.4%. They showed that prototypal HYPXIM images appear to be a good option for urban vegetation compared to HySpex or Sentinel2 images because it has 192 spectral bands/4 m resolution. In this context, expanding hyperspectral image datasets is necessary and urgent, being important to improve both the reconstruction accuracy and the algorithm robustness (Zhang; Lin; Wang, 2022).

Figure 12- Characteristics of the remote sensing data used in the studies analyzed.



Fonte: Author (2024).

Based on articles that identified species of trees, we identified the use of the following data: WorldView-2 and Worldview-3 (Hartling *et al.*, 2019), Sentinel 2 (Brabant *et al.*, 2019), Pléiades-1A and Pléiades-1B (Le Louarn *et al.*, 2017), Google Maps (Branson *et al.*, 2018), Google Street View (Choi *et al.*, 2022; Rodríguez-Puerta *et al.*, 2022) and Google Earth (Yang *et al.*, 2022b). Other studies used UAV images, hyperspectral cameras or other types of cameras

(Abbas *et al.*, 2021; Mozgeris *et al.*, 2018; Brabant *et al.*, 2019; Martins *et al.*, 2021; Guan *et al.*, 2015; Pearse *et al.*, 2021; Wang *et al.*, 2021; Santos *et al.*, 2019; Lobo Torres *et al.*, 2020; Xia *et al.*, 2021; Helman *et al.*, 2022).

Sun *et al.* (2021), showed in their study that street view imagery coupled with deep learning could measure eye-level street green space and distinguish vegetation types. They showed that street view data reflect different aspects of natural environments compared to satellite imagery, which corresponds to another option for studies.

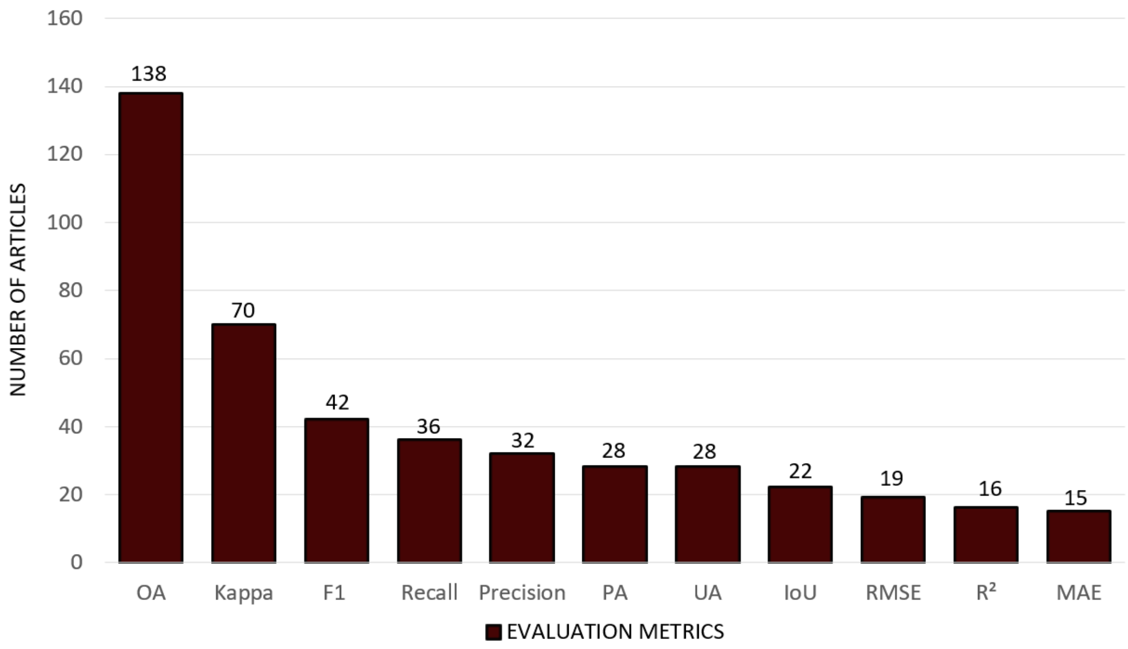
3.6. Evaluation Metrics used in the Studies

Articles that work with classification or regression can use different evaluation metrics. In this study, we also analyzed the metrics used in all articles of interest. We found a total of 37 metrics across all articles. Table 6 shows the abbreviation of the most used metrics. Figure 13 shows the number of articles that used the most frequent metrics.

Table 6- The most used metrics

Metric	Abbreviation
Overall Accuracy	OA
Kappa	-
F1-score	F1
Recall	-
Precision	-
Producer Accuracy	PA
User Accuracy	UA
Intersection over Union	IoU
Root Mean Square Error	RMSE
R-Squared	R^2
Mean Absolute Error	MAE

Figure 13- The most used evaluation metrics in articles.



Fonte: Author (2024).

The most used metrics were Overall Accuracy (OA) and the Kappa coefficient, present in 138 and 70 articles respectively.

4 Conclusion

In this study, we reviewed articles available in the Web of Science on the topics: urban vegetation, urban tree species, remote sensing, machine learning, deep learning. A total of 221 articles were included in our analysis. Studies in this area are increasing every year. The "Remote Sensing" Journal leads the ranking with the most published articles, having 70 articles out of a total of 221. China has the highest number of articles with the first author and co-authors, followed by the United States.

Of the total 221 articles, 79 used machine learning, 87 deep learning and 45 used both. Only 45 articles studied one or more species of trees, but 149 species, genera and families were found. We identified 92 different neural networks and 37 evaluation metrics. We also identify different types of images, satellites and reasons for studying and mapping tree species according to the authors of the 221 papers.

The field of artificial intelligence and remote sensing techniques has demonstrated satisfactory and accurate results in identifying tree species in urban areas, playing a crucial role in diagnosis, monitoring and urban planning. However, as we move into the future, it is necessary to recognize constant technological evolution and the demand for even more refined solutions. It is expected that the accuracy and efficiency in mapping urban vegetation will reach even higher levels with the incorporation of more advanced and adaptable algorithms, combined with improvements in image acquisition techniques (with higher spatial resolutions and new types of sensors).

Vegetation plays a fundamental role in quality of life, including in urban areas. Therefore, investing in research that explores new approaches, high-resolution images and more sophisticated algorithms is essential to meet the growing demands of diagnosis and urban planning. These innovations not only improve the efficiency of existing techniques but also open doors to new possibilities, showing the importance of research and development in this area.

REFERENCES

ABBAS, S.; PENG, Q.; WONG, M. S.; LI, Z.; WANG, J.; KWUN NG, K. T.; KWOK, C. Y. T.; HUI, K. K. W. Characterizing and classifying urban tree species using bi-monthly terrestrial hyperspec- tral images in hong kong. **ISPRS Journal of Photogrammetry and Remote Sensing**, v.177, p. 204–216, 2021.

AGAPIOU, A. Land cover mapping from colorized corona archived greyscale satellite data and feature extraction classification. **Land**, v. 10, 2021.

ALI, U.; ESAU, T. J.; FAROOQUE, A. A.; ZAMAN, Q. U.; ABBAS, F.; BILODEAU, M. F. Limiting the collection of ground truth data for land use and land cover maps with machine learning algorithms. **ISPRS Inter- national Journal of Geo-Information**, v.11, 2022.

ARCE, L. S. D.; OSCO, L. P.; ARRUDA, M. D. S. D.; FURUYA, D. E. G.; RAMOS, A. P. M.; AOKI, C.; POTT, A.; FATHOLAHI, S.; LI, J.; ARAÚJO, F. F. D.; GONÇALVES, W. N.; JUNIOR, J. M. Mauritia flexuosa palm trees airborne mapping with deep convolutional neural network. **Scientific Reports**, v. 11, n. 1, p. 19619, 2021.

AYHAN, B.; KWAN, C.; BUDAVARI, B.; KWAN, L.; LU, Y.; PEREZ, D.; LI, J.; SKARLATOS, D.; VLACHOS, M. Vegetation detection using deep learning and conventional methods. **Remote Sensing**, v.12, 2020.

BRABANT, C.; ALVAREZ-VANHARD, E.; LARIBI, A.; MORIN, G.; NGUYEN, K. T.; THOMAS, A.; HOUET, T. Comparison of hyperspectral techniques for urban tree diversity classification. **Remote Sensing**, v.11, 2019.

BRANSON, S.; WEGNER, J. D.; HALL, D.; LANG, N.; SCHINDLER, K.; PERONA, P. From google maps to a fine-grained catalog of street trees. **ISPRS Journal of Photogrammetry and Remote Sensing**, v.135, p.13–30, 2018.

CAMARGO, F. F.; SANO, E. E.; ALMEIDA, C. M.; MURA, J. C.; ALMEIDA, TATI.; A comparative assessment of machine-learning techniques for land use and land cover classification of the brazilian tropical savanna using alos-2/palsar-2 polarimetric images. **Remote Sensing**, v.11, 2019.

CETIN, Z.; YASTIKLI, N. The use of machine learning algorithms in urban tree species classification. **ISPRS International Journal of Geo- Information**, v.11, 2022.

CHOI, K.; Lim, W.; CHANG, B.; JEONG, J.; KIM, I.; PARK, C. R.; KO, D. W. An automatic approach for tree species detec- tion and profile estimation of urban street trees using deep learning and google street view images. **ISPRS Journal of Photogrammetry and Remote Sensing**, v.190, p.165–180, 2022.

DEGERICKX, J.; HERMY, M.; SOMERS, B. Mapping functional urban green types using high resolution remote sensing data. **Sustainability**, v.12, 2020.

DE ARAÚJO CARVALHO, MARIO.; JUNIOR, J. M.; MARTINS, J. A. C.; ZAMBONI, P.; COSTA, C. S.; SIQUEIRA, H. L.; ARAUJO, M. S.; GONÇALVES, D. N.; FURUYA, D. E. G.; OSCO, L. P.; RAMOS, A. P. M.; LI, JONATHAN.; JUNIOR, A. A. D. C.; GONÇALVES, W. N. A deep learning-based mobile application for tree species mapping in RGB images. **International Journal of Applied Earth Observation and Geoinformation**, v. 114, p. 103045, 2022.

FARKAS, Jenő Zsolt.; HOYK, E.; MORAIS, M. B. D.; CSOMÓS, G. A systematic review of urban green space research over the last 30 years: A bibliometric analysis. **Heliyon**, v. 9, n. 2, 2023.

FURUYA, D. E. G.; AGUIAR, J. A. F.; ESTRABIS, N. B.; PINHEIRO, M. M. F.; FURUYA, M. T. G.; PEREIRA, D. R.; GONÇALVES, W. N.; LIESENBERG, V.; LI, J.; JUNIOR, J. M.; OSCO, L. P.; RAMOS, A. P. M. A machine learning approach for mapping forest vegetation in riparian zones in an atlantic biome environment using sentinel-2 imagery. **Remote Sensing**, v.12, 2020.

FURUYA, M. T. G.; AGUIAR, J. A. F.; ESTRABIS, N. B.; PINHEIRO, M. M. F.; FURUYA, D. E. G.; PEREIRA, D. R.; GOLÇANVES, W. N.; LISENBERG, V,; LI, J.; JUNIOR, J. M.; OSCO, L. P.; RAMOS, A. P. M. A machine learning approach for mapping surface urban heat island using environmental and socioeconomic variables: a case study in a medium-sized Brazilian city. **Environmental Earth Sciences**, v. 82, n. 13, p. 325, 2023.

GARCÍA-PARDO, K.; RANGEL, D. M.; AMARILLO, S. D.; CHAVEZ, J. R. G. Remote sensing for the assess- ment of ecosystem services provided by urban vegetation: A review of the methods applied. **Urban Forestry Urban Greening**, v.74, 127636, 2022.

GAŠPAROVIĆ, M.; DOBRINIĆ. Comparative assessment of ma- chine learning methods for urban vegetation mapping using multitemporal sentinel-1 imagery. **Remote Sensing**, v.12, 2020.

GUAN, H.; YU, Y.; JI, Z.; LI, J.; ZHANG, Q. Deep learning-based tree classification using mobile lidar data. **Remote Sensing Letters**, v.6, p.864–873, 2015.

HARTLING, S., SAGAN, V.; MAIMAITIJIANG, M. Urban tree species classification using uav-based multi-sensor data fusion and machine learning. **GIScience & Remote Sensing**, v.58, p.1250–1275, 2021.

HARTLING, S.; SAGAN, V.; SIDIKE, P.; MAIMAITIJIANG, M.; CARRON, J. Urban tree species classification using a worldview-2/3 and lidar data fu- sion approach and deep learning. **Sensors**, 19, 2019.

HELMAN, D.; YUNGSTEIN, Y.; MULERO, G.; MICHAEL, Y. High- throughput remote sensing of vertical green living walls (vgws) in work- places. **Remote Sensing**, 14, 2022.

HOANG NHAT-DUC, T. X.-L. Remote sensing–based urban green space detection using marine predators algorithm optimized machine learn- ing approach. **Mathematical Problems in Engineering**, 22, 2021.

HOSSAIN, M. T.; ZARIN, T.; SAHRIAR, M. N.; HAQUE, M. N. Machine learning based modeling for future prospects of land use land cover change in gopalganj district, bangladesh. **Physics and Chemistry of the Earth**, Parts A/B/C, v.126, 103022, 2022.

HUERTA, R. E.; YEPEZ, F. D.; LOZANO-GARCIA, D, F.; COBIAN, V. H. G.; FIERRO, A. L. F.; GOMEZ, H. D. L.; GONZALEZ, R. A. C.; VARGAZ-MARTINEZ, A. Mapping urban green spaces at the metropoli- tan level using very high resolution satellite imagery and deep learning techniques for semantic segmentation. **Remote Sensing**, v.13, 2021.

JUTRAS, P., PRASHER, S. O.; MEHUYS, G. R.. Prediction of street tree morphological parameters using artificial neural networks. **Computers and Electronics in Agriculture**, v.67, p.9–17, 2009.

KATTENBORN, T.; LEITLOFF, J.; SCHIEFER, F.; HINZ, S. Review on convolutional neural networks (cnn) in vegetation remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v.173, p.24–49, 2021.

KURAS, A.; BRELL, M., RIZZI, J.; BURUD, I. Perspectral and lidar data applied to the urban land cover machine learning and neural-network-based classification: A review. **Remote Sensing**, v. 13, n. 17, p. 3393, 2021.

LE LOUARN, M.; CLERGEAU, P.; BRICHE.; DESCHAMPS-COTTIN, M. “kill two birds with one stone”: Urban tree species classification using bitemporal pléiades images to study nesting preferences of an invasive bird. **Remote Sensing**, v. 9, 2017.

LI, X., YUAN, W.; DONG, W. A machine learning method for predicting vegetation indices in china. **Remote Sensing**, v. 13, 2021.

LOBO TORRES, D.; FEITOSA, R. Q.; HAPP, P. N.; ROSA, L. E. C. L. R.; JUNIOR, J. M.; MARTINS, J.; BRESSAN, P. O.; GONÇALVES, W. N.; LIESENBERG, V. Applying fully convolutional architectures for semantic segmentation of a single tree species in urban environment on high resolution uav optical imagery. **Sensors**, v. 20, 2020.

LU, Y.; FERRANTI, E. J. S.; CHAPMAN, L.; PFRANG, C. Assessing urban greenery by harvesting street view data: A review. **Urban Forestry & Urban Greening**, p. 127917, 2023.

MARQUES RAMOS, A. P.; OSCO, L. P.; FURUYA, D. E. G.; GONÇALVES, W. N.; SANTANA, D. C.; TEODORO, L. P. R.; JUNIOR, C. A. C. S.; CAPRISTO-SILVA, G. F.; LI, J.; BAIO, F. H. R.; JUNIOR, J. M.; TEODORO, P. E.; PISTORI, H. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. **Computers and Electronics in Agriculture**, v.178, 105791, 2020.

MARTINS, G. B.; ROSA, L. E. C. L.; HAPP, P. N.; FILHO, L. C. T. C.; SANTOS, C. J. F.; FEITOSA, R. Q.; FERREIRA, M. P. Deep learning-based tree species mapping in a highly diverse tropical urban setting. **Urban Forestry Urban Greening**, v.64, 127241, 2021.

MOAYEDI, H.; JAMALI, A.; GIBRIL, M. B. A.; FOONG, L. K.; BAHIRAEI, M. Evaluation of tree-base data mining algorithms in land used/land cover mapping in a semi-arid environment through landsat 8 oli image; shiraz, iran. **Geomatics, Natural Hazards and Risk**, v.11, p.,724–741, 2020.

MOZGERIS, G.; JUODIKIENE, V.; JONIKAVICIUS, D.; STRAIGYTE, L.; GADAL, S.; OUERGHEMMI, W. Ultra-light aircraft-based hyperspectral and colour-infrared imaging to identify deciduous tree species in an urban environment. **Remote Sensing**, v.10, 2018.

NEYNS, R.; CANTERS, F. Mapping of urban vegetation with high- resolution remote sensing: A review. **Remote Sensing**, v.14, 2022.

NGUYEN, U.; GLENN, E. P.; DANG, T. D.; PHAM, L. T. Mapping vegetation types in semi-arid riparian regions using ran- dom forest and object-based image approach: A case study of the Colorado river ecosystem, grand canyon, arizona. **Ecological Informatics**, v.50, p.43–50, 2019.

NITOSLAWSKI, S.; WONG-STEVENS, K.; STEENBERG, J. W. N.; WITHERSPONN, K.; NESBITT, L.; BOSCH, C. C. K. V. D. The digital forest: Mapping a decade of knowledge on technological applications for forest ecosystems. **Earth's Future**, v. 9, n. 8, p. e2021EF002123, 2021.

OKUJENI, A.; LINDEN, S. V. D.; JAKIMOW, B.; RABE, A.; VERRELST, J.; HOSTERT, P. A comparison of advanced regression algorithms for quantifying urban land cover. **Remote Sensing**, v.6, p.6324–6346, 2014.

OSCO, L.; JUNIOR, J. M.; RAMOS, A. P. M.; JORGE, L. A. D. C.; FATHOLAHI, S. N.; SILVA, J. D. A.; MATSUBARA, E. T.; PISTORI, H.; GONÇALVES, W. N.; LI, J. A review on deep learning in UAV remote sensing. **International Journal of Applied Earth Observation and Geoinformation**, v.102 ,102456, 2021.

OSCO, Lucas. P.; FURUYA, D. E. G.; FURUYA, M. T. G.; CORREA, D. V.; GONÇALVES, W. N.; JUNIOR, J. M.; BORGES, M.; BLASSIOLI-MORAES, M. C.; MICHEREFF, M. F. F.; AQUINO, M. F. S.; LAUMANN, R. A.; LISENBERG, V.; RAMOS. A, P. M.; JORGE, L. A. D. C. An impact analysis of pre-processing techniques in spectroscopy data to classify insect-damaged in soybean plants with machine and deep learning methods. **Infrared Physics & Technology**, v. 123, p. 104203, 2022.

PAOLETTI, M.; HAUT, J. M.; PLAZA, J.; PLAZA, A. Deep learning classifiers for hyperspectral imaging: A review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v.158, p.279–317, 2019.

PEARSE, G. D.; WATT, M.; SOEWARTO, J.; TAN , A. Y. S. Deep learning and phenology enhance large-scale tree species classification in aerial imagery during a biosecurity response. **Remote Sensing**, 13, 2021.

RODRÍGUEZ-PUERTA, F.; BARRERA, C.; GARCIA, B.; PEREZ-RODRIGUEZ, F.; GARCIA-PEDRERO, A.M. Mapping tree canopy in urban environments using point clouds from airborne laser scanning and street level imagery. **Sensors**, v. 22, 2022.

ROUSSET, G. DESPINOY, M.; SCHINDLER, K.; MANGEAS, M. Assessment of deep learning techniques for land use land cover classification in southern new caledonia. **Remote Sensing**, v.13, 2021.

SANTOS, A. A.; JUNIOR, J. M.; ARAUJO, M. S.; MARTINI, D. R. D.; TETILA, E. C.; SIQUEIRA, H. L.; AOKI, C.; ELTNER, A.; MATSUBARA, E. T.; PISTORI, H.; FEITOSA, R. Q.; LIESENBERG, V.; GONÇALVES, W. N. Assessment of cnn-based methods for individual tree detection on images captured by RGB cameras attached to uavs. **Sensors**, v.19, 2019.

SCHMOHL, S., NARVÁEZ VALLEJO, A.; SOERGEL, U. Individual tree de- tection in urban als point clouds with 3d convolutional networks. **Remote Sensing**, v.14, 2022.

SHAHTAHMASSEBI, A. R.; LI, C.; FAN, Y.; WU, Y.; LIN, Y.; GAN, M.; WANG, K.; MALIK, A.; BLACKBURN, G. A. Remote sensing of urban green spaces: A review. **Urban Forestry Urban Greening**, v.57, 126946, 2021.

SHI, D.; YANG, X. A relative evaluation of random forests for land cover mapping in an urban area. **Photogrammetric Engineering Remote Sensing**, v.83, p.541–552, 2017.

SHI, Y.; HAN, L.; HUANG, W.; CHANG, S.; DONG, Y.; DANCEY, D. A biologically interpretable two-stage deep neural network (bit-dnn) for vegetation recognition from hyperspectral imagery. **IEEE Transactions on Geoscience and Remote Sensing**, v.60, p.1–20, 2022

SILVA, A. K. V. da. Predicting eucalyptus diameter at breast height and total height with uav-based spectral indices and machine learning. **Forests**, v.12, 2021.

STILLA, Uwe; XU, Yusheng. Change detection of urban objects using 3D point clouds: A review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 197, p. 228-255, 2023.

SUN, Y.; WANG, X.; ZHU, J.; CHEN, L.; JIA, Y.; LAWRENCE, J. M.; JIANG, L. H.; XIE, X.; WU, J. Using machine learning to examine street green space types at a high spatial resolution: Application in los angeles county on socioeconomic disparities in exposure. **Science of The Total Environment**, v.787, p.147653, 2021.

TALIB, M. A.; MAJZOUB, S.; NASIR, Q.; JAMAL, D. A systematic literature review on hardware implementation of artifi- cial intelligence algorithms. **The Journal of**

Supercomputing, v.77, 2021.

TIMILSINA, S.; ARYAL, J.; KIRKPATRICK, J. B. Mapping urban tree cover changes using object-based convolution neural network (ob-cnn). **Remote Sensing**, v.12, 2020.

VELASQUEZ-CAMACHO, L.; CARDIL, A.; MOHAN, M.; ETXEGARAI.; ANZALDI, G.; DE MIGUEL, S. Remotely sensed tree characterization in urban areas: a review. **Remote Sensing**, v. 13, n. 23, p. 4889, 2021.

WANG, Y. QUEN, CHI.; ZHU, Q.; LIU, L.; LI, C.; ZHENG, D. A survey of mobile laser scanning applications and key techniques over urban areas. **Remote Sensing**, v. 11, n. 13, p. 1540, 2019.

WANG, T.; CHANDRA, A.; JUNG, J.; CHANG, A. UAV remote sensing based estimation of green cover during turfgrass establishment. **Computers and Electronics in Agriculture**, v.194, 106721, 2022.

WANG, Y.; WANG, J.; CHANG, S.; SUN, L.; AN, L.; CHEN, Y.; XU, J. Classification of street tree species using UAV tilt photogrammetry. **Remote Sensing**, 13, 2021.

WANG, Y. Detecting occluded and dense trees in urban terrestrial views with a high-quality tree detection dataset. **IEEE Transactions on Geoscience and Remote Sensing**, v.60, p.1–12, 2022.

WANG, Y.; YU, Y.; CAO, S.; ZANG, X.; GAO, S. A review of applications of artificial intelligent algorithms in wind farms. **Artificial Intelligence Review**, v.53, 2020.

WANG, H.; SHAO, Y.; KENNEDY, L. M.; Temporal generalization of sub-pixel vegetation mapping with multiple machine learning and atmospheric correction algorithms. **International Journal of Remote Sensing**, v.35, 2014.

XIA K.; WANG, H.; YANG, Y.; DU, X. Automatic detection and parameter estimation of ginkgo biloba in urban environment based on rgb images. **Journal of Sensors**, 2021.

XU, Z.; ZHOU, Y.; WANG, S.; WANG, L.; LI, F.; WANG, S.; WANG, Z. A novel intelligent classification method for urban green space based on high-resolution remote sensing images. **Remote Sensing**, v.12, 2020.

YANG, M.; ZHOU, X.; LIU, Z.; LI, P.; TANG, J.; XIE, B.; PENG, C. A review of general methods for quantifying and estimating urban trees and biomass. **Forests**, v. 13, n. 4, p. 616, 2022a.

YANG, M.; MOU, Y.; LIU, S.; MENG, Y.; LIU, Z.; LI, P.; XIANG, W.; ZHOU, X.; PENG, C. Detecting and map- ping tree crowns based on convolutional neural network and google earth images. **International Journal of Applied Earth Observation and Geoinformation**, v.108, 102764, 2022b.

YIN, J.; DONG, J.; HAMM, N. A. S.; LI, Z.; WANG, J.; XING, H.; FU, P. Integrating remote sensing and geospatial big data for urban land use mapping: A review. **International Journal of Applied Earth Observation and Geoinformation**, v.103, 102514, 2021.

YUAN, Q.; SHEN, H.; LI, T.; LI, Z.; LI, S.; JIANG, Y.; XU, H.; TAN, W.; YANG, Q.; WANG, J.; GAO, J.; ZHANG, L. Deep learning in environmental remote sensing: Achievements and challenges. **Remote Sensing of Environment**, v.241, 111716, 2020.

ZHANG, L.; LIN, H.; WANG, F. Individual tree detection based on high-resolution RGB images for urban forestry applications. **IEEE Access**, v.10, p.46589–46598, 2022.

ZHANG, J.; SU, R.; REN, W.; FU, Q.; HEIDE, F.; NIE, F. A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. **Scientific Reports**, v.12,11905, 2022.

CAPÍTULO 2 - SEGMENTAÇÃO DE VEGETAÇÃO ARBÓREA EM ÁREA URBANA UTILIZANDO O SEGMENT ANYTHING MODEL (SAM) E IMAGENS DO STREET VIEW.

Resumo: O diagnóstico e monitoramento da vegetação arbórea são elementos importantes para o planejamento urbano. O sensoriamento remoto é uma opção viável para analisar áreas amplas e possibilita o estudo de determinada área sem a necessidade de ir até o local. As imagens do Street View podem fornecer mais detalhes na identificação de um alvo quando comparadas a imagens de satélite. A utilização de modelos e técnicas de Inteligência Artificial podem contribuir e otimizar os procedimentos ambientais. Nesse contexto, o Segment Anything Model (SAM) é um modelo de segmentação de imagens recente disponibilizado em abril de 2023 e possui diversas vantagens quando comparado aos métodos de segmentação tradicionais. O objetivo desse estudo foi analisar a capacidade do SAM em segmentar vegetação arbórea em área urbana com imagens do Street View da cidade de Presidente Prudente. Os resultados foram satisfatórios e mostraram que a utilização e aplicação de Inteligência Artificial traz vantagens para aplicações ambientais. A utilização do SAM para detectar vegetação otimiza as etapas de rotulação e facilita a segmentação de objetos na imagem.

Palavras-chave: Vegetação Arbórea; SAM; Inteligência Artificial; Street View.

Abstract: The diagnosis and monitoring of tree vegetation are important elements for urban planning. Remote sensing is a viable option to analyze large areas and allows the study of a certain area without the need to go to the location. Street View imagery can provide more detail in identifying a target when compared to satellite imagery. The use of Artificial Intelligence models and techniques can contribute and optimize environmental procedures. In this context, the Segment Anything Model (SAM) is a recent image segmentation model available in April 2023 and has several advantages when compared to traditional segmentation methods. The objective of this study was to analyze the ability of SAM to segment tree vegetation in an urban area with Street View images of the city of Presidente Prudente. The results were satisfactory and showed that the use and application of Artificial Intelligence brings advantages to environmental applications. The use of SAM to detect vegetation optimizes the labeling steps and facilitates the segmentation of objects in the image.

Keywords: Tree Vegetation; SAM; Artificial Intelligence; Street View.

1 INTRODUÇÃO

A vegetação em áreas urbanas fornece serviços ecossistêmicos que são importantes para a biodiversidade e o bem estar. As áreas verdes são consideradas essenciais na vida do ser humano (O'brien *et al.*, 2017; Ferrini *et al.*, 2020) e é possível encontrar diferentes tipos de vegetação, como árvores, arbustos e grama, nas áreas urbanas (Drillet *et al.*, 2020).

Alguns dos exemplos das vantagens da vegetação arbórea nas áreas urbanas é o controle da temperatura, resfriamento do ar, produção de biomassa, absorção de poluentes do ar, redução de ruído (Pauleit *et al.*, 2011; Gatto *et al.*, 2020; Livesley; Mcpherson; Calfapietra, 2016) além de ser um elemento central para mitigar as mudanças climáticas (Ferrini *et al.*, 2020).

O sensoriamento remoto tornou-se uma ferramenta importante para análises ambientais e tem potencial para ser combinado com outros métodos de monitoramento e avaliação (Wellmann *et al.*, 2020; Dong *et al.*, 2019). O mapeamento da vegetação em áreas urbanas com técnicas de sensoriamento remoto pode ser útil para o planejamento e para lidar com os desafios urbanos (Gašparović; Dobrinić, 2020). Para determinado estudo, as imagens de satélite podem ser insuficientes para uma análise, como no caso de identificação de espécies arbóreas. Nesses casos, a visão aérea de imagens de satélite não disponibiliza o detalhamento necessário para o reconhecimento de vegetação arbórea.

As imagens do Street View fornecem uma nova opção para analisar e estudar as áreas urbanas com mais detalhes (Larkin *et al.*, 2021). Um estudo feito por Barbierato *et al.* (2020), propôs um conjunto de métricas para áreas verdes urbanas com a combinação de sensoriamento remoto e sensoriamento próximo (Street View). Os autores classificaram as áreas verdes por segmentação semântica com a utilização de redes neurais profundas e obtiveram acurácia de 83%.

Outro estudo feito por Li, Ratti, Seiferling (2018), apresentou um novo método para identificar sombras de árvores nas ruas do centro da cidade de Boston com o auxílio de imagens do Street View para calcular o fator de visão do céu. De acordo com os autores, as imagens no nível da rua possibilitaram considerar obstruções sem depender de simulações do ambiente ou simplificações.

A combinação de métodos ajuda na tomada de decisão, contribui para futuros projetos de esverdeamento urbano para moderação do clima urbano e tornam o sensoriamento remoto mais acessível para planejadores urbanos e pesquisadores que necessitam de dados qualitativos (Li; Ratti; Seiferling, 2018; Wellmann *et al.*, 2020).

A Inteligência Artificial (IA) é uma tecnologia robusta e promissora. As aplicações e desenvolvimento da IA crescem a cada ano em diversas áreas, inclusive nas áreas urbanas e ambiental (Yigitcanlar *et al.*, 2020; Allam; Dhunny, 2019). A utilização de Inteligência Artificial na arborização urbana contribui para a tomada de decisões a fim de melhorar a qualidade das cidades, pois o bem estar dos cidadãos depende em grande parte das florestas urbanas (Araújo *et al.*, 2021).

A segmentação de imagens é uma parte importante do processamento de imagens e diversos métodos são desenvolvidos rapidamente para realizar este processo (Liu *et al.*, 2021). A segmentação divide uma imagem inteira em diversas regiões que possam ter propriedades semelhantes, separando um alvo de interesse (Liu *et al.*, 2021). Novas tecnologias incluindo a Inteligência Artificial contribuem para otimizar o processo de segmentação de imagens com a aplicação de um modelo, algoritmo, entre outros (Lateef; Ruichek, 2019; Liu *et al.*, 2021). A segmentação de imagens com a utilização de técnicas de machine e deep learning atingiram resultados satisfatórios e mostraram que a utilização de IA auxilia na segmentação de imagens (Hesamian *et al.*, 2019; Bhatnagar; Gill; Ghosh, 2020).

Um estudo feito por Lassiter e Darbari (2020), avaliou três métodos não supervisionados para segmentar vegetação em área urbana, incluindo o k-means e a Convolutional Neural Network (CNN). Os autores verificaram que cada método possui um melhor desempenho na segmentação dependendo do alvo, como por exemplo a presença de sombras e vegetação secundária. Apesar de avaliarem mais de um método, os autores verificaram que são opções viáveis para segmentar vegetação urbana.

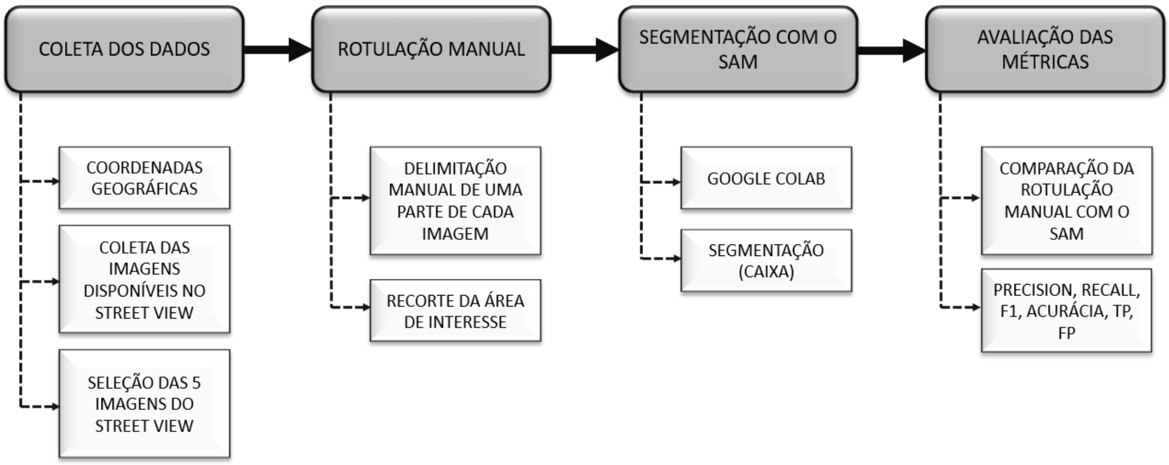
O Segment Anything Model (SAM) é uma ferramenta atual disponibilizada em abril de 2023, e foi projetado para segmentar objetos de interesse em imagens a partir de prompts fornecidos por pontos, caixas ou texto (Mazurowski *et al.*, 2023). A recente contribuição do SAM é segmentar qualquer objeto de uma imagem sem a necessidade de realizar treinamento anteriormente (Meta AI, 2023).

No contexto atual de Inteligência Artificial e segmentação de imagens, o objetivo deste estudo foi avaliar a performance do Segment Anything Model (SAM) na segmentação de vegetação arbórea em imagens do Street View da cidade de Presidente Prudente em comparação com a rotulação manual.

2 METODOLOGIA

A realização deste estudo foi feita em quatro etapas principais: coleta dos dados, rotulação manual, segmentação de vegetação arbórea com o SAM e avaliação das métricas. As etapas deste estudo estão descritas na Figura 1.

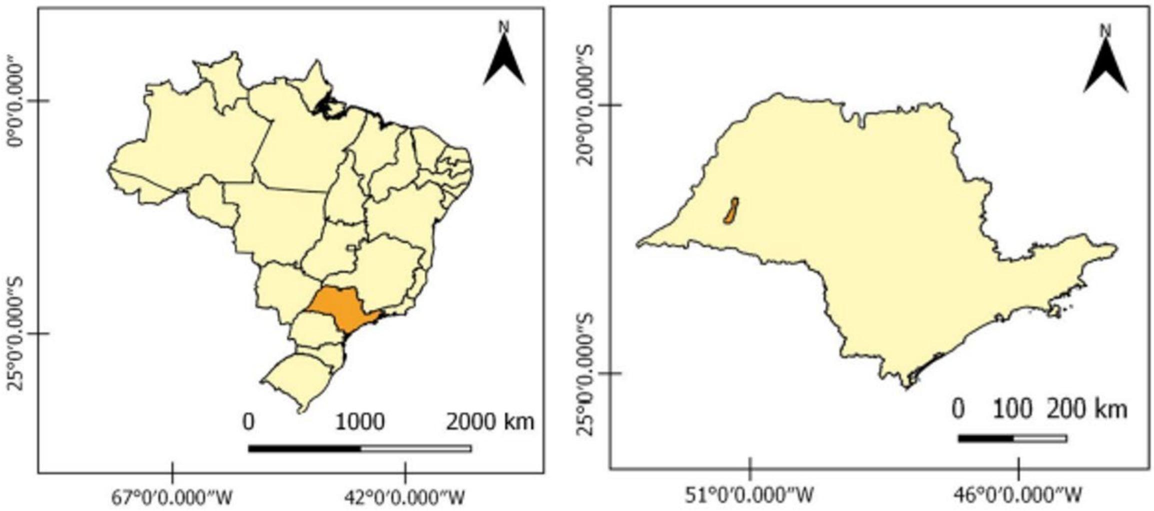
Figura 1- Fluxograma com as principais etapas da metodologia.



Fonte: Autora (2024).

A área de estudo corresponde a cidade de Presidente Prudente, localizada no estado de São Paulo, Brasil. A cidade possui vegetação arbórea composta de vegetação nativa e exótica nas áreas urbanas. A Figura 2 mostra a localização da área de estudo.

Figura 2- Localização da cidade de Presidente Prudente.



Fonte: (Furuya *et al*., 2023)

Para a aquisição das imagens do Street View de Presidente Prudente foi selecionada uma área urbana (Figura 3) que anteriormente foi identificada a presença de espécies arbóreas nativas e exóticas. A partir da seleção, as coordenadas geográficas da área foram identificadas e posteriormente foi realizada a coleta das imagens do Street View disponíveis da área escolhida.

Para a cidade de Presidente Prudente foram encontradas 72 imagens do Street View na área urbana escolhida. Apesar da quantidade, diversas imagens correspondem a imagens com o mesmo ângulo, mas de anos diferentes, sendo visualmente similares. Portanto, para este estudo foram selecionadas 5 imagens. A Figura 4 mostra alguns exemplos das imagens do Street View encontradas para a cidade de Presidente Prudente.

Figura 3- Área urbana com presença de vegetação arbórea escolhida para o estudo.

Fonte: Autora (2024).

Figura 4. Exemplos de imagens do Street View encontradas na área selecionada de Presidente Prudente com base nas coordenadas geográficas.

Fonte: Autora (2024).

Para os testes de segmentação com o SAM, foram selecionadas 5 imagens do Street View: 2 imagens do ano de 2012; 1 imagem de 2015 e 2 imagens do ano de 2017. O código e os testes utilizados foram feitos no Google Colab.

Para verificar o desempenho do SAM na segmentação de vegetação arbórea, foi realizada a rotulação manual de partes de vegetação arbórea em cada uma das imagens selecionadas para posteriormente comparar os resultados na segmentação. A rotulação manual foi realizada no GIMP e o alvo de interesse foi delimitado para cada imagem conforme a Figura 5.

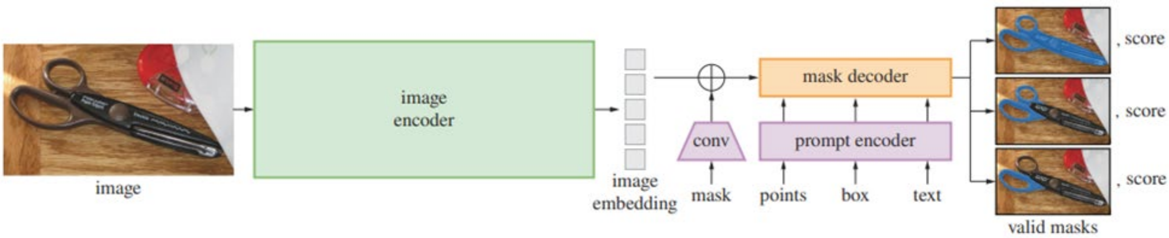
Figura 5- Exemplo da delimitação de vegetação arbórea por meio da rotulação manual.

Fonte: Autora (2024).

Após a delimitação de uma área com vegetação arbórea, foi feito o recorte da rotulação para ser comparado com a segmentação do SAM no Google Colab.

O SAM produz uma incorporação de imagem que pode ser consultado com eficiência por uma variedade de prompts de entrada para produzir máscaras de objeto em velocidade amortizada em tempo real (Kirillov *et al.*, 2023). Para prompts correspondentes a mais de um objeto, o SAM pode produzir diversas máscaras válidas e pontuações de confiança associadas (Figura 6).

Figura 6- Exemplo de segmentação feita pelo SAM.



Fonte: (Kirillov *et al.*, 2023)

Para a realização dos testes é necessário inserir a imagem do Street View escolhida no código utilizado. O Segment Anything Model (SAM) permite selecionar uma área pelo formato de caixas, pontos ou texto, nesse estudo foram feitos testes com o formato de caixas.

Após inserir a imagem no Google Colab, é possível delimitar uma caixa na área escolhida, conforme a Figura 7. Nos testes realizados foi feita a delimitação de uma caixa por imagem, dessa forma, o SAM não irá segmentar toda a vegetação arbórea de cada imagem, somente a que estiver dentro da caixa. Esse procedimento foi realizado pois o objetivo é avaliar o desempenho do SAM comparado a rotulação manual. Portanto, para cada imagem foi delimitada somente uma parte para comparar os dois métodos.

Figura 7- Exemplo da delimitação da área de interesse pelo formato de caixa do SAM.

Fonte: Autora (2024).

A partir da delimitação da caixa (na cor azul) no próprio código no Google Colab conforme a Figura 7, o SAM realiza a segmentação do objeto encontrado dentro da caixa. O SAM também é utilizado para segmentar uma imagem inteira sem precisar de treinamento (Figura 8), mas como o objetivo deste estudo é analisar a capacidade do SAM em segmentar um alvo específico (vegetação arbórea), o procedimento é diferente.

Figura 8- Exemplo de segmentação realizada pelo SAM em toda a imagem fornecida.

Fonte: (Meta AI, 2023)

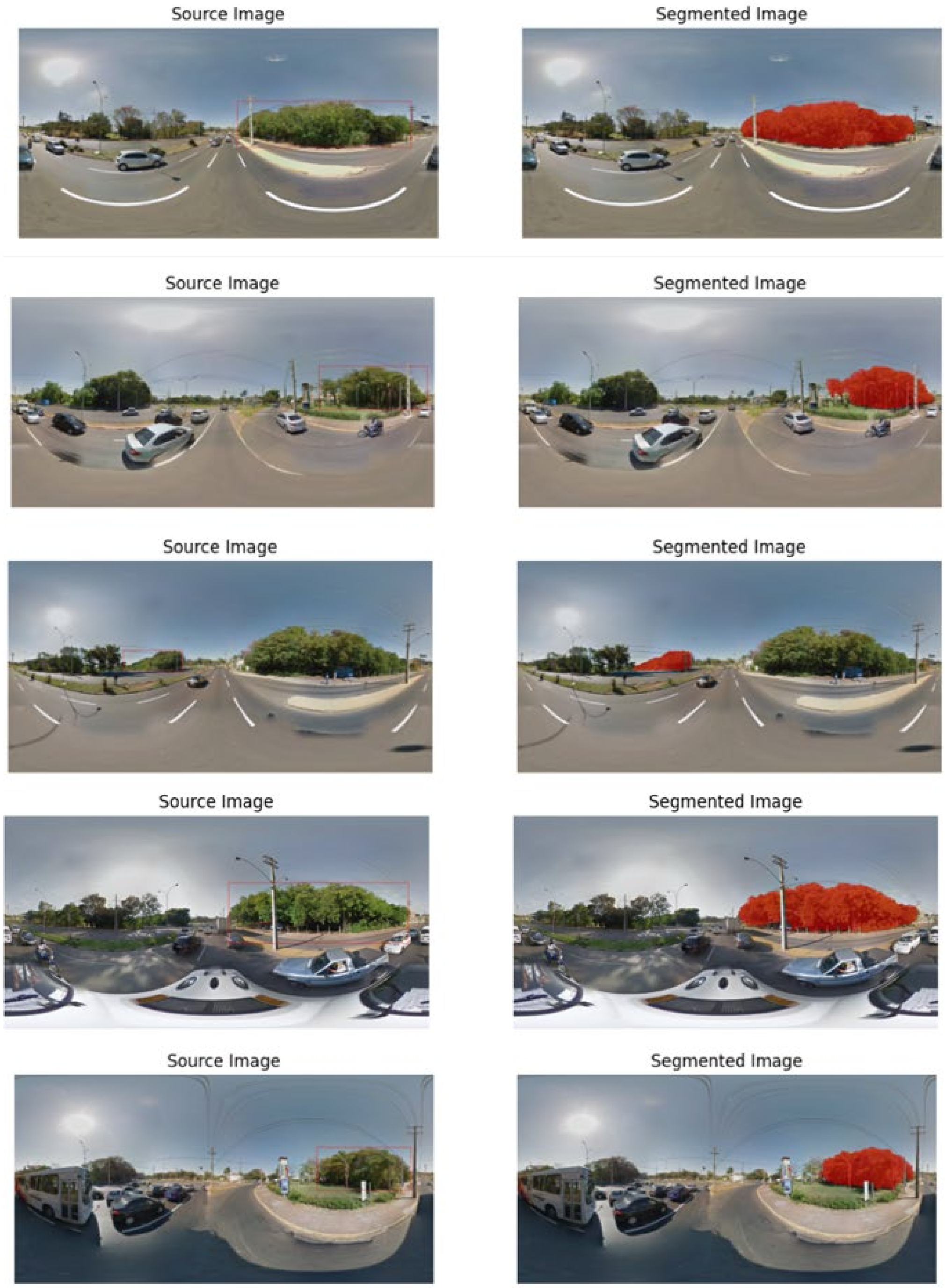
Assim que o SAM realizou a segmentação das áreas delimitadas, foi solicitado por meio do código a comparação da rotulação manual e a segmentação feita pelo SAM, os dois arquivos foram fornecidos e métricas de avaliação foram calculadas.

Neste estudo foi solicitado o cálculo das seguintes métricas: precision, recall, F-score, acurácia, verdadeiro positivo (TP) e falso positivo (FP).

3 RESULTADOS

Conforme descrito na metodologia, o SAM realizou a segmentação dos objetos encontrados dentro das caixas que foram selecionadas (Figura 9).

Figura 9- Segmentação realizada pelo SAM nas 5 imagens do Street View com base no formato da caixa.



Fonte: Autora (2024).

Após a segmentação feita pelo SAM, foi solicitado a comparação dos dois métodos (rotulação manual e o SAM). As métricas de avaliação foram calculadas com base no recorte dos dois arquivos (Figuras 10 e 11).

Figura 10- Cálculo das métricas precision, recall, F1 e acurácia para uma das imagens.

```
# Calculate the precision, recall, F-score, and accuracy
precision = precision_score(img1_flat, img2_flat, average='weighted')
recall = recall_score(img1_flat, img2_flat, average='weighted', zero_division=1)
f1 = f1_score(img1_flat, img2_flat, average='weighted')
accuracy = accuracy_score(img1_flat, img2_flat)

print("Precision: ", precision)
print("Recall: ", recall)
print("F-score: ", f1)
print("Accuracy: ", accuracy)
```

```
Precision:  0.9917624274341269
Recall:  0.9906314646703958
F-score:  0.990940660413282
Accuracy:  0.9906314646703958
```

Fonte: Google Colaboratory (2024).

Figura 11- Cálculo das métricas verdadeiro positivo (TP) e falso positivo (FP) para uma das imagens.

```
# Assuming y_true and y_pred are the true and predicted labels for a binary classification problem
fpr, tpr, thresholds = roc_curve(img1_flat, img2_flat)

# Calculate the area under the ROC curve
roc_auc = auc(fpr, tpr)

print("True Positive Rate (TPR): ", tpr)
print("False Positive Rate (FPR): ", fpr)
print("Area Under the ROC Curve (AUC): ", roc_auc)
```

```
True Positive Rate (TPR):  [0.         0.98661512 1.        ]
False Positive Rate (FPR):  [0.         0.00914518 1.        ]
Area Under the ROC Curve (AUC):  0.9887349700358143
```

Fonte: Google Colaboratory (2024).

Este procedimento foi feito para cada uma das imagens, a Tabela 1 mostra a média de cada métrica calculada para as 5 imagens do Street View.

Tabela 1- Média das métricas obtidas nas 5 imagens com o SAM.

Métrica	Média
Precision	0.9951
Recall	0.9947
F1	0.9948
Acurácia	0.9947
Verdadeiro Positivo (TP)	0.9608
Falso Positivo (FP)	0.0298

4 DISCUSSÃO

A utilização de Inteligência Artificial (IA) em pesquisas e projetos está sendo aplicada em diversas áreas a fim de otimizar e facilitar a realização de uma tarefa. A Inteligência Artificial é um processo pelo qual o ser humano pode construir uma máquina inteligente que trabalha com base em aprendizados passados (Jha *et al.*, 2019).

Dentro da área conhecida como IA, existem subgrupos como o aprendizado de máquina (machine learning) e o aprendizado profundo (deep learning) que auxiliam projetos e tarefas (Benos *et al.*, 2021; Helm *et al.*, 2020). Diversos algoritmos e redes neurais são implantados em todas as áreas do conhecimento. O objetivo do aprendizado de máquina é alimentar a máquina com dados de experiências passadas e dados estatísticos para que ela possa executar uma tarefa atribuída para resolver um problema específico (Jha *et al.*, 2019).

Estudos em diferentes áreas mostraram resultados satisfatórios com a combinação de ML/DL e segmentação de imagens (Piralilou *et al.*, 2019; Haque; Neubert, 2020; Raei *et al.*, 2022; Lee; Le; Kim, 2023). O processo de detecção de objetos a partir de uma imagem é uma etapa fundamental de qualquer técnica de processamento de imagem, pois a acurácia da classificação dependerá da qualidade dos resultados obtidos na etapa de segmentação (Kheradmandi; Mehranfar, 2022).

A segmentação de imagens é importante para organizar uma imagem em pequenos números de regiões para entender “o que” e “onde” é/está o alvo na imagem (Ghiasi *et al.*, 2022). A rotulação de imagens demanda tempo e cautela para identificar corretamente o alvo de interesse. Apesar das técnicas de IA contribuírem com o avanço na classificação de imagens, novas técnicas e modelos podem acelerar as etapas anteriores de rotulação e segmentação.

Nos testes realizados neste estudo, verificou-se que o SAM foi capaz de segmentar a vegetação arbórea (alvo de interesse) com facilidade, apenas com a delimitação de uma caixa. A rotulação realizada anteriormente foi utilizada para comparar a precisão da segmentação realizada pelo SAM. A média das métricas utilizadas para as 5 imagens do Street View mostrou que a segmentação de vegetação arbórea com o SAM foi satisfatória, pois atingiu mais de 99%.

A utilização do Segment Anything Model (SAM) é promissora e tende a crescer nos próximos anos. Estudos estão iniciando os testes com a utilização do SAM pois o modelo foi disponibilizado pela META AI em abril de 2023 (Meta AI, 2023). Até o momento, estudos que realizaram a aplicação do SAM para segmentação de imagens correspondem à área da saúde (análise de imagens médicas). Futuramente será possível comparar resultados mais amplos e complexos em diversas áreas do conhecimento.

5 CONCLUSÃO

A identificação e monitoramento de vegetação arbórea em áreas urbanas é de extrema importância para o planejamento urbano e para o bem estar da população. A utilização de técnicas de Inteligência Artificial na área ambiental tem crescido cada vez mais. O Segment Anything Model (SAM) é um modelo recente de segmentação de imagens disponibilizado em abril de 2023 pela Meta AI e otimiza o processo de identificação de imagens, pois não necessita de treinamento anterior. Os testes realizados nesse estudo com imagens do Street View mostraram que a segmentação de vegetação arbórea com o SAM é viável e atingiu boas acurácias (acima de 99%). A utilização do SAM ainda é pouco explorada devido ao curto tempo de lançamento desse modelo. Futuros estudos podem explorar a capacidade do modelo em segmentar espécies arbóreas e analisar as contribuições do SAM para a identificação de objetos em diferentes imagens.

REFERÊNCIAS

ALLAM, Z; DHUNNY, Z. A. On big data, artificial intelligence and smart cities. **Cities**, v. 89, p. 80-91, 2019.

ARAÚJO, H. C. L; MARTINS, F. S; CORTESE, T. T. P; LOCOSSELLI, G. M. Artificial intelligence in urban forestry—A systematic review. **Urban Forestry & Urban Greening**, v. 66, p. 127410, 2021.

BARBIERATO, E; BERNETTI, I; CAPECCHI, I; SARAGOSA, C. Integrating remote sensing and street view images to quantify urban forest ecosystem services. **Remote sensing**, v. 12, n. 2, p. 329, 2020.

BENOS, L; TAGARAKIS, A. C; DOLIAS, G; BERRUTO, R; KATERIS, D; BOCHTIS, D. Machine learning in agriculture: A comprehensive updated review. **Sensors**, v. 21, n. 11, p. 3758, 2021.

BHATNAGAR, S; GILL, L; GHOSH, B. Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities. **Remote Sensing**, v. 12, n. 16, p. 2602, 2020.

DONG, J; METTERNICHT, G.; HOSTERT, P.; FENSHOLT, R.; CHOWDHURY, R.R. Remote sensing and geospatial technologies in support of a normative land system science: Status and prospects. **Current Opinion in Environmental Sustainability**, v. 38, p. 44-52, 2019.

DRILLET, Z; FUNG, T.K; LEONG, R.A.T; SACHIDHANANDAM, U; EDWARDS, P; RICHARDS, D. Urban vegetation types are not perceived equally in providing ecosystem services and disservices. **Sustainability**, v. 12, n. 5, p. 2076, 2020.

FERRINI, F; FINI, A; MORI, J; GORI, A. Role of vegetation as a mitigating factor in the urban context. **Sustainability**, v. 12, n. 10, p. 4247, 2020.

FURUYA, M. T. G; FURUYA, D. E. G; OLIVEIRA, L. Y. D; SILVA, P. A; CICERELLI, R. E; GONÇALVES, W. N; MARCATO JUNIOR, J; OSCO, L.P; RAMOS, A. P. M. A machine learning approach for mapping surface urban heat island using environmental and socioeconomic variables: a case study in a medium-sized Brazilian city. **Environmental Earth Sciences**, v. 82, n. 13, p. 325, 2023.

GAŠPAROVIĆ, M; DOBRINIĆ, D. Comparative assessment of machine learning methods for urban vegetation mapping using multitemporal sentinel-1 imagery. **Remote Sensing**, v. 12, n. 12, p. 1952, 2020.

GATTO, E; BUCCOLIERI, R; AARREVAARA, E; IPPOLITO, F; EMMANUEL, R; PERRONACE, L; SANTIAGO, J. L. Impact of urban vegetation on outdoor thermal comfort: Comparison between a mediterranean city (Lecce, Italy) and a northern European city (Lahti, Finland). **Forests**, v. 11, n. 2, p. 228, 2020.

GHIASI, G; GU, X; CUI, Y; LIN, T. Y. Scaling open-vocabulary image segmentation with image-level labels. *In*: **European Conference on Computer Vision. Cham: Springer Nature Switzerland**, 2022. p. 540-557.

HAQUE, I. R. I; NEUBERT, J. Deep learning approaches to biomedical image segmentation. **Informatics in Medicine Unlocked**, v. 18, p. 100297, 2020.

HELM, J. M; SWIERGOSZ, A. M; HAEBERLE, H. S; KARNUTA, J. M; SCHAFFER, J. L; KREBS, V. E; SPITZER. A. I; RAMKUMAR, P. N. Machine learning and artificial intelligence: definitions, applications, and future directions. **Current reviews in musculoskeletal medicine**, v. 13, p. 69-76, 2020.

HESAMIAN, M. H; JIA, W; HE, X; KENNEDY, P. Deep learning techniques for medical image segmentation: achievements and challenges. **Journal of digital imaging**, v. 32, p. 582-596, 2019.

JHA, K; DOSHI, A; PATEL, P; SHAH, M. A comprehensive review on automation in agriculture using artificial intelligence. **Artificial Intelligence in Agriculture**, v. 2, p. 1-12, 2019.

KHERADMANDI, N; MEHRANFAR, V. A critical review and comparative study on image segmentation-based techniques for pavement crack detection. **Construction and Building Materials**, v. 321, p. 126162, 2022.

KIRILLOV, A; MINTUN, E; RAVI, N; MAO, H; ROLLAND, C; GUSTAFSON, L; XIAO, T; WHITEHEAD, S; BERG, A .C; LO, W.Y; DOLLÁR, P; GIRSHICK, R. Segment anything. **arXiv,** 2304.02643, 05 abr. 2023.

LARKIN, A; GU, X; CHEN, L; HYSTAD, P. Predicting perceptions of the built environment using GIS, satellite and street view image approaches. **Landscape and urban planning**, v. 216, p. 104257, 2021.

LASSITER, A; DARBARI, M. Assessing alternative methods for unsupervised segmentation of urban vegetation in very high-resolution multispectral aerial imagery. **PLoS One**, v. 15, n. 5, p. e0230856, 2020.

LATEEF, F; RUICHEK, Y. Survey on semantic segmentation using deep learning techniques. **Neurocomputing**, v. 338, p. 321-348, 2019.

LEE, S.Y; LE, T.H. M.; KIM, Y-M. Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches. **Developments in the Built Environment**, v. 13, p. 100109, 2023.

LI, X; RATTI, C; SEIFERLING, I. Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. **Landscape and Urban Planning**, v. 169, p. 81-91, 2018.

LIU, X; SONG, L; LIU, S; ZHANG, Y. A review of deep-learning-based medical image segmentation methods. **Sustainability**, v. 13, n. 3, p. 1224, 2021.

LIVESLEY, S.J.; MCPHERSON, E.G.; CALFAPIETRA, C. The urban forest and ecosystem services: Impacts on urban water, heat, and pollution cycles at the tree, street, and city scale. **Journal of Environmental Quality**, v.45, p.119–124, 2016.

MAZUROWSKI, M.A.; DONG, H.; GU, H.; YANG, J.; KONZ, N.; ZHANG, Y. Segment anything model for medical image analysis: an experimental study. **Medical Image Analysis**, p. 102918, 2023.

META AI. **Segment Anything: Research by Meta AI**, 2023. Disponível em: https://segment-anything.com/.

O'BRIEN, L; VREESE, R.D; KERN, M; SIEVÄNEN, T; STOJANOVA, B; ATMIŞ, E. Cultural ecosystem benefits of urban and peri-urban green infrastructure across different European countries. **Urban Forestry & Urban Greening**, v. 24, p. 236-248, 2017.

PAULEIT, S.; LIU, L.; AHERN, J.; KAZMIERCZAK, A. Multifunctional green infrastructure planning to promote ecological services in the city. *In*: Niemela, J.; Breuste, J.; Guntenspergen, G.; McIntyre, N.; Elmqvist, T.; James, P. (ed.). **Urban Ecology: Patterns, Processes, and Applications**. Oxford University Press: Oxford, UK, 2011. p. 272–286.

PIRALILOU, S. T; SHAHABI, H; JARIHANI, B GHORBANZADEH, O; BLASCHKE, T; GHOLAMNIA, K; MEENA. S.R; ARYAL, J. Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas. **Remote Sensing**, v. 11, n. 21, p. 2575, 2019.

RAEI, E; ASANJAN, A. A; NIKOO, M. R; SADEGH, M; POURSHAHABI, S; ADAMOWSKI, J. F. A deep learning image segmentation model for agricultural irrigation system classification. **Computers and Electronics in Agriculture**, v. 198, p. 106977, 2022.

WELLMANN, T; LAUSCH, A; ANDERSSON, E; KNAPP, S; CORTINOVIS, C; JACHE, J; SCHEUER, S; KREMER, P; MASCARENHAS, A; KRAEMER, R; HAASE, A; SCHUG, F; HAASE, D. Remote sensing in urban planning: Contributions towards ecologically sound policies? **Landscape and urban planning**, v. 204, p. 103921, 2020.

YIGITCANLAR, T.; KANKANAMGE, N.; REGONA, M.; MALDONADO, A.R.; ROWAN, B.; RYU, A.; DESOUZA, K. C.; CORCHADO, J.M.; MEHMOOD, R.; LI, R.Y.M. Artificial intelligence technologies and related urban planning and development concepts: How are they perceived and utilized in Australia? **Journal of Open Innovation: Technology, Market, and Complexity**, v. 6, n. 4, p. 187, 2020.

CAPÍTULO 3 - SEGMENTATION OF POTENTIALLY INVASIVE TREE SPECIES IN URBAN AREA USING THE SEGMENT ANYTHING MODEL (SAM) AND PANORAMIC IMAGES OF A CITY IN BRAZIL

Abstract: The introduction of invasive tree species disrupts ecological balance through habitat modification and alteration of nutrient cycles by competing with native species. Leucaena leucocephala, originating from Central America and widespread across multiple countries, exhibits rapid growth, rendering it potentially invasive in regions where it has been introduced. This underscores the necessity for ongoing surveillance of this species. While deep learning algorithms using remote sensing data are cutting-edge for mapping tree species in urban settings, their application to Leucaena remains untested. This species poses identification challenges due to its non-distinct visual features such as leaf and canopy structures, which may be indistinguishable in aerial or satellite imagery. This study explores the use of panoramic imagery to map Leucaena in urban areas. This study utilized a recent deep learning foundation model, the Segment Anything Model (SAM), to detect Leucaena species in panoramic Street View and mobile phone images. Performance metrics for SAM indicated a significant discrepancy between the two image types: Street View images achieved 0.92 accuracy, whereas mobile phone images scored 0.63 in identifying Leucaena. The superior performance with Street View images may be attributed to their higher quality, consistency, and better coverage of relevant features, facilitating more effective model analysis. In contrast, mobile phone images often suffer from inconsistencies in angles, lighting, and other environmental factors, which impair model performance. Employing panoramic Street View imagery and deep learning models not only supports urban planning and management but also streamlines data analysis, reducing the need for expensive and labor-intensive field visits.

Keywords: *Leucaena leucocephala*; Street View images; semantic segmentation; tree mapping.

INTRODUCTION

The presence of trees in urban areas plays a fundamental role in the quality of life and contributes to several factors in cities. Among the advantages of having tree vegetation, it is possible to mention air quality improvement, temperature control, biodiversity increase, as well as well-being promotion and urban aesthetics improvement (Wood; Esaian, 2020; Shi *et al.*, 2023a). One key factor for efficient urban vegetation planning is the appropriate selection of tree species, opting mainly for native one. Native trees are adapted to local soil, climate, and fauna conditions, and provide essential environmental and ecological benefits (Delavaux *et al.*, 2023; Sacco *et al.*, 2021; Liu; Slik, 2022). The introduction of exotic species into urban areas can lead to negative consequences, as these species have the potential to become invasive, competing with native one for resources and space. As they quickly establish and reproduce, the invasive species can suppress native vegetation, decreasing biological diversity and altering local ecosystems (Dyderski; Jagodziński, 2020; Delavaux *et al.*, 2023).

The Leucaena leucocephala is a tree species native from Mexico and Central America and is easily found in several countries. In places where it is considered exotic, the Leucaena leucocephala is classified as a potentially invasive species and brings several consequences to the environment (Kato-Noguchi; Kurniadie, 2022; Sharma *et al.*, 2022; Obiakara; Olubode; Chukwuka, 2023). The Leucaena leucocephala was listed among the 100 worst invasive alien species in the world, highlighting the importance of identifying and monitoring this species in urban areas to maintain ecological balance, protect natural resources and promote healthier environments (Kato-Noguchi; Kurniadie, 2022; Sharma *et al.*, 2022).

Remote sensing images play a crucial role in environmental studies by enabling extensive and continuous monitoring of ecosystems. This technology facilitates data collection in remote or hard-to-reach areas, significantly contributing to environmental conservation and the planning of sustainable natural resource use (Jensen, 2014). Remote sensing datasets allow for the effective identification of climate change, deforestation, and other environmental impacts on a large scale (Wellmann *et al.*, 2020; Song *et al.*, 2020). For mapping tree vegetation in urban areas efficiently, remote sensing provides accurate and up-to-date data that assist in urban planning and the creation of more efficient and sustainable green spaces (Wan *et al.*, 2021).

Different platforms can be used to capture remote sensing data, such as orbital, aerial, and terrestrial platforms. Orbital and aerial platforms produce nadir (or off-nadir) images that offer a top-down view, ideal for comprehensive mapping and analysis. Terrestrial platforms, on the other hand, can provide perspective images that support detailed and angular views, which are crucial for in-depth local studies and specific applications. In this context, various types of imagery have been employed to assess the effectiveness of remote sensing techniques in different applications. The study of Iqbal, Balzter and Shabbir (2023), analyzed the ability of machine learning algorithms to map invasive species, including Leucaena leucocephala in different types of images: PlanetScope (3 meters of spatial resolution) and Sentinel (10 meters of spatial resolution). They achieved 64% accuracy with the random forest model applied to PlanetScope imagery.

There is a significant difference in perspective between a street-level view and an aerial or orbital-level view (Xia; Yabuki; Fukuda, 2021). For effective mapping, the use of aerial or satellite imagery requires that the target of interest display distinct visual characteristics, such as color, texture, size, shape, and pattern, which distinguish it from surrounding elements in the images. The tree species Leucaena leucocephala presents a challenge for remote sensing using orthogonal imagery due to its non-distinctive visual features like leaf and canopy structures, which can blend into the surrounding vegetation when viewed from above. Consequently, panoramic street view images are an alternative for collecting geospatial data on this species, as these images offer a more detailed and ground-level perspective that can better capture the distinct characteristics of individual trees. Street view images cover urban areas and can be useful for a variety of applications, such as route mapping, visualization of specific locations, urban planning, and to offering a virtual experience of urban streets and resembling the vision of individuals (Li *et al.*, 2022; Liang; Zhao; Biljecki, 2023; Xia; Yabuki; Fukuda, 2021). The potential of street view images for mapping plant species was evaluated by Ringland *et al.* (2021). They adopted a deep learning model, the RetinaNet network, to detect on Google Street View images six plant species and showed significant results for carrying out large-scale research. Another study (Branson *et al*., 2018) proposed an automated system for tree detection and species recognition in urban environments, using publicly available aerial and street view images from Google Maps. Experiments conducted in the USA show that the automated system based on convolutional neural networks can detect >70% of the street trees and assign the correct species to >80% for 40 different species.

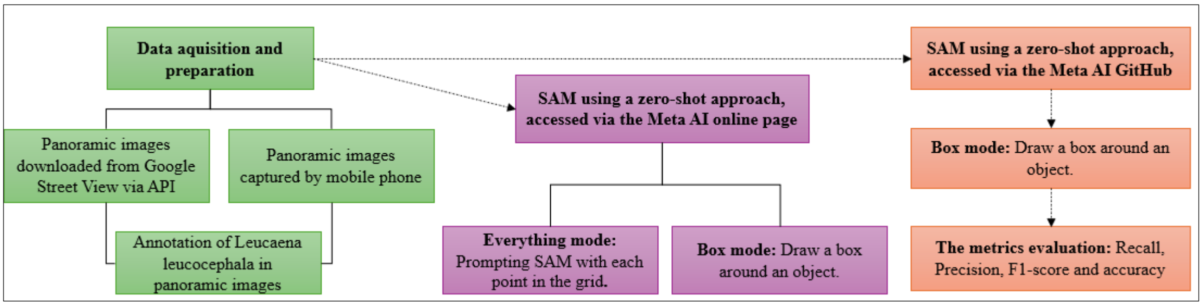
Machine learning algorithms, both shallow and deep learning, represent the state-of-the-art in processing remote sensing data for various applications, including environmental studies. One key application in this field is computer vision, where an important task is semantic segmentation, which involves assigning each pixel of an image to known classes. In remote sensing area, this task is referred to as image classification (Osco *et al.* 2021). A recent and promising deep learning model for image segmentation is the Segment Anything Model (SAM). Released by Meta AI (2023), SAM is driven by prompts, and it has already shown remarkable adaptability across various image datasets without the need for retraining on new objects (Kirillov *et al.*, 2023; Zhang; Shen; Jiao, 2024). SAM can be effectively utilized with minimal training data (one-shot or few-shot) or no training data (zero-shot) using a variety of input prompts, including point, box, or text (Osco *et al.*, 2023). In the zero-shot approach, the model processes input data that it has not explicitly encountered during training (Alayrac *et al*., 2022; Sun *et al.*, 2021). Conversely, the one-shot and few-shot learning approaches enable the model to make accurate inferences from just a single example or a few examples of a new class, respectively (Osco *et al*., 2023; Zhang *et al*., 2023). However, the one-shot and few-shot learning approaches require a retraining phase for the SAM model, which can demand significant computational power and might take a longer time (Osco *et al*., 2023).

The capabilities of SAM have been explored in various fields, including medical imaging (Zhang; Shen; Jiao, 2024; Huang *et al.*, 2024; Mazurowski *et al.*, 2023) and remote sensing (Wang *et al.*, 2024; Osco *et al.*, 2023; Ding *et al.*, 2023; Alagialoglou *et al.*, 2023). One study conducted by Alagialoglou *et al*. (2023) compared the efficiency of different shallow learners (random forest and logistic regression) and the Segment Anything Model within the few-shot learning regime for mapping underwater aquatic vegetation using orbital images (WorldView-2 and Sentinel-2). They noted that it remains a challenge for SAM to outperform shallow learners in segmentation tasks with lower spatial resolution images. In addition to this, the exploration of panoramic images captured by ground-based platforms for identifying the tree species in urban environments using the advanced SAM represents a promising area for exploration. Once the effectiveness of this approach is validated, it will provide a low-cost method for accurate and detailed detection of the tree species, which is crucial for urban management and environmental planning. In this context, this study aims to evaluate the performance of the Segment Anything Model using a zero-shot approach to segment Leucaena leucocephala in panoramic images from Google Street View. Additionally, tests with mobile phone images were included to determine whether the source of the collected data affects the model's ability to segment the species of interest.

MATERIALS AND METHOD

Figure 1 shows the flowchart with the steps taken in this study to verify the efficiency of the SAM zero-shot in segmenting the tree species Leucaena leucocephala in urban panoramic images.

Figure 1- A Flowchart with the steps performed in this study to verify the performance of SAM zero-shot to segment the tree species Leucaena leucocephala in panoramic images.

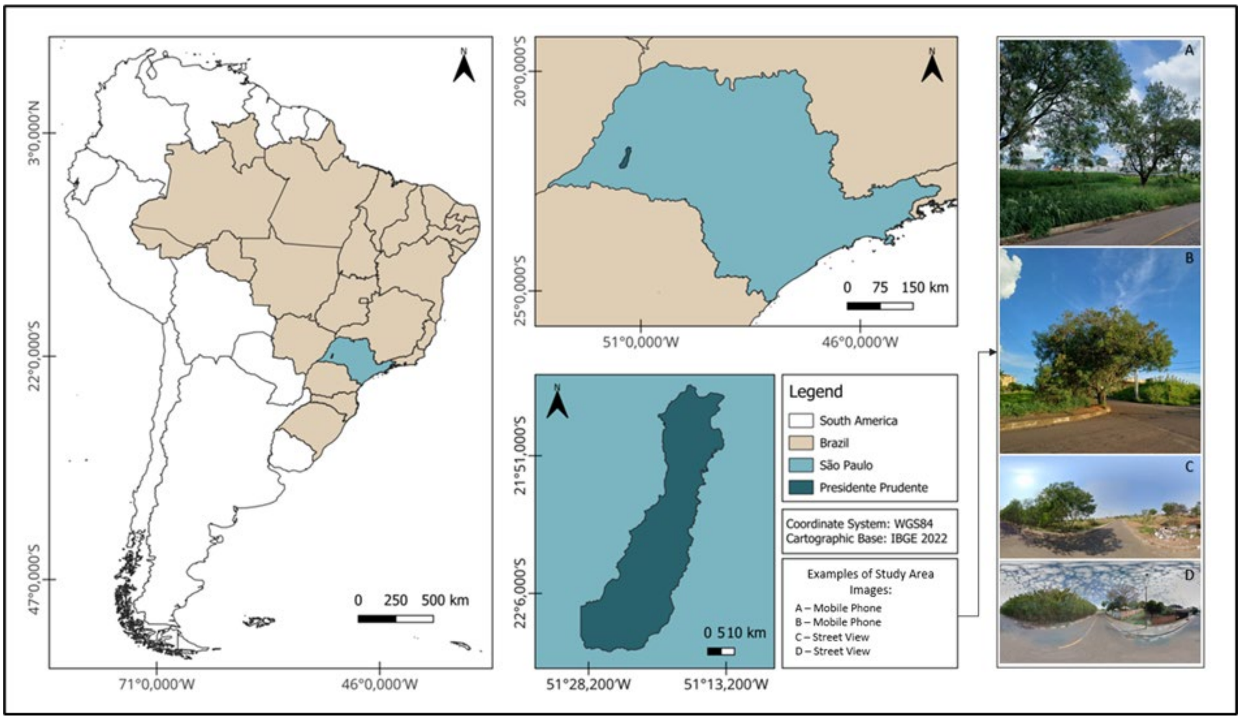


Fonte: Author (2024).

Study Area

To analyze the performance of SAM zero-shot, panoramic images of the city of Presidente Prudente (Figure 2), in the state of São Paulo, in Brazil (Latitude 22º07'21.06''S and Longitude 51º23' 17.71''W) were collected. Presidente Prudente is a medium-sized city with an estimated population of 225,668 in 2022 (IBGE, 2022). This city has an area of approximately 230.05 km², including the urban and rural perimeter (Furuya *et al.*, 2023) and 60.83 km² of urbanized area (IBGE, 2019).

Figure 2- Location of the city of Presidente Prudente, in the state of São Paulo, Brazil. The Figure shows examples of collected images where the presence of the Leucaena leucocephala tree species was identified in the city.



Fonte: Author (2024).

Collection and Annotation of Panoramic Images for Leucaena leucocephala Segmentation Study

A total of 20 panoramic images (3328 x 1664) from Google Street View and 20 panoramic images (3000 x 3000) from mobile phone were collected at points in the city where the presence of the Leucaena leucocephala species had previously been detected by the authors. Street View images come from a Google mapping platform and are obtained using cars equipped with 360-degree cameras. Street View images are made available by Google in different years and are constantly updated in urban areas around the world (Li; Yao, 2020). The panoramic images were downloaded from Google Street View via API for the city of Presidente Prudente covering the years 2021 and 2022, as they are the most recent years with available data. Regarding mobile phones images, there were taken by the authors with their mobile devices in 2023. The smartphone images were captured under varied conditions to enhance data diversity, including different dates and times of the day to capture a range of lighting conditions, and various distances to Leucaena leucocephala trees. The camera was orientated to vertical position and set up to 48 megapixels resolutions.

Each of the 40 collected images was annotated to enable the calculation of evaluation metrics, facilitating an assessment of the performance of the SAM zero-shot model after it performed image segmentation. Image annotation was manually done in the LabelMe software. The annotation process is crucial for successful analysis and pattern recognition. It involves identifying and grouping the pixels that constitute an object within the image, thereby forming distinct regions for further analysis (Jasim; Mohammed, 2021; Wu; Castleman, 2023). A common method used is polygon labeling, in which the contours of an object are traced with geometric shapes. This approach simplifies both the description and analysis of the object (Ibrahim; El-Kenawy, 2020). This technique provides an efficient and compact representation of the object, enhancing feature extraction and decision-making in image processing and computer vision tasks. Consequently, the Leucaena leucocephala trees were labeled using polygon shapes in the LabelMe, as illustrated in Figure 3.

Figure 3- Examples of polygon made in the LabelMe software for the 40 panoramic images collected delimiting the tree species Leucaena leucocephala (street view images on the left and mobile phone images on the right).

Fonte: Author (2024).

Assessment of Zero-shot SAM for Panoramic Image Segmentation

On the online Meta AI page (https://segment-anything.com/demo), it is possible to access the SAM in a zero-shot regime using various prompts, such as point, boundary box, and general, to segment images available on the Meta AI website or images uploaded by the user. In the general case, referred to by Meta AI as 'everything,' the entire image is segmented prompting SAM with each point in the grid. However, when using the point or box prompts, the model segments only the regions indicated by these elements (Osco *et al*., 2023).

Firstly, this study evaluates the performance of the zero-shot SAM available on the Meta AI online page. Tests were conducted using all images collected in this study, considering both the general case and the box prompt (Figure 4). The purpose of employing the zero-shot SAM with a general prompt was to assess its ability to distinguish arboreal vegetation from other elements in panoramic images. Subsequently, this study evaluates the SAM's performance using a box prompt to specifically identify Leucaena species among other trees in the image.

However, on the Meta AI website, it is not possible to perform a quantitative analysis of SAM's performance, only qualitative (visual) results. To conduct both qualitative and quantitative analyses with SAM zero-shot, the SAM zero-shot model via the Meta AI GitHub was downloaded and adapted it according to the method described by Osco *et al.* (2023). This experiment aimed to segment all panoramic images using the SAM zero-shot model guided by box prompts on the Google Colaboratory (Colab) platform and to quantitatively analyze the model's performance in segmenting the species Leucaena leucocephala. The adaptation involved reading files in JSON format (polygon labeling format created in LabelMe) and calculating evaluation metrics.

Two inputs were required for this process: the image in PNG or JPG format (from Street View or a mobile phone) and a JSON file containing the labeling/delimitation of the Leucaena leucocephala species in polygon format created in LabelMe. Using these JSON files, it was possible to compare the segmentation performed by SAM and calculate the evaluation metrics.

Evaluation Metrics

For each of the 40 panoramic images, four different evaluation metrics were calculated: Accuracy, Precision, Recall and F1-score (Table 1). These metrics are commonly used in studies involving remote sensing and artificial intelligence, particularly in semantic segmentation tasks (Zhang *et al.*, 2020; Maxwell; Pourmohammadi; Poyner, 2020; Mahmoud *et al.*, 2023; Carvalho *et al.*, 2022; Osco *et al.*, 2022; Arce *et al.*, 2021).

Table 1- Evaluation metrics used to evaluate SAM performance in Google Colaboratory code.

METRIC	**EQUATION**
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F1-score	$F1 - score = \frac{2\,(Precision * Recall)}{Precision + Recall}$

Where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative (Minaee *et al.*, 2021).

The accuracy metric measures the proportion of examples correctly classified by the model regarding the total number of examples (Hicks *et al.*, 2022; Gencturk *et al.*, 2024). Precision indicates the proportion of positive examples correctly classified regarding the total number of examples classified as positive by the model (Gencturk *et al.*, 2024). The recall represents the proportion of correctly classified positive examples regarding the total number of truly positive examples (Gencturk *et al.*, 2024). Finally, the F1-score is a metric that combines precision and recall, providing a single measure that takes both metrics into account, being especially useful when there is an imbalance between classes (Gencturk *et al.*, 2024).

RESULTS

Figure 4 shows examples of the qualitative results obtained with the SAM zero-shot model on the Meta AI website using different prompts. For the box prompt (Figure

4b), the model correctly segmented the tree species Leucaena leucocephala, even when multiple trees of the species were included in the box or the presence of a streetlamp in front of the species of interest. In the general case, i.e, the segmentation of entire images (Figure 4a), despite the model segmenting parts of the sky, street, and sidewalk into different classes and colors, the model accurately delimited the trees present in the images (without specifying the species).

Following these initial online tests, further tests were conducted with the SAM zero-shot model using code in Google Colaboratory. It is important to highlight that the tests were conducted exclusively with the box prompt to guide the SAM. Evaluation metrics were calculated for each of the 40 images, and Table 2 shows their averages.

A superior performance of SAM was observed in segmenting panoramic images from the Google Street View platform compared to mobile phone panoramic images (Table 2). The average accuracy and F1-score were 29% and 27% higher, respectively, for SAM zero-shot applied to Street View images. Figure 5 provides visual examples of the segmentation performed by the SAM zero-shot model using code in Google Colaboratory.

In the visual examples (Figure 5), it is evident that the model was able to correctly segment the species of interest, including cases where more than one tree of interest was present in the image. Another notable detail is the images where Leucaena leucocephala was present in areas with other species or types of vegetation, such as shrubs and grasses; the model was also able to correctly segment the species.

For mobile phone image segmentation, some metric results were around 0.35 (Figure 6). When delimiting the box where the tree species Leucaena leucocephala was in the image, parts of other objects, such as buildings and cars, were sometimes segmented as the species of interest. These difficulties encountered by the SAM zero-shot in segmenting the Leucaena leucocephala species, as shown in Figure 6, could be due to the presence of complex backgrounds and overlapping objects like buildings and other vegetation. These elements likely caused the model to misclassify portions of the image, leading to lower accuracy and F1-score metrics compared to those obtained from Street View images (Table 2).

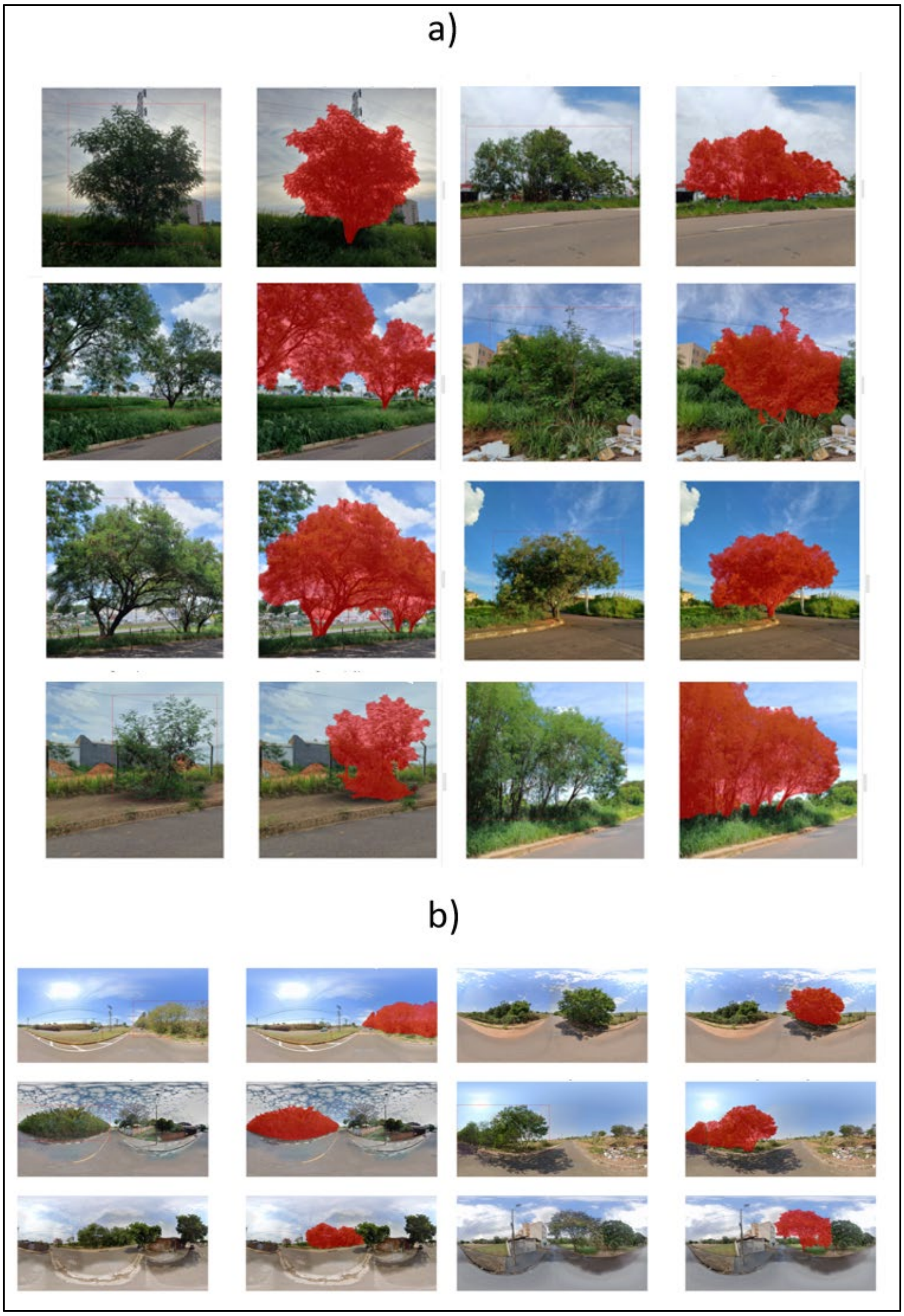
Figure 4- Examples of segmentation performed by SAM zero-shot on the Meta AI website with panoramic images. The first four images (up-down) are from mobile phone captures, while the last four are from Street View. (a) Segmentation of the entire image; (b) Segmentation of the object of interest (Leucaena leucocephala) within the bounded box.

Fonte: Author (2024).

Table 2- Average of the evaluation metrics calculated using 40 panoramic images segmented by SAM zero-shot using the box prompt.

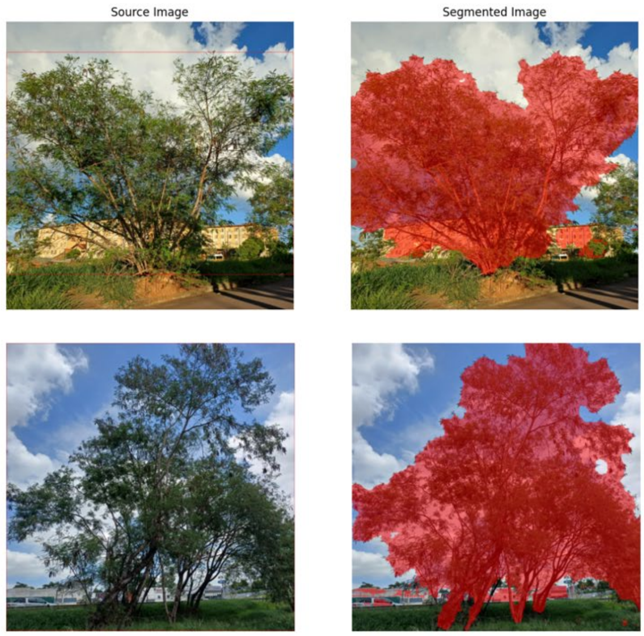
	Street View Images (20 Images)	Mobile Phone Images (20 Images)	All Images (40 Images)
Accuracy	0.9263	0.6341	0.7763
F1-score	0.9394	0.6618	0.7969
Precision	0.9540	0.6940	0.8205
Recall	0.9263	0.6341	0.7763

Figure 5- Segmentation (in red) of the Leucaena leucocephala tree species performed by SAM zero-shot code on Google Colaboratory using the box prompt. (a) Mobile phone images. (b) Street View images.



Fonte: Author (2024).

Figure 6- Examples of difficulties encountered by the SAM zero-shot in segmenting the Leucaena leucocephala species, where parts of buildings and other objects were incorrectly segmented as trees, resulting in metrics of 0.30 and 0.40.



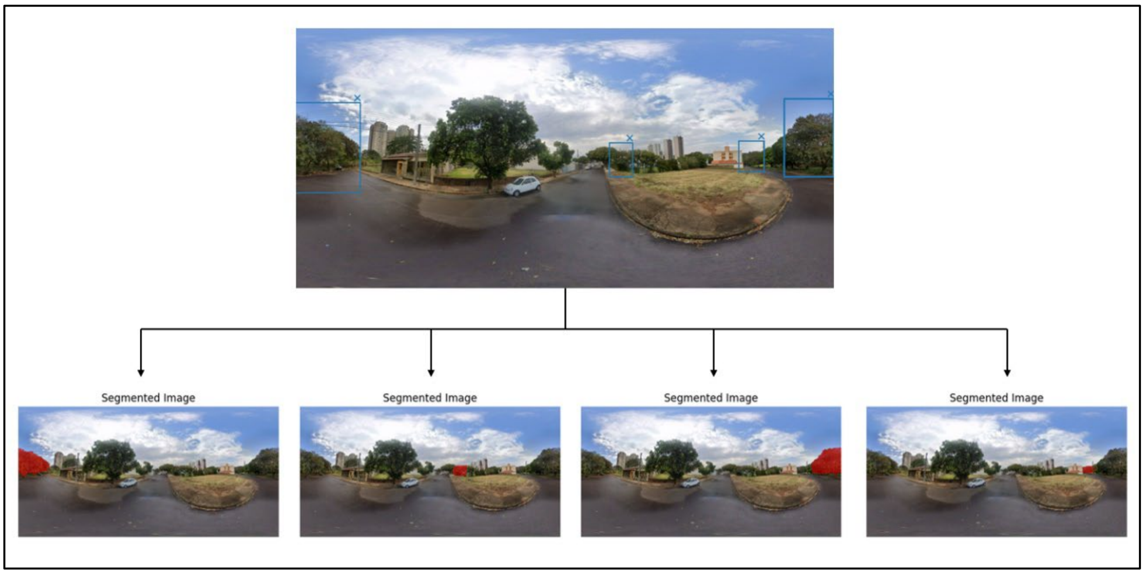
Accuracy	0.35
F1-score	0.41
Precision	0.48
Recall	0.35

Accuracy	0.30
F1-score	0.32
Precision	0.36
Recall	0.30

Fonte: Author (2024).

Another observed case involved images containing the Leucaena leucocephala species in different parts of the image, requiring two or more boxes to be delimited for segmentation with SAM. In these instances, SAM returned a segmented image for each box, as shown in Figure 7. The processing time for segmentation on panoramic Street View and mobile phone images with SAM zero-shot averaged five minutes, depending on the number of boxes. Although some images may take a few additional minutes for segmentation, the model can quickly segment the requested boxes and significantly contribute to the identification of Leucaena leucocephala in urban panoramic images.

Figure 7- Example of a panoramic image with more than one box containing the species Leucaena leucocephala. SAM zero-shot segmented each box into an image separately.



Fonte: Author (2024).

DISCUSSION

The use of artificial intelligence has proven fundamental in segmenting images representing urban areas (Sohail *et al.*, 2022; Boonpook *et al.*, 2023; Soylu *et al.*, 2023). These models allow the identification and delimitation of different elements, including types of vegetation, and provide a more accurate and detailed view of the urban environment (Li *et al.*, 2022; Xia; Yabuki; Fukuda, 2021). Through image analysis, the artificial intelligence can distinguish specific characteristics and segment different types of objects, facilitating urban planning, environmental monitoring, and resource management (Xia; Yabuki; Fukuda, 2021).

Our results with the foundational model ‘Segment Anything Model’ (SAM) applied to Google Street View images for segmenting Leucaena leucocephala validate the effectiveness of the proposed approach in providing a low-cost method for accurately identifying tree species in urban areas, which is crucial for urban management and environmental planning. Street View images offer the advantage of being accessible in many cities worldwide. The box prompt proved to be an effective option to guide SAM zero-shot, consistent with findings from other studies (Carraro; Sozzi; Marinello, 2023; Mazurowski *et al*., 2023).

Despite presenting innovative and accurate potential, SAM presents some obstacles in image segmentation (Alagialoglou *et al.*, 2023; Baziak; Bodziony; Szczepanek, 2024), including in the tests carried out in this study with SAM zero-shot. Situations like those illustrated in Figure 6 demonstrate that the model is not error-free. SAM presented limitations in several areas (Ji *et al.*, 2023; Kirillov *et al.*, 2023), including remote sensing, as it was trained mainly with RGB images and may have difficulty with multispectral or hyperspectral data (Osco *et al.*, 2023). The study by Baziak, Bodziony, and Szczepanek (2024) applied SAM zero-shot to the field of hydrology. According to these authors, the advantages of SAM are its simplicity, low cost, and scalability. They conducted tests using both the online version Meta AI (2023), website and code on Colab, encountering some limitations. Baziak, Bodziony, and Szczepanek (2024) argue that, despite being an innovative model, SAM has limitations when using standard parameters, resulting in unsatisfactory outcomes. In the medical field, studies have also shown that SAM has limitations (Ma *et al.*, 2024; Shi *et al*., 2023b), such as modality imbalance in the training set and ambiguity in segmentation. One of the challenges encountered by SAM in previous studies is the presence of shadows in the images that can confuse the model and segment an object as being another class. The study by Kim and Kim (2024), used SAM and spectral index-based labeling to increase accuracy in an unsupervised environment to segment buildings. The authors found that despite presenting superior F1-score and Intersection over Union (IoU), SAM presented errors when there were shadows in the images, which corresponds to a challenge that needs to be solved for the foundational model.

Despite having limitations, SAM represents an innovative approach in image segmentation and has the potential to optimize tasks in several areas (Osco *et al.*, 2023; Baziak; Bodziony; Szczepanek, 2024; Ma *et al.*, 2024; Shi *et al.*, 2023b). SAM promotes the development of basic models for computer vision, shows good generalization performance, and can reduce data annotation (Zhang *et al.*, 2023; Yan *et al.*, 2023; Osco *et al.*, 2023).

The SAM zero-shot represents a significant advance in images (Osco *et al.*, 2023) including in the identification of tree species, such as Leucaena leucocephala, a potentially invasive species common in many tropical and subtropical regions. The combination of urban image segmentation with the SAM in the identification of tree species allows the identification of priority green areas for conservation and the development of effective strategies for managing and preserving urban environments in constant transformation.

CONCLUSION

The Segment Anything Model (SAM) zero-shot, introduced by Meta AI (2023), demonstrates effective performance in segmenting the tree species Leucaena leucocephala in panoramic images from Google Street View and mobile phones. The model achieved higher accuracy with Street View images compared to mobile phone images. Although SAM zero-shot has some limitations, its advantage lies in not requiring additional training for unknown objects. This makes SAM a valuable tool for environmental management and urban planning to monitor Leucaena leucocephala and assess its potential impact as an invasive species. Future research should explore applying SAM to satellite and aerial images and consider retraining the model to improve its accuracy and reduce limitations.

REFERENCES

ALAGIALOGLOU, L.; MANAKOS, I.; PAPADOPOULOU, S.; CHADOULIS, R. T. The Case of Polyphitos Lake. **Remote Sensing**, v. 15, n. 16, p. 4001, 2023

ALAYRAC J.-B., DONAHUE J., LUC P., MIECH A., BARR I., HASSON Y., LENC K., MENSCH A., MILLICAN K., REYNOLDS M., RING R., RUTHERFORD E., CABI S., HAN T., GONG Z., SAMANGOOEI S., MONTEIRO M., MENICK J., BORGEAUD S., BROCK A., NEMATZADEH A., SHARIFZADEH S., BINKOWSKI M., BARREIRA R., VINYALS O., ZISSERMAN A., SIMONYAN K. Flamingo: a visual language model for few-shot learning, **arXiv:2204.14198,** 2022.

ARCE, L. S. D.; OSCO, L. P.; ARRUDA, M. D. S. D.; FURUYA, D. E. G.; RAMOS, A. P. M.; AOKI, C.; POTT, ARNILDO.; FATHOLAHI, SARAH.; LI, JONATHAN.; ARAUJO, F. F. D.; GONÇALVES, W. N.; JUNIOR, J. M. Mauritia flexuosa palm trees airborne mapping with deep convolutional neural network. **Scientific Reports**, v. 11, n. 1, p. 19619, 2021.

BAZIAK, B.; BODZIONY, M.; SZCZEPANEK, R. Mountain Streambed Roughness and Flood Extent Estimation from Imagery Using the Segment Anything Model (SAM). **Hydrology,** v. 11, n. 2, p. 17, 2024.

BOONPOOK, W.; TAN, Y.; NARDKULPAT, A.; TORSRI, K.; TORTEEKA, P.; KAMSING, P.; SAWANGWIT, U.; PENA, J.; JAINAEN, M. Deep learning semantic segmentation for land use and land cover types using Landsat 8 imagery. **ISPRS International Journal of Geo-Information**, v. 12, n. 1, p. 14, 2023.

BRANSON, S.; WEGNER, J. D.; HALL, D.; LANG, N.; SCHINDLER, K.; PERONA, P. from Google Maps to a fine-grained catalog of street trees. In: **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 135, p. 13-30, 2018.

CARRARO, A.; SOZZI, M.; MARINELLO, F. The Segment Anything Model (SAM) for accelerating the smart farming revolution. **Smart Agricultural Technology**, v. 6, p. 100367, 2023.

CARVALHO, M. D. A.; JUNIOR, J. M.; MARTINS, J. A. C.; ZAMBONI, P.; COSTA, C. S.; SIQUEIRA, H. L.; ARAUJO, M. S.; GONÇALVES, D. N.; FURUYA, D. E. G.; OSCO, L. P.; RAMOS, A. P. M.; LI, J.; JUNIOR, A. A. D. C.; GONÇALVES. W. N. A deep learning-based mobile application for tree species mapping in RGB images. **International Journal of Applied Earth Observation and Geoinformation**, v. 114, p. 103045, 2022.

DELAVAUX, C. S.; CROWTHER, C. M.; ZOHNER, C. M.; ROBMANN, N. M.; LAUBER, T.; HOOGEN, J. V. D.; KUEBBING, S.; LIANG, J.; DE MIGUEL, S.; NABUURS, G. J.; REICH, P. B.; ABEGG, M.; YAO, Y. C. A.; ALBERTI, G.; ZAMBRANO, A. M. A.; ALVORADA, V. B.; ALVAREZ-DAVILA, ESTEBAN.;

ALVAREZ-LOAYZA, PATRICK.; ALVEZ, L. F.; AMMER, C.; ANTONFERNANDEZ, C.; ARAUJO-MURAKAMI, A.; ARROYO, L.; AVITABILLE, V.; MAYNARD, D. S. Native diversity buffers against severity of non-native tree invasions. **Nature**, v. 621, n. 7980, p. 773-781, 2023.

DING, L.; ZHU, K.; PENG, D.; TANG, H.; YANG, K.; BRUZZONE, L. Adapting Segment Anything Model for Change Detection in VHR Remote Sensing Images. **IEEE Transactions on Geoscience and Remote Sensing**, 2024.

DYDERSKI, M. K.; JAGODZIŃSKI, A. M. Impact of invasive tree species on natural regeneration species composition, diversity, and density. **Forests**, v. 11, n. 4, p. 456, 2020.

FURUYA, M. T. G.; FURUYA, D. E. G.; OLIVEIRA, L. Y. D. E.; SILVA, P. A. D.; CICERELLI, R. E.; GONÇALVES, W. N.; JUNIOR, J. M.; OSCO, L. P.; RAMOS, A. P. M. A machine learning approach for mapping surface urban heat island using environmental and socioeconomic variables: a case study in a medium-sized Brazilian city. **Environmental Earth Sciences**, v. 82, n. 13, p. 325, 2023.

GENCTURK, B. ARSOY, S.; TASPINAR, Y. S.; CINAR, I.; KURSUN, R.; YASIN, E. T.; KOKLU, M. Detection of hazelnut varieties and development of mobile application with CNN data fusion feature reduction-based models. **European Food Research and Technology**, v. 250, n. 1, p. 97-110, 2024.

HICKS, S.; SRUMKE, I.; THAMBAWITA, V.; HAMMOU, M.; RIEGLER, M. A.; HALVORSEN, P.; PARASA, S. On evaluation metrics for medical applications of artificial intelligence. **Scientific reports**, v. 12, n. 1, p. 5979, 2022.

HUANG, Y.; HE, Y.; LI, F.; HAN, L.; YOU, C.; WANG, B. Segment anything model for medical images? **Medical Image Analysis**, v. 92, p. 103061, 2024.

IBGE. Brazilian Institute of Geography and Statistics. **População Presidente Prudente 2019**. 2019. Disponível em: https://cidades.ibge.gov.br/brasil/sp/presidente-prudente/panorama. Acesso em: 28 mar. 2024.

IBGE. Brazilian Institute of Geography and Statistics. **População Presidente Prudente 2022**. 2022. Disponível em: https://cidades.ibge.gov.br/brasil/sp/presidente-prudente/panorama. Acesso em: 28 mar. 2024.

IBRAHIM, A.; EL-KENAWY, E-S. M. Image segmentation methods based on superpixel techniques: A survey. **Journal of Computer Science and Information Systems**, v. 15, n. 3, p. 1-11, 2020.

IQBAL, I. M.; BALZTER, H.; SHABBIR, A. Mapping Lantana camara and Leucaena leucocephala in protected areas of Pakistan: a geo-spatial approach. **Remote Sensing**, v. 15, n. 4, p. 1020, 2023.

JASIM, W. N.; MOHAMMED, R. A Survey on Segmentation Techniques for Image Processing. **Iraqi Journal for Electrical & Electronic Engineering**, v. 17, n. 2, 2021.

JENSEN, R. J. Remote Sensing of the Environment: An Earth Resource Perspective. 2. ed. Harlow: **Pearson New International Edition**, Pearson Education Limited Edinburgh Gate, p. 619, 2014.

JI, W.; LI, J.; BI, Q.; LIU, T.; LI, W.; CHENG, L. Segment anything is not always perfect: An investigation of sam on different real-world applications. **arXiv preprint** arXiv:2304.05750, 2023.

KATO-NOGUCHI, H.; KURNIADIE, D. Allelopathy and allelochemicals of Leucaena leucocephala as an invasive plant species. **Plants**, v. 11, n. 13, p. 1672, 2022.

KIM, J.; KIM, Y. Integrated Framework for Unsupervised Building Segmentation with Segment Anything Model-Based Pseudo-Labeling and Weakly Supervised Learning. **Remote Sensing**, v. 16, n. 3, p. 526, 2024.

KIRILLOV, A.; MINTUN, E.; RAVI, N.; MAO, H.; ROLLAND, C.; GUSTAFSON, L.; XIAO, T. Segment anything. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision**. p. 4015-4026, 2023.

LI, Y.; PENG, L.; WU, C.; ZHANG, J. Street View Imagery (SVI) in the built environment: A theoretical and systematic review. **Buildings**, v. 12, n. 8, p. 1167, 2022.

LI, M.; YAO, W. 3D map system for tree monitoring in hong kong using google street view imagery and deep learning. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 3, p. 765-772, 2020.

LIANG, X.; ZHAO, T.; BILJECKI, F. Revealing spatio-temporal evolution of urban visual environments with street view imagery. **Landscape and Urban Planning**, v. 237, p. 104802, 2023.

LIU, J.; SLIK, F. Are street trees friendly to biodiversity?. **Landscape and Urban Planning**, v. 218, p. 104304, 2022.

MA, J. H. Y.; LI, F.; HAN, L.; YOU, C.; WANG, B. Segment anything in medical images. **Nature Communications**, v. 15, n. 1, p. 654, 2024.

MAHMOUD, R.; HASSANIN, M.; FELL, H. A.; BADRY, R. M. Machine learning-based land use and land cover mapping using multi-spectral satellite imagery: A case study in Egypt. **Sustainability**, v. 15, n. 12, p. 9467, 2023.

MAXWELL, A. E.; POURMOHAMMADI, P.; POYNER, J. D. Mapping the topographic features of mining-related valley fills using mask R-CNN deep learning and digital elevation data. **Remote Sensing**, v. 12, n. 3, p. 547, 2020.

MAZUROWSKI, M.; DONG, H.; GU, H.; YANG, J.; KONZ, N.; ZHANG, Y. Segment anything model for medical image analysis: an experimental study. **Medical Image Analysis**, v. 89, p. 102918, 2023.

META AI. **Segment Anything: Research by Meta AI**, 2023. Disponível em: https://segment-anything.com/.

MINAEE, S.; BOYKOV, Y.; PORIKLI, F.; PLAZA, A.; KEHTARNAVAZ, N.; TERZOPOULOS, D. Image segmentation using deep learning: A survey. **IEEE transactions on pattern analysis and machine intelligence**, v. 44, n. 7, p. 3523-3542, 2021.

OBIAKARA, M. C.; OLUBODE, O. S.; CHUKWUKA, K. S. Climate change and the potential distribution of the invasive shrub, Leucaena leucocephala (Lam.) De Wit in Africa. **Tropical Ecology**, v. 64, n. 4, p. 698-711, 2023.

OSCO, L. P.; FURUYA, D. E. G.; FURUYA, M. T. G.; CORREA, D. V.; GONÇALVES, W. N.; JUNIOR, J. M.; BORGES, M.; BLASSIOLI-MORAES, M. C.; MICHEREFF, M. F. F.; AQUINO, M. F. S.; LAUMANN, R. A.; LISENBERG, V.; RAMOS, A. P. M.; JORGE, L. A. D. C. An impact analysis of pre-processing techniques in spectroscopy data to classify insect-damaged in soybean plants with machine and deep learning methods. **Infrared Physics & Technology**, v. 123, p. 104203, 2022.

OSCO, L. P.; WU, Q.; LEMOS. E. L. D.; GONÇALVES, W. N.; RAMOS, A. P. M.; LI, J.; JUNIOR, J. M. The segment anything model (sam) for remote sensing applications: From zero to one shot. **International Journal of Applied Earth Observation and Geoinformation**, v. 124, p. 103540, 2023.

RINGLAND, J.; BOHM, M.; BAEK, S. R.; EICHHORN, M. Automated survey of selected common plant species in Thai homegardens using Google Street View imagery and a deep neural network. **Earth Science Informatics**, v. 14, p. 179-191, 2021.

SACCO, A. D.; HARDWICK, A. D.; BLACKESLEY, DAVID.; BRANCALION, P. H. S.; BREMAN, E.; REBOLA, L. C.; CHOMBA, S.; DIXON, K.; ELLIOT, S.; RUYONGA, G. Ten golden rules for reforestation to optimize carbon sequestration, biodiversity recovery and livelihood benefits. **Global Change Biology**, v. 27, n. 7, p. 1328-1348, 2021.

SHARMA, P.; BATISH, D.; KAUR, A.; KAUR, S.; CHAUHAN, B. S. Critical insights into the ecological and invasive attributes of Leucaena leucocephala, a tropical agroforestry species. **Frontiers in Agronomy**, v. 4, p. 890992, 2022.

SHI, F.; MENG, Q.; PAN, L.; WANG, J. Root damage of street trees in urban environments: An overview of its hazards, causes, and prevention and control measures. **Science of the Total Environment**, p. 166728, 2023a.

SHI, P.; QIU, J.; ABAXI, S. M. D.; WEI, H.; LO, F. P. W., YUAN, WU. Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation. **Diagnostics**, v. 13, n. 11, p. 1947, 2023b.

SOHAIL, A. NAWAZ, N. A.; SHAH, A. A.; RASHEED, S.; IIYAS, SHEEBA.; EHSAN, M. K. A systematic literature review on machine learning and deep learning methods for semantic segmentation. **IEEE Access**, v. 10, p. 134557-134570, 2022.

SOYLU, B. E.; GUZEL, M. S.; BOSTANCI, G. E.; EKINCI, F.; ASUROGLU, T.; ACICI, K. Deep-learning-based approaches for semantic segmentation of natural scene images: A review. **Electronics**, v. 12, n. 12, p. 2730, 2023.

SONG, W.; SONG, W.; GU, H.; LI, F. Progress in the remote sensing monitoring of the ecological environment in mining areas. **International Journal of Environmental Research and Public Health**, v. 17, n. 6, p. 1846, 2020.

SUN, Y.; WANG, X.; ZHU, J.; CHEN, L.; JIA, Y.; LAWRENCE, J. M.; JIANG, L. H.; XIE, X.; WU, J. Using machine learning to examine street green space types at a high spatial resolution: Application in los angeles county on socioeconomic disparities in exposure. **Science of The Total Environment**, v.787, 147653, 2021.

WAN, H.; TANG, Y.; JING, L.; LI, H.; QIU, F.; WU, W. Tree species classification of forest stands using multisource remote sensing data. **Remote Sensing**, v. 13, n. 1, p. 144, 2021.

WANG, D.; ZHANG, J.; DO, B.; XU, M.; LIU, L.; TAO, D.; ZHANG, L. Samrs: Scaling-up remote sensing segmentation dataset with segment anything model. **Advances in Neural Information Processing Systems**, v. 36, 2024.

WELLMANN, T.; LAUSCH, A.; ARDERSSON, E.; KNAPP, S.; CORTINOVIS, C.; JACHE, J.; SCHEUER, S.; KREMER, P.; MASCARENHAS, A.; KRAEMER, R.; HAASE, A.; SCHUG, F.; HAASE, D. Remote sensing in urban planning: Contributions towards ecologically sound policies? **Landscape and urban planning**, v. 204, p. 103921, 2020.

WOOD, E. M.; ESAIAN, S. The importance of street trees to urban avifauna. **Ecological Applications**, v. 30, n. 7, p. e02149, 2020.

WU, Q.; CASTLEMAN, K. R. Image segmentation. *In*: Microscope Image Processing. **Academic Press**. p. 119-152, 2023.

XIA, Y.; YABUKI, N.; FUKUDA, T. Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. **Urban Forestry & Urban Greening**, v. 59, p. 126995, 2021.

YAN, Z.; LI, J.; LI, X.; ZHOU, R.; ZHANG, W.; FENG, Y.; DIAO, W.; FU, K.; SUN, X.

Ringmo-sam: A foundation model for segment anything in multimodal remote-sensing images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 61, p. 1-16, 2023.

ZHANG, C. LIU, L.; CUI, Y.; HUANG, G.; LIN, W.; YANG, Y.; HU, Y. A comprehensive survey on segment anything model for vision and beyond. **arXiv preprint** arXiv:2305.08196, 2023.

ZHANG, X. HAN, L.; HAN, L.; ZHU, L. How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?. **Remote Sensing**, v. 12, n. 3, p. 417, 2020.

ZHANG, Y.; SHEN, Z.; JIAO, R. Segment anything model for medical image segmentation: Current applications and future directions. **Computers in Biology and Medicine**, p. 108238, 2024.

CAPÍTULO 4 - DETECTION OF POTENTIALLY INVASIVE TREE SPECIES LEUCAENA LEUCOCEPHALA IN PANORAMIC IMAGES WITH DIFFERENT DEEP LEARNING NETWORKS

Abstract: Mapping tree species helps in the management and planning of urban areas. Several exotic species can become invasive and harm the environment including the green areas. Remote sensing and artificial intelligence techniques have shown favorable potential for carrying out and optimizing environmental tasks. In this context, the objective of this study was to evaluate the ability of the YOLOv8 (You Only Look Once) and DETR (Detection Transformer) networks to detect the tree species Leucaena leucocephala in panoramic Street View and mobile phone images. The tests in this study considered panoramic Street View and mobile phone images from the city of Presidente Prudente, with a dataset of 205 Street View images and 217 mobile phone images. In this study, tests were done with three datasets: 1- Street View images; 2- mobile phone images; 3- Street View and mobile phone images (datasets 1 + dataset 2). The average of YOLOv8 was higher in mobile phone images with Recall = 0.75; mAP50 = 0.84 and mAP50-95 = 0.61. The DETR network also obtained superior results with mobile phone images, presenting Recall = 0.68; mAP50 = 0.84 and mAP50-95 = 0.56. Although mobile phone images presented higher metrics, Street View images also showed satisfactory results in tests with YOLOv8 (Recall = 0.74 and mAP50 = 0.83). Despite the limitations found in this study, the results obtained showed that panoramic images can be used to detect the Leucaena leucocephala tree species in urban areas and can contribute to environmental monitoring and planning.

Keywords: Leucaena leucocephala; YOLOv8; Detection Transformer (DETR); Deep Learning; Street View.

INTRODUCTION

The presence of tree vegetation in urban areas plays a fundamental role in quality of life. In addition to the aesthetic issue, trees improve air quality, reduce heat islands and conserve biodiversity (Zheng *et al.*, 2023; Yilmaz; Irmak; Qaid, 2022; Liu; Slik, 2022). To guarantee ecological balance and the advantages provided by tree vegetation, it is necessary to analyze the choice of tree species inserted in cities (Liu;

Slik,, 2022).

The introduction of exotic species can result in considerable environmental consequences (Zhao *et al.*, 2022). An example of an exotic tree species that has invasive potential is Leucaena leucocephala, originally from Central America. This species has been introduced in tropical and subtropical regions in several countries as an option for urban afforestation, as it grows quickly. The Leucaena species competes with native species and alters local ecosystems. Therefore, it is necessary to monitor the presence of this species in urban areas to conserve biodiversity and ensure a sustainable environment (Kato-Noguchi; Kurniadie, 2022; Iqbal *et al.*, 2023).

To map tree vegetation in urban areas, remote sensing and other types of images have been applied in several locations., as they provide a comprehensive and detailed view to monitor and propose improvements to environments (Sishodia; Ray; Singh, 2020; Li *et al.*, 2023a). In addition to the use of remote sensing data and images, several artificial intelligence algorithms are generally applied to mapping to optimize the procedure. The combination of different images with artificial intelligence can provide more accurate results in mapping vegetation as they can automate procedures.

Deep learning is a branch of machine learning and has been applied in studies to analyze and map complex data (Moein *et al.*, 2023; Li *et al.*, 2023b). Deep learning allows models to analyze and extract features from large datasets and is effective in pattern recognition, language processing and computer vision tasks (Maxwell; Warner; Guillen, 2021; Yuan; Shi; Gu, 2021). Several studies have used deep learning networks to analyze vegetation and different tree species with remote sensing data (Arce *et al.*, 2021; Bolyn *et al.*, 2022; Nezami *et al.*, 2020; Carvalho *et al.*, 2022; Xu *et al.*, 2023; Ferreira *et al.*, 2024). One of the networks used is YOLO (You Only Look Once) and is known for its innovative approach to object detection in images. YOLO can efficiently detect objects in real time and allows a fast, agile and accurate implementation (Jiang *et al.*, 2022; Terven; Cordova-Esparza, 2023; Chen *et al.*, 2023).

The study carried out by Itakura and Hosoi (2020), proposed a method for measuring trees using 360° spherical cameras and another automatic method for detecting trees from 3D images. The trees included in the 360° spherical images were detected using YOLOv2 and then the trees were detected with the information obtained from the 3D images. The authors obtained an F-measure of 0.94. Another study by

Wang *et al.* (2023a), proposed an object detection algorithm based on YOLOv8 to detect diseased pine trees. The results obtained by the authors with the improved YOLOv8s-GAM model achieved 81%, 67.2% and 76.4% optimal detection performance on mAP50, mAP50-95 and average evaluation metrics. The results are 4.5%, 4.5% and 2.7% higher than the original YOLOv8s model. The study carried out by Choi *et al.* (2022), classified tree species with YOLO and street view images. The proposed method achieved an average precision (Map) of 0.564. According to the authors, the results demonstrate the potential for creating inventories of urban tree profiles.

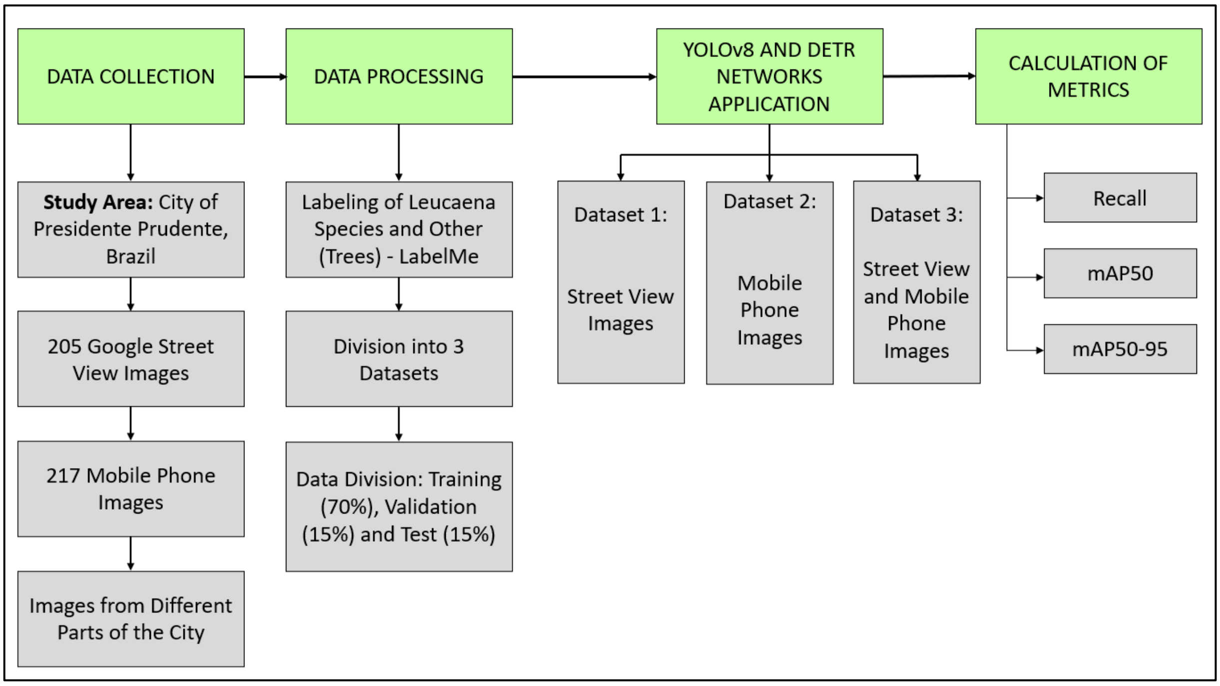
Another network with potential for object detection is the Detection Transformer (DETR). DETR is a simple model that does not require a specialized library, unlike other detectors. In addition to showing accuracy and good performance, DETR can be easily generalized (Carion *et al.*, 2020). The study of Dersch *et al.* (2023), applied two methods: Mask R – CNN and DETR to detect tree canopy. The authors used multispectral images and images from UAV (unmanned aerial vehicle) lidar data. The results obtained by the authors showed that multispectral images performed better than lidar data, and DETR with color infrared images reached 90% in coniferous, 81% in deciduous and 84% in mixed stands.

In the context of tree species in urban areas, this study analyzed the performance of the YOLOv8 and DETR networks to detect the tree species Leucaena leucocephala in panoramic images from Google Street View and mobile phones.

MATERIALS AND METHOD

Figure 1 shows the flowchart with the steps taken in this study to verify the performance of YOLOv8 and DETR to detect the tree species Leucaena leucocephala in urban panoramic images.

Figure 1- A Flowchart with the steps taken with YOLOv8 and DETR on a total of 422 Street View and cell phone images to detect the Leucaena leucocephala.

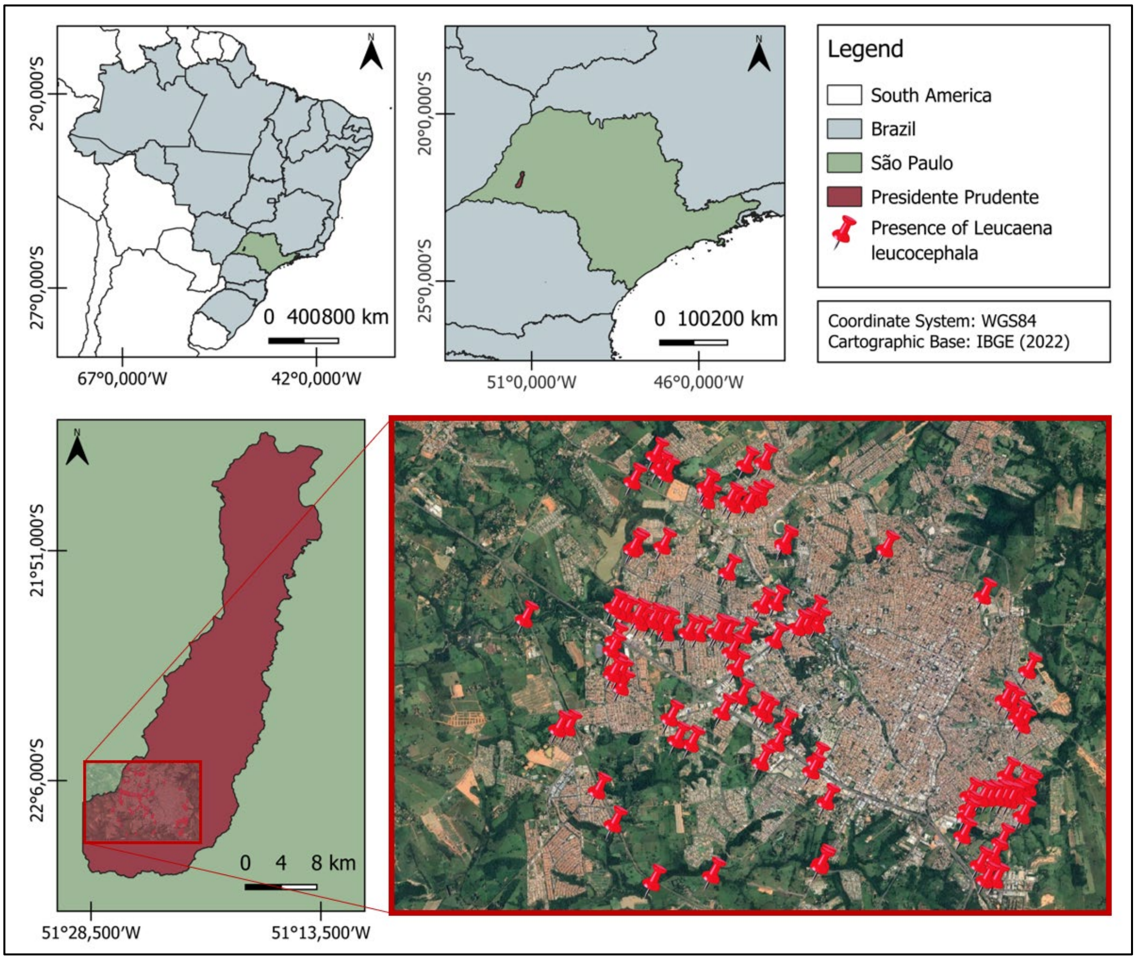


Fonte: Author (2024).

Study Area

To carry out the tests in this study, images of the city of Presidente Prudente (Figure 2), in the state of São Paulo, in Brazil (Latitude 22º07'21.06"S and Longitude 51º23' 17.71"W) were used. Presidente Prudente is a medium-sized city with an estimated population of 225,668 in 2022 (IBGE, 2022). The city chosen for the study has a high quantity of the tree species Leucaena leucocephala distributed throughout the urban area. In this study, it was possible to identify points in Presidente Prudente where Leucaena leucocephala is present to later collect images from Google Street View and mobile phones.

Figure 2- Location of the Study Area: Presidente Prudente, Brazil. The Figure also shows the points with the presence of Leucaena leucocephala in the city where the 422 Street View and Mobile Phone Images were collected.



Fonte: Author (2024).

Data Collection and Preparation

A total of 422 images were collected. Of the total images, 205 correspond to Google Street View images (3328 x 1664) and 217 correspond to mobile phone images (3000 x 3000). Street View images are obtained by cars equipped with 360-degree cameras and have advantages as they are frequently updated by the Google platform (Li; Yao, 2020). In addition to facilitating the collection of images of urban areas around the world, Street View (SV) helps to identify the characteristics of trees in more detail, including the species Leucaena leucocephala.

In this study, the images collected from Google Street View of the city of Presidente Prudente correspond to the years 2021 and 2022 (the most recent years at the time of collection). The mobile phone images were collected by the authors and corresponded to the year 2023. The camera was oriented in a vertical position and configured with a resolution of 48 megapixels. Furthermore, images were taken at different periods and days to collect a dataset with different light and shadow characteristics.

For the tests in this study, labeling was made on the targets of interest to later verify the performance of two neural networks. The labeling was done using the LabelMe software and the tree vegetation was divided into two classes in this study: Leucaena and others (which include the rest of the trees). Figure 3 shows some examples of the labeling carried out for this study considering the two classes in the 422 images. The images were divided into 3 datasets: Dataset 1 = 205 SV images; Dataset 2 = 217 mobile phone images; Dataset 3 = 422 images (sum of Dataset 1 and Dataset 2).

Figure 3- Example of labeling done in LabelMe with SV and mobile phone images considering 2 classes (green boxes = Leucaena; red boxes = other). The horizontal images correspond to the SV image examples. The square images correspond to mobile phone image examples.

Fonte: Author (2024).

Networks

The YOLOv8 algorithm is a fast object detection method and has stood out in the field of computer vision (Wang *et al.*, 2023b; Wang *et al.*, 2023c). YOLOv8 is an algorithm released by the company Ultralytics on January 10, 2023, and compared to previous models in the YOLO series (such as YOLOv5 and YOLOv7) it is an advanced and cutting-edge model that offers greater accuracy and detection speed (Wang *et al.*, 2023b). YOLOv8 comprises an input segment, a backbone, a neck and an output segment. The input segment performs mosaic data augmentation and adaptive grayscale filling on the input image. The backbone network and the neck module form the core structures of the YOLOv8 network. In the YOLO network, C2 and C3 refer to specific convolutional layers in the network architecture. C2 refers to the second convolutional layer and C3 refers to the third convolutional layer These convolutional layers are crucial for object detection. The input image is processed by various Conv and C2f modules to extract feature maps at different scales. The C2f module is an improved version of the original C3 module and functions as the main residual learning module (Wang *et al.*, 2023c).

The Detection Transformer (DETR) can be considered a Transformer's pioneering work in the field of target detection and contains three main components: CNN backbone network; Encoder-decoder transformer and simple feedforward networks. The DETR simplifies the object detection pipeline by removing hand-crafted designs and hyperparameters as employed in conventional two-stage object detectors (Zhang *et al.*, 2023a).

In this study, the performance of the YOLOv8 and DETR networks was analyzed. The YOLOv8 model has five different scaled models by adjusting two parameters: width and depth. These models are referred to as YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large) and YOLOv8x (extra-large) (Wang *et al.*, 2023b; Terven; Cordova-Esparza, 2023). The version of YOLOv8 used in this study was Yolov8m.

The DETR backbone in this study was ResNet50. Tables 1 and 2 show the hyperparameters for each of the networks used in this study. The tests were carried out with the following hardware conditions: CPU = Ryzen 7 5700X; GPU = Nvidia RTX 3090; RAM = 32GB DDR4. The training time for the networks in this study was 2 hours for YOLOv8 and 3 and a half hours for DETR. In this study, 70% of the data was considered for training, 15% for validation and 15% for testing.

Table 2- YOLOv8 Hyperparameters

YOLOv8	
Batch Size	16
Image Size	640 x 640
Epochs	100
Learning Rate	0,01
Optimizer	AdamW
Inference Time	0,01 s

Table 2- DETR Hyperparameters

DETR	
Batch Size	8
Image Size	640 x 640
Epochs	100
Learning Rate	0,0005
Optimizer	AdamW
Inference Time	0,7 s

Evaluation Metrics

In this study, five evaluation metrics were used for YOLOv8 and DETR: recall, precision, F1, mAP50 and mAP50-95 (Huang *et al.*, 2023):

- Recall (R): measures the model's ability to detect all objects of a specific class. It is calculated as the number of true positives (TP) divided by the sum of true positives and false negatives (FN). A high Recall score indicates that the model is capturing most objects of the desired class (Huang *et al.*, 2023; Arifando; Eto; Wada, 2023):

$$\text{Recall} = TP/(TP+FN) \quad (1)$$

- Precision (P): measures the proportion of positive instances correctly identified about the total instances identified as positive by the model (Huang *et al.*, 2023; Arifando; Eto; Wada, 2023):

$$\text{Precision} = TP/(TP+FP) \quad (2)$$

- F-measure (F1): The F1 measure combines precision and recall into a single value. It is especially useful when there is an imbalance between classes in a dataset (Huang *et al.*, 2023);

- mAP50 (Mean Average Precision at 50% IoU): is a metric that evaluates the average accuracy of the model over multiple classes of objects. Calculates the average precision considering different Intersection-over-Union (IoU) Threshold values between the predicted bounding boxes and the true bounding boxes. The value "50" indicates that an IoU of 50% is used as a criterion to determine whether a detection is considered correct (Huang *et al.*, 2023; Su *et al.*, 2024).

- mAP50-95 (Mean Average Precision from 50% to 95% IoU): is similar to mAP50, but instead of considering just a specific IoU value, mAP50-95 averages accuracy over a range of IoU values, from 50% to 95%, increasing by 5% each time. This provides a more comprehensive assessment of model accuracy under various conditions of overlap between the predicted and true bounding boxes (Huang *et al.*, 2023; Su *et al.*, 2024).

These metrics together provide a comprehensive view of the object detection model's performance: how it balances between finding all relevant objects (Recall), the accuracy of detections made (mAP50), and the accuracy in locating objects (mAP50-95).

RESULTS

Tests were carried out with YOLOv8 and DETR. The average metrics for the three datasets are shown in Table 3. Dataset 1 considered the 205 Street View (SV) images; Dataset 2 considered the 217 mobile phone images (MP); and Dataset 3 considered the 422 images, that is, the combination of Dataset 1 with Dataset 2 (SV +

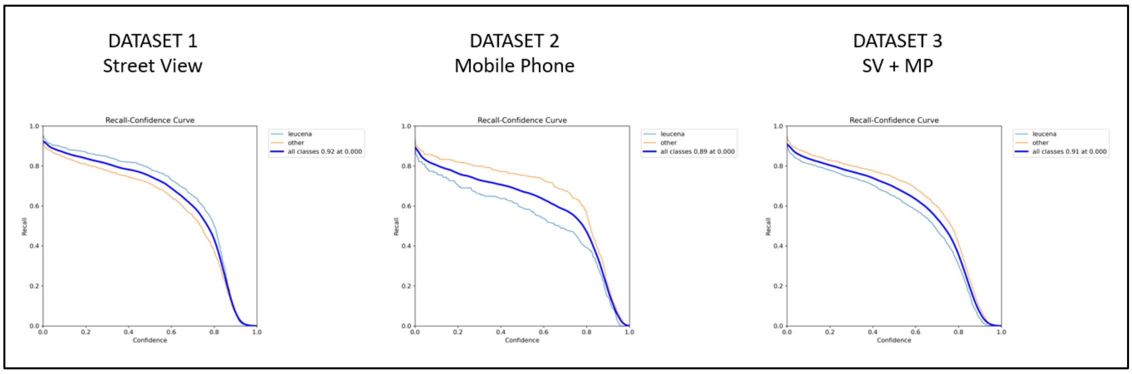
MP).

Table 3- Results of recall (R), mAP50 and mAP50-95 for the tests with YOLOv8 and DETR with 3 different datasets.

YOLOv8			
	R	**mAP50**	**mAP50-95**
Dataset 1 - SV	0.749	0.830	0.517
Dataset 2 - MP	0.751	0.849	0.614
Dataset 3 – SV + MP	0.766	0.844	0.540
DETR			
	R	**mAP50**	**mAP50-95**
Dataset 1 - SV	0.218	0.455	0.169
Dataset 2 - MP	0.684	0.847	0.563
Dataset 3 – SV + MP	0.219	0.355	0.130

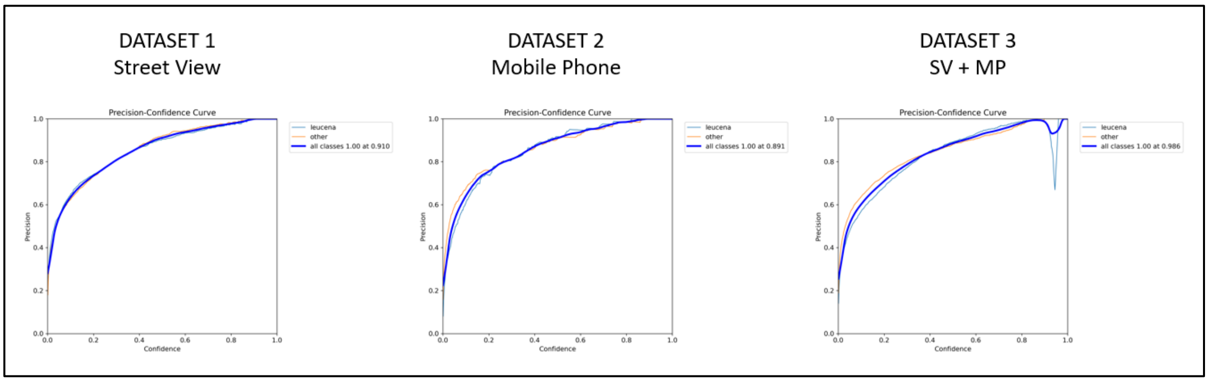
Concerning dataset 3, the networks studied presented opposite results. In the case of YOLOv8, mixing the datasets (dataset 3) generated advances in the average metrics and proved to be a model capable of extracting the characteristics that define the Leucaena tree species regardless of distortion. In the case of DETR, there was a significant worsening when trained with dataset 3. The combination of data from different domains caused the model to enter a state of underfit, with less generalization of learning. Figures 4, 5, 6 and 7 show comparisons of the metrics graphs for the three datasets with the YOLOv8 network.

Figure 4- Comparison of the Graphs with the Recall Confidence Curve for the 3 Datasets with YOLOv8.



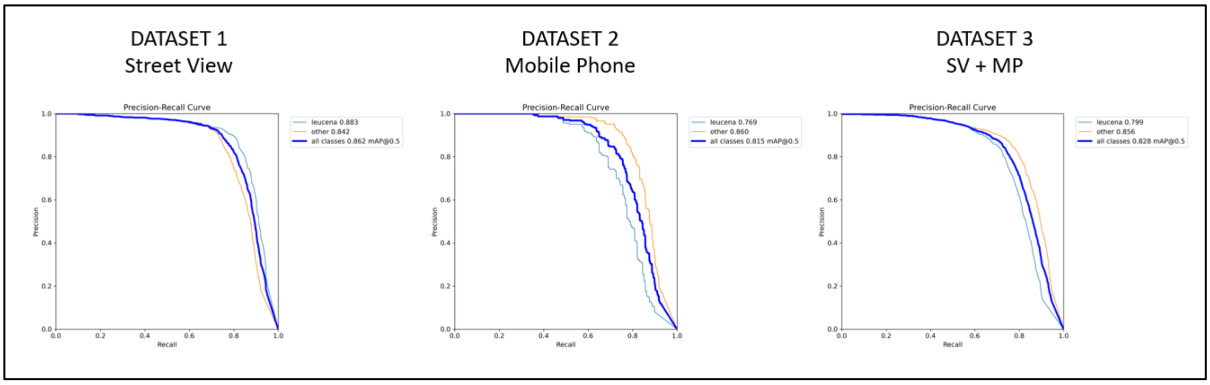
Fonte: Author (2024).

Figure 5- Comparison of the Graphs with the Precision Confidence Curve for the 3 Datasets with YOLOv8.



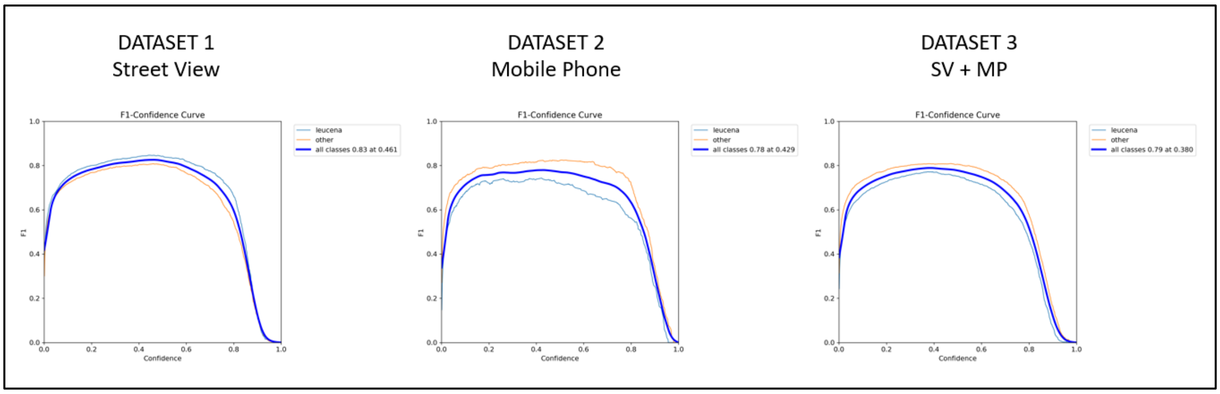
Fonte: Author (2024).

Figure 6- Comparison of the Graphs with the Precision-Recall Confidence Curve for the 3 Datasets with YOLOv8.



Fonte: Author (2024).

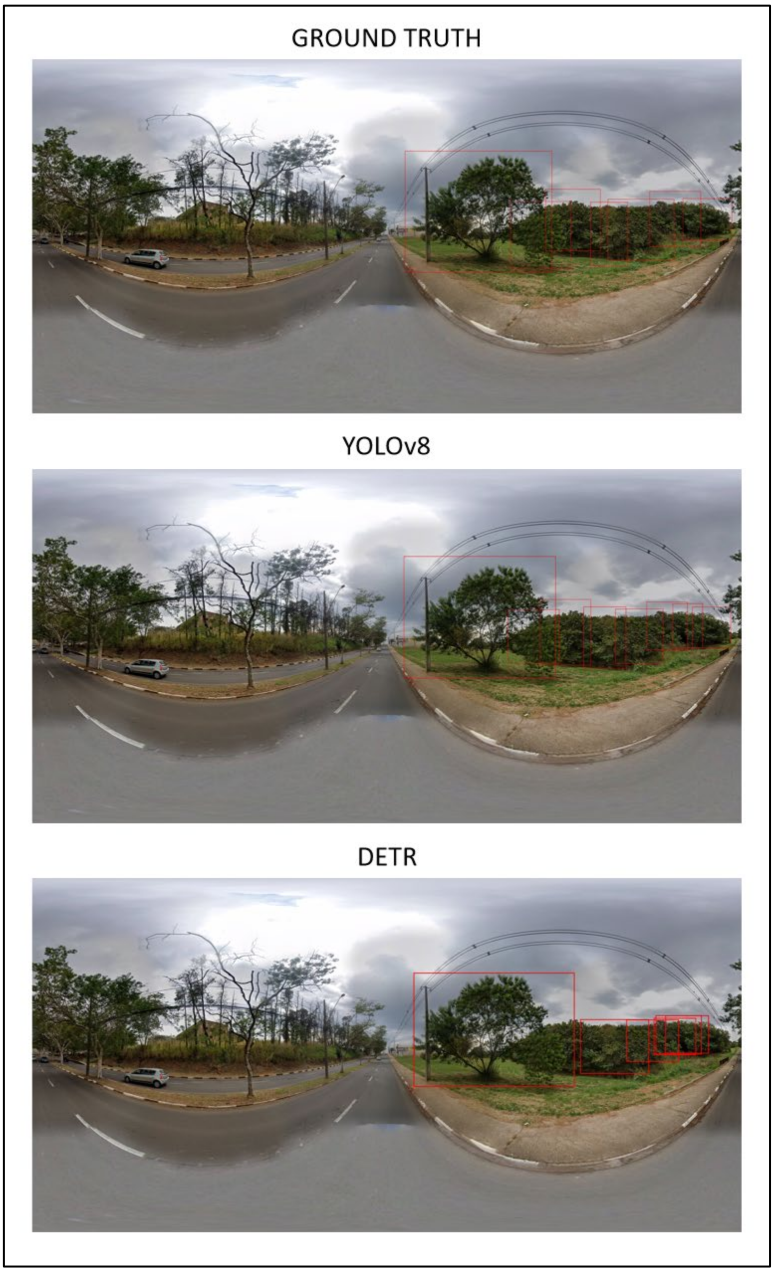
Figure 7- Comparison of the Graphs with the F1 Confidence Curve for the 3 Datasets with YOLOv8.



Fonte: Author (2024).

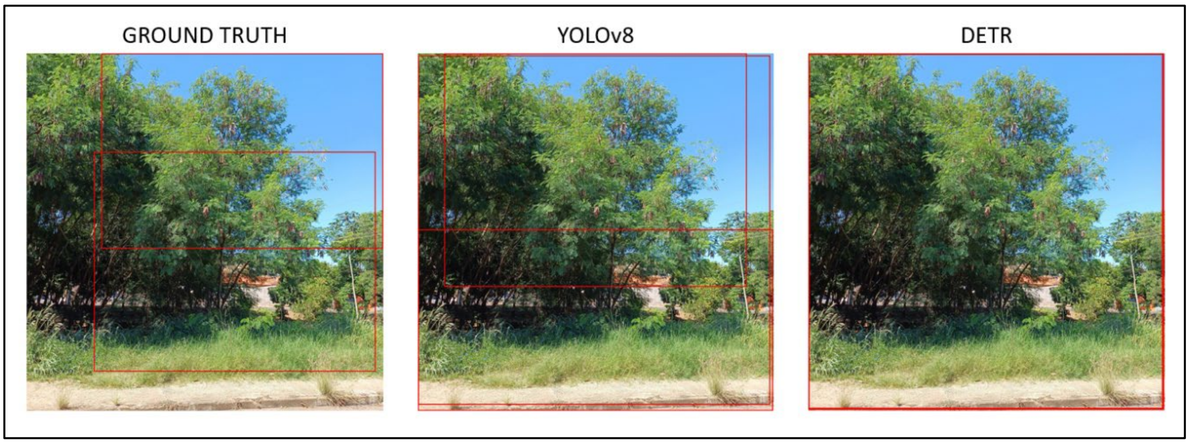
Figures 8, 9 and 10 show the visual detection comparisons obtained with the YOLOv8 and DETR networks on the three datasets tested.

Figure 8- Visual comparison of the detection obtained with YOLOv8 and DETR in dataset 1 with Street View images.



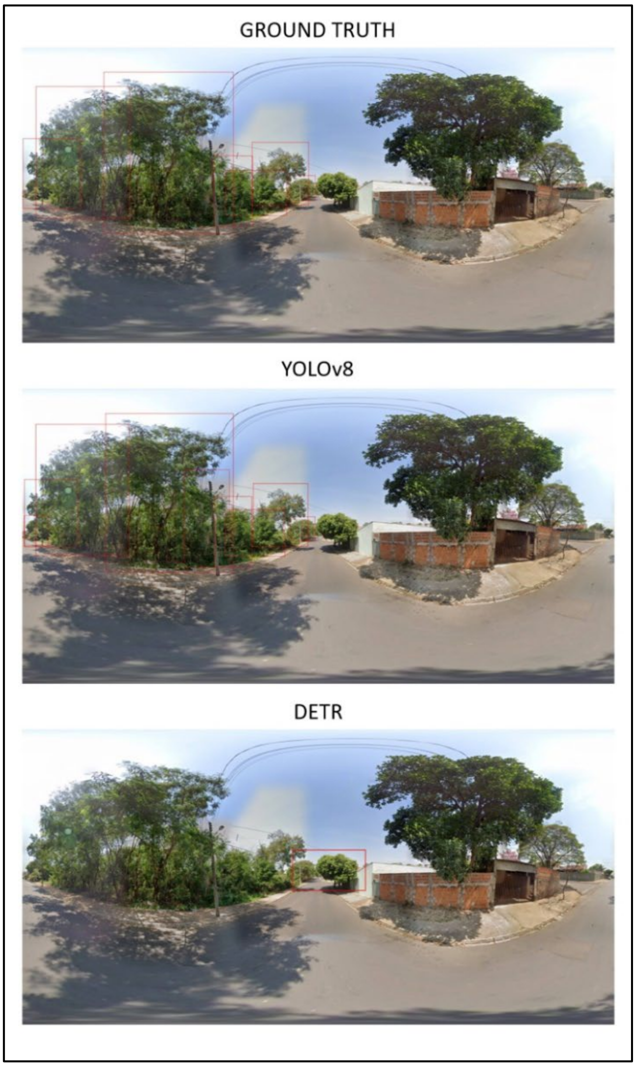
Fonte: Author (2024).

Figure 9- Visual comparison of the detection obtained with YOLOv8 and DETR in dataset 2 with cell phone images.



Fonte: Author (2024).

Figure 10- Visual comparison of the detection obtained with YOLOv8 and DETR in dataset 3 with Street View and cell phone images.



Fonte: Author (2024).

DISCUSSION

In this study, YOLOv8 was shown to be 70 times faster in inference time than DETR and outperformed DETR (Table 3). YOLOv8 is an advanced, state-of-the-art model launched in January 2023 and has stood out for its greater accuracy and detection speed (Wang *et al.*, 2023b; Hussain, 2023).

YOLOv8 and derived models have proven efficient in the environmental and agricultural areas to detect leaves, trees and crops (Zhong *et al.*, 2024; Wang *et al.*, 2023d; Guo *et al.*, 2023; Solimani *et al.*, 2024; Guan *et al.*, 2024; Korznikov *et al.*, 2023). Zhong *et al.* (2024), used YOLOv8 to identify individual trees with RGB and LIDAR data. The authors found that YOLOv8 presented better performance compared to other models. Korznikov *et al.* (2023), analyzed the performance of YOLOv8 to detect two species of evergreen coniferous trees. The authors obtained an F1 of 0.990 and overall accuracy of 0.981, which allowed a detailed analysis of species distribution.

One of the obstacles encountered by the networks was the characteristics of Street View images. Google Street View images present considerable distortion as they are 360° images and the fact that Google vehicles only travel in one direction, causes the other side of the street to present the distortion (Hipp *et al.*, 2021). This feature in the images made it difficult to identify the object of interest (Leucaena) in our tests. Correction attempts have been made to reduce distortion in Street View images, such as conversion using a rotation matrix and cropping a portion of the image. Despite attempts, none showed an improvement in the model's performance.

In the field of computer vision, object detection has been a challenge. Object detection seeks to predict a bounding box and a class label (Zhang *et al.*, 2023a). Despite its innovative designs, DETR also has limitations, including slow convergence and lower performance (Liu *et al.*, 2023). In this study, tests with DETR proved to be less accurate when compared to the YOLOv8 detection network (Table 3). The most accurate result from DETR was in the test with dataset 2, where the network obtained a mAP50 of 0.847. Despite presenting a higher average for this network, it was still lower than YOLOv8, which obtained a mAP50 of 0.849 with dataset 2. The improvement in DETR in this case can be explained by the images that make up dataset 2 (only mobile phone images). The use of images that do not present distortion proved to be more efficient for the DETR network, which justifies the low accuracy of the network in datasets 1 and 3, which were composed of Street View images.

New transformer-based approaches have been applied in studies to DETR to improve model accuracy: DA-DETR: Domain Adaptive Detection Transformer (Zhang *et al.*, 2023); OW-DETR: Open-world Detection Transformer (Gupta *et al.*, 2022); AO2-DETR: Arbitrary-Oriented Object Detection Transformer (Dai *et al.*, 2022); Semi-DETR: Semi-Supervised Object Detection with Detection Transformers (Zhang *et al.*, 2023b). In all the cases mentioned, the authors obtained a performance improvement when compared to DETR, which presents limitations. These results indicate that new approaches can improve DETR performance.

Traditional object detection methods are based on shallow, trainable architectures. The performance of traditional methods tends to stagnate when they have a complex data set (Zhao *et al.*, 2019). Deep learning has been introduced to solve existing problems in traditional architectures and improve object detection results, as these models can learn low- and high-level image features (Kaur; Singh, 2023; Xiao *et al.*, 2020). Furthermore, deep learning behaves differently in network architecture, training strategy and optimization function (Zhao *et al.*, 2019).

Morera *et al.* (2020), used YOLO to detect outdoor advertising panels in Street View images. The authors found that YOLO produced better panel localization results detecting a higher number of True Positive (TP) panels with a higher accuracy compared to Single Shot MultiBox Detector (SSD) network. Another study by Shu, Yan and Xu (2021), applied YOLOv5 to detect cracks in pavements with street view images. The authors obtained mAP greater than 70% and a detection time of 152 ms. YOLOv5 proved to be more efficient than YOLOv3, which shows the improvement of the model as new versions are released. In addition to the precise result, the authors mention that the ease of collecting street view data corresponds to an advantage for detection.

YOLOv8 has been shown to be more accurate than other models including the old versions of YOLO (Sary; Andromeda; Armin, 2023; Diwan; Anirudh; Tembhurne, 2023; Talaat; Zaineldin, 2023). The study of Bayrak, Erdem and Uzar (2023), used a convolutional neural network (CNN) and variants of YOLOv8 to classify tree species with aerial images from TreeSatAI. The authors showed that YOLOv8 outperformed other YOLO variants with scores above 70%. The tests carried out in this study showed that street-level images contribute to the detection of the Leucaena tree species.

CONCLUSIONS

This study conducted out tests to verify the performance of the YOLOv8 and DETR networks in detecting the tree species Leucaena leucocephala considered a potentially invasive species in several regions, including Brazil. Panoramic Street View and mobile phone images allow a detailed view at street level, making it possible to observe tree components such as foliage, fruits, flowers and trunks. We carried out tests with a dataset with 422 images (205 from Street View and 217 from mobile phone) from different parts of the city chosen as the study area. The images were divided into three datasets (dataset 1 = Street View images; dataset 2 = mobile phone images; dataset 3 = dataset 1 + dataset 2). YOLOv8 was launched in 2023 and presented better results than DETR, presenting mAP50 of 0.83 for Street View and 0.84 for mobile phone images, while DETR presented mAP50 of 0.45 for Street View and 0.84 for mobile phones. YOLOv8 also presented higher recall than DETR, with an average of 0.74 for SV images and 0.75 for MP images, while DETR presented an average of 0.21 for SV and 0.68 for MP. The networks presented difficulties with certain types of images (such as the distortions present in Street View images). Despite the limitations, the networks studied presented interesting metrics and contribute to urban planning, as they will provide a quick procedure to detect the invasive species, which can speed up decision making. Future studies can analyze the performance of networks with new approaches and also the application of methods that reduce distortion in Street View images, to minimize errors and achieve more accurate metrics.

REFERENCES

ARCE, L. S. D.; OSCO, L. P.; ARRUDA, M. D. S. D.; FURUYA, D. E. G.; RAMOS, A. P. M.; AOKI, C.; POTT, A.; FATHOLAHI, S.; LI, J.; ARAUJO, F. F. D.; GONÇALVES, W. N.; JUNIOR, J. M. Mauritia flexuosa palm trees airborne mapping with deep convolutional neural network. **Scientific Reports**, v. 11, n. 1, p. 19619, 2021.

ARIFANDO, R.; ETO, S.; WADA, C. Improved YOLOv5-based lightweight object detection algorithm for people with visual impairment to detect buses. **Applied Sciences**, v. 13, n. 9, p. 5802, 2023.

BAYRAK, O. C.; ERDEM, F.; UZAR, M. Deep Learning Based Aerial Imagery Classification for Tree Species Identification. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 48, p. 471-476, 2023.

BOLYN, C., LEJEUNE, P.; MICHEZ, A.; LATTE, N. Mapping tree species proportions from satellite imagery using spectral–spatial deep learning. **Remote Sensing of Environment**, v. 280, p. 113205, 2022.

CARION, N.; MASSA, F.; SYNNAEVE, G.; USUNIER, N.; KIRILLOV, A.; ZAGORUYKO, S. End-to-end object detection with transformers. In: European conference on computer vision. **Cham: Springer International Publishing**, p. 213-229, 2020.

CARVALHO, M. D. A.; JUNIOR, J. M.; MARTINS, J. A. C.; ZAMBONI, PEDRO.; COSTA, C. S.; SIQUEIRA, H. L.; ARAUJO, M. S.; GONÇALVES, D. N.; FURUYA, D. E. G.; OSCO, L. P.; RAMOS, A. P. M.; LI, J.; JUNIOR, A. A. D. C.; GONÇALVES, W. N. A deep learning-based mobile application for tree species mapping in RGB images. **International Journal of Applied Earth Observation and Geoinformation**, v. 114, p. 103045, 2022.

CHEN, C.; ZHENG Z.; XU, T.; GUO, S.; FENG, S.; YAO, W.; LAN, Y. YOLO-Based UAV Technology: A Review of the Research and Its Applications. **Drones**, v. 7, n. 3, p. 190, 2023.

CHOI, K.; LIM, W.; CHANG, B.; JEONG, J.; KIM, I.; PARK, C. R.; KO, D. W. An automatic approach for tree species detection and profile estimation of urban street trees using deep learning and Google street view images. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 190, p. 165-180, 2022.

DAI, L.; LIU, H.; TANG, H.; WU, Z.; SONG, P. Ao2-detr: Arbitrary-oriented object detection transformer. **IEEE Transactions on Circuits and Systems for Video Technology**, 2022.

DERSCH, S.; SCHOTTL, A.; KRZYSTEK, P.; HEURICH, M. Towards complete tree crown delineation by instance segmentation with Mask R–CNN and DETR using UAV-based multispectral imagery and lidar data. **ISPRS Open Journal of Photogrammetry and Remote Sensing**, v. 8, p. 100037, 2023.

DIWAN, T.; ANIRUDH, G.; TEMBHURNE, J. V. Object detection using YOLO: Challenges, architectural successors, datasets and applications. **Multimedia Tools and Applications**, v. 82, n. 6, p. 9243-9275, 2023.

FERREIRA, M. P.; SANTOS, D. R. D.; FERRARI, F.; FILHO, L. C. T. C.; MARTINS, G. B.; FEITOSA, R. Q.; Improving urban tree species classification by deep-learning based fusion of digital aerial images and LiDAR. **Urban Forestry & Urban Greening**, p. 128240, 2024.

GUAN, H.; DENG, H.; MA, X.; ZHANG, T.; ZHANG, Y.; ZHU, T.; ZHOU, H.; GU, Z.; LU, Y. A corn canopy organs detection method based on improved DBi-YOLOv8 network. **European Journal of Agronomy**, v. 154, p. 127076, 2024.

GUO, B.; LING, S.; TAN, H.; WANG, S.; WU, CAILAN.; YANG, D. Detection of the Grassland Weed Phlomoides umbrosa Using Multi-Source Imagery and an Improved YOLOv8 Network. **Agronomy**, v. 13, n. 12, p. 3001, 2023.

GUPTA, A.; NARAYAN, S.; JOSEPH, K. J.; KHAN, S.; KHAN, F. S.; SHAH, M. Ow-detr: Open-world detection transformer. *In*: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. 2022. p. 9235-9244.

HIPP, J. R.; LEE, S.; KI, D.; KIM, J. H. Measuring the built environment with Google Street View and machine learning: Consequences for crime on street segments. **Journal of Quantitative Criminology**, p. 1-29, 2021.

HUANG, H.; LA, G.; WEI, J.; ZHONG, Z.; XU, Z.; LI, D.; ZOU, F. TLI-YOLOv5: A Lightweight Object Detection Framework for Transmission Line Inspection by Unmanned Aerial Vehicle. **Electronics**, v. 12, n. 15, p. 3340, 2023.

HUSSAIN, M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. **Machines**, v. 11, n. 7, p. 677, 2023.

IBGE. Brazilian Institute of Geography and Statistics. **População Presidente Prudente 2022**. 2022. Disponível em: https://cidades.ibge.gov.br/brasil/sp/presidente-prudente/panorama.

IQBAL, I. M.; HEIKO, B.; BAREEN, F. E.; SHABBIR, A. Mapping Lantana camara and Leucaena leucocephala in Protected Areas of Pakistan: A Geo-Spatial Approach. **Remote Sensing**, v. 15, n. 4, p. 1020, 2023.

ITAKURA, K; HOSOI, F. Automatic tree detection from three-dimensional images reconstructed from 360 spherical camera using YOLO v2. **Remote Sensing**, v. 12, n. 6, p. 988, 2020.

JIANG, P.; ERGU, D.; LIU, F.; CAI, Y.; MA, B. A Review of Yolo algorithm developments. **Procedia Computer Science**, v. 199, p. 1066-1073, 2022.

KATO-NOGUCHI, H; KURNIADIE, D. Allelopathy and allelochemicals of Leucaena leucocephala as an invasive plant species. **Plants**, v. 11, n. 13, p. 1672, 2022.

KAUR, R.; SINGH, S. A comprehensive review of object detection with deep learning. **Digital Signal Processing**, v. 132, p. 103812, 2023.

KORZNIKOV, K.; KISLOV, D.; PETRENKO, T.; DZIZYUROVA, V.; DOLEZAL, J.; KRESTOV, P.; ALTMAN, J. Unveiling the Potential of Drone-Borne Optical Imagery in Forest Ecology: A Study on the Recognition and Mapping of Two Evergreen Coniferous Species. **Remote Sensing**, v. 15, n. 18, p. 4394, 2023.

LI, M.; YAO, W. 3D map system for tree monitoring in hong kong using google street view imagery and deep learning. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 3, p. 765-772, 2020.

LI, F.; YAGITCANLAR, T.; NEPAL, M.; NGUYEN, K.; DUR, F. Machine Learning and Remote Sensing Integration for Leveraging Urban Sustainability: A Review and Framework. **Sustainable Cities and Society**, p. 104653, 2023a.

LI, E.; WANG, L.; XIE, Q.; GAO, R.; SU, Z.; LI, Y. A novel deep learning method for maize disease identification based on small sample-size and complex background datasets. **Ecological Informatics**, v. 75, p. 102011, 2023b.

LIU, JIAJIA; SLIK, FERRY. Are street trees friendly to biodiversity?. **Landscape and Urban Planning**, v. 218, p. 104304, 2022.

LIU, S.; REN, T.; CHEN, J.; ZENG, Z.; ZHANG, H.; LI, F.; LI, H.; HUANG, J.; SU, H.; ZHU, J.; ZHANG, L. Detection transformer with stable matching. *In*: **Proceedings of the IEEE/CVF International Conference on Computer Vision.** 2023. p. 6491-6500.

MAXWELL, A.; WARNER, T.; GUILLEN, L. A. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review. **Remote Sensing**, v. 13, n. 13, p. 2450, 2021.

MOEIN, M. M.; SARADAR, A.; RAHMATI, K.; MOUSAVINEJAD, S. H. G.; BRISTOW, J.; ARAMALI, V.; KARAKOUZIAN, M. Predictive models for concrete properties using machine learning and deep learning approaches: A review. **Journal of Building Engineering**, v. 63, p. 105444, 2023.

MORERA, A. SANCHEZ, A.; MORENO, A. B.; SAPPA, A. D.; VÉLEZ, J. F. SSD vs. YOLO for detection of outdoor urban advertising panels under multiple variabilities.

Sensors, v. 20, n. 16, p. 4587, 2020.

NEZAMI, S.; KHORAMSHAHI, E.; NEVALAINEN, O.; POLONEN, I.; HONKAVAARA, E. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks. **Remote Sensing**, v. 12, n. 7, p. 1070, 2020.

SARY, I. P.; ANDROMEDA, S.; ARMIN, E. U. Performance Comparison of YOLOv5 and YOLOv8 Architectures in Human Detection using Aerial Images. Ultima Computing: **Jurnal Sistem Komputer**, v. 15, n. 1, p. 8-13, 2023.

SHU, Z.; YAN, Z.; XU, X. Pavement crack detection method of street view images based on deep learning. *In*: **Journal of Physics: Conference Series**. IOP Publishing, p. 022043, 2021.

SISHODIA, R. P.; RAY, R. L.; SINGH, S. K. Applications of remote sensing in precision agriculture: A review. **Remote Sensing**, v. 12, n. 19, p. 3136, 2020.

SOLIMANI, F.; CARDELLICCHIO, A.; DIMAURO, G.; PETROZZA, A.; SUMMERER, S.; CELLINI, F.; RENÓ, V. Optimizing tomato plant phenotyping detection: Boosting YOLOv8 architecture to tackle data complexity. **Computers and Electronics in Agriculture**, v. 218, p. 108728, 2024.

SU, X.; ZHANG, J.; MA, Z.; DONG, Y.; ZI, J.; XU, N.; ZANG, H.; XU, F.; CHEN, F. Identification of Rare Wildlife in the Field Environment Based on the Improved YOLOv5 Model. **Remote Sensing**, v. 16, n. 9, p. 1535, 2024.

TALAAT, F. M.; ZAINELDIN, H. An improved fire detection approach based on YOLO-v8 for smart cities. **Neural Computing and Applications**, v. 35, n. 28, p. 20939-20954, 2023.

TERVEN, J.; CORDOVA-ESPARZA, D. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. **Machine Learning and Knowledge Extraction**, v. 5, n. 4, p. 1680-1716, 2023.

WANG, S.; CAO, X.; WU, M.; YI, C.; ZHANG, Z,; FEI, H.; ZHENG, H.; JIANG, H.; JIANG, Y.; ZHAO, X.; ZHAO, X.; YANG, P. Detection of Pine Wilt Disease Using Drone Remote Sensing Imagery and Improved YOLOv8 Algorithm: A Case Study in Weihai, China. **Forests**, v. 14, n. 10, p. 2052, 2023a.

WANG, G.; CHEN, Y.; AN, P.; HONG, H.; HU, J.; HUANG, T. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. **Sensors**, v. 23, n. 16, p. 7190, 2023b.

WANG, X.; GAO, H.; JIA, Z.; LI, Z. BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8. **Sensors**, v. 23, n. 20, p. 8361, 2023c.

WANG, N.; LIU, H.; LI, Y.; ZHOU, W.; DING, M. Segmentation and phenotype calculation of rapeseed pods based on YOLO v8 and mask R-convolution neural networks. **Plants**, v. 12, n. 18, p. 3328, 2023d.

XIAO, Y.; ZHIQIANG, T.; JYACHEN, Y.; YINCHU, Z.; SHUAI, L.; SHAOYI, D.; XUGUANG, L. A review of object detection based on deep learning. **Multimedia Tools and Applications**, v. 79, p. 23729-23791, 2020.

XU, S.; WANG, R.; SHI, W.; WANG, X. Classification of tree species in transmission line corridors based on YOLO v7. **Forests**, v. 15, n. 1, p. 61, 2023.

YILMAZ, S.; IRMAK, M. A.; QAID, A. Assessing the effects of different urban landscapes and built environment patterns on thermal comfort and air pollution in Erzurum city, Turkey. **Building and Environment**, v. 219, p. 109210, 2022.

YUAN, X.; SHI, J.; GU, L. A review of deep learning methods for semantic segmentation of remote sensing imagery. **Expert Systems with Applications**, v. 169, p. 114417, 2021.

ZHANG, J.; HUANG, J.; LUO, Z.; ZHANG, G.; ZHANG, X.; LU, S. Da-detr: Domain adaptive detection transformer with information fusion. *In*: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. 2023a. p. 23787-23798.

ZHANG, J.; LIN, X.; ZHANG, W.; WANG, K.; TAN, X.; HAN, J.; DING, E.; WANG, J.; LI, G. Semi-supervised object detection with detection transformers. *In*: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**. 2023b. p. 23809-23818.

ZHAO, Z.; ZHENG, P.; XU, S.; WU, X. Object detection with deep learning: A review. **IEEE transactions on neural networks and learning systems**, v. 30, n. 11, p. 3212-3232, 2019.

ZHAO, H; ZHONG, Y; WANG, X; HU, X; LUO, C; BOITT, M.; PIIROINEN, R.; ZHANG, L; HEISKANEN, J.; PELLIKKA, P. Mapping the distribution of invasive tree species using deep one-class classification in the tropical montane landscape of Kenya. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 187, p. 328-344, 2022.

ZHENG, S.; HE, C.; GULDMANN, J.; XU, H.; LIU, XIAO. Heat Mitigation Benefits of Urban Trees: A Review of Mechanisms, Modeling, Validation and Simulation. **Forests**, v. 14, n. 12, p. 2280, 2023.

ZHONG, H.; ZHANG, Z.; LIU, H.; WU, J.; LIN, W Individual Tree Species Identification for Complex Coniferous and Broad-Leaved Mixed Forests Based on Deep Learning Combined with UAV LiDAR Data and RGB Images. **Forests**, v. 15, n. 2, p. 293, 2024.

3 CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desempenho de redes de aprendizagem profunda na identificação da espécie arbórea Leucaena leucocephala em imagens panorâmicas do Google Street View e de celular. No capítulo 1 foi apresentado uma revisão da literatura na área de interesse com base nos estudos disponíveis no Web of Science. A revisão mostra os dados quantitativos que combina a utilização de sensoriamento remoto e inteligência artificial (machine e deep learning) para mapear vegetação (geral, arbórea e espécies). As análises mostraram dados importantes da área, foi possível identificar 149 espécies arbóreas (inclusive a Leucaena leucocephala) já mapeadas em estudos anteriores com a utilização de sensoriamento remoto e diferentes algoritmos e redes. Além disso, a análise mostra os satélites já utilizados, métricas de avaliação, redes de aprendizagem profunda mais utilizadas, entre outras informações relevantes.

No segundo capítulo é apresentado um artigo com a utilização do Segment Anything Model (SAM) para segmentar vegetação arbórea em geral com imagens do Google Street View. Este artigo foi apresentado no Encontro Nacional de Ensino, Pesquisa e Extensão da Universidade do Oeste Paulista (UNOESTE) em 2023 e obteve o 2° lugar na categoria de Pós Graduação do 12° Prêmio Científico Unoeste. Após o evento, o artigo foi publicado nos Anais de Evento do ENEPE. Este artigo apresenta os testes iniciais com o SAM e as métricas possuem média de 99% de acurácia.

No terceiro capítulo é apresentado um novo estudo com o Segment Anything Model (SAM) e imagens panorâmicas. Neste estudo, o objetivo foi avaliar a capacidade do SAM em segmentar a espécie arbórea Leucaena leucocephala em imagens do Street View e de celular. É importante identificar a Leucaena de forma otimizada pois essa espécie é comum em diversos países e pode se tornar invasora. Neste estudo, o SAM obteve resultados superiores para imagens do Street View do que de celular. Enquanto a média F1 do Street View foi de 93%, as de celular tiveram média F1 de 66%.

No quarto capítulo, é apresentado um estudo do desempenho de duas redes de detecção de objetos: YOLOv8 (You Only Look Once) e DETR (Detection Transformer) para detectar a espécie Leucaena leucocephala em imagens panorâmicas do Street View e de celular. Foram realizados testes com três datasets (com 205 imagens do Street View e 217 de celular da cidade de Presidente Prudente) e as duas redes. Os resultados mostraram que a YOLOv8 superou a DETR e apresentou mAP50 em torno de 80% para os três datasets testados (que são compostos por imagens do SV, de celular e ambas).

A utilização de redes de aprendizagem profunda para detecção e segmentação da espécie arbórea Leucaena leucocephala em imagens panorâmicas do Street View e de celular apresentou resultados precisos em alguns testes deste projeto. O SAM, YOLOv8 e DETR correspondem a opções relevantes para auxiliar na identificação da espécie arbórea Leucaena em áreas urbanas e podem otimizar o planejamento ambiental a fim de verificar a presença de espécie invasora nas cidades. Estudos futuros podem analisar formas de minimizar as distorções das imagens do Google Street View e melhorar o desempenho na detecção da Leucaena.

REFERÊNCIAS

BARTHOLOMEU, C. S.; DE SOUSA, C. S. M; BRAZOLIN, S. De árvore invasora à matéria-prima–pesquisa sobre o potencial de uso da leucena para o design de produtos. **Estudos em Design**, v. 28, n. 2, 2020.

BILJECKI, F; ITO, K. Street view imagery in urban analytics and GIS: A review. **Landscape and Urban Planning**, v. 215, p. 104217, 2021.

BODEVAN, D. I. L; COELHO, C. W. G. A; MARQUES, A. R. Ocorrência da invasora Leucaena Leucocephala na regional pampulha em Belo Horizonte. **VII Congea**. Campina Grande, 2016.

DEMATTÊ, J. A; SAYÃO, V. M; RIZZO, R; FONGARO, C. T. Soil class and attribute dynamics and their relationship with natural vegetation based on satellite remote sensing. **Geoderma**, v. 302, p. 39-51, 2017.

KAMILARIS, A; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. **Computers and electronics in agriculture**, v. 147, p. 70-90, 2018.

LOWE, S.; BROWNE, M.; BOUDJELAS, S.; DE POORTER, M. 100 of the world's worst invasive alien species: a selection from the global invasive species database. **Invasive Species Specialist Group**, Auckland, Nova Zelândia, 2000.

MINAEE, S.; BOYKOV, Y.; PORIKLI, F.; PLAZA, A.; KEHTARNAVAZ, N.; TERZOPOULOS, D. Image Segmentation Using Deep Learning: A Survey. **arXiv preprint** arXiv:2001.05566, 2020.

NESBITT, L.; MEITNER, M.J.; GIRLING, C.; SHEPPARD, S.R.J.; LU, Y. Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities. **Landscape Urban Plann,** v. 181, p. 51–79, 2019.

XIE, Y; SHA, Z; YU, M. Remote sensing imagery in vegetation mapping: a review. **Journal of plant ecology**, v. 1, n. 1, p. 9-23, 2008.