

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**ALICE CAMARGO RODRIGUES DA SILVA**

**USO DE VISÃO COMPUTACIONAL PARA INVENTARIAR ÁRVORES URBANAS  
NA UTFPR CAMPUS LONDRINA**

**LONDRINA  
2025**

**ALICE CAMARGO RODRIGUES DA SILVA**

**USO DE VISÃO COMPUTACIONAL PARA INVENTARIAR ÁRVORES URBANAS  
NA UTFPR CAMPUS LONDRINA**

**Use of computer vision to inventory urban trees at UTFPR Londrina Campus**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Produção, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador(a): Rafael Henrique Palma Lima.

**LONDRINA**

**2025**



Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**ALICE CAMARGO RODRIGUES DA SILVA**

**USO DE VISÃO COMPUTACIONAL PARA INVENTARIAR ÁRVORES URBANAS  
NA UTFPR CAMPUS LONDRINA**

Trabalho de Conclusão de Curso de Graduação  
apresentado como requisito para obtenção do título  
de Bacharel em Engenharia de Produção, da  
Universidade Tecnológica Federal do Paraná  
(UTFPR).

Data de aprovação: 24/novembro/2025

---

Rafael Henrique Palma Lima  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Bruno Samways dos Santos  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Ligia Flávia Antunes Batista  
Doutorado  
Universidade Tecnológica Federal do Paraná

**LONDRINA**

**2025**

Dedico este trabalho aos meus pais, que me ensinaram o valor da educação e nunca deixaram de acreditar no meu potencial. Todo este caminho só foi possível graças ao amor e à força que recebi deles.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, por me conceder força, sabedoria e perseverança para superar cada etapa desta jornada acadêmica.

À minha família, deixo meu profundo agradecimento pelo apoio incondicional, pelas palavras de incentivo e pela compreensão nos momentos de ausência. Sem vocês, esta conquista não seria possível.

Aos meus professores, agradeço por todo o conhecimento compartilhado, pela dedicação e pela orientação que contribuíram diretamente para a construção deste trabalho e da minha formação.

Por fim, agradeço aos meus colegas de curso, que tornaram a caminhada mais leve, pelas conversas, trocas de experiência e apoio mútuo ao longo do caminho.

Meu sincero agradecimento a todos que, de alguma forma, contribuíram para a realização deste trabalho.

## RESUMO

Este trabalho apresenta uma pesquisa sobre a implementação de técnicas de visão computacional no mapeamento de árvores urbanas, utilizando imagens capturadas por drones e modelos de deep learning. O foco principal foi analisar a viabilidade e as restrições dos algoritmos DeepForest e YOLOv8 na detecção e contagem automática de árvores no campus da Universidade Tecnológica Federal do Paraná - Londrina. A abordagem metodológica incluiu a anotação manual de 177 árvores, o pré-processamento da imagem em tiles de diferentes dimensões e sobreposições, a eliminação de tiles sem informações relevantes e a aplicação de técnicas de aumento artificial de dados, permitindo o treinamento dos modelos sob diferentes configurações. O desempenho foi avaliado por meio de métricas tradicionais de visão computacional, como precisão, recall, F1 Score e mean Average Precision (mAP). Os resultados indicaram diferenças significativas: o DeepForest exibiu maior sensibilidade, com um recall superior a 60%, mas apresentou um alto número de falsos positivos, enquanto o YOLOv8 demonstrou uma seletividade superior, alcançando até 69% de precisão, embora com recall muito baixo, resultando em uma subcontagem considerável de árvores. Essa discrepância evidencia que nenhum dos modelos, de forma isolada, conseguiu alcançar um equilíbrio adequado entre confiabilidade e abrangência, embora ambos tenham mostrado ser promissores para aplicações práticas. As principais limitações identificadas no estudo foram o número reduzido de árvores anotadas e o número limitado de imagens disponíveis, o que restringiu a diversidade de exemplos para o treinamento e a validação dos modelos. Mesmo assim, o trabalho representa uma iniciativa inaugural na UTFPR - Londrina, evidenciando que a utilização de inteligência artificial pode reduzir custos e tempo em inventários de árvores urbanas, servindo como um ponto de partida para futuras investigações que abranjam bases de dados mais extensas, arquiteturas mais recentes, estratégias de aprendizado semi-supervisionado e a integração dos resultados a sistemas de georreferenciamento.

Palavras-chave: inventário arbóreo urbano; segmentação de imagens; *deep learning*; imagens de drone.

## ABSTRACT

This work presents research on the implementation of computer vision techniques in urban tree mapping, using images captured by drones and deep learning models. The main focus was to analyze the predictions and limitations of the DeepForest and YOLOv8 algorithms in the automatic detection and counting of trees on the campus of the Federal Technological University of Paraná - Londrina. The methodological approach included the manual annotation of 177 trees, image preprocessing in blocks of different dimensions and overlaps, elimination of blocks without relevant information, and the application of artificial data augmentation techniques, allowing the training of models under different configurations. Performance was evaluated through traditional computer vision analyses, such as accuracy, recall, F1 Score, and mean accuracy (mAP). The results indicated significant differences: DeepForest exhibited greater sensitivity, with a recall exceeding 60%, but presented a high number of false positives, while YOLOv8 demonstrated superior selectivity, reaching up to 69% accuracy, although with very low recall, resulting in a specific undercount of trees. This discrepancy highlights that neither model, in isolation, managed to achieve an adequate balance between reliability and comprehensiveness, although both showed promise for practical applications. The main limitations identified in the study were the limited number of annotated trees and the limited number of available images, which restricted the diversity of examples for training and validating the models. Even so, the work represents a pioneering initiative at UTFPR - Londrina, demonstrating that the use of artificial intelligence can reduce costs and time in urban tree inventories, implemented as a starting point for future investigations encompassing more extensive databases, more recent architectures, semi-supervised development strategies, and the integration of results into georeferencing systems.

Keywords: urban tree inventory; image segmentation; deep learning; drone imagery.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>6</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>8</b>
<b>2.1</b>	<b>Segmentação de imagens .....</b>	<b>8</b>
2.1.1	Tipos de Segmentação.....	9
<b>2.2</b>	<b>Inventário de árvores urbanas .....</b>	<b>10</b>
<b>2.3</b>	<b>Machine learning para inventário de árvores urbanas.....</b>	<b>13</b>
<b>3</b>	<b>MÉTODO DE PESQUISA .....</b>	<b>19</b>
<b>3.1</b>	<b>Área de Estudo e Conjunto de Dados .....</b>	<b>20</b>
<b>3.2</b>	<b>Anotação dos dados .....</b>	<b>22</b>
<b>3.3</b>	<b>Pré-processamento das Imagens .....</b>	<b>25</b>
<b>3.4</b>	<b>Modelos Utilizados .....</b>	<b>26</b>
3.4.1	DeepForest.....	26
3.4.2	YOLOv8.....	26
<b>3.5</b>	<b>Avaliação de Desempenho .....</b>	<b>27</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>28</b>
<b>4.1</b>	<b>Resultados com o <i>DeepForest</i>.....</b>	<b>32</b>
<b>4.2</b>	<b>Resultados com o YOLOv8 .....</b>	<b>39</b>
<b>4.3</b>	<b>Comparação dos resultados do <i>DeepForest</i> e YOLOv8 .....</b>	<b>42</b>
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>44</b>
	<b>REFERÊNCIAS.....</b>	<b>46</b>

## 1 INTRODUÇÃO

As árvores e florestas urbanas são elementos centrais da paisagem das cidades, não só por sua presença visual marcante, mas também pelos diversos benefícios ambientais, sociais e econômicos que proporcionam. Elas contribuem para a regulação do clima urbano, melhoria da qualidade do ar, redução da poluição sonora e promoção da biodiversidade, além de oferecerem espaços agradáveis para lazer e recreação. Do ponto de vista social, sua presença está ligada ao bem-estar psicológico da população, à valorização estética dos bairros e ao fortalecimento do vínculo das pessoas com a natureza. Florestas urbanas maduras, com árvores de grande porte, oferecem oportunidades de descanso, contato com o ambiente natural e até mesmo educação ambiental (TYRVÄINEN et al., 2005).

A relevância desses benefícios também é reconhecida em documentos oficiais, como o Plano Diretor de Arborização Urbana de Londrina (LONDRINA, 2013), que estabelece objetivos voltados à promoção, conservação e expansão das áreas verdes do município. Entre seus principais propósitos estão manter densidade arbórea adequada, assegurar a arborização como elemento de infraestrutura urbana, reduzir impactos climáticos, preservar a biodiversidade e orientar ações de planejamento, manejo e participação social. Esse plano evidencia que a arborização urbana deve ser tratada como componente estratégico da qualidade ambiental e do desenvolvimento sustentável da cidade.

A gestão eficiente das árvores urbanas é essencial para o desenvolvimento sustentável das cidades, considerando os múltiplos benefícios que essas estruturas verdes proporcionam. Desde a melhoria da qualidade do ar e conforto térmico até o valor simbólico e cultural que representam para as comunidades. No entanto, à medida que as árvores envelhecem e se tornam maiores, crescem também os riscos associados à sua manutenção, exigindo um equilíbrio entre segurança pública e preservação ambiental.

A realização de inventários arbóreos é uma ferramenta estratégica nesse processo, pois fornece dados fundamentais como localização, espécie, diâmetro do tronco, altura e condição geral das árvores. Essas informações subsidiam decisões orçamentárias, ações de mitigação de risco e políticas públicas voltadas à arborização urbana. Apesar disso, muitos municípios ainda enfrentam dificuldades em manter

inventários atualizados devido ao alto custo, à escassez de pessoal técnico e à complexidade logística das coletas manuais (ESPERON-RODRIGUEZ et al., 2023).

Nesse contexto, torna-se evidente a importância do inventário arbóreo como ferramenta contínua de gestão. Quando realizado de maneira sistemática, o inventário permite manter dados atualizados sobre o patrimônio arbóreo urbano, tornando o processo de tomada de decisão mais preciso e eficiente. Essa prática também contribui para reduzir custos operacionais, otimizar intervenções e apoiar estratégias de manejo, conservação e replanejamento da arborização (SILVA et al., 2007, apud OLIVEIRA et al., 2018).

Nesse cenário, o uso de tecnologias digitais, como sistemas móveis de coleta de dados e plataformas de mapeamento colaborativo, tem se mostrado promissor ao ampliar a escala e a eficiência do monitoramento arbóreo nas cidades (ROLOFF, 2016).

Este trabalho tem como objetivo principal investigar a aplicação de técnicas de segmentação de imagens com *deep learning* para identificar e quantificar automaticamente árvores urbanas em imagens aéreas obtidas por drone. A utilização de métodos automáticos para esse tipo de tarefa apresenta vantagens relevantes, como a redução do tempo de execução, menor necessidade de mão de obra especializada, diminuição de erros humanos e possibilidade de padronização dos levantamentos, além de viabilizar análises em áreas mais amplas com maior rapidez e menor custo operacional (VELASQUEZ-CAMACHO et al., 2023; WEINSTEIN et al., 2020).

Diante desse cenário, este trabalho se propõe a alcançar dois objetivos específicos: avaliar o desempenho de dois modelos de *deep learning* (*DeepForest* e *YOLOv8*) na detecção e contagem automática de árvores, e analisar as vantagens e limitações práticas de cada abordagem quando aplicadas ao contexto urbano do campus da UTFPR - Londrina. A proposta visa contribuir para o desenvolvimento de métodos mais eficientes e automatizados de inventário arbóreo, com foco em ambientes urbanos.

## 2 REFERENCIAL TEÓRICO

### 2.1 Segmentação de imagens

O desenvolvimento do processamento digital de imagens iniciou no começo do século XX, tendo como finalidade aprimorar a qualidade das imagens transmitidas pelo sistema Bartlane, um método de transmissão de imagens por cabogramas que utilizava técnicas de codificação e decodificação para enviar fotografias por cabos submarinos entre Londres e Nova Iorque. Posteriormente, na década de 1960, ganhou grande impulso com o uso de computadores para processar imagens da Lua transmitidas por sondas espaciais. Desde 1964, a área de processamento de imagens tem experimentado um crescimento notável, com usos que abrangem quase todas as áreas da atividade humana (MARQUES E NETO, 1999).

Nesse contexto, pode-se dizer que a visão, um dos sentidos mais aprimorados nos seres humanos, desempenha um papel fundamental na interpretação do ambiente e de imagens. Com base nessa habilidade, o processamento digital de imagens procura extrair informações relevantes de representações visuais obtidas por diversos aparelhos. De acordo com Gonzalez e Woods (2008), essa área envolve desde procedimentos básicos, como a correção do contraste, até atividades mais complexas, como a segmentação e o reconhecimento de objetos, sendo que uma imagem digital é formada por um conjunto organizado de píxeis.

A segmentação de imagens é um passo fundamental no processamento de imagens, responsável por identificar e separar áreas significativas em uma cena. Segundo Nikhil e Sankar (1993), a segmentação é o procedimento de dividir a imagem em áreas não sobrepostas, assegurando que cada área seja uniforme internamente e que a combinação de regiões adjacentes não compartilhe essa homogeneidade. Este princípio está diretamente ligado à noção de que a segmentação tem como objetivo identificar áreas diferentes na imagem, o que pode ser interpretado como a procura por uniformidade e contraste. Nesse sentido, segundo Moik (1980), os objetos exibem duas propriedades essenciais: uma uniformidade interna em relação a alguma característica da imagem e um contraste perceptível em relação ao seu entorno.

Em relação a isso, a segmentação é uma das etapas mais importantes na análise de dados de imagens processadas. O objetivo principal dessa etapa é dividir uma imagem em partes que se relacionam com objetos ou regiões do mundo real

presentes na cena. Essa divisão pode resultar em uma segmentação completa, caracterizada por regiões distintas que correspondem exclusivamente aos objetos presentes, ou em uma segmentação parcial, na qual as regiões não representam diretamente os objetos. A segmentação completa requer cooperação com níveis de processamento mais elevados, que utilizam conhecimento específico do domínio, enquanto problemas de segmentação mais simples podem ser resolvidos apenas com processamento de nível inferior (SONKA et al., 2013).

### 2.1.1 Tipos de Segmentação

De acordo com Anjna e Kaur (2017), as técnicas de segmentação de imagens podem ser classificadas em diferentes tipos, entre os quais se destacam:

- Limiarização (*Thresholding*): Separa a imagem de acordo com um valor de intensidade do limiar, que pode ser global, variável (local ou adaptável) ou múltiplo.
- Segmentação Baseada em Bordas: Concentra-se na identificação das bordas dos objetos, por meio da variação de intensidade entre os píxeis.
- Segmentação por Regiões: Fundamenta-se na homogeneidade dos píxeis, utilizando técnicas como a expansão de regiões e a divisão e fusão de regiões.
- Clusterização: Segmenta a imagem em clusters de píxeis semelhantes, que podem ser do tipo duro (cada píxel pertence a um único agrupamento) ou suave (um píxel pode pertencer a vários agrupamentos).

Essas técnicas variam em complexidade e aplicação, sendo escolhidas de acordo com as características específicas das imagens e os objetivos do processamento.

Além dessas técnicas clássicas, métodos mais recentes têm ganhado destaque devido aos avanços em *Deep Learning* e Redes Neurais. Alguns desses novos métodos incluem:

- *Segment Anything Model (SAM)*: desenvolvido pelo projeto *Segment Anything* da Meta em 2023, é uma abordagem inovadora para segmentação de imagens, projetada para ser adaptável a uma variedade de tarefas de segmentação. O SAM utiliza um modelo fundacional que pode ser facilmente ajustado a diferentes tipos de segmentação, como detecção de bordas, proposição de objetos e segmentação por texto, oferecendo uma flexibilidade única. A principal característica

do SAM é a capacidade de aceitar múltiplos tipos de entradas de prompt, como pontos, caixas delimitadoras, máscaras e até descrições textuais, permitindo uma segmentação interativa (ZHANG et al., 2023).

- Redes Neurais Convolucionais (CNNs): são eficazes no reconhecimento e classificação de imagens, aplicando filtros em regiões específicas da imagem para extrair características, como bordas e texturas. Ao contrário das redes tradicionais, as CNNs utilizam campos receptivos locais, onde os neurônios processam apenas partes da imagem, o que reduz o número de parâmetros e acelera o treinamento. Essa capacidade de aprender automaticamente as características relevantes torna as CNNs amplamente utilizadas em tarefas de visão computacional, como detecção de objetos e reconhecimento de imagens (VARGAS et al., 2016).

## 2.2 Inventário de árvores urbanas

Segundo Santos et al. (2025), a arborização urbana desempenha um papel fundamental na promoção da qualidade de vida nos centros urbanos, oferecendo uma série de benefícios ambientais, sociais e econômicos. No entanto, para que esses benefícios sejam alcançados de forma eficaz, é necessário um planejamento adequado, aliado a ações de manutenção contínua e à conscientização da população. Os autores destacam:

A arborização urbana é um dos elementos mais fundamentais no planejamento de cidades sustentáveis. Além de contribuir diretamente para a qualidade ambiental, as árvores nas áreas urbanas desempenham um papel essencial na promoção da saúde pública, no fortalecimento das relações sociais e na melhoria do ambiente econômico. (LOURENÇO et al., 2016, apud SANTOS et al., 2025)

Dessa forma, o investimento em estratégias sustentáveis e em uma gestão eficiente das áreas verdes torna-se essencial para superar os desafios existentes e promover cidades mais saudáveis e sustentáveis para as futuras gerações.

Diversos municípios do Brasil têm adotado o inventário de árvores como uma prática importante para auxiliar no planejamento urbano e na administração da arborização pública. Um exemplo é o município de Salto de Pirapora (SP), onde foi conduzido um levantamento em 24 bairros para avaliar qualitativa e quantitativamente

as espécies de árvores existentes nas vias públicas. Trata-se de um inventário tradicional, realizado sem o uso de técnicas de automação ou modelos de Machine Learning, baseado em observação direta e registros manuais. A pesquisa trouxe dados relevantes acerca da diversidade, origem e distribuição das espécies, bem como evidenciou questões estruturais e fitossanitárias, fundamentais para a criação de estratégias de gestão (BENATTI et al., 2012).

Outro exemplo é a cidade de Rio Branco (AC), que realizou um inventário da arborização urbana com o objetivo de não só identificar as espécies, mas também avaliar a condição geral das árvores, incluindo a condição física das copas e a altura. A pesquisa revelou que muitos indivíduos tinham comprometimento físico, além da necessidade de plantio, enfatizando a relevância da manutenção constante e da seleção cuidadosa das espécies cultivadas (PAIVA et al., 2010).

No município de Rorainópolis (RR), Santana et al. (2024) conduziram um estudo com o propósito de avaliar a condição da arborização em sete praças da cidade, por meio da identificação e inventário das variedades arbóreas existentes. Os autores ressaltam que o crescimento da população urbana aumenta a necessidade de áreas verdes, lançando desafios aos administradores públicos na tarefa de projetar espaços que favoreçam tanto a preservação ambiental quanto a saúde humana. O estudo incluiu avaliações fitossanitárias e fitossociológicas, bem como a avaliação das áreas disponíveis para novos plantios.

Esses inventários, além de proporcionarem vantagens operacionais, contribuem para a conscientização ambiental da população e para a elaboração de políticas públicas mais eficientes e sustentáveis. Isso possibilita o planejamento da arborização urbana com base em informações atualizadas, o que favorece a diminuição de riscos e a construção de um ambiente urbano mais seguro e saudável. Em Botelhos (MG), realizou-se um inventário florestal urbano com a finalidade de inventariar e caracterizar as espécies de árvores existentes nas vias públicas. O estudo apresentou informações acerca do número de árvores, diversidade, tamanho e a presença de conflitos com a infraestrutura urbana, servindo como base técnica para futuras iniciativas de gestão e planejamento ambiental (MORAIS et al., 2024).

A Lei nº 12.651/2012, que dispõe sobre a proteção da vegetação nativa no Brasil, prevê em seu Art. 71 a instituição de um Inventário Florestal Nacional, que será realizado pela União em colaboração com os Estados, o Distrito Federal e os

Municípios. O objetivo é fornecer informações para a avaliação da presença e qualidade das florestas no território nacional, incluindo tanto propriedades privadas quanto terras públicas. Embora o enfoque principal da lei seja voltado às áreas rurais e de vegetação nativa, sua diretriz de monitoramento florestal se estende ao ambiente urbano, na medida em que prevê a uniformização da coleta, manutenção e atualização das informações sobre a cobertura vegetal. Assim, os inventários de arborização urbana podem ser compreendidos como instrumentos contribuintes para o cumprimento da legislação nacional e para a formulação de políticas públicas voltadas à gestão sustentável das árvores nas cidades.

Várias cidades brasileiras vêm implementando legislações específicas que tornam obrigatório o inventário arbóreo como parte do planejamento urbano. Essas leis estabelecem diretrizes técnicas para o levantamento, mapeamento, cadastro e monitoramento das árvores, garantindo que o manejo da arborização urbana siga critérios sustentáveis e técnicos. Um exemplo é a cidade de São Paulo (SP), que regulamenta sua arborização por meio do Plano Municipal de Arborização Urbana (PMAU). Este plano está estabelecido como ação prioritária no Art. 288 da Lei Municipal nº 16.050/2014 - Plano Diretor Estratégico do Município de São Paulo. O PMAU é um instrumento para definir o planejamento e a gestão da arborização no município, visando o aumento da resiliência da cidade às mudanças climáticas, a qualificação da paisagem e a satisfação da população. Para isso, o PMAU propõe ainda a realização de inventários arbóreos sistemáticos.

Outro exemplo relevante é o município do Rio de Janeiro, que instituiu o Plano Diretor de Arborização Urbana (PDAU) por meio do Decreto nº 42.685/2016. O PDAU é um documento técnico que define as diretrizes necessárias para a implantação, monitoramento, avaliação, conservação e expansão da arborização urbana na cidade. Ele representa a primeira consolidação do planejamento da arborização do município e foi elaborado por uma equipe multidisciplinar com a participação social por meio de oficinas e consulta pública.

Em Porto Alegre (RS), a gestão da arborização urbana também tem avançado com o uso de tecnologias. Em abril de 2023, o município implementou o *Arbolink*, um sistema digital para o gerenciamento da arborização da cidade. Entre suas funcionalidades, destaca-se a realização do inventário quali-quantitativo das árvores do município. Esse tipo de inventário busca não apenas estimar a quantidade de

indivíduos arbóreos (avaliação quantitativa), mas também registrar informações sobre suas características e condições (avaliação qualitativa), como espécie, altura total e do tronco, diâmetro do tronco e da copa, e estado fitossanitário (PREFEITURA DE PORTO ALEGRE, [s.d.]).

### 2.3 Machine learning para inventário de árvores urbanas

Com o avanço das técnicas de *machine learning*, novas ferramentas têm sido desenvolvidas para automatizar o processo de inventário de árvores urbanas. Uma dessas ferramentas é o *DetecTree*, uma biblioteca em *Python*, de código aberto, voltada para a detecção de árvores individuais a partir de imagens aéreas. A ferramenta utiliza modelos baseados em redes neurais convolucionais treinadas com imagens RGB e permite a segmentação de copas de árvores com boa acurácia, mesmo sem a necessidade de dados multiespectrais ou sensores LiDAR. O *DetecTree* foi projetado com foco em acessibilidade, sendo facilmente adaptável a diferentes conjuntos de dados, oferecendo uma alternativa acessível e eficaz aos métodos tradicionais (BOSCH, 2020).

Além do *DetecTree*, existem outros modelos para identificação de copas de árvores, como o *DeepForest*, que é uma ferramenta também desenvolvida em *Python* que utiliza inteligência artificial para identificar automaticamente as copas das árvores em imagens aéreas comuns (RGB). De acordo com Weinstein *et al.* (2020), o modelo foi treinado com milhares de imagens de diferentes florestas e tipos de vegetação, o que permite que ele funcione bem em várias regiões do mundo. Uma das grandes vantagens do *DeepForest* é que ele já vem com um modelo pré-treinado, mas também permite que o usuário melhore os resultados ajustando o modelo com dados locais. Isso é feito por meio de aprendizado supervisionado, onde as copas reais são desenhadas nas imagens RGB e usadas para ajustar o modelo. O treinamento personalizado permite que o *DeepForest* se adapte a diferentes tipos de vegetação, ambientes urbanos e condições de iluminação, aumentando a precisão da segmentação.

Sendo assim, o uso de algoritmos de *deep learning* tem se mostrado eficaz para superar os altos custos e a complexidade dos inventários arbóreos urbanos tradicionais. No estudo de Velasquez-Camacho *et al.* (2023), foi desenvolvida uma metodologia baseada em modelos de *deep learning*, como YOLOv5, *DeepForest* e

Faster R-CNN, com o objetivo de detectar, contar e geolocalizar árvores urbanas a partir de diferentes fontes de imagens. A pesquisa foi aplicada na cidade de Lleida, na Espanha, e utilizou imagens do Google Street View e imagens aéreas/satélite.

Para validar o desempenho dos modelos de detecção, os pesquisadores utilizaram diferentes estratégias para estabelecer a quantidade real de árvores presentes nas áreas analisadas. Em primeiro lugar, foi utilizado o inventário arbóreo oficial da cidade de Lleida, na Espanha, contendo a localização georreferenciada das árvores urbanas. Além disso, mais de 40.000 árvores foram anotadas manualmente com o uso da ferramenta Labellmg, a partir de imagens de satélite, aéreas e de nível de solo. Por fim, uma validação visual foi realizada em um subconjunto amostral de 9,25 hectares, onde foram inspecionadas manualmente 267 árvores, possibilitando a construção de matrizes de confusão e a análise detalhada de acertos e erros dos modelos.

Entre os resultados para a detecção de árvores em imagens de nível aéreo/satélite, o modelo YOLOv5x atingiu a maior precisão, sendo esta igual a 86,9%. Enquanto os demais modelos, Deepforest e Faster R-CNN, alcançaram precisões de 79,6% e 74,6%, respectivamente. Além disso, o Deepforest e o YOLOv5 apresentaram tempos de processamento mais rápidos que o Faster R-CNN. As métricas utilizadas incluíram precisão, recall e F-score. O estudo demonstrou ser possível produzir inventários urbanos com alta acurácia e baixo custo, integrando dados de múltiplas origens com modelos de deep learning e ferramentas de rotulagem acessíveis (VELASQUEZ-CAMACHO et al., 2023).

Avançando nas abordagens baseadas em inventários de árvores urbanas com o uso de machine learning, é destacado o estudo de Onishi e Ise (2021), que propuseram uma metodologia de identificação e mapeamento de árvores a partir de imagens RGB obtidas por veículos aéreos não tripulados (UAVs). A pesquisa foi conduzida em uma área experimental da Universidade de Kyoto, no Japão, e teve como principal diferencial a utilização de técnicas de segmentação orientadas a objetos para a extração individual das copas das árvores, etapa essencial para garantir a precisão do mapeamento.

O método utilizado foi a segmentação orientada a objetos, mais especificamente o algoritmo Multiresolution Segmentation, executado na plataforma eCognition Developer v9.0.0. O algoritmo foi parametrizado manualmente por

tentativa e erro, com o objetivo de otimizar a separação individual das copas. Além disso, os autores destacam que os resultados da segmentação foram avaliados visualmente com base em cinco categorias de correspondência entre os objetos segmentados e as copas reais observadas: correspondente (sobreposição superior a 80%), quase igual (sobreposição entre 60% e 80%), divisão (sobreposição entre 20% e 60%), mesclado (quando múltiplas copas reais foram agrupadas em um único polígono segmentado) e fragmentado (sobreposição inferior a 20%). Essa avaliação qualitativa foi aplicada em quatro áreas diferentes, permitindo verificar a acurácia da segmentação das copas antes da etapa de classificação das espécies (ONISHI E ISE, 2021).

Com as copas segmentadas, os autores treinaram diferentes tipos de redes neurais convolucionais (CNNs) para classificar as árvores por espécie. A CNN ResNet152 obteve o melhor desempenho, com acurácia de 97,6% e F1-score médio de 0,962. O uso de imagens segmentadas corretamente foi essencial para garantir que as CNNs aprendessem padrões relevantes sem interferência de ruídos do solo ou de sobreposição de vegetação.

Outro exemplo para o inventário automatizado de árvores urbanas foi desenvolvido por Ventura *et al.* (2024), que propuseram um método de detecção de árvores baseado em *deep learning*, utilizando imagens multiespectrais de alta resolução do programa norte-americano NAIP (*National Agricultural Imagery Program*). A pesquisa foi conduzida em larga escala no estado da Califórnia, abrangendo oito cidades e seis zonas climáticas diferentes. O objetivo principal do estudo foi identificar e mapear automaticamente as localizações de árvores individuais em ambientes urbanos, cobrindo tanto áreas públicas quanto privadas, com o intuito de apoiar o planejamento urbano e o monitoramento ambiental.

O modelo utilizado pelos autores é baseado na arquitetura HR-SFANet, uma rede neural convolucional desenvolvida para detectar árvores com alta precisão. Essa rede combina mecanismos de atenção (que auxiliam o modelo a focar nas partes mais relevantes da imagem) com uma técnica chamada regressão de mapas de confiança, que indica os pontos mais prováveis onde há uma árvore. O treinamento foi feito com imagens aéreas multiespectrais, contendo quatro bandas (vermelho, verde, azul e infravermelho próximo), além do índice de vegetação NDVI, usado para identificar

vegetação verde viva. A base de dados contou com mais de 1500 imagens e quase 100 mil árvores anotadas manualmente.

A arquitetura do modelo usa como base a VGG-16, uma rede neural conhecida e adaptada para receber essas imagens com mais canais. O modelo foi treinado com a biblioteca *TensorFlow* e *Keras*, usando técnicas como normalização dos dados, aumento artificial do conjunto de imagens e o otimizador *Adam* para melhorar o desempenho. Esse método gera um mapa de calor, onde os pontos mais intensos indicam maior probabilidade de presença de uma árvore. Para cada árvore anotada, é criada uma representação em forma de função Gaussiana. Durante a inferência, o modelo detecta os pontos mais fortes do mapa como posições das árvores. Além disso, o processo de treinamento foi ajustado automaticamente com a ferramenta *Optuna*, que busca os melhores hiperparâmetros para maximizar o desempenho da rede.

Quanto ao desempenho, os resultados demonstraram uma precisão de 76,5% e recall de 73,3% quando o modelo foi treinado e testado em toda a Califórnia. A pontuação *F-Score* atingiu 0,735 e o erro quadrático médio (RMSE) foi de apenas 2,157 metros. Essas métricas indicam um bom desempenho, especialmente quando é considerada a diversidade climática e topográfica da região avaliada. O estudo também comparou o desempenho do seu método com outras abordagens, como *PyCrown* e *DeepForest*, demonstrando que o HR-SFANet superou ambas em precisão, tempo de inferência e capacidade de generalização.

Além dos modelos mencionados, novas abordagens continuam sendo exploradas, como por exemplo o estudo recente de Teng *et al.* (2025), que avaliou o uso do modelo *Segment Anything* (SAM), desenvolvido pela Meta, na tarefa de segmentar copas de árvores em imagens de drone. As imagens foram coletadas em uma área de floresta no Canadá, com resolução de 5 mm por píxel, incluindo imagens RGB e modelos de elevação (DSM). Os autores testaram diferentes formas de usar o SAM, tanto de maneira automática quanto com o auxílio de informações de altura para orientar a segmentação. No entanto, os resultados evidenciaram que, quando usado sem adaptações, o SAM teve dificuldades para separar copas muito próximas umas das outras e também confundiu arbustos ou vegetação baixa com árvores. Para melhorar o desempenho, os autores propuseram ajustes usando dados de altura e uma abordagem chamada *DSMPrompter*, que melhorou bastante os resultados. A

pesquisa mostrou que o SAM tem potencial, mas precisa ser adaptado quando aplicado a tarefas específicas, como o inventário de árvores a partir de imagens aéreas (TENG *et al.*, 2025).

Um exemplo notável do uso de *deep learning* em inventário arbóreo urbano é o estudo de Korznikov *et al.* (2021), conduzido em florestas temperadas mistas do Extremo Oriente Russo, utilizando imagens do satélite *GeoEye-1*. Os autores aplicaram uma arquitetura do tipo *U-Net* para o reconhecimento preciso de copas de árvores em imagens de satélite RGB de altíssima resolução (0,46 m/píxel). O modelo foi treinado para classificar cada píxel como pertencente ou não a uma copa arbórea, utilizando máscaras binárias previamente construídas como referência. Para a criação dessas máscaras, os autores realizaram uma delimitação manual das copas das árvores, com base em imagens capturadas no início da primavera, o que permitiu diferenciar coníferas (com folhagem permanente) das árvores decíduas da espécie *Populus suaveolens* (sem folhas nessa estação). Cada copa delimitada manualmente foi considerada uma árvore individual, possibilitando a definição da quantidade real de árvores no local.

A avaliação do modelo foi realizada com base nas métricas de acurácia média, *F1-Score* e *Intersection over Union* (IoU), atingindo desempenho superior aos métodos tradicionais como *Random Forest*, KNN e *Naive Bayes*. O *U-Net* atingiu acurácia média de 0,96, IoU de até 0,94 e F1 Score de até 0,97, evidenciando sua alta capacidade de segmentação precisa com base somente em dados RGB. Um dos diferenciais do método foi a capacidade de distinguir padrões espaciais e geométricos das copas das árvores, reduzindo significativamente os falsos positivos causados por objetos com cores semelhantes (como telhados ou corpos d'água). Além disso, os autores destacaram que a escolha da época do ano (início da primavera) foi essencial para diferenciar coníferas (verdes o ano todo) de árvores decíduas ainda sem folhas.

O quadro a seguir apresenta um resumo das principais referências abordadas neste tópico, destacando os métodos utilizados para o inventário arbóreo, bem como os resultados obtidos em cada estudo. Essa síntese facilita a compreensão dos avanços e das diferentes abordagens aplicadas na área.

**Quadro 1 - Resumo dos métodos utilizados e resultados das referências**

Referência	Método Utilizado	Principais Resultados
------------	------------------	-----------------------

Bosch (2020)	Biblioteca <i>DetecTree</i> com redes neurais convolucionais em <i>Python</i> para segmentação de copas usando imagens RGB.	Boa acurácia na segmentação sem uso de dados multiespectrais ou LiDAR; ferramenta acessível e adaptável.
Weinstein et al. (2020)	<i>DeepForest (Python)</i> , rede pré-treinada ajustável via aprendizado supervisionado para identificar copas em imagens RGB.	Modelo pré-treinado com alta flexibilidade e boa precisão em diferentes regiões e tipos de vegetação.
Korznikov et al. (2021)	<i>U-Net</i> para segmentação de copas em imagens de satélite <i>GeoEye-1</i> (RGB, 0,46 m/píxel).	Alta acurácia (0,96), IoU (0,94), <i>F1 Score</i> (0,97); diferenciou coníferas de decíduas usando época do ano.
Onishi e Ise (2021)	Segmentação orientada a objetos ( <i>Multiresolution Segmentation</i> ) e CNN ResNet152 para classificação por espécie com imagens UAV RGB.	Segmentação com avaliação qualitativa detalhada; CNN ResNet152 alcançou 97,6% de acurácia e F1-score 0,962.
Velasquez-Camacho et al. (2023)	Modelos YOLOv5, <i>DeepForest</i> e <i>Faster R-CNN</i> para detecção e geolocalização com imagens aéreas e <i>Google Street View</i> .	YOLOv5x atingiu 86,9% de precisão; <i>DeepForest</i> 79,6%; <i>Faster R-CNN</i> 74,6%. Processamento rápido e alta acurácia.
Ventura et al. (2024)	HR-SFANet (CNN com atenção + regressão de mapas de confiança) com imagens multiespectrais NAIP.	Precisão 76,5%, recall 73,3%, F-score 0,735 e RMSE 2,157 m; desempenho superior a <i>PyCrown</i> e <i>DeepForest</i> .
Teng et al. (2025)	Modelo <i>Segment Anything</i> (SAM) adaptado com dados de altura (DSM) para segmentação de copas em imagens de drone RGB.	SAM mostrou potencial, mas precisou de adaptações para separar copas próximas; melhorias com <i>DSMPrompter</i> .

**Fonte: Autoria própria (2025).**

### 3 MÉTODO DE PESQUISA

O presente trabalho foi desenvolvido de forma exploratória e aplicada, com o objetivo de testar e avaliar a viabilidade do uso das técnicas de *deep learning*, especificamente os modelos *DeepForest* e *YOLOv8*, para identificação e quantificação automática de árvores em uma imagem aérea capturada por drone do campus da Universidade Tecnológica Federal do Paraná - Câmpus Londrina. O processo metodológico foi conduzido em etapas sequenciais, conforme descrito a seguir. A Figura 1 apresenta o fluxograma geral do pipeline adotado, incluindo o pré-processamento, a anotação manual e os caminhos específicos para o *DeepForest* e *YOLOv8*.

**Figura 1 - Fluxograma geral do pipeline metodológico**



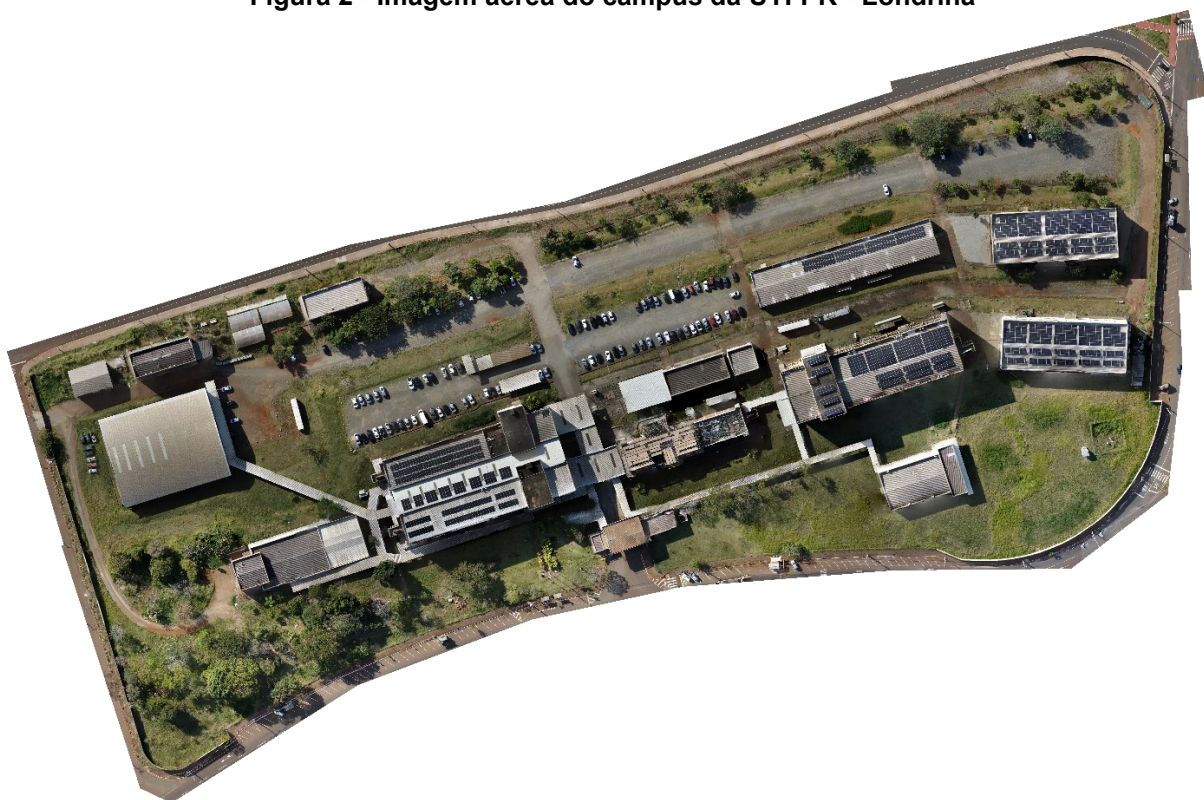
Fonte: Autoria própria (2025).

Além disso, todo o desenvolvimento da versão final do trabalho foi realizado em linguagem *Python*, utilizando a biblioteca *Ultralytics YOLOv8* para o treinamento e a detecção, além de ferramentas auxiliares para o pré-processamento das imagens. Com o objetivo de garantir a reprodutibilidade e transparência da pesquisa, o código-fonte completo está disponibilizado publicamente em um repositório no GitHub (SILVA, 2025).

### 3.1 Área de Estudo e Conjunto de Dados

A área de estudo corresponde ao campus da UTFPR - Londrina, para o qual foi disponibilizada uma imagem aérea em alta resolução, aproximadamente 5 centímetros por píxel (0,04999 x 0,04999 m). A imagem, no formato .tif, não foi capturada pela autora deste trabalho, sendo um dado secundário fornecido pela instituição, obtido por meio de um drone institucional. Essa imagem, apresentada na Figura 2, serviu como base principal para o treinamento e a validação dos modelos, seu formato compunha 6370 píxeis de altura x 9656 píxeis de largura x 4 canais.

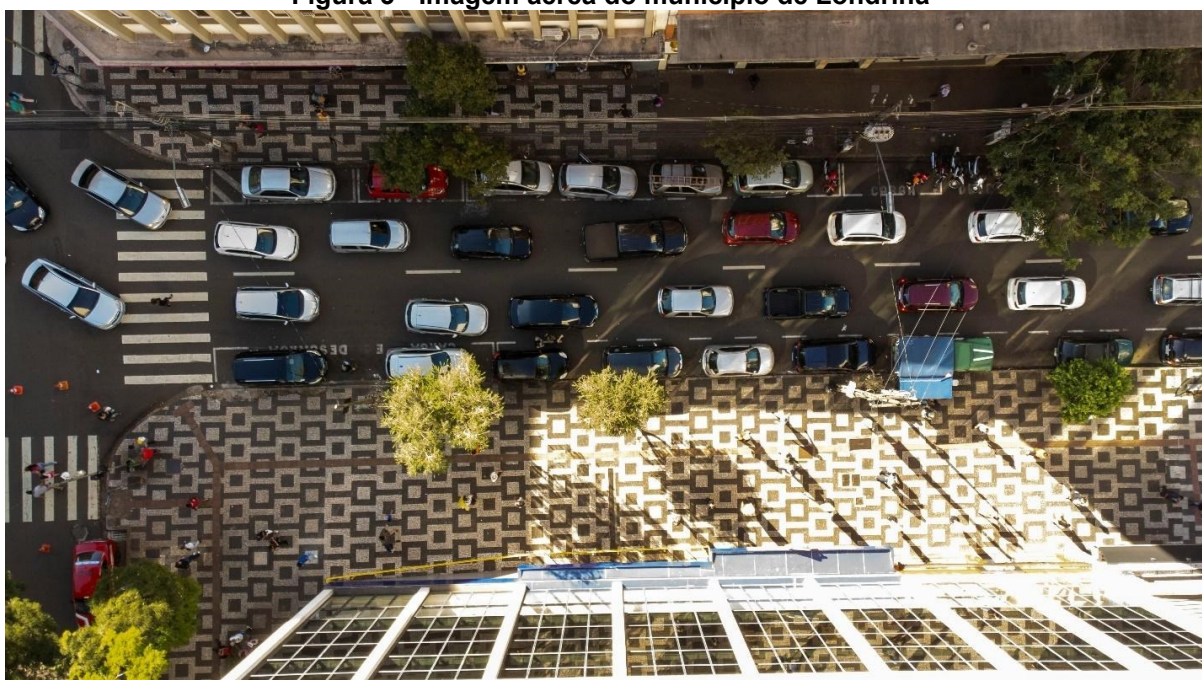
Figura 2 - Imagem aérea do campus da UTFPR - Londrina



Fonte: Nuvem UTFPR (2025).

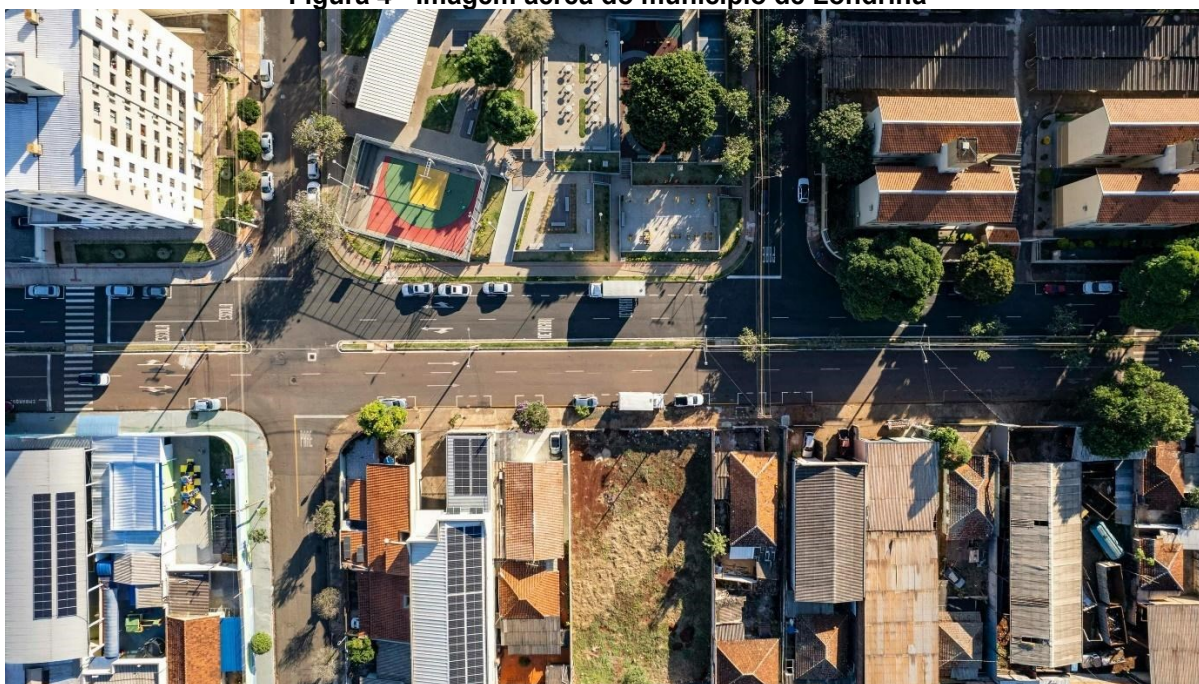
Complementarmente, em determinados experimentos, foram utilizadas imagens de áreas urbanas do município de Londrina obtidas no site Pexels, capturadas por drone com ângulo de 90°, como nas Figuras 3 e 4. A utilização dessas imagens teve como objetivo ampliar a diversidade do conjunto de dados e avaliar a capacidade de generalização do modelo em diferentes contextos urbanos.

**Figura 3 - Imagem aérea do município de Londrina**



**Fonte: Rodolfo Gaion (2025).**

**Figura 4 - Imagem aérea do município de Londrina**



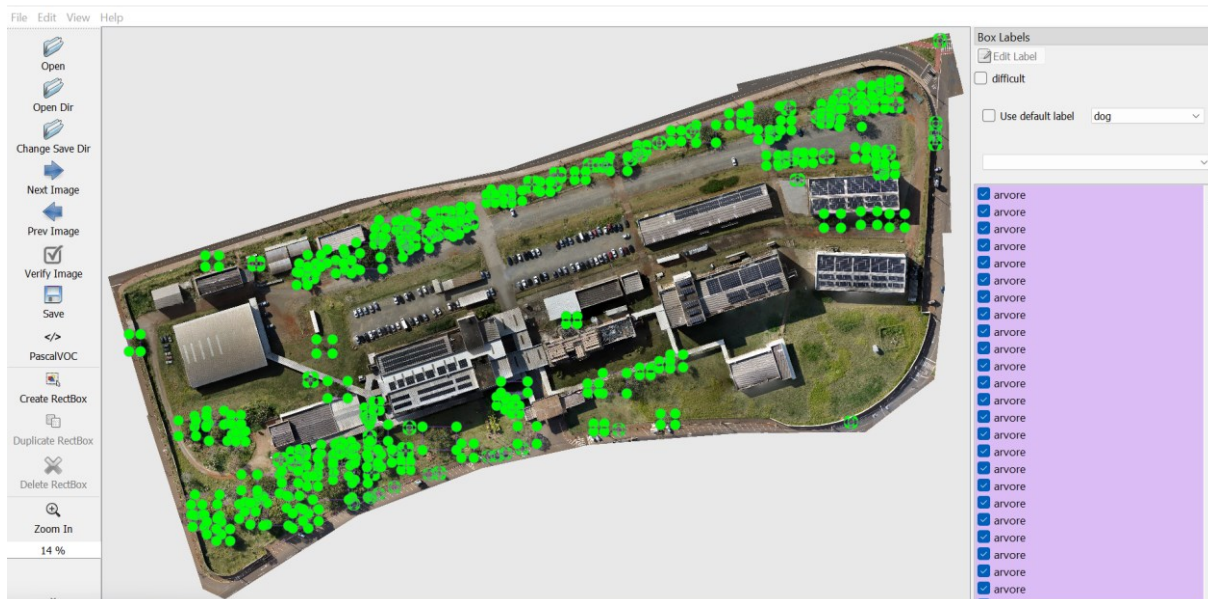
**Fonte: Rodolfo Gaion (2025).**

### **3.2 Anotação dos dados**

Para possibilitar o treinamento supervisionado dos modelos, foi necessário criar um conjunto de anotações manuais das árvores presentes na imagem do campus. Essa tarefa foi realizada utilizando o *software Labellmg*, ferramenta de código aberto amplamente utilizada em aplicações de visão computacional.

No *Labellmg*, cada árvore visível na imagem do campus foi delimitada manualmente por meio de caixas delimitadoras (*bounding boxes*), assim como na Figura 5, totalizando 177 árvores.

**Figura 5 - Delimitação manual das árvores através do *Labellmg***



Fonte: Autoria própria (2025).

Da mesma maneira com que foi feito na imagem do campus, as imagens do site Pexels passaram pelo processo de rotulação manual através do *LabelImg*. Ao todo, foram utilizadas 8 imagens do fotógrafo Rodolfo Gaion, retiradas do site, contendo 118 anotações no total.

As anotações foram salvas em arquivos no formato XML (PASCAL VOC), contendo as informações de posição (coordenadas  $(x, y)$  dos vértices da caixa) e da classe atribuída (“árvore” ou “*tree*”), assim como mostra a Figura 6 abaixo.

Figura 6 - Arquivo XML com as anotações das árvores

```

▼ <annotation>
  <folder>londrina</folder>
  <filename>pexels-gaion-23709254.jpg</filename>
  <path>C:\Users\alice\Downloads\teste\londrina\pexels-gaion-23709254.jpg</path>
  ▼ <source>
    <database>Unknown</database>
  </source>
  ▼ <size>
    <width>4000</width>
    <height>2250</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  ▼ <object>
    <name>Tree</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    ▼ <bndbox>
      <xmin>1177</xmin>
      <ymin>1654</ymin>
      <xmax>1448</xmax>
      <ymin>1915</ymin>
    </bndbox>
  </object>
  ▼ <object>
    <name>Tree</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    ▼ <bndbox>
      <xmin>1777</xmin>
      <ymin>1604</ymin>
      <xmax>2483</xmax>
      <ymin>2033</ymin>
    </bndbox>
  </object>

```

Fonte: Autoria própria (2025).

Esses arquivos XML foram fundamentais em duas etapas do trabalho. A primeira consistiu no treinamento, em que as anotações serviram como rótulos para os modelos poderem aprender a reconhecer o padrão visual das árvores. A segunda etapa foi a avaliação, pois permitiu a comparação entre as predições dos modelos e as anotações reais, possibilitando o cálculo das métricas de desempenho.

Ao longo do processo de anotação e preparação dos dados, foi necessário utilizar diferentes ferramentas computacionais que desempenharam papéis fundamentais na execução dos experimentos. A Tabela 1 apresenta um resumo das principais ferramentas empregadas, suas finalidades e observações relevantes, de modo a documentar o ambiente de trabalho e facilitar a reprodutibilidade deste estudo por outros pesquisadores.

Tabela 1 - Principais ferramentas utilizadas

Ferramenta	Finalidade	Observações
<i>LabelImg</i>	Anotação manual das árvores.	Exportação no formato XML (PASCAL VOC).
<i>DeepForest</i>	Detecção de copas de árvores.	Utilizado em versão pré-treinada e com <i>fine-tuning</i> .
YOLOv8	Detecção de objetos em imagens.	Aplicado exclusivamente à imagem do campus.
<i>PyTorch</i>	Plataforma de aprendizado profundo.	Execução dos experimentos com bibliotecas auxiliares como NumPy e Pandas.

Fonte: Autoria própria (2025).

### 3.3 Pré-processamento das Imagens

Antes do treinamento, a imagem original do campus foi convertida de quatro canais para três canais (RGB) e, em seguida, subdividida em blocos menores (*tiles*) de diferentes tamanhos: 1024 × 1024 píxeis, 640 × 640 píxeis e 600 × 600 píxeis. Cada configuração foi testada com diferentes valores de sobreposição (*overlap*). Essa subdivisão buscou equilibrar a captura de detalhes das copas das árvores com a quantidade de informações por bloco, além de aumentar o número de exemplos disponíveis para o treinamento. A sobreposição entre *tiles* foi adotada para evitar perdas de objetos que se encontravam nas bordas das subdivisões. Valores baixos de *overlap* reduzem a redundância, mas podem perder árvores cortadas na borda, e valores maiores aumentam a cobertura.

Além disso, após a etapa de subdivisão, foi realizada uma filtragem automática para remover *tiles* que continham apenas píxeis pretos (sem informação útil). Essa situação ocorreu porque, como as dimensões da imagem original não eram múltiplos exatos do tamanho dos *tiles*, algumas subdivisões correspondiam a áreas fora da região de interesse. A exclusão desses blocos evitou que exemplos irrelevantes fossem utilizados no treinamento, o que poderia introduzir ruído nos dados e comprometer o aprendizado dos modelos.

Adicionalmente, foram aplicadas técnicas de aumento artificial de dados (*data augmentation*), incluindo rotações, espelhamentos horizontais e verticais e variações

de brilho e contraste. Essas estratégias visaram aumentar a diversidade do conjunto de treinamento e reduzir o risco de sobreajuste (*overfitting*), situação em que o modelo se adapta excessivamente aos dados de treino e perde capacidade de generalização.

### 3.4 Modelos Utilizados

#### 3.4.1 DeepForest

O primeiro modelo selecionado para os experimentos foi o *DeepForest*, uma biblioteca em *Python* projetada especificamente para a detecção de copas de árvores a partir de imagens RGB. Foram conduzidas três abordagens distintas: uso do modelo pré-treinado disponibilizado pelo *DeepForest*; treinamento personalizado (*fine-tuning*), utilizando exclusivamente os *tiles* gerados da imagem do campus; e treinamento personalizado combinando *tiles* do campus com imagens adicionais obtidas no site Pexels. O modelo foi selecionado por já ser validado em estudos prévios e por possuir um pré-treino robusto, que poderia ser adaptado ao cenário urbano de Londrina.

O treinamento foi realizado em diferentes cenários: somente com imagens do campus, apenas com imagens do Pexels, de forma mista (campus e Pexels em treino e teste) e de forma híbrida (100% das imagens do Pexels e apenas 50% das imagens do campus no treino, reservando o restante do campus para validação). O número de épocas variou entre 17, 40 e 60, ajustando-se parâmetros como a taxa de aprendizado, o limiar de pontuação (*score threshold*) e o tamanho do lote de treinamento (*batch size*).

#### 3.4.2 YOLOv8

O segundo modelo testado foi o *YOLOv8*, uma rede de detecção de objetos de última geração. Diferentemente do *DeepForest*, esse modelo foi aplicado exclusivamente à imagem do campus porque, embora inicialmente tenham sido testadas imagens adicionais obtidas na plataforma Pexels, os resultados foram insatisfatórios. As diferenças de resolução, iluminação e ângulo de captura entre essas imagens e a imagem principal do campus prejudicaram significativamente o desempenho do modelo, resultando em falhas de detecção e *overfitting*. Por esse

motivo, optou-se por não incluir essas imagens no conjunto final de treinamento. Foram realizados dois experimentos principais. O primeiro com *tiles* de  $1024 \times 1024$  píxeis, sobreposição 0,4 e 100 épocas, e o segundo seguindo a configuração de *tiles* de  $640 \times 640$  píxeis, sobreposição de 0,4 e 100 épocas. Ambas as configurações utilizaram as anotações geradas no *Labellmg* em formato XML como base para o treinamento.

Durante a preparação do conjunto de treinamento para o *YOLOv8*, além da etapa de geração dos *tiles*, foi aplicada uma filtragem adicional para excluir todas as imagens que não continham nenhuma árvore anotada. Essa decisão teve como objetivo evitar o desequilíbrio entre classes e reduzir a presença de exemplos negativos irrelevantes, que poderiam induzir o modelo a aprender padrões sem relação com o objeto de interesse. Assim, apenas *tiles* contendo pelo menos uma anotação de árvore foram utilizados no processo de treino, o que contribuiu para aumentar a eficiência do aprendizado e a consistência dos resultados obtidos.

### 3.5 Avaliação de Desempenho

Os modelos foram avaliados a partir de duas métricas principais: precisão e *recall*. A precisão corresponde à proporção de árvores detectadas corretamente em relação ao total de detecções realizadas pelo modelo, enquanto o *recall* indica a proporção de árvores corretamente identificadas em relação ao total de árvores existentes nas imagens anotadas manualmente. O cálculo dessas métricas foi realizado a partir da comparação entre as detecções produzidas pelos modelos e os arquivos de referência gerados no processo de anotação com o *software Labellmg*, que serviram como base de dados utilizada como modelo para a validação dos resultados.

## 4 RESULTADOS E DISCUSSÃO

O presente capítulo apresenta os resultados obtidos a partir da aplicação dos modelos *DeepForest* e YOLOv8 na tarefa de detecção automática de árvores em imagens aéreas do campus da UTFPR - Londrina. Os experimentos foram realizados com diferentes configurações de parâmetros, tais como tamanho dos *tiles*, valores de sobreposição, número de épocas de treinamento e uso de técnicas de *data augmentation*, como mostra a Tabela 2.

Tabela 2 - Desempenho obtido nos diferentes testes realizados antes da aplicação na imagem completa

Análise	Modelo	Dataset usado	Tamanho dos tiles	Overlap	Épocas	Precisão (%)	Recall (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)
1	<i>DeepForest</i> (pré-treinado)	Campus	-	-	-	21	46	-	-
2	<i>DeepForest</i> (fine-tuning)	Campus	1024 × 1024	0,1	40	45,25	80,54	-	-
3	<i>DeepForest</i> (fine-tuning)	Campus + Pexels	1024 × 1024	0,1	60	38,08	79,57	-	-
4	<i>DeepForest</i> (fine-tuning)	Campus + Pexels	1024 × 1024	0,1	17	40,09	65,46	-	-
5	<i>DeepForest</i> (fine-tuning)	Campus + <i>Augmentations</i>	1024 × 1024	0,1	40	21,63	72,34	-	-
6	<i>DeepForest</i> (fine-tuning)	Campus	600 × 600	0,1	40	34,05	83,99	-	-
7	<i>DeepForest</i> (fine-tuning)	Campus	1024 × 1024	0,2	40	23,73	78,04	-	-
8	YOLOv8	Campus	1024 × 1024	0,4	100	79,16	61,70	73,81	43,88
9	YOLOv8	Campus	640 × 640	0,4	100	71,57	63,43	68,36	39,53

Fonte: Autoria própria (2025).

Os resultados obtidos contemplam tanto métricas quantitativas (como precisão, *recall*, *F1 Score* e *mean Average Precision* (mAP)) quanto análises qualitativas sobre os acertos e falhas de cada modelo. Observa-se que, quando aplicados à imagem completa do campus, os valores de desempenho foram ligeiramente inferiores em relação à validação realizada com *tiles*, conforme mostra a Tabela 3. Nessa etapa, definiu-se o valor de IoU (*Intersection over Union*) em 0,5 para o *DeepForest* e em 0,4 para o YOLOv8. A IoU, também chamada de índice de Jaccard, é amplamente empregada para mensurar a similaridade entre formas, como as caixas delimitadoras em tarefas de detecção de objetos (REZATOFIGHI *et al.*, 2019). Já a métrica *Average Precision* (AP) corresponde à área sob a curva de precisão em função do *recall*, expressando o equilíbrio entre essas duas medidas ao longo de diferentes limiares de detecção. A *mean Average Precision* (mAP), por sua vez, é obtida pela média dos valores de AP calculados para todas as classes e se consolidou como a principal métrica para avaliar o desempenho de modelos de detecção de objetos (PADILLA *et al.*, 2021).

Tabela 3 - Desempenho obtido nos diferentes testes realizados

Análise	Modelo	Dataset usado	Tamanho dos tiles	Overlap	Épocas	Precisão (%)	Recall (%)	F1 Score (%)	Observações
1	<i>DeepForest</i> (pré-treinado)	Campus	-	-	-	21	46	29	Resultado inicial sem ajustes; baixo desempenho geral
2	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus	1024 × 1024	0,1	40	31	66	43	Melhor equilíbrio no <i>DeepForest</i> ; <i>recall</i> elevado, mas precisão limitada
3	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus + Pexels	1024 × 1024	0,1	60	31	62	41	Inclusão do Pexels manteve <i>recall</i> , mas não melhorou a precisão
4	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus + Pexels	1024 × 1024	0,1	17	30	41	34	Treino curto (17 épocas) reduziu <i>recall</i> e F1 Score
5	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus + <i>Augmentations</i>	1024 × 1024	0,1	40	17	62	27	<i>Augmentation</i> elevou <i>recall</i> , mas reduziu drasticamente a precisão
6	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus	600 × 600	0,1	40	20	66	31	<i>Tiles</i> menores mantiveram <i>recall</i> alto, mas com baixa precisão
7	<i>DeepForest</i> ( <i>fine-tuning</i> )	Campus	1024 × 1024	0,2	40	9	70	15	<i>Overlap</i> maior aumentou <i>recall</i> , mas causou excesso de falsos positivos
8	YOLOv8	Campus	1024 × 1024	0,4	100	69	14	23	Alta precisão, mas <i>recall</i> muito baixo; elevado número de falsos negativos
9	YOLOv8	Campus	640 × 640	0,4	100	61	19	29	Redução da precisão sem muitos ganhos de <i>recall</i>

Fonte: Autoria própria (2025).

#### 4.1 Resultados com o *DeepForest*

Os experimentos realizados com o modelo *DeepForest* revelaram diferenças importantes entre o uso direto do modelo pré-treinado e os ajustes realizados por meio de *fine-tuning* com dados do campus da UTFPR. Na configuração inicial (Análise 1), em que o modelo foi aplicado sem ajustes, os resultados foram bastante limitados quando aplicados na imagem completa: apenas 21% de precisão, 46% de *recall* e *F1 Score* de 29%.

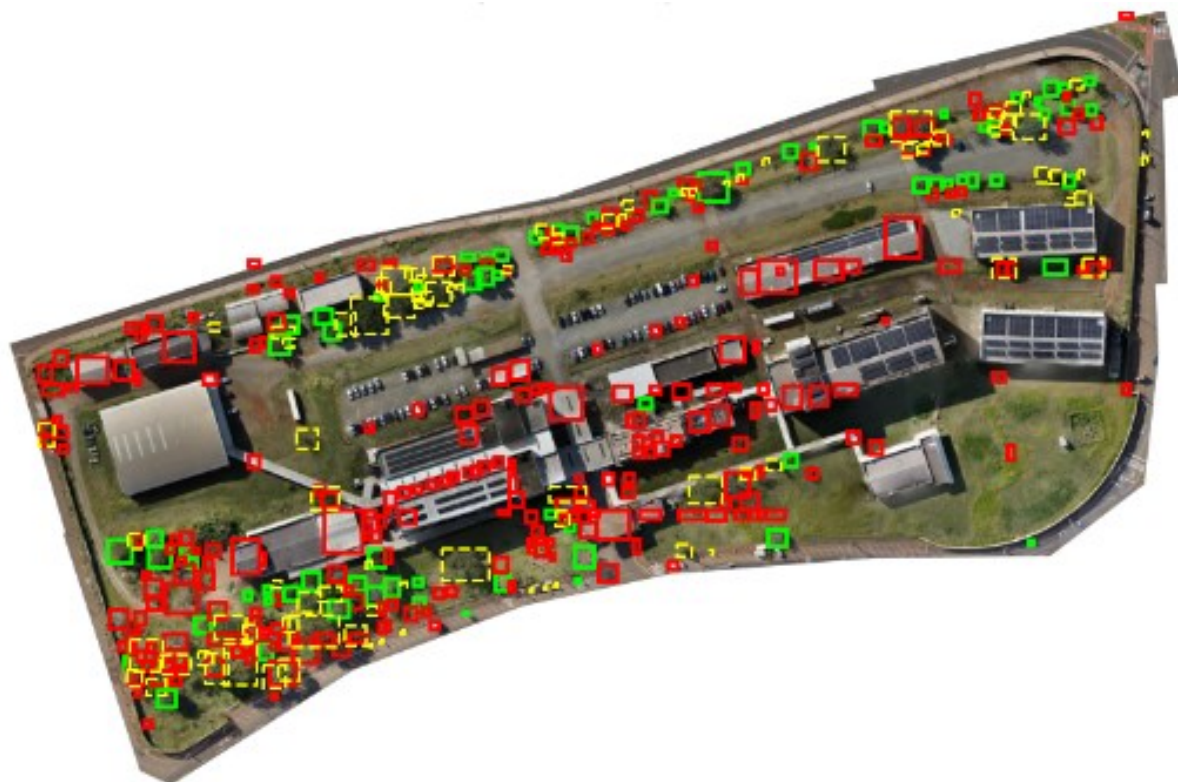
Na prática, esse desempenho significa que, de cada 100 detecções realizadas pelo modelo, apenas 21 corresponderam a árvores reais, enquanto 79 foram falsos positivos. Em um contexto de inventário urbano, esse nível de erro poderia induzir gestores públicos a superestimarem de forma expressiva a quantidade de árvores, comprometendo a confiabilidade das informações levantadas.

Esses valores confirmam a necessidade de adaptação local do modelo, já que o pré-treino original não refletia as características específicas da vegetação do campus. De acordo com Weinstein *et al.* (2020), o modelo *DeepForest* foi pré-treinado com milhões de copas de árvores geradas a partir de dados LiDAR e imagens RGB coletadas em florestas naturais pelo *National Ecological Observatory Network* (NEON), nos Estados Unidos. Essas imagens apresentam características distintas da cena estudada neste trabalho.

A Figura 7 mostra os efeitos do modelo pré-treinado do *DeepForest*. Nessa imagem, e nas demais presentes neste trabalho, caixas verdes indicam detecções corretas (verdadeiros positivos), vermelhas representam falsos positivos e amarelas marcam árvores que não foram detectadas (falsos negativos). Nota-se que, embora algumas copas mais isoladas tenham sido identificadas corretamente, regiões com sombras ou sobreposição de árvores geraram confusões significativas.

Portanto, a Análise 1 evidencia que, para contextos específicos como o da UTFPR - Londrina, o uso direto do modelo pré-treinado não é adequado. É indispensável realizar ajustes finos (*fine-tuning*) com dados locais para que o modelo possa capturar as características visuais particulares da vegetação urbana da região.

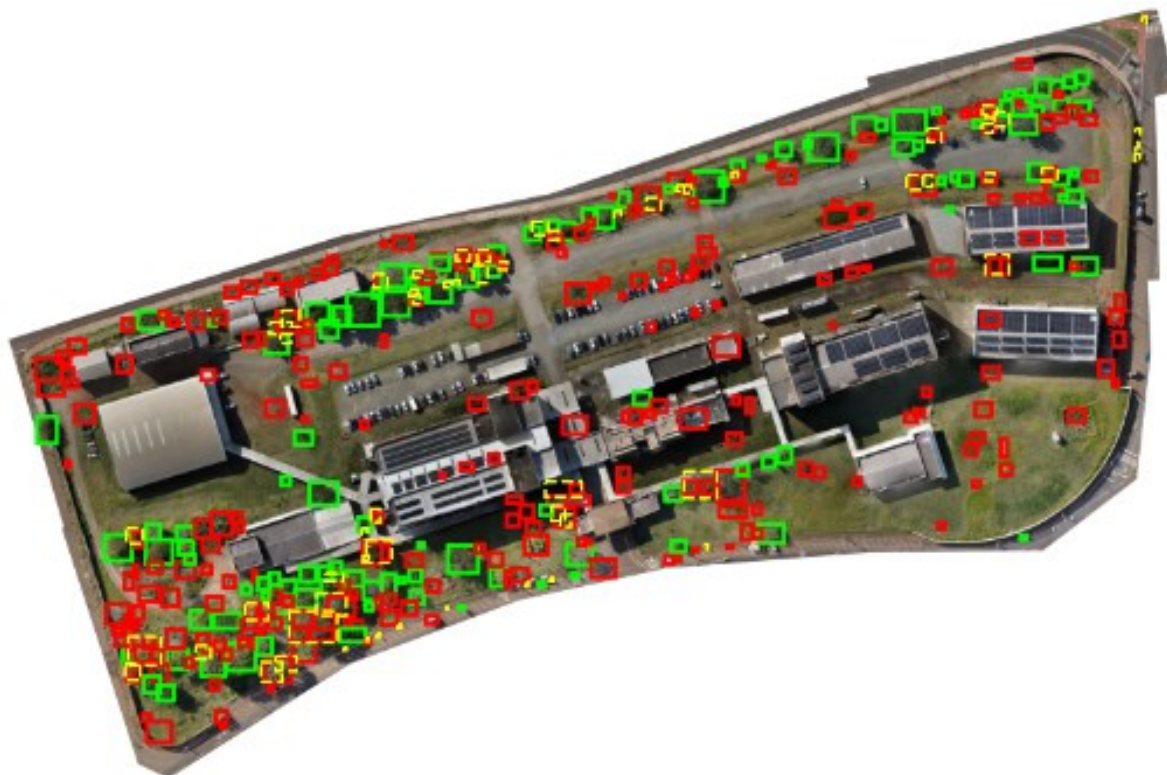
**Figura 7 - Análise 1 utilizando *Deepforest***



Fonte: Autoria própria (2025).

Com o *fine-tuning* em *tiles* de  $1024 \times 1024$  píxeis, sobreposição de 0,1 e 40 épocas (Análise 2), houve uma melhora significativa nos resultados: a precisão subiu para 31%, o *recall* atingiu 66% e o *F1 Score* chegou a 43%. Nesse cenário, ainda que a precisão permanecesse relativamente baixa, a sensibilidade do modelo aumentou consideravelmente, permitindo detectar uma quantidade maior de árvores.

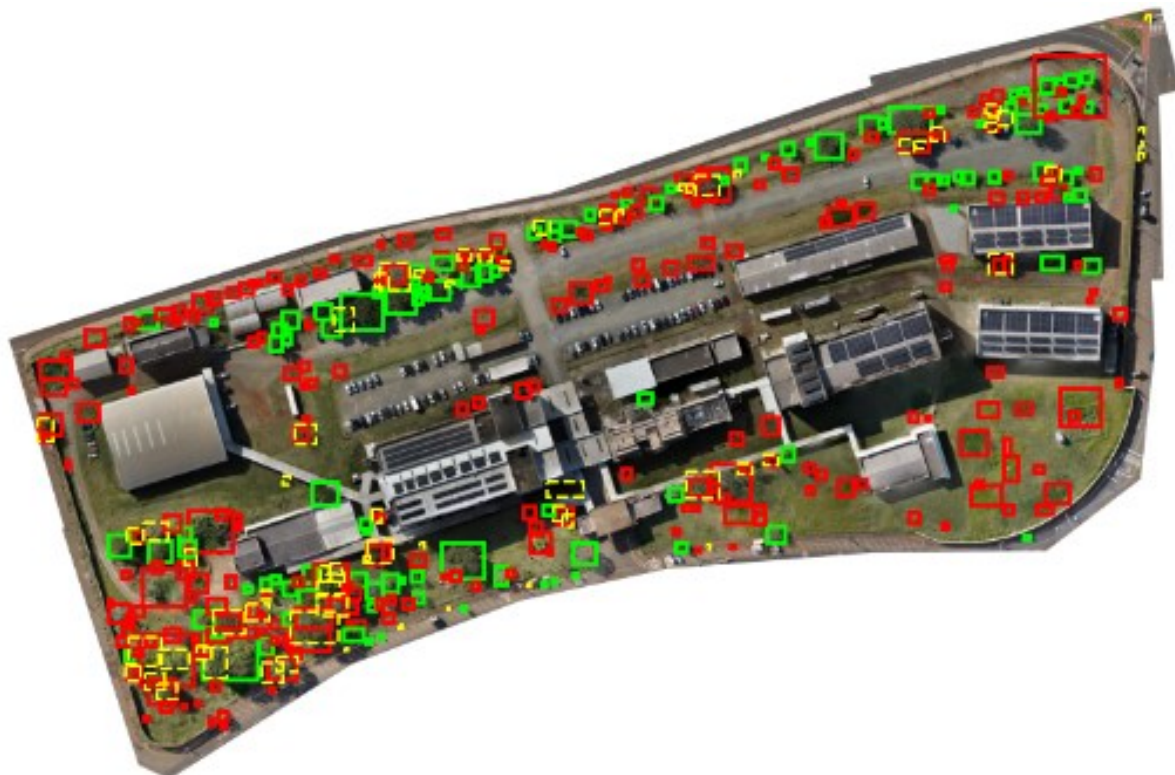
Em termos de inventário urbano, esse comportamento pode ser considerado positivo, pois um *recall* elevado significa que o risco de subestimar a quantidade de árvores é menor. No entanto, o aumento de falsos positivos implica que parte das detecções exigiria revisão manual, o que reduz a eficiência do processo automatizado. A Figura 8 demonstra visualmente essa melhoria: nota-se que áreas antes ignoradas pelo modelo passaram a ser detectadas, embora ao custo de novas detecções incorretas.

**Figura 8 - Análise 2 utilizando Deepforest**

Fonte: Autoria própria (2025).

Na Análise 3, evidenciada na Figura 9, foram adicionadas imagens do *dataset* Pexels ao conjunto de treino, com o objetivo de diversificar os dados e avaliar se a inclusão de cenários urbanos distintos poderia contribuir para melhorar a generalização do modelo. Os resultados demonstraram que o *recall* permaneceu relativamente alto (62%), mas a precisão continuou em 31%, sem avanços expressivos. Esse comportamento sugere que a introdução de imagens externas não melhora necessariamente a capacidade do modelo de reconhecer árvores no contexto específico da UTFPR, podendo inclusive gerar ruídos decorrentes das diferenças de iluminação, ângulo de captura e vegetação local.

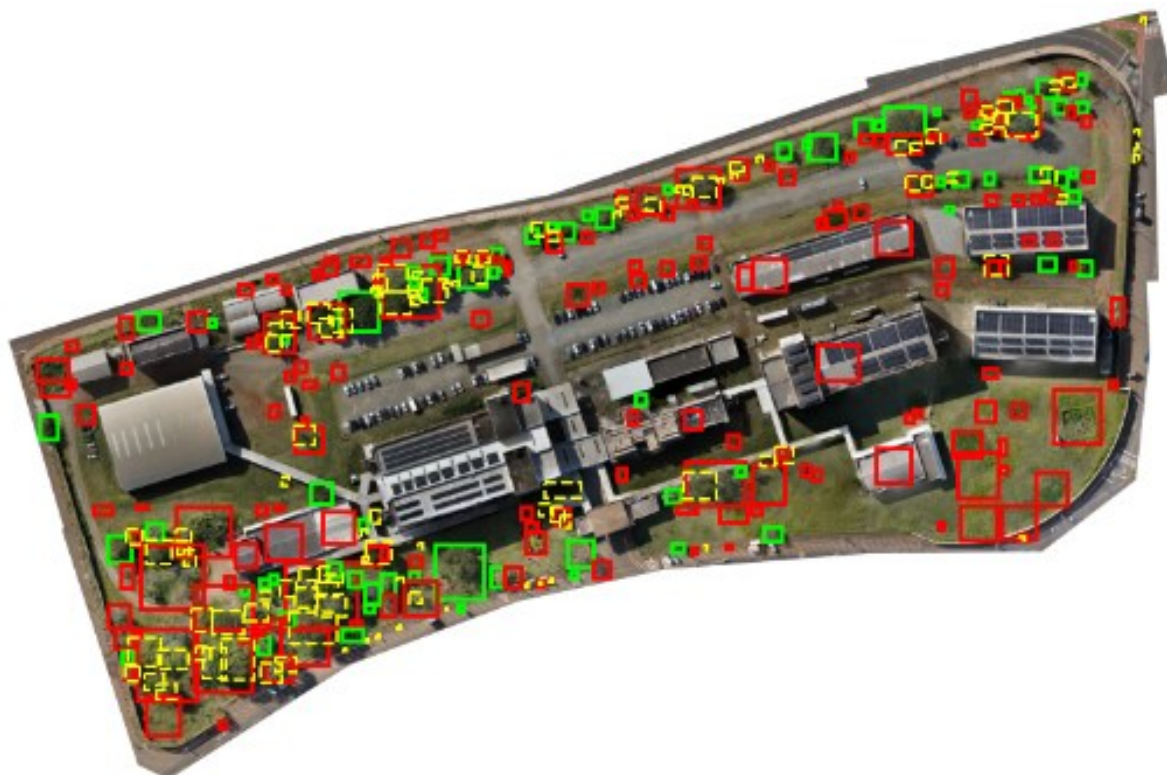
**Figura 9 - Análise 3 utilizando Deepforest**



Fonte: Autoria própria (2025).

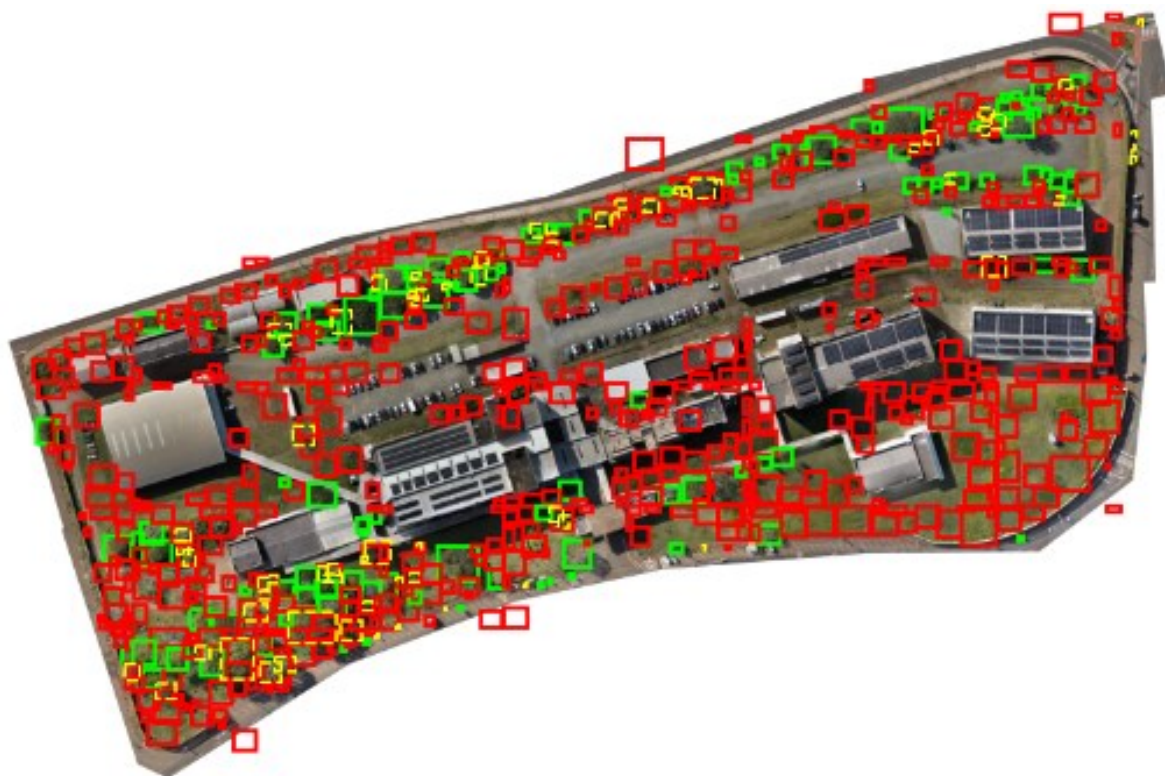
Um resultado semelhante foi observado na Análise 4 (Figura 10), na qual o treinamento foi interrompido precocemente, com apenas 17 épocas, resultando em queda do *recall* para 41% e do *F1 Score* para 34%, mostrando que a quantidade de épocas de treino exerce influência direta sobre a capacidade de generalização do modelo. Em outras palavras, o treinamento incompleto resultou em uma rede que não conseguiu captar de forma adequada os padrões visuais das copas de árvores. Esse resultado comprova o entendimento de que, em modelos baseados em redes neurais profundas, o número de épocas deve ser suficiente para garantir a harmonia do aprendizado, sob risco de comprometer a generalização.

Figura 10 - Análise 4 utilizando *Deepforest*



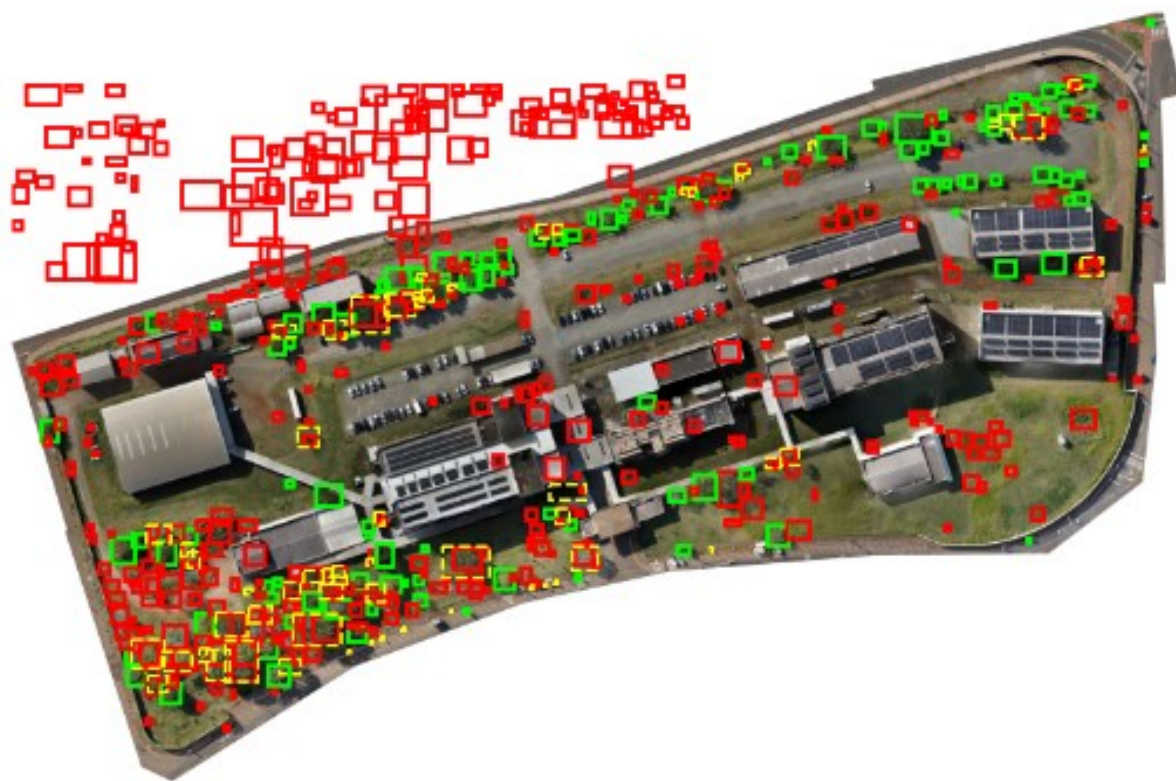
Fonte: Autoria própria (2025).

Na Análise 5, foram aplicadas técnicas de *data augmentation*, como rotações, espelhamentos e variações de brilho e contraste. Embora o *recall* tenha atingido 62%, a precisão caiu drasticamente para 17%, refletindo em um *F1 Score* de apenas 27%. Isso significa que o modelo detectou mais árvores, mas também classificou incorretamente muitas regiões como copas, aumentando o número de falsos positivos, assim como mostra a Figura 11.

Figura 11 - Análise 5 utilizando *Deepforest*

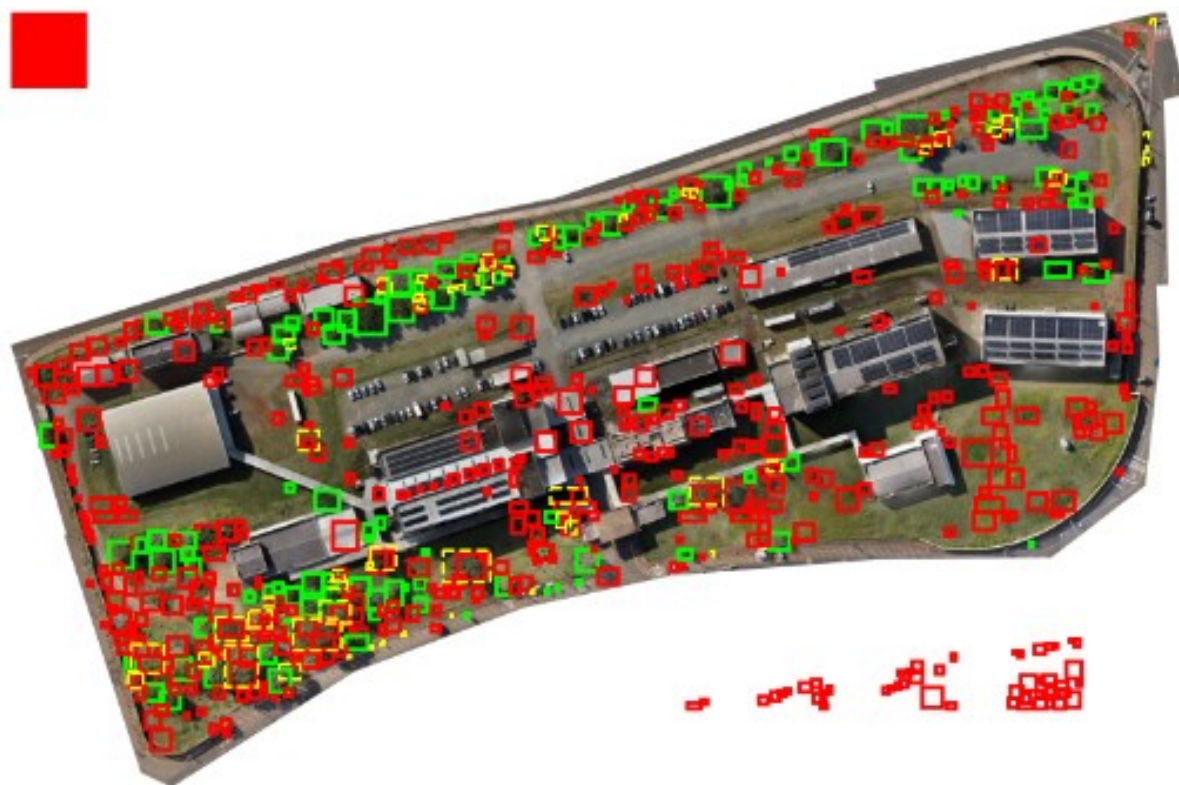
Fonte: Autoria própria (2025).

Por outro lado, na Análise 6, ao utilizar *tiles* menores ( $600 \times 600$  píxeis) com sobreposição de 0,1, obteve-se 20% de precisão e 66% de *recall*, o que demonstra que a redução do tamanho dos blocos não trouxe ganhos relevantes de desempenho. A falta de melhora nos resultados indica que o tamanho de *tiles* mais restrito acabou reduzindo o contexto espacial, o que dificultou a identificação das copas das árvores (Figura 12). Em outras palavras, apesar de os *tiles* menores fornecerem mais detalhes locais, o modelo deixou de ter a visão global necessária para distinguir corretamente as árvores de outros elementos urbanos parecidos.

Figura 12 - Análise 6 utilizando *Deepforest*

Fonte: Autoria própria (2025).

Finalmente, na Análise 7, foi adotada uma sobreposição maior, de 0,2. Nesse caso, o *recall* atingiu 70%, mas a precisão caiu para apenas 9%, resultando em um *F1 Score* de 15%. Esse cenário reforça a tendência de que valores mais altos de sobreposição aumentam a sensibilidade do modelo, permitindo identificar mais árvores, mas comprometem excessivamente a confiabilidade devido ao número elevado de falsos positivos (Figura 13). Para um inventário urbano, essa configuração seria inviável, pois exigiria uma etapa de correção manual ainda maior, anulando os ganhos de automação.

Figura 13 - Análise 7 utilizando *Deepforest*

Fonte: Autoria própria (2025).

De forma geral, os resultados confirmam que o *DeepForest* apresenta desempenho significativamente melhor quando adaptado ao contexto local, mas ainda enfrenta dificuldades para alcançar precisão elevada. Isso indica que, embora o modelo seja eficiente para identificar árvores em maior quantidade (*recall* elevado), sua aplicação prática em inventários urbanos ainda necessita de complementação com métodos de filtragem ou combinação com outros modelos, a fim de alcançar um equilíbrio mais adequado entre sensibilidade e confiabilidade.

#### 4.2 Resultados com o YOLOv8

O modelo YOLOv8 foi avaliado em duas configurações distintas, ambas aplicadas exclusivamente ao *dataset* do campus da UTFPR - Londrina. Na oitava análise de treinamento, utilizando *tiles* de  $1024 \times 1024$  píxeis, sobreposição de 0,4 e 100 épocas, o modelo YOLOv8 apresentou 69% de precisão e apenas 14% de *recall*. Esses valores indicam que o modelo conseguiu identificar corretamente a maior parte

das árvores que foram classificadas como positivas, porém deixou de detectar muitas outras que estavam presentes na cena.

Esse comportamento revela um viés conservador do modelo: ele evita gerar muitas previsões incorretas, mas, por outro lado, não consegue abranger a totalidade dos objetos de interesse. Em termos práticos, significa que quando o modelo indica a presença de uma árvore, a probabilidade de estar correto é alta, entretanto, uma grande quantidade de árvores não é reconhecida.

A Figura 14 ilustra visualmente esse cenário no ortomosaico do campus da UTFPR - Londrina, onde se observa a predominância de objetos não detectados, principalmente em regiões sombreadas ou em áreas com elevadas quantidades de elementos urbanos. Essa análise evidencia a necessidade de ajustes adicionais na etapa de treinamento, de modo a melhorar o equilíbrio entre precisão e *recall* e ampliar a capacidade de generalização do modelo para diferentes condições do ambiente estudado.

Figura 14 - Análise 8 utilizando YOLOv8



Fonte: Autoria própria (2025).

Na segunda configuração de treinamento (Análise 9), utilizando *tiles* menores de  $640 \times 640$  píxeis, mas mantendo os mesmos parâmetros de sobreposição (0,4) e número de épocas (100), o YOLOv8 apresentou 39% de precisão e 13% de *recall*. Relativamente à configuração anterior, observa-se uma considerável redução na precisão, sem ganhos expressivos no *recall*, o que indica que a diminuição do tamanho dos *tiles* comprometeu a capacidade do modelo em detectar corretamente as árvores, ao mesmo tempo em que manteve elevada a taxa de omissões. A Figura 15 evidencia esse comportamento, destacando um número ainda maior de falsos positivos em regiões próximas a construções, além da permanência de falsos negativos em áreas de vegetação densa ou sombreadas. Esses resultados sugerem que a utilização de *tiles* menores pode introduzir excesso de ruído contextual, dificultando a generalização do modelo e reforçando a necessidade de ajustes mais robustos nos parâmetros de treinamento.

Figura 15 - Análise 9 utilizando YOLOv8



Fonte: Autoria própria (2025).

### 4.3 Comparação dos resultados do *DeepForest* e YOLOv8

A comparação entre os modelos *DeepForest* e YOLOv8 revela diferenças significativas em termos de desempenho e limitações. O *DeepForest*, mesmo após várias etapas de *fine-tuning*, manteve níveis relativamente baixos de precisão, embora em algumas configurações tenha alcançado valores de *recall* superiores a 60%. Isso mostra que o modelo conseguiu identificar a maior parte das árvores presentes na cena, mas ao custo de um número elevado de falsos positivos, o que reduz sua utilidade em aplicações que exigem maior confiabilidade nas detecções.

Por outro lado, o YOLOv8 demonstrou comportamento oposto: nas duas configurações avaliadas, destacou-se pela maior precisão em comparação ao *DeepForest*, mas apresentou *recall* bastante reduzido, revelando uma tendência de omitir objetos reais presentes no ortomosaico. Na Análise 8 (*tiles* de 1024 × 1024 píxeis), o modelo atingiu 69% de precisão, porém apenas 14% de *recall*, enquanto na Análise 9 (640 × 640 píxeis) a precisão caiu para 39%, sem ganhos significativos no *recall* (13%).

Esses resultados revelam que, embora o YOLOv8 tenha se mostrado mais eficiente em evitar falsos positivos, não foi capaz de representar adequadamente a distribuição real das árvores no campus, devido ao grande número de falsos negativos. Em compensação, o *DeepForest*, ainda que menos seletivo, capturou melhor a diversidade de árvores, embora acompanhado de erros de classificação.

Cabe ressaltar que esses valores foram obtidos na aplicação sobre a imagem completa do campus. Nos *tiles* de validação, utilizados em condições controladas, as métricas de desempenho foram superiores, tanto para o *DeepForest* quanto para o YOLOv8. Essa diferença evidencia que o processo de recomposição da cena e as variações contextuais presentes no ortomosaico (sombras, construções, elementos urbanos) impactam negativamente os modelos, reduzindo sua capacidade de generalização.

De forma geral, pode-se afirmar que o *DeepForest* apresentou maior sensibilidade, favorecendo aplicações em que seja preferível identificar o maior número possível de árvores mesmo com algum ruído adicional, enquanto o YOLOv8 apresentou maior seletividade, sendo mais indicado em contextos que demandam maior confiabilidade nas detecções individuais, ainda que com perda de abrangência. Essa comparação ressalta a importância de ajustar não apenas os parâmetros de

treinamento, mas também de considerar as características específicas do ambiente de aplicação na escolha do modelo mais adequado.

Cabe destacar que todo o processamento, treinamento e avaliação dos modelos foram realizados em um computador com processador Intel Core i3 (10ª geração), 4 GB de memória RAM e armazenamento SSD de 128 GB.

O Quadro 2 resume a comparação entre os dois modelos aplicados neste trabalho.

**Quadro 2 - Comparação entre os modelos *DeepForest* e *YOLOv8***

<b>Critério</b>	<b><i>DeepForest</i></b>	<b><i>YOLOv8</i></b>
Abordagem	Modelo específico para detecção de copas de árvores; já possui versão pré-treinada com dados globais.	Detector de objetos de uso geral, ajustado para identificar árvores em <i>tiles</i> .
Desempenho em <i>tiles</i> (validação)	<i>Recall</i> elevado (até 83,9%), mas precisão limitada (entre 20–45%).	Métricas mais equilibradas, com precisão superior (até 79,1%) e <i>recall</i> intermediário (61,7%).
Desempenho na imagem completa	<i>Recall</i> manteve-se mais alto (até 66–70%), mas com grande número de falsos positivos.	Precisão elevada (até 69%), mas <i>recall</i> muito baixo (13–14%), com forte tendência a omitir árvores.
Sensibilidade (detecção do maior número de árvores)	Maior sensibilidade; reduz risco de subcontagem, mas exige revisão manual devido a falsos positivos.	Baixa sensibilidade; elevada subcontagem de árvores devido a falsos negativos.
Seletividade (confiabilidade das detecções)	Baixa seletividade; muitas árvores detectadas incorretamente.	Alta seletividade; quando detecta, há maior chance de ser correto.
Vantagens	Detecta mais árvores, incluindo em áreas sombreadas ou complexas.	Reduz falsos positivos; adequado para aplicações que demandam confiança nas detecções individuais.
Limitações	Excesso de falsos positivos prejudica a aplicabilidade direta.	<i>Recall</i> baixo compromete a representatividade da vegetação.
Aplicações mais adequadas	Inventários exploratórios, monitoramento de cobertura e análises que priorizam abrangência.	Inventários seletivos, situações em que é preferível detectar menos árvores, mas com maior confiabilidade.

**Fonte: Autoria própria (2025).**

## 5 CONCLUSÃO

Este trabalho buscou avaliar o uso de modelos de aprendizado profundo para auxiliar no inventário de árvores urbanas, aplicando os algoritmos *DeepForest* e YOLOv8 sobre imagens aéreas de drone do campus da UTFPR - Londrina. A proposta foi verificar se tais ferramentas, originalmente desenvolvidas em outros contextos, poderiam ser adaptadas de forma prática à realidade local.

Os resultados mostraram diferenças marcantes entre os modelos. O *DeepForest*, após ajustes, alcançou maior capacidade de detectar árvores (*recall* elevado), mas com grande número de falsos positivos. O YOLOv8, por sua vez, apresentou maior precisão, porém deixou de identificar muitas árvores. Essa oposição indica que nenhum dos modelos, isoladamente, conseguiu oferecer um equilíbrio satisfatório entre sensibilidade e seletividade, mas ambos demonstraram potencial para aplicações futuras.

Outro ponto importante é que os resultados foram mais consistentes na etapa de validação com *tiles* do que na aplicação sobre a imagem completa. Isso ocorre porque, na validação, o modelo é avaliado apenas em uma amostra limitada dos *tiles*, geralmente mais equilibrada e representativa, o que favorece métricas mais altas. Já na aplicação sobre todos os *tiles* e na recomposição da cena, surgem novos desafios: há maior diversidade de situações, presença de áreas mais complexas ou com menos informação, além da possibilidade de sobreposições e inconsistências entre *tiles* vizinhos. Esses fatores acabam elevando o número de falsos positivos e falsos negativos, o que reduz os valores finais das métricas de desempenho.

Diante disso, não é possível afirmar que um único modelo seja suficiente para atender todas as necessidades de um inventário arbóreo urbano. O *DeepForest* pode ser útil quando se busca não perder árvores na contagem, ainda que seja necessário revisar manualmente as detecções incorretas. Já o YOLOv8 tende a ser mais indicado quando a prioridade é a confiabilidade em cada detecção, mesmo que isso implique em subcontagem.

As limitações encontradas não se restringiram às imagens utilizadas, mas envolveram também o número reduzido de árvores anotadas manualmente e de imagens disponíveis, o que limitou o treinamento. Esses fatores ajudam a explicar por que os resultados, embora válidos, ainda não atingiram o mesmo nível de desempenho de pesquisas mais consolidadas.

Para trabalhos futuros, abrem-se diversas possibilidades, como ampliar o conjunto de dados anotados com a colaboração de mais pesquisadores ou alunos; testar novas arquiteturas de redes, como variantes do YOLOv9 ou modelos baseados em segmentação; aplicar técnicas de *transfer learning* mais robustas, aproveitando modelos já treinados em vegetações semelhantes; e explorar aprendizado semi-supervisionado ou ativo, reduzindo a necessidade de anotações manuais.

Além disso, uma perspectiva promissora seria avaliar combinações entre os modelos utilizados neste estudo. Considerando que o YOLOv8 apresentou maior precisão e o *DeepForest* obteve melhor desempenho em *recall*, um *ensemble* entre os dois pode potencialmente equilibrar essas características, produzindo *bounding boxes* mais consistentes e reduzindo tanto falsos positivos quanto falsos negativos. Essa estratégia poderia permitir a construção de inventários mais confiáveis e adaptáveis a diferentes cenários urbanos.

Outra possibilidade é integrar os resultados a sistemas de georreferenciamento para criar mapas úteis a gestores públicos, bem como realizar comparações em diferentes ambientes urbanos para avaliar a generalização dos modelos além do campus.

É importante destacar que este trabalho representou uma primeira iniciativa dentro da UTFPR - Londrina em aplicar modelos de visão computacional ao inventário arbóreo urbano. Por esse caráter inicial e exploratório, os resultados ainda não alcançaram níveis de desempenho próximos aos de pesquisas consolidadas, mas servem como ponto de partida para novas investigações na instituição.

Em síntese, conclui-se que o uso de *deep learning* em inventários arbóreos urbanos é uma alternativa promissora para reduzir custos e tempo em comparação aos métodos tradicionais. Apesar das limitações encontradas, este estudo contribuiu para abrir caminho no uso dessas tecnologias no contexto local, apontando tanto os desafios quanto as oportunidades que podem guiar trabalhos futuros na área.

## REFERÊNCIAS

- ANJNA, E.; KAUR, E. R. Review of image segmentation technique. **International Journal of Advanced Research in Computer Science**, v. 8, n. 4, p. 36-39, 2017. Disponível em: [https://www.researchgate.net/profile/Rajanddeep-Kaur/publication/340739499\\_Review\\_of\\_Image\\_Segmentation\\_Technique/links/641d421da1b72772e4228f93/Review-of-Image-Segmentation-Technique.pdf](https://www.researchgate.net/profile/Rajanddeep-Kaur/publication/340739499_Review_of_Image_Segmentation_Technique/links/641d421da1b72772e4228f93/Review-of-Image-Segmentation-Technique.pdf). Acesso em: 27 abril 2025.
- BENATTI, D. P. et al. Inventário arbóreo-urbano do município de Salto de Pirapora, SP. **Revista Árvore**, v. 36, p. 887-894, 2012. Disponível em: <https://www.scielo.br/j/rarv/a/B5MwFzYqnCghFMKKkp9f5vS/?format=html>. Acesso em: 07 maio 2025.
- BRASIL. **Lei nº 12.651, de 25 de maio de 2012**. Dispõe sobre a proteção da vegetação nativa; altera as Leis nos 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as Leis nos 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória nº 2.166-67, de 24 de agosto de 2001. *Diário Oficial da União: seção 1*, Brasília, DF, ano 149, n. 100, p. 1, 28 maio 2012. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2012/lei/l12651.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12651.htm). Acesso em: 15 maio 2025.
- BOSCH, M. DetecTree: Tree detection from aerial imagery in Python. **Journal of Open Source Software**, v. 5, n. 50, p. 2172, 2020. Disponível em: <https://joss.theoj.org/papers/10.21105/joss.02172.pdf>. Acesso em: 21 maio 2025.
- DE OLIVEIRA, L. M. et al. **Inventário da arborização urbana e descrição das características físicas das principais vias do setor central de Gurupi (TO)**. *Ambiência*, v. 14, n. 3, 2018. Disponível em: <https://revistas.unicentro.br/index.php/ambiencia/article/view/3502>. Acesso em: 06 maio 2025.
- DE PAIVA, A. V. et al. Inventário e diagnóstico da arborização urbana viária de Rio Branco, AC. **Revista da Sociedade Brasileira de Arborização Urbana**, v. 5, n. 1, 2010. Disponível em: <https://agris.fao.org/search/en/providers/122277/records/64745b4313d110e4e7ac41d7>. Acesso em: 07 maio 2025.
- DE SANTANA, D. K. S. et al. Análise da arborização e de áreas disponíveis para plantio arbóreo em praças públicas do município de Rorainópolis, RR, Brasil. **Advances in Forestry Science**, v. 11, n. 4, 2024. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/afor/article/view/17494/14446>. Acesso em: 12 maio 2025.
- ESPERON-RODRIGUEZ, M. et al. Urban tree inventories as a tool to assess tree growth and failure: The case for Australian cities. **Landscape and Urban Planning**, v. 233, p. 104705, 2023. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0169204623000245>. Acesso em: 29 novembro 2025.

GAION, R. **Perfil do fotógrafo no Pexels**. Disponível em: <https://www.pexels.com/pt-br/@gaion/>. Acesso em: 02 outubro 2025.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 2. ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2008. Acesso em: 26 abril 2025.

KORZNIKOV, K. A. et al. Using U-Net-like deep convolutional neural networks for precise tree recognition in very high resolution RGB (red, green, blue) satellite images. **Forests**, v. 12, n. 1, p. 66, 2021. Disponível em: <https://www.mdpi.com/1999-4907/12/1/66>. Acesso em: 31 maio 2025.

MARQUES FILHO, O.; NETO, H. V. **Processamento digital de imagens**. Brasport, 1999. Disponível em: <http://projetoaprendizagemgrupo4.pbworks.com/w/file/attach/96395952/Processamento%20Digital%20de%20Imagens.pdf>. Acesso em: 27 abril 2025.

MOIK, J. G. **Digital processing of remotely sensed images**. Scientific and Technical Information Branch, National Aeronautics and Space Administration, 1980. Disponível em: [https://books.google.com.br/books?hl=pt-BR&lr=&id=9UICAAAIAAJ&oi=fnd&pg=PA1&dq=MOIK,+J.+G.+Digital+processing+of+remotely+sensed+images.+Scientific+and+Technical+Information+Branch,+National+Aeronautics+and+Space+Administration,+1980.&ots=9oQi27QBvb&sig=IzDvqEe3pIVJ2IZmUJ\\_dKN7GO8s](https://books.google.com.br/books?hl=pt-BR&lr=&id=9UICAAAIAAJ&oi=fnd&pg=PA1&dq=MOIK,+J.+G.+Digital+processing+of+remotely+sensed+images.+Scientific+and+Technical+Information+Branch,+National+Aeronautics+and+Space+Administration,+1980.&ots=9oQi27QBvb&sig=IzDvqEe3pIVJ2IZmUJ_dKN7GO8s). Acesso em: 26 abril 2025.

MORAIS, S. M. F.; PEREIRA, A. A.; OLIVEIRA, U. F. de. Inventário Florestal Urbano do município de Botelhos, MG. **Ciência Florestal**, v. 34, n. 1, p. e71628, 2024. Disponível em: <https://www.scielo.br/j/cflo/a/VBCzngR5q9Wwpf9YrKdf7nM/>. Acesso em: 12 maio 2025.

ONISHI, M.; ISE, T.. Explainable identification and mapping of trees using UAV RGB image and deep learning. **Scientific reports**, v. 11, n. 1, p. 903, 2021. Disponível em: <https://www.nature.com/articles/s41598-020-79653-9>. Acesso em: 19 maio 2025.

PADILLA, R. et al. A comparative analysis of object detection metrics with a companion open-source toolkit. **Electronics**, v. 10, n. 3, p. 279, 2021. Disponível em: <https://www.mdpi.com/2079-9292/10/3/279>. Acesso em: 29 setembro 2025.

PAL, N. R.; PAL, S. K. **A review on image segmentation techniques**. Pattern recognition, v. 26, n. 9, p. 1277-1294, 1993. Disponível em: <https://www.sciencedirect.com/science/article/pii/003132039390135J>. Acesso em: 26 abril 2025.

Prefeitura de Porto Alegre. **Planejamento e cuidado com árvores**. Disponível em: <https://prefeitura.poa.br/smamus/planejamento-e-cuidado-com-arvores>. Acesso em: 31 maio 2025.

Prefeitura do Rio de Janeiro. **Plano de arborização urbana do Rio começa a sair do papel.** Disponível em: <https://prefeitura.rio/cidade/plano-de-arborizacao-urbana-do-rio-comeca-a-sair-do-papel/>. Acesso em: 31 maio 2025.

Prefeitura de São Paulo. **Plano Municipal de Arborização Urbana.** Disponível em: [https://capital.sp.gov.br/web/meio\\_ambiente/w/projetos\\_e\\_programas/284680](https://capital.sp.gov.br/web/meio_ambiente/w/projetos_e_programas/284680). Acesso em: 31 maio 2025.

Prefeitura Municipal de Londrina. **Plano Diretor de Arborização Urbana do Município de Londrina.** Disponível em: [https://ippul.londrina.pr.gov.br/images/legislacao-urbanistica/pd\\_arborizacao.pdf](https://ippul.londrina.pr.gov.br/images/legislacao-urbanistica/pd_arborizacao.pdf). Acesso em: 29 novembro 2025.

REZATOFIGHI, H. et al. Generalized intersection over union: A metric and a loss for bounding box regression. **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**, p. 658-666, 2019. Disponível em: <https://arxiv.org/abs/1902.09630>. Acesso em: 29 setembro 2025.

ROLOFF, A. (Ed.). **Urban tree management: for the sustainable development of green cities.** John Wiley & Sons, 2016. Disponível em: [https://books.google.com.br/books?hl=pt-BR&lr=&id=h4ibCgAAQBAJ&oi=fnd&pg=PR13&dq=ROLOFF,+A.+\(Ed.\).+Urban+tree+management:+for+the+sustainable+development+of+green+cities.+John+Wiley+%26+Sons,+2016.&ots=7yJM71cLQk&sig=Wz4qqgfPclBRoWWWIt-c0\\_M-4YM](https://books.google.com.br/books?hl=pt-BR&lr=&id=h4ibCgAAQBAJ&oi=fnd&pg=PR13&dq=ROLOFF,+A.+(Ed.).+Urban+tree+management:+for+the+sustainable+development+of+green+cities.+John+Wiley+%26+Sons,+2016.&ots=7yJM71cLQk&sig=Wz4qqgfPclBRoWWWIt-c0_M-4YM). Acesso em: 01 junho 2025.

SANTOS, G. P. dos; PAPKE, I. H.; ARAÚJO, G. B. F.; SILVEIRA, D. D. da; SOUZA, H. da C. de; ROSA, G. de O. da. A importância da arborização urbana e seus desafios para a implantação de qualidade. **Brazilian Journal of Development**, [S. l.], v. 11, n. 2, p. e77859, 2025. DOI: 10.34117/bjdv11n2-060. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/77859>. Acesso em: 06 maio 2025.

SILVA, A. **Detecção de árvores com YOLOv8.** GitHub, 2025. Disponível em: [https://github.com/alicecarosi/deteccao\\_de\\_arvores\\_com\\_YOLOv8](https://github.com/alicecarosi/deteccao_de_arvores_com_YOLOv8). Acesso em: 02 outubro 2025.

SONKA, M.; HLAVAC, V.; BOYLE, R. **Image processing, analysis and machine vision.** Springer, 2013. Disponível em: [https://books.google.com.br/books?hl=pt-BR&lr=&id=yA7yBwAAQBAJ&oi=fnd&pg=PR11&dq=SONKA,+M.%3B+HLAVAC,+V.%3B+BOYLE,+R..+Image+processing,+analysis+and+machine+vision.+Springer,+2013.&ots=l\\_VB-mYonq&sig=FbGAXdrWn1tNi9w\\_v\\_kgfR28BpQ](https://books.google.com.br/books?hl=pt-BR&lr=&id=yA7yBwAAQBAJ&oi=fnd&pg=PR11&dq=SONKA,+M.%3B+HLAVAC,+V.%3B+BOYLE,+R..+Image+processing,+analysis+and+machine+vision.+Springer,+2013.&ots=l_VB-mYonq&sig=FbGAXdrWn1tNi9w_v_kgfR28BpQ). Acesso em: 26 abril 2025.

TENG, M. et al. Assessing SAM for Tree Crown Instance Segmentation from Drone Imagery. arXiv preprint **arXiv:2503.20199**, 2025. Disponível em: <https://arxiv.org/abs/2503.20199>. Acesso em: 22 maio 2025.

TYRVÄINEN, L. et al. Benefits and uses of urban forests and trees. **Urban forests and trees: A reference book**, p. 81-114, 2005. Disponível em:

[https://link.springer.com/chapter/10.1007/3-540-27684-X\\_5](https://link.springer.com/chapter/10.1007/3-540-27684-X_5). Acesso em: 01 junho 2025.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the xxix conference on graphics, patterns and images**. sn, 2016. Disponível em:

<http://sibgrapi.sid.inpe.br/attachment.cgi/sid.inpe.br/sibgrapi/2016/09.12.15.44/doc/um-estudo-sobre.pdf>. Acesso em: 27 abril 2025.

VELASQUEZ-CAMACHO, L.; ETXEGARAI, M.; DE-MIGUEL, S. Implementing Deep Learning algorithms for urban tree detection and geolocation with high-resolution aerial, satellite, and ground-level images. **Computers, Environment and Urban Systems**, v. 105, p. 102025, 2023. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0198971523000881>. Acesso em: 15 maio 2025.

VENTURA, J. et al. Individual tree detection in large-scale urban environments using high-resolution multispectral imagery. **International Journal of Applied Earth Observation and Geoinformation**, v. 130, p. 103848, 2024. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S1569843224002024>. Acesso em: 20 maio 2025.

WEINSTEIN, B. G. et al. DeepForest: A Python package for RGB deep learning tree crown delineation. **Methods in Ecology and Evolution**, v. 11, n. 12, p. 1743-1751, 2020. Disponível em:

<https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13472>. Acesso em: 22 maio 2025.

ZHANG, C. et al. A comprehensive survey on segment anything model for vision and beyond. **arXiv preprint arXiv:2305.08196**, 2023. Disponível em:

<https://arxiv.org/abs/2305.08196>. Acesso em: 27 abril 2025.