

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

RÔMULO DIÓGENES MORO

**SEGMENTAÇÃO E CLASSIFICAÇÃO DE ESPÉCIES FLORESTAIS DO BIOMA
AMAZÔNICO POR MEIO DE IMAGENS DIGITAIS MACROSCÓPICAS**

MEDIANEIRA

2025

RÔMULO DIÓGENES MORO

**SEGMENTAÇÃO E CLASSIFICAÇÃO DE ESPÉCIES FLORESTAIS DO BIOMA
AMAZÔNICO POR MEIO DE IMAGENS DIGITAIS MACROSCÓPICAS**

**Segmentation and Classification of Forest Species from the Amazon Biome
Using Macroscopic Digital Images**

Proposta de Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Ciência da Computação do Curso de Bacharelado em Ciência da Computação da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Pedro Luiz de Paula Filho

MEDIANEIRA

2025



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

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.

RÔMULO DIÓGENES MORO

**SEGMENTAÇÃO E CLASSIFICAÇÃO DE ESPÉCIES FLORESTAIS DO BIOMA
AMAZÔNICO POR MEIO DE IMAGENS DIGITAIS MACROSCÓPICAS**

Proposta de Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Ciência da Computação do Curso de Bacharelado em Ciência da Computação da Universidade Tecnológica Federal do Paraná.

Data de aprovação: 17/fevereiro/2025

Pedro Luiz de Paula Filho
Prof. Dr
Universidade Tecnológica Federal do Paraná

Alessandra Bortoletto Garbelotti Hoffmann
Profª. Dra
Universidade Tecnológica Federal do Paraná

Kelyn Schenatto
Profª. Dra
Universidade Tecnológica Federal do Paraná

MEDIANEIRA

2025

Dedico este trabalho à minha família, pelo apoio incondicional e por sempre fornecer as condições necessárias; aos amigos e a todos que, de alguma forma, fizeram parte da minha caminhada, contribuindo para este momento tão especial.

AGRADECIMENTOS

Agradeço à minha família, pelo suporte e pelas condições necessárias para que eu pudesse chegar até aqui; aos amigos, pela companhia e motivação nos momentos desafiadores; e aos colegas de trabalho, pela contribuição para o meu crescimento pessoal e profissional.

Agradeço ao Prof. Dr. Pedro Luiz de Paula Filho, pela paciência, orientação e disposição durante o desenvolvimento deste trabalho, sendo essencial para a sua concretização.

Agradeço também aos demais professores e servidores da universidade, cujo esforço e dedicação são fundamentais, e ao Departamento de Ciência da Computação, pelo suporte e estrutura oferecidos ao longo do curso.

Por fim, minha gratidão a todos que fizeram parte dessa caminhada, deixando sua marca neste importante capítulo da minha vida.

RESUMO

O desmatamento e a degradação das florestas contribuem para o aquecimento global, tornando a sua preservação essencial para que os níveis do efeito estufa sejam reduzidos. Uma das formas para haver essa diminuição, é a fiscalização das florestas através de agentes fiscalizadores. No entanto, está é uma tarefa complexa devida à extensa biodiversidade das florestas, nesse quesito a floresta Amazônica se destaca. Para auxiliar na identificação de espécies florestais, diversas técnicas computacionais tem sido utilizadas, em sua maioria técnicas com imagens da espécies com outros elementos, o que pode reduzir a precisão da classificação. Com o objetivo de otimizar esta tarefa, este trabalho propõe o uso de técnicas de segmentação de imagens como YOLO11, para isolar a madeira antes da classificação usando o modelo Resnet-50. Para realizar o treinamento, a segmentação, foi desenvolvido um conjunto de imagens com 20 espécies, cada uma com 50 imagens. A metodologia deste trabalho envolveu a seleção do conjunto de dados, a escolha e treinamento do modelo de segmentação, a anotação das imagens, a classificação das espécies com as imagens originais, sem a segmentação, e as imagens segmentadas, geradas a partir das máscaras de segmentação do YOLO11, e a análise dos resultados utilizando métricas de avaliação como interseção sobre união, precisão, *recall*, *F1-score* e coeficiente dice. Os experimentos demonstraram que, apesar do YOLO11 ter um bom desempenho na segmentação, a remoção do fundo não resultou em uma melhora na acurácia final do modelo de classificação. Assim, o modelo com as imagens originais teve um desempenho superior ao modelo com imagens segmentadas, com 91% de acurácia contra 87%.

Palavras-chave: aprendizado profundo; redes neurais; floresta amazônica; reconhecimento de espécies.

ABSTRACT

Deforestation and forest degradation contribute to global warming, making their preservation essential for reducing greenhouse gas levels. One way to achieve this reduction is through forest monitoring by inspection agents. However, this is a complex task due to the extensive biodiversity of forests, with the Amazon rainforest standing out in this regard. To assist in the identification of forest species, various computational techniques have been used, mostly involving images of species along with other elements, which can reduce classification accuracy. To optimize this task, this study proposes the use of image segmentation techniques, such as YOLO11, to isolate the wood before classification using the ResNet-50 model. For training and segmentation, a dataset of 20 species was developed, each containing 50 images. The methodology of this study involved dataset selection, the choice and training of the segmentation model, image annotation, species classification using both the original images (without segmentation) and the segmented images (generated from YOLO11 segmentation masks), and the analysis of results using evaluation metrics such as Intersection over Union (IoU), Precision, Recall, F1-score, and Dice coefficient. The experiments demonstrated that, although YOLO11 performed well in segmentation, background removal did not result in an improvement in the final classification accuracy. Thus, the model using the original images achieved superior performance compared to the segmented images, with an accuracy of 91% versus 87%.

Keywords: deep learning; neural networks; amazon rainforest; species recognition.

LISTA DE FIGURAS

Figura 1 – Estrutura macroscópica da madeira	15
Figura 2 – Estrutura microscópica da madeira	16
Figura 3 – Exemplo de imagens de madeira macroscópica e microscópica	17
Figura 4 – Tipos de cortes	17
Figura 5 – Detecção de objetos	20
Figura 6 – Tipos de segmentação de imagem	21
Figura 7 – Comportamento YOLO	22
Figura 8 – Fluxo para predição YOLO	22
Figura 9 – Arquitetura simplificada YOLO	23
Figura 10 – Estrutura básica e componentes principais de um neurônio artificial	24
Figura 11 – Rede neural totalmente conectada com 2 camadas ocultas	24
Figura 12 – Bloco residual da <i>ResNet</i>	27
Figura 13 – Fluxograma da organização geral do projeto	33
Figura 14 – Exemplo de imagens do conjunto de dados	35
Figura 15 – Tipos de máscara	36
Figura 16 – Exemplos de imagens de entrada	36
Figura 17 – IoU	37
Figura 18 – Evolução das métricas durante o treinamento	39
Figura 19 – Matriz de confusão - teste	42
Figura 20 – Evolução das métricas durante o treinamento	42
Figura 21 – Imagens segmentadas	44
Figura 22 – Evolução das métricas durante o treinamento	44
Figura 23 – Crop	45
Figura 24 – Evolução das métricas durante o treinamento	46
Figura 25 – Matriz de confusão - Teste	47

LISTA DE TABELAS

Tabela 1 – Resultados artigos de classificação de espécies florestais	28
Tabela 2 – Resultados detecção de objetos aplicada a produtos em geral	29
Tabela 3 – Resultados modelos de segmentação de imagens em geral	30
Tabela 4 – Síntese de informações dos conjuntos de dados usados no trabalho	34
Tabela 5 – Relatório de classificação - Teste	41
Tabela 6 – Métricas de treinamento YOLO11	43
Tabela 7 – Métricas de avaliação YOLO11	43
Tabela 8 – Relatório de classificação - Teste	45
Tabela 9 – Teste - Relatório de classificação	46
Tabela 10 – Comparação entre resultados com e sem segmentação	48

LISTA DE ABREVIATURAS E SIGLAS

Siglas

CNN	Redes Neurais Convolucionais
Colab	<i>Google Colaboratory</i>
CPU	<i>Central Processing Unit</i>
FAO	Organização das Nações Unidas para Alimentação e Agricultura
FURB	Fundação Universidade Regional de Blumenau
GLCM	<i>Gray Level Co-occurrence Matrix</i>
GPU	<i>Graphics Processing Unit</i>
Grad-CAM	<i>Gradient-weighted Class Activation Mapping</i>
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
INPA	Instituto Nacional de Pesquisas da Amazônia
IPT	Instituto de Pesquisas Tecnológicas do Estado de São Paulo
LBP	<i>Local Binary Patterns</i>
LPQ	<i>Local Phase Quantisation</i>
PF	Polícia Federal
RAM	<i>Random Access Memory</i>
ResNet-50	<i>50-layer convolutional neural network</i>
RNA	Redes Neurais Artificiais
SAM	<i>Segment Anything Model</i>
SVM	<i>Support Vector Machine</i>
UFPR	Universidade Federal do Paraná
UTFPR	Universidade Tecnológica Federal do Paraná
ViT	<i>Vision Transformer</i>

YOLO

You Only Look Once

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Objetivos	13
1.2	Justificativa	13
1.3	Estrutura do trabalho	14
2	REFERENCIAL TEÓRICO	15
2.1	Anatomia da madeira	15
2.2	Metodologia tradicional de identificação de espécies florestais	18
2.3	Visão Computacional	18
2.3.1	Detecção de objetos	19
2.3.2	Segmentação de imagens	19
2.3.3	You Only Look Once	21
2.4	Aprendizado de Máquina	23
2.5	Redes neurais artificiais	23
2.6	Aprendizado profundo	25
2.6.1	Resnet-50	26
2.7	Trabalhos Relacionados	26
2.7.1	Classificação de Espécies Florestais	26
2.7.2	Detecção de Objetos em Geral	28
2.7.3	Modelos de Segmentação de Imagens em geral	29
3	MATERIAIS E MÉTODOS	31
3.1	Materiais	31
3.1.1	Ambiente de desenvolvimento	31
3.1.2	Bibliotecas e ferramentas	32
3.2	Métodos	33
3.2.1	Primeira etapa: Seleção conjunto de dados	33
3.2.2	Segunda etapa: Seleção do modelo de segmentação	34
3.2.3	Terceira etapa: Preparação das anotações	35
3.2.4	Quarta etapa: Treinamento dos modelos	35
3.2.5	Quinta etapa: Análise dos resultados	37
3.2.6	Sexta etapa: Aplicação do classificador	38

3.2.7	Sétima etapa: Avaliação da solução desenvolvida	38
4	RESULTADOS E DISCUSSÃO	39
4.1	Classificação de Imagens com <i>ResNet-50</i> Sem Segmentação	39
4.2	Segmentação das Imagens no Conjunto de Dados YOLO11	41
4.3	Classificação das Imagens Segmentadas com ResNet-50	43
4.4	Classificação das Imagens Segmentadas com ResNet-50 usando Crop	44
4.5	Avaliação	47
5	CONCLUSÃO	49
5.1	Trabalhos futuros	50
	REFERÊNCIAS	51

1 INTRODUÇÃO

O aquecimento global tem se tornado uma preocupação crescente, sendo uma das principais causas a emissão de gases do efeito estufa, causada pela degradação das florestas (DICKMANN; LIOTTI, 2024). Segundo o Ministério do Meio Ambiente (2024), as florestas desempenham funções importantes para o equilíbrio ecossistêmico da Terra, pois são o lar de inúmeras espécies de plantas e animais, auxiliam na qualidade dos recursos hídricos e contribuem capturando o dióxido de carbono (CO_2). A retenção desse composto químico ajuda a reduzir o efeito estufa, diminuindo o aquecimento global (ANDREZA *et al.*, 2024).

Dentre as florestas, a Amazônica é a maior floresta tropical do mundo, abrigando uma grande biodiversidade (CASTRO; ANDRADE, 2016; SA *et al.*, 2019). No entanto, vem sofrendo degradação nos últimos anos, entre os diversos motivos para sua diminuição estão a pecuária, que tem crescido desde os anos 90 (FEARNSIDE, 2020), e a exploração ilegal de recursos como a madeira (DINIZ, 2017).

Para uma bioeconomia sustentável, a madeira é um recurso renovável com propriedades e características fundamentais (SCHUBERT; PANZARASA; BURGERT, 2022). Por ser um produto valioso, o cultivo, a extração e o comércio da madeira são atividades com forte participação econômica em muitos países. Entretanto, com a alta demanda de madeira, a exploração ilegal das florestas tem aumentado (GEUS *et al.*, 2021).

O desmatamento é uma das principais atividades que levam à diminuição das florestas (BRAGAGNOLO; DA SILVA; GRZYBOWSKI, 2021). De acordo com o relatório de 2018 da Organização das Nações Unidas para Alimentação e Agricultura (FAO), o desmatamento é a segunda maior causa das mudanças climáticas.

Entre as medidas que podem ajudar a diminuir o desmatamento ilegal, está uma maior fiscalização. Porém, os agentes fiscalizadores têm dificuldade em identificar as espécies florestais. Especialistas de diversos órgãos de fiscalização e controle atuam na identificação de espécies. Entretanto, essa identificação não é uma tarefa trivial, visto que a examinação pode ser muito relativa, devido à grande quantidade de espécies encontradas no Brasil (PAULA FILHO, 2012).

Para auxiliar nesta tarefa, diversas soluções de visão computacional foram propostas (HWANG; SUGIYAMA, 2021), utilizando Redes Neurais Artificiais (RNA), *Support Vector Machine* (SVM) (PAULA FILHO, 2012) e Redes Neurais Convolucionais (CNN) (ACHATZ *et al.*, 2024; OLIVEIRA, 2018; IBRAHIM *et al.*, 2017). Porém, essas soluções utilizam a imagem total da madeira, junto a essas imagens, outros objetos podem ser encontrados, como rochas, solo, entre outros. Esses objetos podem prejudicar o desempenho. Diante disso, surge a necessidade de soluções que busquem abordagens focadas somente na madeira. Assim, uma ferramenta que segmente a madeira da imagem, permitindo que o classificador possa obter melhores resultados.

Este estudo propõe uma abordagem que utiliza um modelo de segmentação, e posteriormente, as imagens resultantes do modelo segmentado serão testadas em um modelo de classificação.

1.1 Objetivos

O objetivo geral deste trabalho consiste em identificar espécies florestais por meio de técnicas de aprendizado profundo utilizando técnicas de segmentação de imagens. O objetivo geral pode ser dividido nos seguintes objetivos específicos:

- Construir um conjunto de imagens de espécies do bioma amazônico;
- Treinar e aplicar modelo de segmentação de madeiras;
- Treinar e aplicar um classificador na área segmentada com diferentes abordagens;
- Avaliar o desempenho da solução desenvolvida.

1.2 Justificativa

A madeira contribuiu para o desenvolvimento humano ao longo da história (ACHATZ *et al.*, 2024). O reconhecimento de espécies através da madeira tem impacto em diversas áreas, como comércio, construção civil, indústria, infraestrutura, entre outros. A correta classificação das espécies é importante para o manejo florestal adequado (OLIVEIRA, 2018).

A dificuldade de identificar adequadamente as espécies florestais afetadas pelo desmatamento ilegal cria obstáculos significativos para os esforços de conservação (SOUZA *et al.*, 2020a). Sem a capacidade de identificar com precisão as espécies, dificulta-se implementar medidas de proteção direcionadas e eficazes. Além disso, a precisão na identificação das espécies é fundamental para o manejo sustentável das florestas e a preservação da biodiversidade (SHARMA; GARG; SHARMA, 2024).

Conforme Paula Filho (2012), no Brasil, existe uma pequena quantidade de laboratórios especializados no reconhecimento de espécies florestais, como o Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT), Fundação Universidade Regional de Blumenau (FURB), Instituto Nacional de Pesquisas da Amazônia (INPA), Universidade Federal do Paraná (UFPR). Essa limitação afeta diretamente a fiscalização, que apesar do trabalho desenvolvido por órgãos como Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) e a Polícia Federal (PF), ainda faltam ferramentas eficazes para uso em campo. Assim, uma ferramenta computacional ajudaria os fiscais a tomar medidas adequadas.

Devido a grande variedade de espécies da flora brasileira, composta por cerca de 22.000 espécies (Jardim Botânico do Rio de Janeiro, 2024), características como sementes, fruto e folha ajudam na identificação. No entanto, ao serem serradas, essas características são perdidas,

dificultando o processo de identificação. Nesse contexto, a análise anatômica da madeira utilizando sistemas computacionais torna-se útil (OLIVEIRA, 2018).

Abordagens para identificação de espécies florestais que utilizam imagens contendo objetos além da madeira, como folhagens, solo, rochas e outros elementos presentes podem apresentar limitações. Assim sendo, há uma lacuna nessas abordagens, com uma escassez de estudos que se concentrem apenas na madeira.

Portanto, a pesquisa deste trabalho busca demonstrar que técnicas de segmentação de imagens podem ou não aumentar a taxa de acertos. Ao separar a madeira de outros elementos presentes nas imagens, essas técnicas tem o potencial de auxiliar em uma identificação com uma maior precisão das espécies florestais. Assim, os interessados em identificar as espécies terão maiores probabilidades de sucesso na identificação, contribuindo para uma melhor gestão dos produtos florestais.

1.3 Estrutura do trabalho

Este trabalho será organizado da seguinte forma. Capítulo 2 apresenta o tema reconhecimento de espécies florestais. No decorrer do capítulo, serão abordados modelos de detecção, segmentação e classificação de imagens, bem como trabalhos relacionados. O capítulo 3 descreve os materiais e os métodos utilizados para obtenção dos dados experimentais. No capítulo 4 são apresentados e discutidos os resultados obtidos através dos materiais e métodos propostos, e por fim o capítulo 5 com as conclusões finais sobre o estudo e sugestões de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os conceitos teóricos para a realização e compreensão deste trabalho. O capítulo inicia com uma descrição da madeira e suas características. Em seguida, os conceitos de visão computacional, detecção e segmentação de objetos são detalhados. Depois, são abordados os conceitos de aprendizado de máquina, aprendizado profundo e redes neurais. Por fim, trabalhos relacionados à área são apresentados.

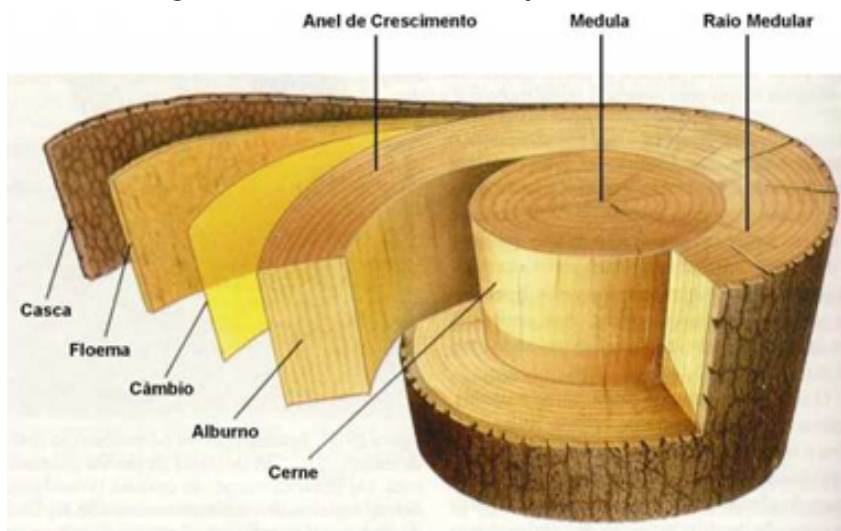
2.1 Anatomia da madeira

A anatomia da madeira é uma área de estudo que busca compreender as estruturas que constituem o lenho, partindo da sua organização, complexidade e peculiaridades (BOTOSSO, 2011). A madeira é um produto florestal natural composta por diferentes tipos de células e tecidos, como vasos, fibras, parênquima, entre outros. Cada uma das propriedades desempenha funções importantes para a planta, desde a absorção de nutrientes até sustentação e crescimento (CARVALHO *et al.*, 2009).

As características macroscópicas da madeira são aquelas que podem ser observadas a olho nu ou com auxílio de uma lupa (BOTOSSO, 2011). Essas características podem ser físicas ou morfológicas, e podem variar conforme a espécie, como também podem variar para um mesmo tipo de árvore. Algumas características físicas são cor e brilho, as características morfológicas são compostas por cerne, albúrnio, raios medulares, anéis de crescimento, entre outros (OLIVEIRA, 2018). A Figura 1 ilustra as características macroscópicas da madeira.

As características microscópicas permitem uma compreensão detalhada da estrutura da madeira, essa estrutura é composta por milhares de células, que estão divididas em grupos

Figura 1 – Estrutura macroscópica da madeira

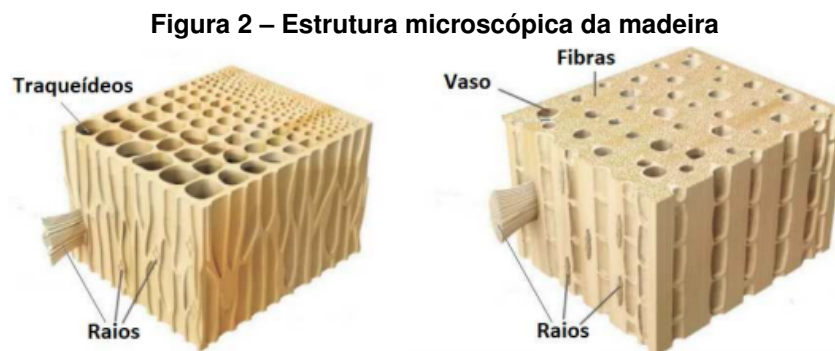


Fonte: Oliveira (2018).

como: traqueóides, vasos, fibras e raios medulares (OLIVEIRA, 2018). Conforme Santos (2015) os grupos podem ser descritos da seguinte maneira:

- Traqueídeos: são células alongadas, suas dimensões aproximadas são 5mm de comprimento e 0,6mm de diâmetro. Dentre suas funções estão transporte de seiva entre as células, e o armazenamento de substâncias polimerizadas;
- Vasos: são células longas, suas dimensões aproximadas são 1mm de comprimento e 0,3mm de diâmetro. Dentre suas funções estão transporte de seiva entre as células, e o armazenamento de substâncias polimerizadas;
- Fibras: são células longas, possuem comprimento de até 1,5mm. Algumas de suas características são parede grossa e uma camada de seção arredondada. Sua função é fornecer rigidez e resistência mecânica ao tronco;
- Raios Medulares: são células longas e achatadas. Sua função é transportar seiva elaborada.

A Figura 2 demonstra a estrutura desses grupos:



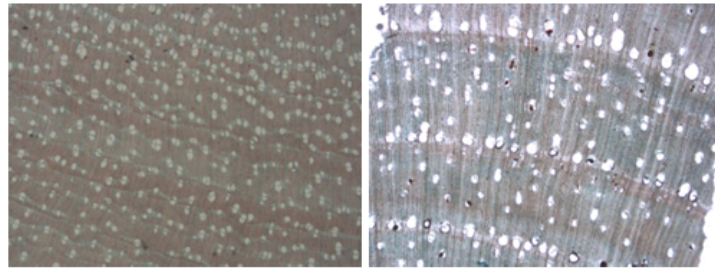
Fonte: Santos (2015).

A Figura 3a consiste em duas imagens com características microscópicas, já a Figura 3b demonstra características macroscópicas.

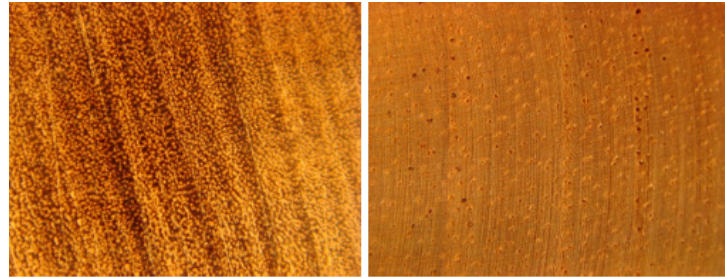
Segundo Paula e Alves (1997), por ser um material anisotrópico, a madeira não apresenta simetria de propriedades em relação aos sentidos dos cortes. Devido a isso, diferentes características da estrutura celular são revelados, portanto, se faz necessário o estudo da anatomia nos diferentes planos de corte. Assim sendo, três planos de referência são utilizados conforme Botosso (2011) :

- Corte transversal (X): O corte transversal, ou corte em rodela, é realizado perpendicularmente ao eixo longitudinal da árvore;
- Corte longitudinal radial (R): O corte radial é feito ao longo do raio da árvore, perpendicular aos anéis de crescimento;

Figura 3 – Exemplo de imagens de madeira macroscópica e microscópica



(a) Descrição microscópica



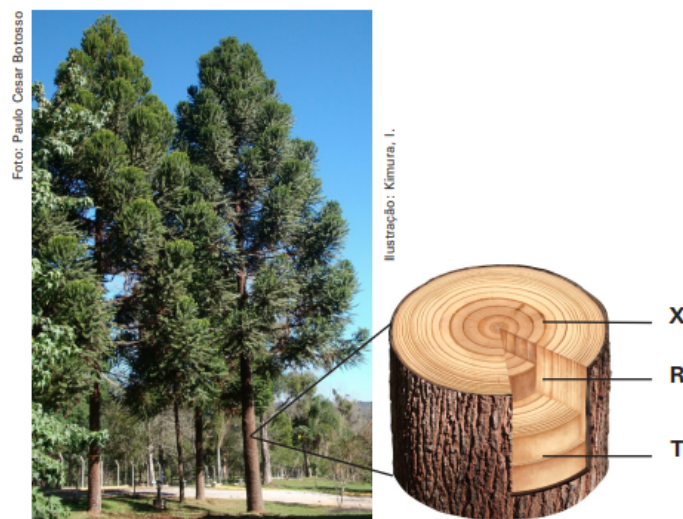
(b) Descrição macroscópica

Fonte: Paula Filho (2012).

- Corte longitudinal tangencial (T): O corte tangencial é realizado paralelo ao eixo longitudinal da árvore, ao longo dos anéis de crescimento.

A Figura 4 demonstra os três planos de cortes para madeira, essa imagem proporciona uma compreensão mais precisa dos tipos de corte.

Figura 4 – Tipos de cortes



Fonte: Botosso (2011).

2.2 Metodologia tradicional de identificação de espécies florestais

De acordo com Oliveira (2018), a observação realizada por especialistas visa os aspectos anatômicos macroscópicos e microscópicos para identificação da madeira. Assim, os aspectos ligados as propriedades podem ser comparados com outras ferramentas, como manuais e bancos de dados.

Segundo Paula Filho (2012) a identificação das espécies quando ainda estão na floresta é fácil, devido às flores, folhas, frutos, entre outros atributos ainda estarem presentes, o que auxilia na comparação com as descrições presentes na literatura. Entretanto, quando esses atributos são perdidos devido à serragem da madeira ou à madeira não estar mais na floresta, torna-se complexo. A identificação correta realizada por humanos pode ser demorada e tediosa, visto que a verificação unitária de cada madeira, além da grande quantidade de espécies existentes, torna o processo difícil. Além disso, para que um humano possa se tornar um especialista, demanda um treinamento e tende a ser um processo demorado, o que torna escasso o número de especialistas.

Paula Filho (2012) também destaca um exemplo no qual uma ferramenta computacional seria útil para auxiliar nesse processo é na exportação, pois um grande número de navios, carregados com madeira que podem ter características semelhantes, porém com propriedades físicas distintas, permitiria aos fiscais uma identificação mais rápida e precisa, evitando danos aos possíveis compradores da madeira.

Na análise macroscópica, as características observadas são anatômicas e organolépticas. Dentre as características anatômicas, os três planos transversal, radial e tangencial podem ser observados. Em cada um desses planos camadas de crescimento, vasos e raios podem ser identificados. Para as organolépticas, são observadas características de odor, cor, brilho, gosto, entre outras. Na análise microscópica, as características observadas são as estruturas celulares e os tecidos, entre elas, a dimensão e composição das células, presença de cristais, entre outros (KOCH; HAAG, 2015).

2.3 Visão Computacional

A visão computacional é um campo de pesquisa que se concentra em desenvolver modelos capazes de compreender informações visuais. Seu objetivo é fazer com que as máquinas possam interpretar e converter os dados da imagem por meio de modelos de processamento e análise de imagens, em informações visuais úteis (ZHANG *et al.*, 2024).

Ensinar uma máquina a interpretar uma imagem não é uma tarefa fácil, mesmo que a imagem tenha uma boa qualidade, por si só a máquina não é capaz de compreender o seu conteúdo. Para que a máquina consiga detectar e compreender objetos presentes em uma imagem, ela precisa do auxílio de técnicas de visão computacional aliadas a técnicas de aprendizado de máquina ou aprendizado profundo (NAFEA *et al.*, 2024).

Conforme Poyser e Breckon (2024), o interesse em visão computacional tem crescido significativamente, pois pode ser aplicada em diversas áreas, como diagnósticos médicos (LI; HU; YANG, 2023; MAGHANAKI *et al.*, 2024), veículos autônomos (TAN *et al.*, 2024) e identificação de espécies florestais (HWANG; SUGIYAMA, 2021).

A compreensão por parte da máquina depende de como ela entende as imagens. E para isso, existem uma variedade de técnicas, que vão desde abordagens tradicionais como extração de características, até modelos mais avançados de aprendizado profundo. A classificação de uma imagem é um processo de atribuir rótulos ou categorizar com base no conteúdo presente na imagem (SINGH; SINGH, 2020). Para realizar a classificação de imagens, dois métodos são utilizados:

O método supervisionado utiliza um conjunto de dados de entrada e um conjunto de dados de saída já rotulados para o treinamento, a rotulagem é feita de forma manual pelo analista (SCHMARJE *et al.*, 2021).

O método não supervisionado recebe um conjunto de dados de entrada, porém, não recebe nenhum dado de saída rotulado. O modelo é encarregado de identificar e reconhecer padrões e agrupar os dados (SCHMARJE *et al.*, 2021).

2.3.1 Detecção de objetos

A detecção de objetos é uma tarefa de localização e classificação, seu objetivo é encontrar objetos em uma determinada imagem e associar cada objeto a um rótulo de classe. A demarcação do objeto é realizada fazendo uso de uma caixa delimitadora (do inglês, *bounding box*) (AMJOURD; AMROUCH, 2023).

Os avanços recentes na detecção de objetos estão relacionados com o incremento das redes neurais convolucionais e técnicas de aprendizado profundo. Essa evolução permite uma detecção mais eficiente, mesmo em conjuntos de dados com imagens contendo múltiplos objetos e oclusões (AMJOURD; AMROUCH, 2023).

Modelos de detecção de objetos podem ser aplicados em inúmeras áreas, como detecção facial (LIU; HASAN; LIAO, 2023), análise de comportamento (SUN *et al.*, 2023), robótica (LENZ; LEE; SAXENA, 2015), entre outros. A Figura 5 demonstra exemplos de detecção de objetos.

2.3.2 Segmentação de imagens

De maneira geral, segmentação pode ser definida como a busca por elementos semelhantes em um ambiente heterogêneo. Para visão computacional, a segmentação de imagens é uma tarefa importante, no qual se divide imagens em múltiplos segmentos e objetos (LEI *et al.*,

Figura 5 – Detecção de objetos



Fonte: Amjoud e Amrouch (2023).

2024). Essa divisão tem o objetivo de atribuir cada *pixel* de uma imagem a um rótulo, diferente de uma tarefa de classificação, que rotula a imagem inteira (VASILEV *et al.*, 2019).

Diversos problemas e aplicações podem demandar o uso de segmentação. Para isso, três técnicas diferentes podem ser utilizadas, são elas:

- **Segmentação Semântica:** busca classificar cada *pixel* presente em uma imagem em um rótulo específico. *Pixels* de um mesmo objeto ou que apresentem características próximas, recebem a mesma máscara. Pode ser aplicado em áreas que tenham uma alta demanda por velocidade, como realidade aumentada, vigilância e direção autônoma (YU *et al.*, 2018).

A Figura 6b mostra a segmentação resultante, observa-se que não houve uma identificação individual dos elementos, os carros, por exemplo, pertencem a uma mesma máscara, assim como as pessoas;

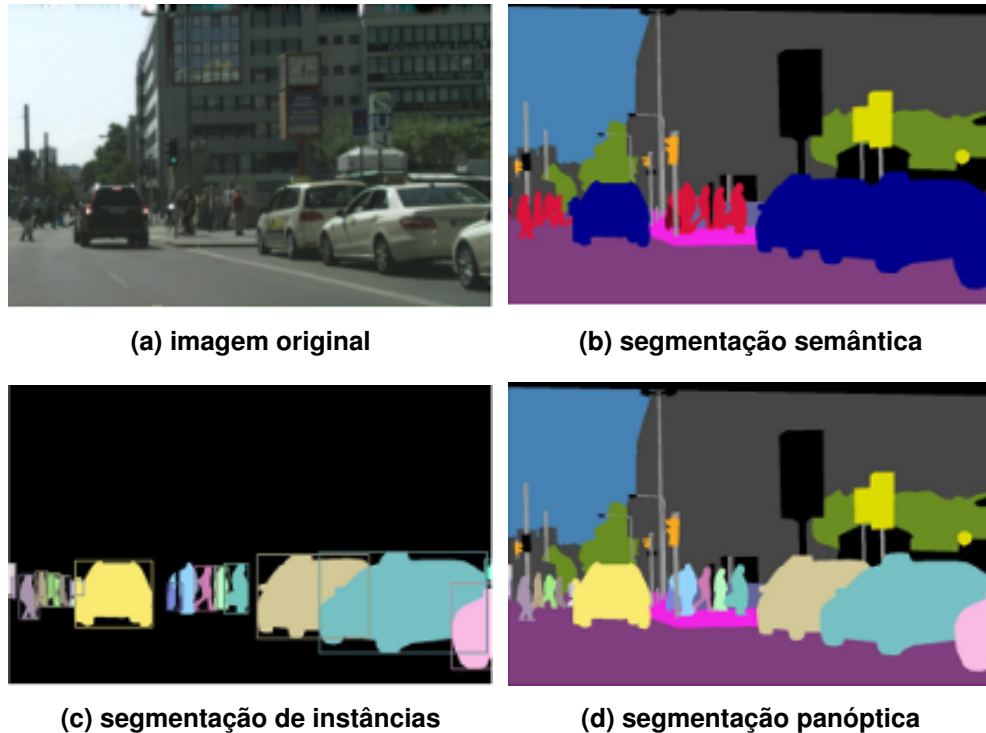
- **Segmentação Instâncias:** A classificação é feita *pixel a pixel*, porém este modelo diferencia diferentes instâncias de objetos que pertencem a mesma classe. É útil em aplicações onde a contagem de objetos em uma imagem é necessária (HAFIZ; BHAT, 2020).

A Figura 6c mostra a segmentação resultante, nela pode-se notar que identificação dos objetos foi realizada de forma individual. O fundo da imagem foi ignorado;

- **Segmentação Panóptica:** esse modelo combina características de segmentação de instância e semântica. A identificação dos objetos é realizada individualmente. Neste modelo, todos os *pixels* são identificados e rotulados. É útil em aplicações de realidade aumentada, análise de imagens médicas e robótica (KIRILLOV *et al.*, 2019).

A Figura 6d mostra a segmentação resultante, todos os objetos foram identificados, cada pixel foi associado a uma classe, e foi possível distinguir entre diferentes instâncias dessa classe.

Figura 6 – Tipos de segmentação de imagem



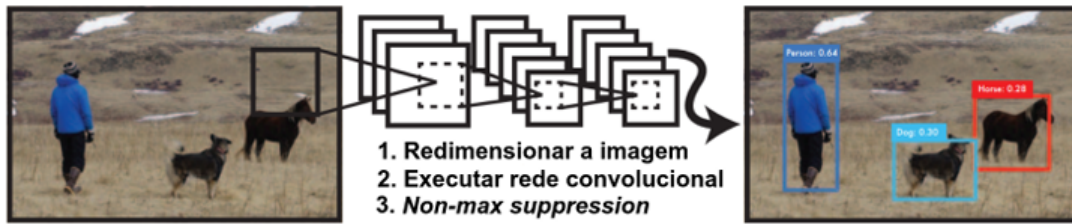
Fonte: Kirillov *et al.* (2019).

2.3.3 You Only Look Once

Conforme Redmon *et al.* (2016), o *You Only Look Once* (YOLO) em sua versão original é um modelo para detecção de objetos que olha apenas uma vez a imagem para enviar para a rede neural extrair as características. Devido a essa única passagem, o modelo apresenta uma velocidade maior do que outros modelos sem perder acurácia. Conforme a Figura 7, o modelo apresenta somente uma rede neural convolucional que prediz múltiplas caixas delimitadoras de forma simultânea, com as caixas, o modelo também mostra as probabilidades das caixas associadas a cada classe.

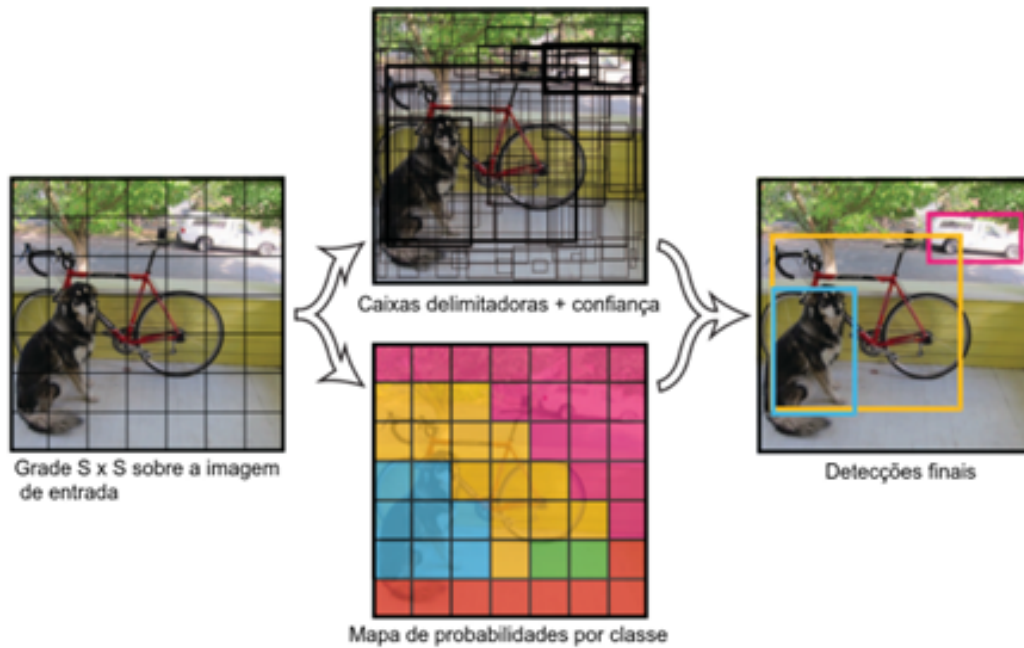
O YOLO trata a detecção como uma tarefa de regressão, e para isso divide a imagem em um grade $S \times S$. Para cada célula é predita B caixas delimitadoras com a métrica de confiança respectiva, para indicar o grau de confiança que o modelo tem de um objeto presente na caixa. A Figura 8 ilustra esse conceito.

Figura 7 – Comportamento YOLO



Fonte: Redmon *et al.* (2016).

Figura 8 – Fluxo para predição YOLO



Fonte: Redmon *et al.* (2016).

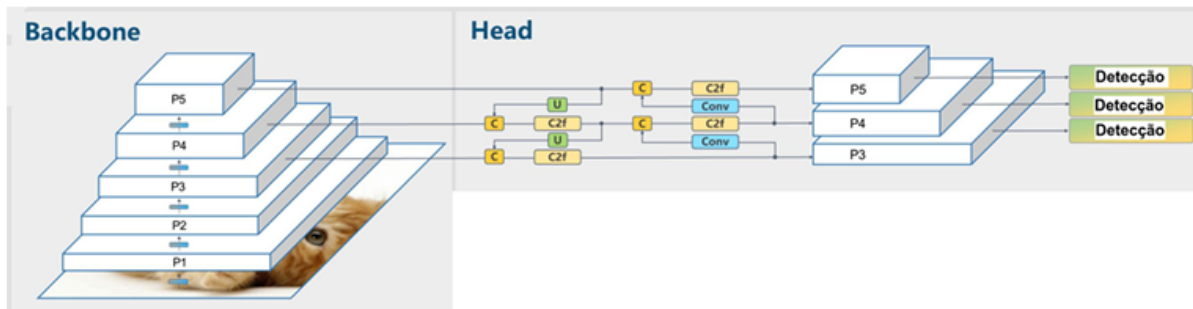
O YOLO a partir da versão 8, é um modelo que pode ser utilizado para tarefas de detecção de objetos, segmentação de instâncias e classificação de imagens. O modelo é composto por duas funções principais, são elas: a *backbone* e a *head*.

A função *backbone* extrai características da imagem de entrada, fornecendo uma representação das informações contidas na imagem. A função *head* recebe as características geradas pela função *backbone* e as processa para gerar uma saída para o modelo. A figura 9 demonstra a arquitetura do YOLO (TERVEN; ESPARZA; ROMERO-GONZALEZ, 2023).

O modelo YOLO está disponível em cinco variantes, cada uma projetada para casos específicos, são elas: YOLOv^{*1}n (*nano*), YOLOv^{*s} (*small*), YOLOv^{*m} (*medium*), YOLOv^{*l} (*large*), YOLOv^{*x} (*extra large*) (JOCHER; CHAURASIA; QIU, 2024).

¹ O asterisco representa variações das diferentes versões do YOLOv8, YOLOv9, YOLOv10 e YOLOv11.

Figura 9 – Arquitetura simplificada YOLO



Fonte: King (2023).

2.4 Aprendizado de Máquina

O aprendizado de máquina (do inglês, *machine learning*) é uma área da computação com o objetivo de tornar os computadores capazes de aprender por meio da experiência para resolver problemas complexos. Algoritmos de aprendizado de máquina podem ser aplicados em diversos problemas, como classificação, regressão, tradução, detecção de anomalias, entre outros (MITCHELL, 1997).

Dentre os muitos métodos de aprendizado, os principais são aprendizado supervisionado, não supervisionado e aprendizado por reforço. O supervisionado utiliza dados com seus rótulos, para que o algoritmo encontre as associações entre entrada e a saída esperada. Para o não supervisionado, não são informados os resultados, o algoritmo deve encontrar os padrões presentes nos dados e descrevê-los. O aprendizado por reforço, busca encontrar o resultado baseado em tentativa e erro, buscando aprender com os dados (VASILEV *et al.*, 2019).

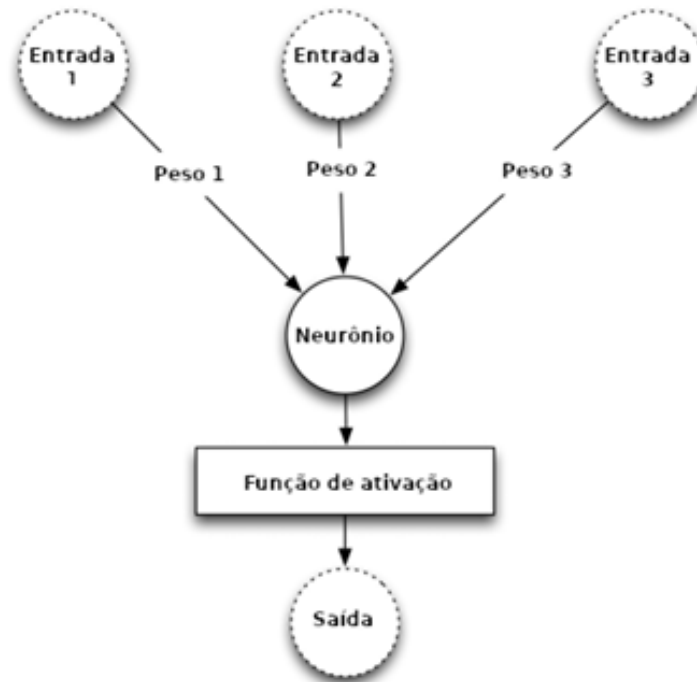
2.5 Redes neurais artificiais

Um neurônio artificial é inspirado no neurônio biológico humano e também é conhecido como nó ou unidade de processamento. Cada nó recebe entradas com pesos associados, após uma computação matemática, o neurônio aplica uma função de ativação. Essa função determina se o nó poderá enviar um sinal para a próxima camada (HEATON, 2015). A Figura 10 apresenta a estrutura de um neurônio artificial.

Inspirado no funcionamento do cérebro humano, uma rede neural é um modelo de computação para dados complexos. Uma rede neural é composta por camadas, que são denominadas de neurônios artificiais, a estrutura básica para um rede inclui, camada de entrada, camada oculta e camada de saída. As camadas estão interconectadas, havendo assim troca de informações entre elas (VASILEV *et al.*, 2019). A Figura 11 apresenta a estrutura de uma rede neural.

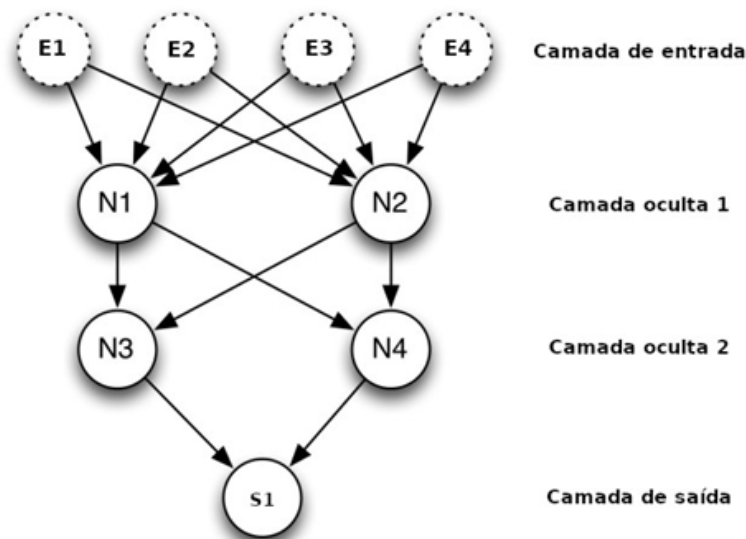
As redes neurais podem ter inúmeras camadas ocultas, com milhares de neurônios artificiais conectados. A camada de entrada é responsável por receber os dados, enquanto as

Figura 10 – Estrutura básica e componentes principais de um neurônio artificial



Fonte: Scheeren (2022).

Figura 11 – Rede neural totalmente conectada com 2 camadas ocultas



Fonte: Scheeren (2022).

camadas intermediárias, denominadas de camadas ocultas, realizam a maior parte do processamento. A camada de saída apresenta o resultado final do processamento (VASILEV *et al.*, 2019).

Uma função de ativação é uma função matemática responsável por determinar se um neurônio será ativado ou não a cada nó da rede neural, se ativado a informação é repassada para a próxima camada. Esse cálculo é realizado com base na entrada que é obtida da camada anterior. As funções de ativação exercem um papel importante em redes neurais, pois

elas permitem que a rede aprenda relações de dados complexos por meio de não-linearidades nos dados (RUSSELL; NORVIG, 2020). Conforme Ding, Qian e Zhou (2018), as funções mais utilizadas são:

- *Sigmoide*: é comumente utilizada em problemas de classificação binária, sua saída é um valor entre zero e um. A função *sigmoide* não ativa se o valor de entrada for negativo, retornando valores próximos de zero, caso o valor seja zero ou positivo, retorna valores próximos de um. A Equação 1 demonstra o cálculo:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

- *Tanh*: a função é usada para suavizar a saída das camadas por meio de normalização, não é ativada se o valor de entrada for negativo, retornando -1, porém, se o valor for zero ou positivo, mantém o valor de entrada. A Equação 2 demonstra o cálculo:

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (2)$$

- *Rectified Linear Unit (ReLU)* é uma das funções mais utilizadas para modelos de redes neurais profundas. Ela não é ativada se o valor de entrada for negativo, retornando zero. Caso o valor de entrada seja positivo ou zero, o valor de entrada é mantido. A Equação 3 demonstra o cálculo:

$$\text{ReLU}(x) = \max(0, x) \quad (3)$$

- *Softplus*: não é ativada se o valor for negativo, retornando 0, caso o valor de entrada seja positivo, retorna valores no intervalo $(0, \infty)$, conforme a Equação 4:

$$\text{softplus}(x) = \log(1 + e^x) \quad (4)$$

2.6 Aprendizado profundo

Aprendizado profundo é uma subárea de aprendizado de máquina, é composta por redes neurais profundas, esse tipo de rede pode conter pelo menos duas camadas ocultas (HEATON, 2015). Por utilizar várias camadas de processamento em conjunto, é capaz de extrair e compreender dados com altos níveis de complexidade. Uma das formas de aplicar aprendizado profundo é utilizando redes neurais convolucionais (VASILEV *et al.*, 2019).

Rede neural convolucional é uma rede neural profunda amplamente utilizada em processamento de imagens (GOODFELLOW; BENGIO; COURVILLE, 2016). Está sendo utilizada em diversas aplicações, como detecção de objetos, reconhecimento facial, entre outras. Utiliza

camadas de convolução para extrair e mapear padrões, esses mapas representam aspectos da imagem, como texturas, bordas, linhas, entre outros elementos (HEATON, 2015).

A arquitetura de de uma rede neural convolucional é composta por camadas convolucionais, camadas de *pooling* e camadas conectadas. As camadas convolucionais extraem as características das imagens, aplicando filtros e operações de convolução. A medida que as informações avançam pela rede, mais características são capturadas (HEATON, 2015).

A camada de *pooling* é responsável por substituir um conjunto de valores por um único valor. Essa substituição permite reduzir o carga computacional, a memória e a quantidade de parâmetros eu serão aprendidos (RUSSELL; NORVIG, 2020). As camadas conectadas são camadas na qual um neurônio está conectado a todos os neurônios da camada anterior, a partir dessa camada, são gerados a classificação final (HEATON, 2015).

2.6.1 Resnet-50

A *50-layer convolutional neural network* (ResNet-50) é uma rede convolucional utilizada para visão computacional, é um modelo que venceu o desafio *ImageNet* em 2015. Essa rede permite treinar com profundidade e eficiência em problemas de classificação de imagens (HE *et al.*, 2015).

Sua abordagem permitiu superar problemas de gradiente evanescente em redes neurais, ao se tornarem mais profundas, a precisão da rede satura e se degrada rapidamente. Por meio de conexões residuais, a propagação entre conexões evita que falhe por esvanecimento (SOUZA *et al.*, 2020b).

A arquitetura da ResNet-50 permite que a função de ativação *ReLU*, descrita na seção 2.5, continue processando valores resultantes da camada anterior, bem como continue recebendo os valores de entrada das funções (SOUZA *et al.*, 2020b). Conforme pode ser observado na Figura 12.

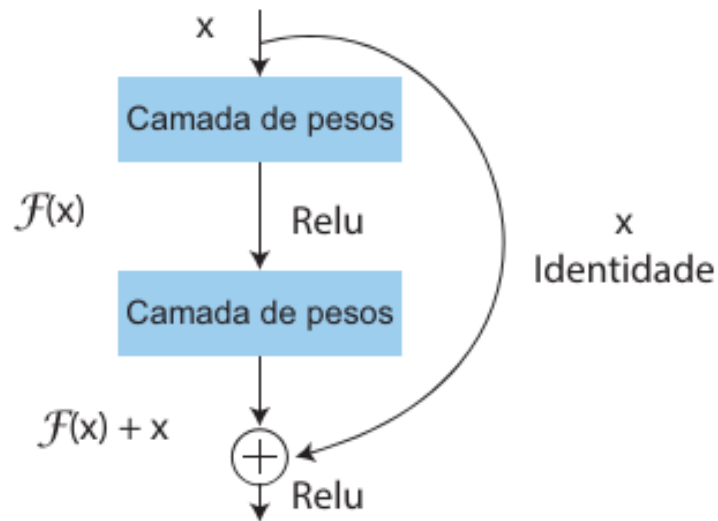
2.7 Trabalhos Relacionados

Nesta seção são apresentados alguns trabalhos relacionados a reconhecimento de espécies florestais, abordando temas como visão computacional, modelos de classificação e técnicas de segmentação de imagens. Os seguintes trabalhos mostram a aplicação dos temas em problemas do mundo real.

2.7.1 Classificação de Espécies Florestais

Paula Filho (2012) usa duas bases de imagens macroscópicas, totalizando 2942 imagens de 41 espécies, sua abordagem divide as imagens em subimagens, isolando problemas

Figura 12 – Bloco residual da *ResNet*



Fonte: Souza *et al.* (2020b).

locais e melhorando a classificação. A extração de atributos, como cor e textura, é feita via análise de cor, *Gray Level Co-occurrence Matrix* (GLCM), histograma de borda, Fractais, *Local Binary Patterns* (LBP), *Local Phase Quantisation* (LPQ) e Gabor. Inicialmente com imagens de uma câmera comercial normal, utilizando o classificador *Support Vector Machine* (SVM) a taxa de reconhecimento atingiu 95,82%. Já para imagens tiradas de uma câmera de laboratório, a taxa aumentou para 99,49%, validando a eficácia do método.

Hwang e Sugiyama (2021) realizaram uma revisão quanto ao reconhecimento de espécies florestais usando apenas rotinas baseadas em visão computacional. Para realizar o estudo a base de dados foi a *Xylarium*¹, contém mais de 3.300 imagens de mais de 130 tipos diferentes de espécies. Para a análise foram usadas apenas imagens com cortes transversais sendo que técnicas como *machine learning* e *deep learning* se mostraram eficazes para identificação de espécies, mostrando que tais tecnologias contribuem para os avanços nos estudos de madeira.

Explorando um conjunto de dados de 312 imagens com 14 espécies de árvores diferentes, utilizando imagens dos núcleos da madeira, Fabijańska, Danek e Barniak (2021) propuseram um estudo utilizando redes neurais convolucionais para identificar cada uma das espécies. Para isso foram divididos em dois grupos de classificação, o primeiro com manchas de madeira e o segundo com núcleos de madeira, os resultados respectivos foram 93% e 98,7%, sendo considerados bons resultados.

Com grande importância comercial e científica, diferentes espécies florestais tem se tornado objeto de estudos. O reconhecimento convencional requer grande experiência e treinamento em anatomia da madeira, e para aprimorar a eficiência desse processo Kirbaş e Çifci (2022) realizaram um estudo utilizando o conjunto de dados *Wood-Auth*², que contém 8544

¹ <https://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/243328>

² <https://zenodo.org/record/1216746>

imagens de 12 diferentes espécies, esse conjunto incluía exemplos de imagens com três tipos de cortes: transversal, radial e tangencial, foram testadas várias arquiteturas de aprendizado, como ResNet-50, Inception v3, Xception e Vgg19, como melhor resultado, a técnica de aprendizagem por transferência, com a arquitetura Xception alcançou 95,88% de precisão.

Vieira *et al.* (2022) desenvolveram um sistema de reconhecimento de espécies geralmente comercializadas, utilizando dez espécies diferentes, três tratamentos de polimentos e vinte imagens de cada espécie, as imagens foram classificadas por redes neurais artificiais, durante o treinamento o modelo alcançou uma taxa de 94%, já em fase de pós-treinamento a taxa foi de 65%, demonstrando um potencial para identificar as espécies estudadas.

A Tabela 1 apresenta um resumo das técnicas utilizadas para classificação de espécies florestais.

Tabela 1 – Resultados artigos de classificação de espécies florestais

Trabalho	Classificador	Img/Esp.	Melhor acerto
Paula Filho (2012)	SVM	2942/41	99,49%
Hwang e Sugiyama (2021)	LeNet	14102/238	99,30%
Fabijańska, Danek e Barniak (2021)	CNN	312/14	98,70%
Kırbaş e Çifci (2022)	Xception	8544/12	95,88%
Vieira <i>et al.</i> (2022)	MLP	600/75	65,00%

Fonte: Autoria própria (2025).

2.7.2 Detecção de Objetos em Geral

Visando detecção precoce da erva daninha *Phlomis umbrosa* em pastagens, essencial para a estabilidade dos ecossistemas e o desenvolvimento pecuário, Guo *et al.* (2023) propuseram o modelo de rede BSS-YOLOv8 com 2180 imagens tiradas de *Veículo Aéreo Não Tripulado (UAV)*. Incorporou o módulo BoTNet na rede backbone, melhorando a extração de recursos e adicionando uma camada de detecção para fortalecer o reconhecimento do alvo pequeno *Phlomis umbrosa*. Com uma precisão de 91,1% e uma taxa de recuperação de 86,7%, o modelo BSS-YOLOv8 mostra-se eficaz na detecção de ervas daninhas em pastagens.

Wang *et al.* (2023) necessidade de detectar rapidamente defeitos superficiais na madeira serrada na produção florestal. Apresentou uma rede de fusão de características bidirecionais, baseada no algoritmo YOLOv8, e propôs um modelo que combina mecanismo de atenção e otimização da função de perda. Essa abordagem permitiu uma detecção mais eficaz de pequenos defeitos na madeira, resultando em um modelo eficiente e de baixo consumo de recursos.

Os resultados destacam que o modelo TSW-YOLOv8n aprimorado alcançou uma eficiência de 91,10% mAP50 na identificação de oito tipos de defeitos em madeira serrada, com tempo médio de detecção de 6 milissegundos. Esses resultados representam um aumento de 5,1% na eficiência em comparação com o modelo original YOLOv8, além de uma redução de 1 milissegundos no tempo médio de detecção.

A Tabela 2 apresenta um resumo referente a detecção de produtos florestais.

Tabela 2 – Resultados detecção de objetos aplicada a produtos em geral

Trabalho	Detecção	Img/Esp.	Melhor acerto
Guo <i>et al.</i> (2023)	YOLOv8	2180/1	91,10%
Wang <i>et al.</i> (2023)	YOLOv8	21672/1	91,10%

Fonte: Autoria própria (2024).

2.7.3 Modelos de Segmentação de Imagens em geral

Por meio de técnicas de reconhecimento florestal, é possível identificar espécies de acordo com suas características anatômicas, no entanto, para uma maior precisão e eficácia desse processo, utilizam-se técnicas de segmentação de imagens.

Scheeren (2022) utiliza alguns modelos de aprendizado profundo como U-Net, *Rede de Pirâmide de Recursos (FPN)*, DeepLabv3+, CFNet e OCRNet em sua pesquisa para segmentação de folhas de café. Neste estudo também utilizou redes extratoras como ResNet-50, ResNet-152, DenseNet-121 e EfficientNet-b3, sendo que todos os modelos descritos aqui, apresentaram desempenhos satisfatórios acima de 83%.

A segmentação de imagens desempenha um papel crucial na visão computacional, buscando investigar a robustez do modelo *Segment Anything Model (SAM)* em cenários no mundo real, Wang, Zhao e Petzold (2023) realizaram um estudo comparativo utilizando nove conjuntos de dados, que totalizam 156268 imagens. As imagens foram submetidas a perturbações como ruído, desfoque, iluminação, cor, compressão, entre outros, em comparação com imagens não perturbadas, os resultados revelaram que o desempenho do modelo geralmente diminui conforme a intensidade das perturbações aumenta.

Visando melhorar a eficiência no diagnóstico de colonoscopia, Li, Hu e Yang (2023) propõem uma variação para o *SAM* utilizando técnica de aprendizagem por transferência. Foram utilizados cinco conjuntos de dados, CVC-ClinicDB, CVC-ColonDB, ETIS e CVC-300, sendo que todos os conjuntos foram obtidos resultados acima de 88%, variação essa com grande potencial de adaptação na área médica.

Chen *et al.* (2023) buscando propor um método para segmentação de feridas de Úlceras de Pé Diabético do inglês *Diabetic Foot Ulcer (DFU)*. Havendo necessidade de separar as feridas individualmente, em imagens nas quais existem mais de um tipo de ferida, o *SAM* possui técnicas avançadas para garantir resultados robustos e confiáveis, o conjunto de dados consiste em 1210 imagens de pés capturadas de 889 pacientes de várias clínicas diferentes, no qual o conjunto foi dividido em três partes de 800 para treinamento, 200 para validação e 200 para teste, resultando em uma eficiência de 83,5%.

A Tabela 3 apresenta um resumo dos trabalhos referentes a modelos de segmentação de imagens.

Tabela 3 – Resultados modelos de segmentação de imagens em geral

Trabalho	Segmentação	Imagens	Melhor acerto
Scheeren (2022)	U-Net	3846	93,23%
Wang, Zhao e Petzold (2023)	SAM	156268	81,00%
Li, Hu e Yang (2023)	SAM	2248	90,60%
Chen <i>et al.</i> (2023)	SAM	1210	83,50%

Fonte: Autoria própria (2024).

3 MATERIAIS E MÉTODOS

Após os estudos realizados sobre a identificação de espécies florestais e as pesquisas referentes às técnicas de segmentação e classificação de imagens, é importante definir as principais tecnologias e as etapas empregadas para alcançar os objetivos propostos. Portanto, neste capítulo são apresentados os materiais e os métodos utilizados para o desenvolvimento deste trabalho.

3.1 Materiais

Nesta seção, são descritas definições referentes ao ambiente de desenvolvimento escolhido, bem como suas configurações, bibliotecas e conjunto de dados.

3.1.1 Ambiente de desenvolvimento

Como ambiente de desenvolvimento foi utilizado o *Google Colaboratory (Colab)*¹ (Colab), este não requer instalação prévia, pois é um serviço em nuvem fornecido pela Google. Com o *Colab* a tarefa de escrever, executar, salvar e compartilhar código e análises é simplificada pois essas ações podem ser realizadas diretamente de um navegador de internet. Outro ponto positivo é a possibilidade de acessar de forma gratuita recursos computacionais importantes para o processamento de dados, como por exemplo *Graphics Processing Unit (GPU)*. Com a *GPU* é possível executar tarefas de processamento em um menor tempo de execução (COLAB, 2024).

A versão gratuita do *Colab* foi utilizada para o estudo. Entretanto, essa versão tem suas limitações, incluindo o tempo de execução, a quantidade de memória e o armazenamento. Essas limitações podem ser superadas adquirindo planos alternativos com custos adicionais (COLAB, 2024).

A escolha do *Colab* como ambiente se deu pela facilidade de uso, como também pela gratuidade dos recursos disponíveis. O plano gratuito oferece as seguintes configurações:

- *Central Processing Unit (CPU)* com dois núcleos *Xeon 2,2 GHz*;
- *Random Access Memory (RAM)* com 12,67 GB;
- 107.7 GB disponível de armazenamento;
- GPU *Tesla P100-PCIE 15 GB*;
- Python 3.

¹ <https://colab.google/>

*Python*² é uma linguagem de programação de alto nível e de fácil aprendizado, fornece uma vasta gama de bibliotecas padrão e diversas bibliotecas externas dedicadas a área de ciência de dados, tais como *PyTorch*, *NumPy*, *Scikit-learn*, entre outras (PYTHON DOCUMENTATION, 2024). A combinação da linguagem com essas ferramentas disponíveis tornam o Python uma escolha muito eficaz para o presente trabalho.

3.1.2 Bibliotecas e ferramentas

Bibliotecas são pacotes de códigos criados para serem reutilizados, evitando que a cada novo projeto tenha que criar códigos e rotinas do zero. Existem diversas opções de bibliotecas com ferramentas úteis para tratamento de dados, manipulação, aprendizado profundo, aprendizado de máquina e visão computacional. Dentre as diversas opções existentes, as seguintes foram escolhidas para serem utilizadas neste estudo:

- ***PyTorch***³ : é uma biblioteca de código aberto utilizada para implementações de aprendizado profundo e visão computacional. Devido a sua capacidade de distribuição de carga de trabalho em diferentes *GPUs*, é uma boa opção para realizar cálculos matemáticos complexos. Pytorch também se destaca por sua facilidade de aprendizagem e uso (PYTORCH, 2024);
- ***Scikit-learn***⁴ : é uma biblioteca desenvolvida para a aplicação prática de aprendizado de máquina. Essa biblioteca permite trabalhar com diferentes aplicações, tais como pré-processamento, classificação, regressão e ajuste de parâmetros (SCIKIT-LEARN, 2024);
- ***Numpy***⁵ : é uma biblioteca desenvolvida para computação numérica e científica. Ela permite trabalhar no tratamento dos dados, utilizando funções matemáticas para operações numéricas (NUMPY, 2024);
- ***Matplotlib***⁶ : é uma biblioteca de código aberto utilizada para criação de gráficos e visualização de dados (MATPLOTLIB, 2024);
- ***Open Source Computer Vision Library***⁷ (**OpenCV**) : é uma biblioteca de código aberto para visão computacional e processamento de imagens, fornece diversas funcionalidades como detecção de objetos, reconhecimento facial, rastreamento de movimento, entre outros (OPENCV, 2024);

² <https://www.python.org/doc/>

³ <https://pytorch.org/>

⁴ <https://scikit-learn.org/>

⁵ <https://numpy.org/>

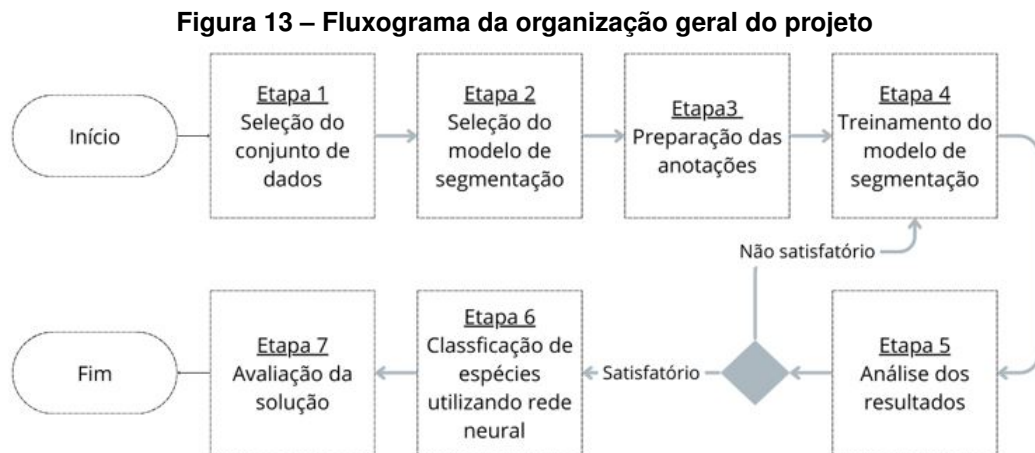
⁶ <https://matplotlib.org/>

⁷ <https://opencv.org/>

- **Computer Vision Annotation Tool⁸ (CVAT)** : é uma ferramenta de código aberto utilizada para projetos de visão computacional. O CVAT permite realizar a anotação de objetos em imagens e vídeos, esta tarefa é essencial para utilizar modelos de visão computacional (CVAT, 2024);
- **Roboflow⁹** : é uma plataforma desenvolvida com o intuito de simplificar o processo de desenvolvimento de projetos de visão computacional e aprendizado profundo. O Roboflow fornece diversos recursos para auxiliar no pré-processamento de dados, análise, visualização, geração de código e treinamentos de modelos (ROBOFLOW, 2024).

3.2 Métodos

Para melhorar a organização e demonstrar de forma objetiva um plano geral do trabalho, é apresentado na Figura 13 o fluxograma para as seis atividades principais realizadas. As etapas estão divididas desde a seleção do conjunto de dados até a análise dos resultados gerados pelos modelos de segmentação e classificação. Em cada etapa serão descritos os passos, o que foi feito e o que se obteve em cada uma das etapas.



Fonte: Autoria própria (2025).

3.2.1 Primeira etapa: Seleção conjunto de dados

A primeira etapa consistiu na busca por um conjunto de dados que contenha espécies florestais. A base de imagens foi desenvolvida em conjunto com o Grupo de Inteligência Computacional (GIC) da Universidade Tecnológica Federal do Paraná (UTFPR), campus Medianeira. As imagens são compostas por espécies do Bioma Amazônico, as fotografias foram retiradas de amostras fornecidas pela PF do Amazonas. As fotografias foram tiradas com diferentes apa-

⁸ <https://www.cvat.ai/>

⁹ <https://roboflow.com/>

relhos celulares, em ambiente com iluminação controlada. O conjunto é formado por vinte espécies diferentes, totalizando 386 imagens. A Tabela 4 sintetiza as informações referente ao conjunto de dados.

Tabela 4 – Síntese de informações dos conjuntos de dados usados no trabalho

Nº	Nome vulgar	Nome científico	Quantidade
1	Açoita cavalo	<i>Luehea divaricata</i>	50
2	<i>Astronium lecointei</i>	<i>Astronium lecointei</i>	50
3	Castanheira	<i>Bertholletia excelsa</i>	50
4	Sucupira-preta	<i>Bowdichia nitida</i>	50
5	Tanimbuca	<i>Buchenavia oxycarpa</i>	50
6	Piquiá	<i>Caryocar villosum</i>	50
7	Tauari	<i>Couratari tauari</i>	50
8	Cumarú	<i>Dipteryx odorata</i>	50
9	Angelim-pedra	<i>Dinizia excelsa</i>	50
10	Baru	<i>Dipterix</i>	50
11	Ducke	<i>Manilkara huberi</i>	50
12	Baru	<i>Eperua oleifera</i>	50
13	Garapeira	<i>Apuleia leiocarpa</i>	50
14	Maracatiara	<i>Astronium lecointei</i>	50
15	Peroba	<i>Aspidosperma polyneuron</i>	50
16	Roxinho	<i>Peltogyne maranhensis</i>	50
17	Sextonia rubra	<i>Sextonia rubra</i>	50
18	Tajuvá/Taiúva	<i>Maclura tinctoria</i>	50
19	Timbuíla	<i>Enterolobium contortisiliquum</i>	50
20	<i>Vochysia</i>	<i>Vochysia</i>	50
Total			1000

Fonte: Autoria própria (2025).

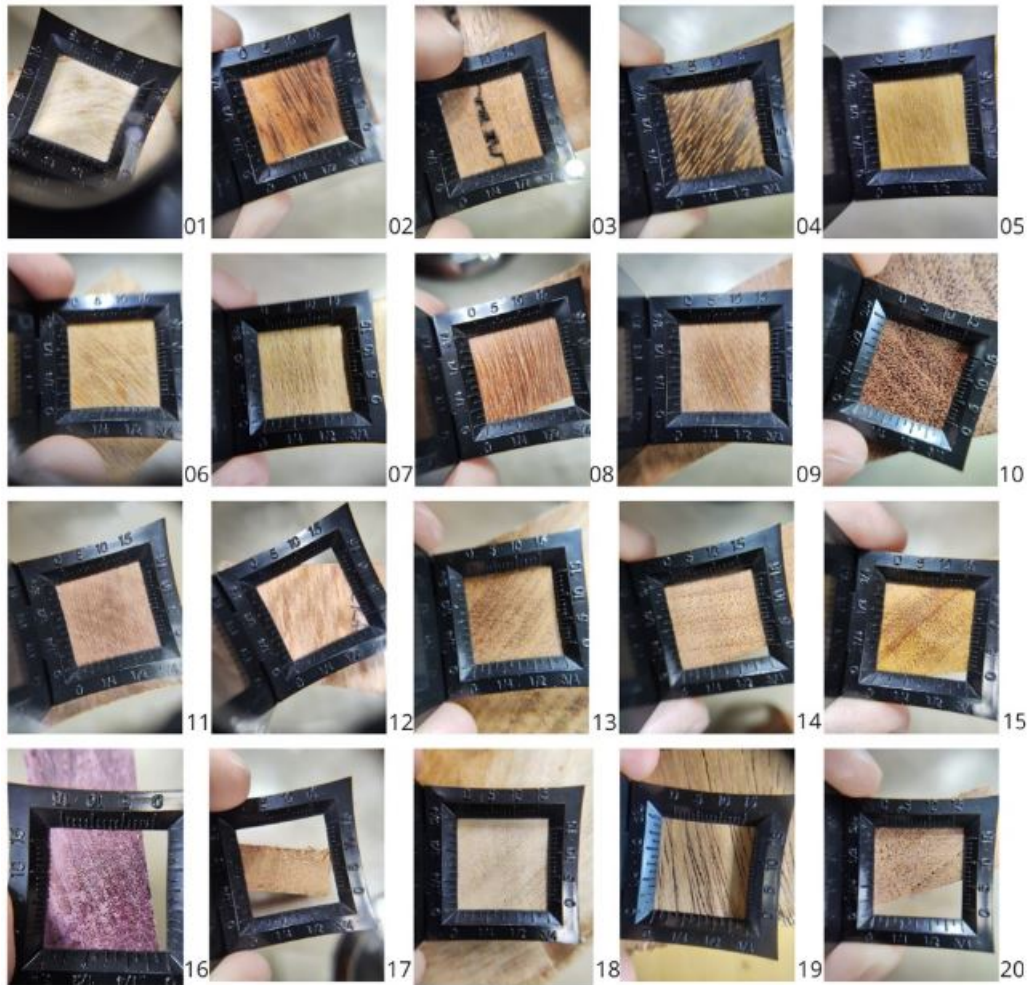
A Figura 14 apresenta exemplos de imagens das vinte espécies que estão presentes no conjunto de dados, no qual são visíveis as características macroscópicas físicas como a cor, e as características morfológicas como os raios medulares, medula, anéis de crescimento, entre outros, que compõem a madeira.

3.2.2 Segunda etapa: Seleção do modelo de segmentação

Nesta etapa é feita a descrição do modelo escolhido para a segmentação de imagens. O modelo é o YOLO na última versão disponível até a realização do trabalho, que é a versão 11, conforme descrito na subseção 2.3.3. A escolha deste modelo se justifica devido a sua alta precisão e facilidade de uso. Conforme Jocher, Chaurasia e Qiu (2024), o YOLO11 se destaca por sua eficiência e sua capacidade de lidar com diferentes tarefas de segmentação.

Devida à baixa quantidade de estudos focados na segmentação de madeira utilizando ambos os modelos até o momento, este estudo visa preencher essa lacuna. Assim, ao explorar e comparar a eficácia dos modelos, espera-se aprimorar o processo de identificação de espécies florestais.

Figura 14 – Exemplo de imagens do conjunto de dados



Fonte: Autoria própria (2025).

3.2.3 Terceira etapa: Preparação das anotações

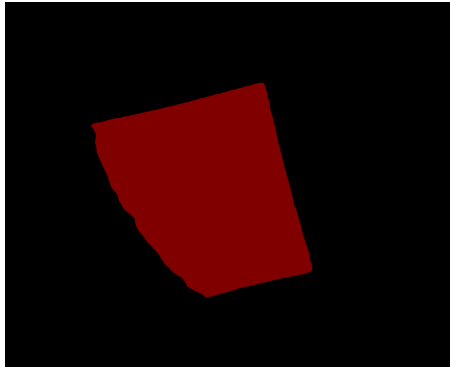
O conjunto de dados em desenvolvimento não possui anotações, estas foram criadas a partir da utilização das ferramenta *Roboflow*. Conforme exemplificado na Figura 15, dois tipos de anotação podem ser realizados. O primeiro tipo é a anotação binária, onde os píxeis podem ser vermelhos para madeira e preto para o fundo ou objetos estranhos. O segundo tipo é a anotação por polígonos, que circundam a madeira. O modelo de anotação binária será escolhido para realizar as anotações.

3.2.4 Quarta etapa: Treinamento dos modelos

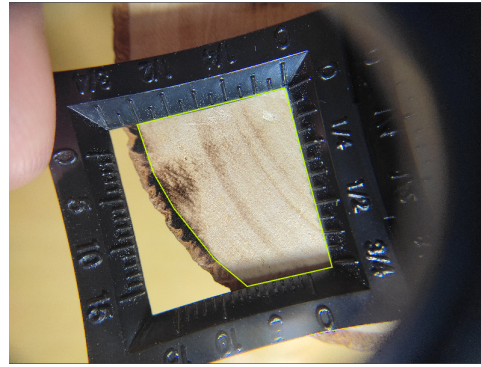
Como forma de comparação, e para ter uma base de performance, o primeiro experimento para a tarefa de classificação, será utilizado um conjunto de dados com imagens sem a segmentação, contanto com imagens originais tiradas da madeira, conforme pode ser observado na Figura 16.

Figura 15 – Tipos de máscara

(a) Máscara binária



(b) Máscara Poligonal



Fonte: Autoria própria (2025).

Figura 16 – Exemplos de imagens de entrada



Fonte: Autoria própria (2025).

Diversos fatores podem influenciar o desempenho de um modelo, entre esses fatores estão os hiperparâmetros, testar todas as combinações não é uma tarefa viável. Em busca dos parâmetros ideais, são necessários experimentos preliminares. Esses experimentos preliminares terão como objetivo encontrar as configurações que melhor se adaptam ao modelo para a tarefa proposta.

Para encontrar esses hiperparâmetros, dois algoritmos serão utilizados, começando com o *Grid Search*, que visa encontrar os melhores hiperparâmetros dentre todas as combinações possíveis dentro de um conjunto previamente especificado. Também, será utilizado o *Random Search*, esse algoritmo testa um número limitado de combinações, esse número é previamente especificado.

3.2.5 Quinta etapa: Análise dos resultados

Os resultados são avaliados segundo algumas medidas quantitativas utilizadas para avaliar o desempenho de um modelo em relação a um determinado problema. A utilização de métricas conhecidas permite a comparação entre técnicas presentes na literatura, a escolha de uma métrica está relacionada com o objetivo do estudo (Rizwan I Haque; NEUBERT, 2020).

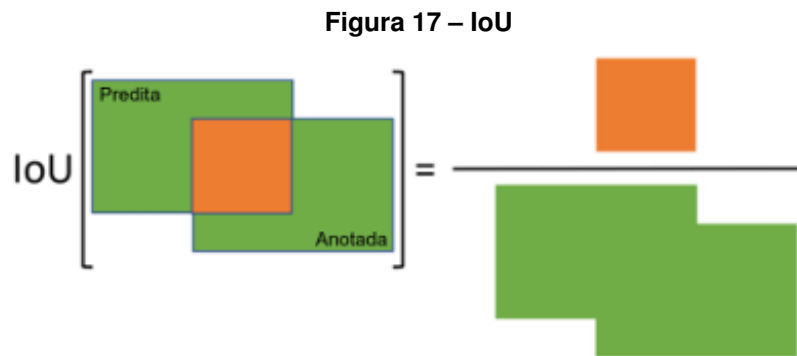
Para realizar o cálculo das métricas, os conjuntos de *pixels* classificados corretamente para classe positiva, são denominados verdadeiros positivos (VP), aqueles atribuídos incorretamente para classe positiva são falsos positivos (FP), não atribuídos a classe positiva, mas que pertencem a ela falsos negativos (FN) e os que foram classificados corretamente como não pertencentes a classe positiva, os verdadeiros negativos (VN).

Algumas métricas de desempenho estão listadas abaixo para avaliar a acurácia dos modelos de segmentação.

- **Intersection over Union (IoU):** é o resultado da divisão da área de intersecção pela área da união entre a caixa delimitadora predita pelo modelo e uma caixa anotada. A Equação 5 demonstra o cálculo, o resultado varia entre 0 e 1 (MINAEE *et al.*, 2022).

$$IoU = \frac{|Preditada \cap Anotada|}{|Preditada \cup Anotada|} \quad (5)$$

A Figura 17 ilustra a caixa delimitadora e a caixa anotada.



Fonte: Junior (2023).

- **Precision:** é a porcentagem de *pixels* classificados corretamente, o resultado varia entre 0 e 1. A Equação 6 demonstra o cálculo (Rizwan I Haque; NEUBERT, 2020).

$$Precision = \frac{VP}{VP + FP} \quad (6)$$

- **Recall:** é o número de *pixels* positivos que foram classificados corretamente em relação a todos os *pixels* da classe positiva. O resultado varia entre 0 e 1, conforme o cálculo da Equação 7 (Rizwan I Haque; NEUBERT, 2020).

$$Recall = \frac{VP}{VP + FN} \quad (7)$$

- **F1-score:** é a média harmônica entre *precision* e o *recall* conforme a Equação 8. O resultado varia entre 0 e 1 (Rizwan I Haque; NEUBERT, 2020).

$$F1 = \frac{2Precisão * Recall}{Precisão + Recall} \quad (8)$$

- **Dice Coefficient:** mede a similaridade entre as áreas predita e anotada, conforme a Equação 9. O resultado varia entre 0 e 1 (Rizwan I Haque; NEUBERT, 2020).

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (9)$$

3.2.6 Sexta etapa: Aplicação do classificador

Depois que os treinamentos forem realizados, se os resultados forem satisfatórios, o próximo passo será avaliar se a tarefa de segmentação de imagens teve impacto em relação a classificação das espécies. Para essa tarefa será utilizado o modelo *ResNet-50*, conforme proposto por He *et al.* (2015), que oferece bom desempenho, baixo consumo de memória e é um modelo eficiente para treinamentos rápidos.

A escolha foi fundamentada através dos resultados obtidos por Silva *et al.* (2022). Para um conjunto de imagens microscópicas, a acurácia alcançou 99,43%. Já para o conjunto com imagens macroscópicas, a acurácia foi de 99,12% após quatro épocas de treinamento.

3.2.7 Sétima etapa: Avaliação da solução desenvolvida

Para esta etapa, será feita uma avaliação da solução desenvolvida, serão consideradas a qualidade e a representatividade das espécies presentes no conjunto de imagens. Para o modelo de segmentação, será analisado a precisão e a eficiência computacional. Também, será avaliado se o desempenho da tarefa de segmentação teve impacto positivo em relação a classificação utilizando as imagens originais. Por fim, será feita uma análise em relação a aplicabilidade da solução desenvolvida e sua eficácia para auxiliar na tarefa de identificação de espécies.

4 RESULTADOS E DISCUSSÃO

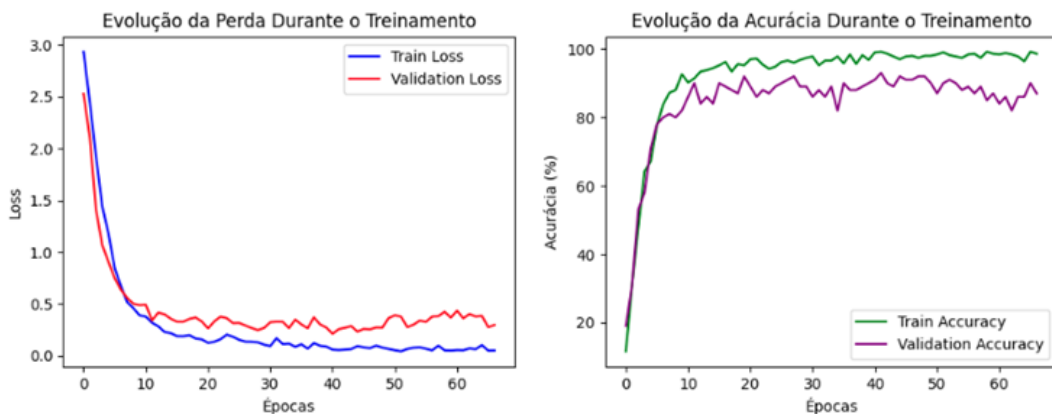
Neste capítulo, serão apresentados os resultados dos experimentos, e uma discussão de cada um deles. Para o treinamento, o conjunto de dados utilizado foi o desenvolvido ao longo deste trabalho, no qual, contém 1000 imagens, de 20 espécies diferentes, com 50 imagens para cada classe. O primeiro experimento utiliza a classificação das imagens do conjunto de dados sem nenhuma segmentação, este treinamento tem como objetivo ser usado como parâmetro de desempenho. O conjunto de dados foi dividido para todos os treinamentos da seguinte forma, 50% para treinamento, 10% para validação e 40% para teste, essa divisão busca garantir uma porcentagem segura para validação dos testes.

4.1 Classificação de Imagens com *ResNet-50* Sem Segmentação

Após testar vários parâmetros, com diferentes épocas e técnica de *early stopping*, foi escolhido realizar o treinamento com 100 épocas. Os resultados obtidos ao longo do treinamento do modelo demonstram uma boa capacidade de classificação. Na primeira época, a perda de treinamento foi de 2,93, com uma acurácia 11,60%. No conjunto de validação, a perda foi de 2,52 e a acurácia de 19,00%. Conforme o treinamento avançou, é possível observar uma melhora significativa nos indicadores.

Na época 10, a acurácia do treinamento atingiu 92,60%, com a diminuição da perda para 0,38. No conjunto de validação, a acurácia chegou a 82,00%, demonstrando uma maior capacidade de generalização do modelo. A acurácia se mantém alta alcançando 99,20% na época 42, com uma perda de 0.05. No conjunto de validação, a acurácia também atingiu um pico de 93,00%. A Figura 18 ilustra o crescimento da acurácia e a diminuição perda ao longo do treinamento, destacando a rápida melhoria inicial e a estabilização nas últimas épocas.

Figura 18 – Evolução das métricas durante o treinamento



Fonte: Autoria própria (2025).

No entanto, é possível verificar sinais de *overfitting* a partir da época 50, pois a perda no conjunto de validação teve uma oscilação e a acurácia diminuiu. A técnica de *early stopping* foi aplicada na época 66, interrompendo do treinamento antes que o modelo tivesse degradação dos resultados. Os parâmetros utilizados para a realização do treinamento:

- **Aumento de Dados (*Data Augmentation*):** Essa técnica auxilia na redução do *overfitting*, pois aumenta a variabilidade das imagens, ajustando configurações de brilho, contraste e rotações. Dessa forma o modelo tende a ter uma maior capacidade de generalização.
- **Regularização com *Dropout*:** Essa técnica auxilia para que o modelo não tenha um grande ajuste sobre os dados do treinamento, equilibrando o aprendizado e a generalização, como valor, foi utilizado 0,5.
- **Balanceamento de Classes:** Essa técnica ajusta a distribuição das amostras para evitar um viés do modelo para classes mais frequentes.
- **Parada Antecipada (*Early Stopping*):** Configurado para interromper o treinamento caso o valor da perda de validação não apresente melhoria nas últimas 25 épocas.
- **Redução Dinâmica da Taxa de Aprendizado:** Ajusta dinamicamente a taxa de aprendizado, reduzindo-a em 50% em caso de a perda de validação não melhorar por três épocas consecutivas. Essa configuração auxilia na melhora dos pesos para as últimas etapas do treinamento, melhorando a convergência.

A Tabela 5 demonstra o relatório de classificação do conjunto de teste. A acurácia global foi de 91%, com métricas de precisão, *recall* e *F1-score* equilibradas. No entanto, observou-se a variação no desempenho de algumas classes, como *Bertolethia excelsa* e *Roxinho*, que apresentaram valores de *recall* mais baixos. Esse comportamento sugere que o modelo teve dificuldade de lidar com as características dessas espécies.

Conforme pode ser observado na Figura 19, a matriz de confusão do modelo demonstra os acertos e erros de classificação e os padrões de semelhança entre algumas classes. A maioria das classes apresenta um grande número de acertos, porém, em algumas ocorre erros de classificação.

A estabilização da perda e da acurácia indica que o modelo atingiu seu melhor desempenho, com uma acurácia global satisfatória tanto no conjunto de treinamento quanto no conjunto de teste. Além disso, a configuração *early stopping* foi importante para evitar ajustes excessivos aos dados de treinamento, economizando recursos computacionais e prevenindo a perda da capacidade de generalização.

Tabela 5 – Relatório de classificação - Teste

Espécie	Precisão	Recall	F1-score	Imagens
Açoita cavalo	1,00	1,00	1,00	20
<i>Astronium lecantei</i>	1,00	0,80	0,89	20
<i>Bertolethia excelsa</i>	1,00	0,75	0,86	20
<i>Bowdichia nitida</i>	0,86	0,90	0,88	20
<i>Buchenavia oxycarpa</i>	0,89	0,85	0,87	20
<i>Cariniana miantha</i>	1,00	1,00	1,00	20
<i>Caryocar villosum</i>	0,89	0,85	0,87	20
<i>Couratari tauari</i>	0,78	0,90	0,84	20
<i>Dinizia excelsa</i>	0,77	0,85	0,81	20
<i>Dipterix</i>	0,87	1,00	0,93	20
<i>Dipteryx odorata</i>	0,87	1,00	0,93	20
<i>Eperua oleifera</i>	0,94	0,80	0,86	20
Garapeira	0,76	0,95	0,84	20
<i>Manilkara huberi</i>	1,00	1,00	1,00	20
Maracatiara	0,74	1,00	0,85	20
Peroba	0,95	0,90	0,92	20
Roxinho	1,00	0,75	0,86	20
<i>Sextonia rubra</i>	1,00	0,90	0,95	20
Tajuva	0,95	1,00	0,98	20
<i>Vochysia</i>	0,94	0,80	0,86	20
Média	0,91	0,90	0,90	

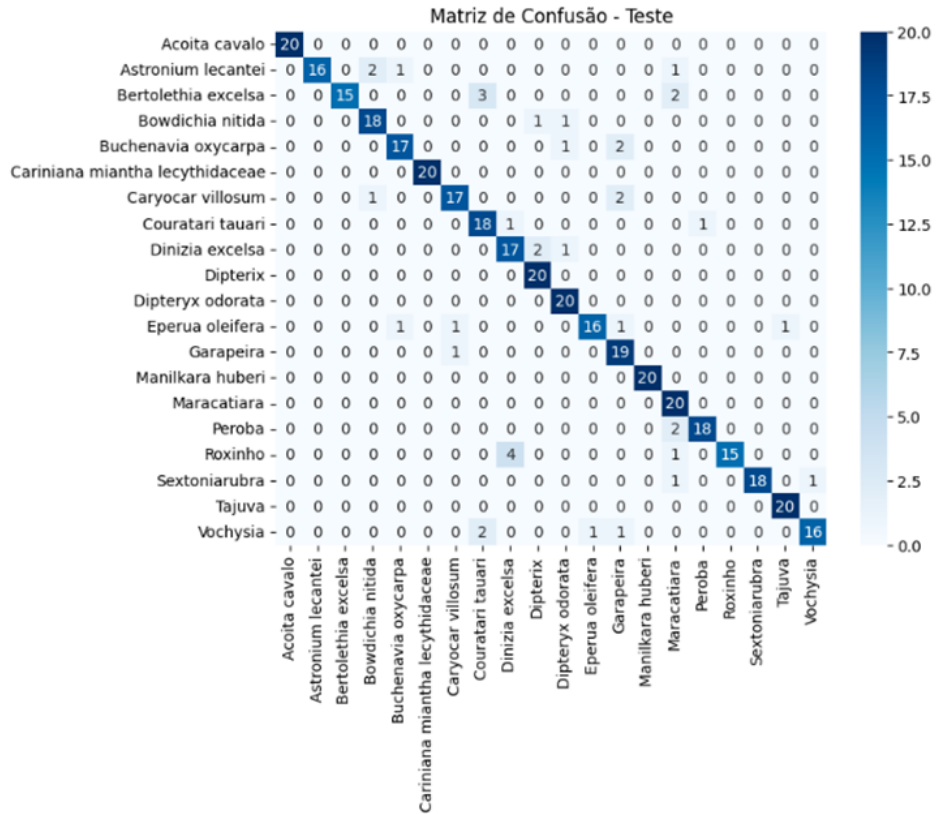
Fonte: Autoria própria (2025).

4.2 Segmentação das Imagens no Conjunto de Dados YOLO11

O treinamento do modelo YOLO11 foi realizado com 30 épocas, utilizando um conjunto de dados dividido em 50% para treinamento, 10% para validação e 40% para teste. O objetivo principal foi avaliar a capacidade do modelo em identificar e segmentar objetos presentes nas imagens, usando métricas como precisão (P), *recall* (R), IoU, *F1-Score* e Coeficiente Dice. A precisão do modelo durante ao longo das épocas, demonstra a capacidade do modelo em acertar os verdadeiros positivos, com um valor de 0,72 ao final do treinamento. Indicando que o modelo é eficiente na identificação correta das instâncias, embora algumas classes, como Açoita cavalo e Garapeira, apresentaram precisão inferior, demonstrando que o modelo teve dificuldade em diferenciar essas instâncias.

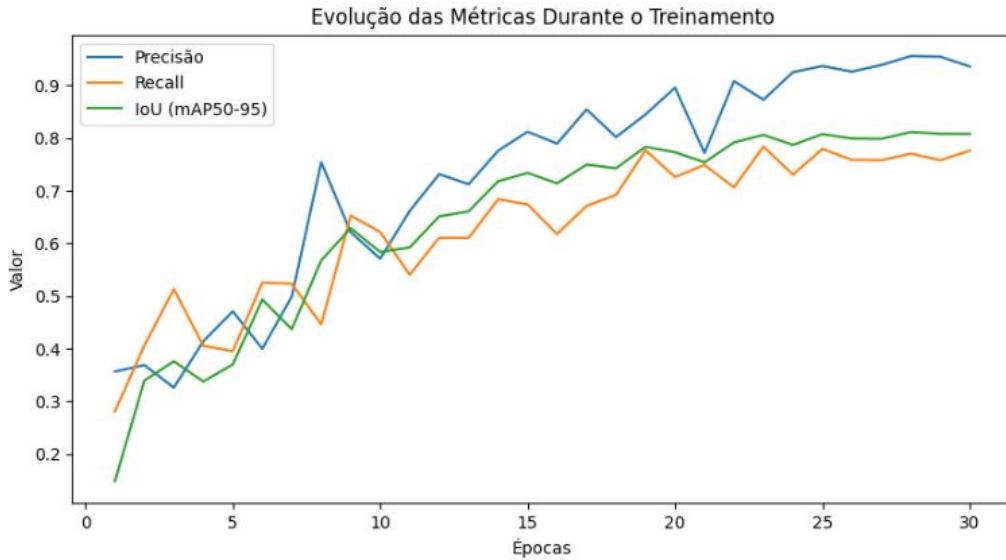
O *recall*, que mede a capacidade do modelo de identificar corretamente os os casos positivos, apresentou uma média de 0,63, para o treinamento. indicando que o modelo foi capaz de detectar a maioria das instâncias, embora algumas classes, como Açoita cavalo, mostraram um *recall* muito baixo 0,161. A evolução do IoU médio foi de 0,64, indicando que o modelo tem uma sobreposição moderada entre as caixas previstas e as reais. A Figura 20 ilustra a evolução das métricas de precisão, recall e IoU ao longo do treinamento, evidenciando como o modelo foi aprimorando seu desempenho conforme o número de épocas aumentava. Já a Tabela 6 apresenta um resumo quantitativo das métricas do treinamento.

Figura 19 – Matriz de confusão - teste



Fonte: Autoria própria (2025).

Figura 20 – Evolução das métricas durante o treinamento



Fonte: Autoria própria (2025).

O modelo YOLO11 apresentou bom desempenho durante a fase de teste, com precisão de 0,87 e *recall* de 0,78, demonstrando uma boa capacidade de identificar e gerar as máscaras de segmentação nas imagens, conforme mostrado na Tabela 7. As métricas de IoU de 0,80,

Tabela 6 – Métricas de treinamento YOLO11

Métrica	Valor
Precisão	0,72
<i>Recall</i>	0,63
Interseção sobre união (IoU)	0,64
<i>F1-Score</i>	0,67
Coeficiente Dice	0,78

Fonte: Autoria própria (2025).

F1-Score de 0,82 e coeficiente dice de 0,81 indicam uma boa sobreposição e qualidade de segmentação.

Tabela 7 – Métricas de avaliação YOLO11

Métrica	Valor
Precisão	0,87
<i>Recall</i>	0,78
Interseção sobre união (IoU)	0,80
<i>F1-Score</i>	0,82
Coeficiente Dice	0,81

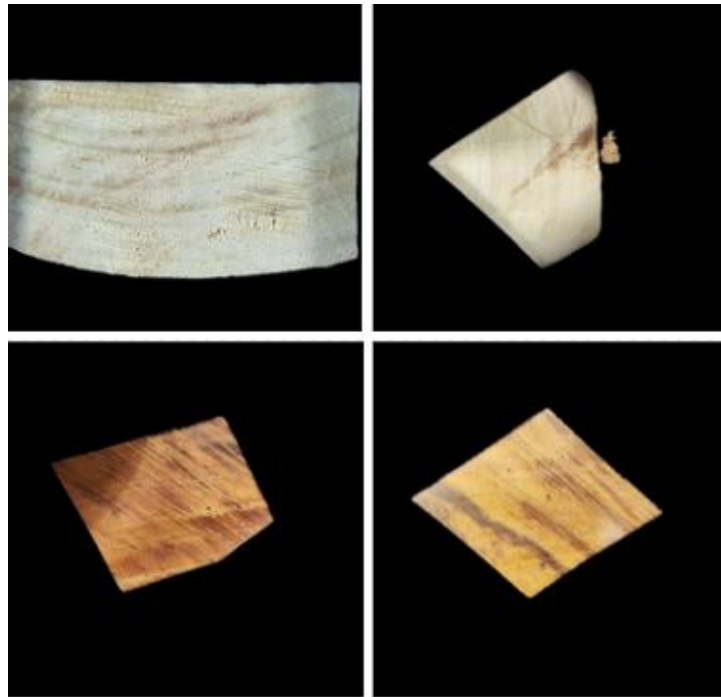
Fonte: Autoria própria (2025).

4.3 Classificação das Imagens Segmentadas com ResNet-50

Após a segmentação das imagens pelo modelo YOLO11, as máscaras geradas foram utilizadas para criar um novo conjunto de dados, composto pelas imagens recortadas dentro da região de interesse. A Figura 21 demonstra alguns exemplos das imagens que fazem parte do novo conjunto.

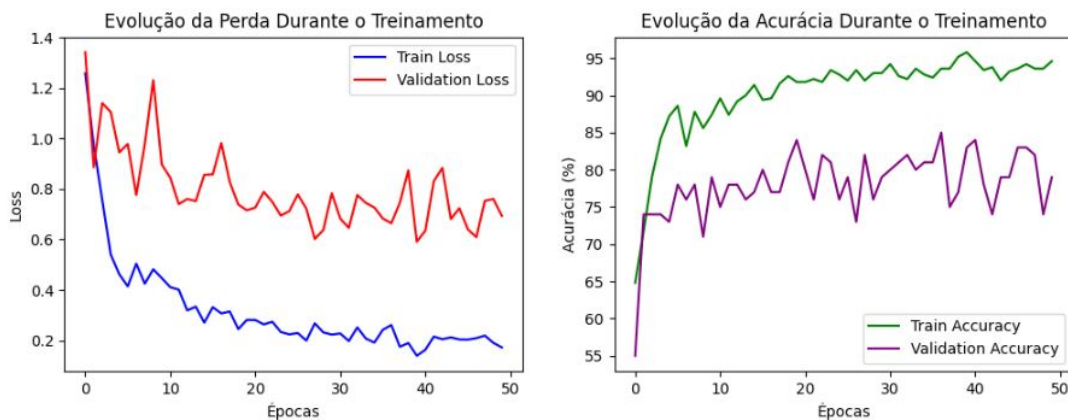
O modelo foi treinado por 50 épocas, a Figura 22 demonstra a evolução das métricas conforme o treinamento foi avançando as épocas, com uma perda do treinamento na época 1 de 1,25 para 0,17 na época 50, com a acurácia do modelo saindo de 64% na fase inicial, até atingir 94% na época 50 para o treinamento. Já para a validação, o modelo apresentou uma variação, com alguns picos de performance na época 20, e algumas quedas nas épocas 43 e 49, indicando um possível sobreajuste. O modelo demonstrou uma boa capacidade de reconhecimento das espécies. Conforme pode ser observado na Tabela 8 a precisão e o *recall* nas classes apresentaram variação, com algumas classes, como Acoita cavalo, *Dipterix* e Roxinho, obtendo 100% de precisão. No entanto, classes como Peroba e Garapeira tiveram desempenhos mais baixos, com *f1-score* de 0,48 e 0,68, respectivamente. A acurácia global foi de 83%.

Figura 21 – Imagens segmentadas



Fonte: Autoria própria (2025).

Figura 22 – Evolução das métricas durante o treinamento



Fonte: Autoria própria (2025).

4.4 Classificação das Imagens Segmentadas com ResNet-50 usando Crop

Com o intuito de melhorar a acurácia do modelo com imagens segmentadas, foi realizado um novo treinamento utilizando a técnica de *Crop*, na qual faz o recorte das imagens em porções menores, para ajustar o tamanho da entrada. Existem diferentes configurações de *crop*, como centralizado, que realiza o recorte da imagem ao redor do centro, e a aleatório, onde a imagem é recortada de forma variada. Para o trabalho foi escolhido o formato centralizado. A Figura 23 ilustra a imagem segmentada, em relação a nova imagem após o *crop*. A aplicação dessa técnica tem impacto no conjunto de dados, pois pode aumentar a variabilidade das imagens, o que auxilia na generalização do modelo. A utilização dessa técnica foi realizada visando

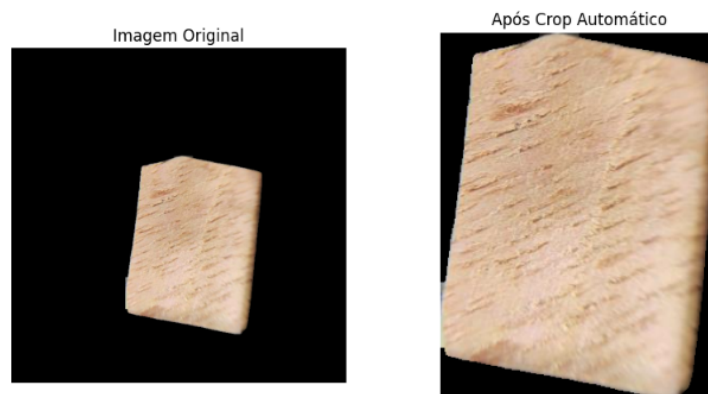
Tabela 8 – Relatório de classificação - Teste

Espécie	Precisão	Recall	F1-score	Imagens
<i>Acoita cavalo</i>	1,00	0,95	0,97	20
<i>Astronium lecantei</i>	0,94	0,85	0,89	20
<i>Bertolethia excelsa</i>	0,75	0,90	0,82	20
<i>Bowdichia nitida</i>	1,00	0,90	0,95	20
<i>Buchenavia oxycarpa</i>	0,76	0,80	0,78	20
<i>Cariniana miantha</i>	0,87	1,00	0,93	20
<i>Caryocar villosum</i>	0,77	0,85	0,81	20
<i>Couratari tauari</i>	0,79	0,55	0,65	20
<i>Dinizia excelsa</i>	0,88	0,75	0,81	20
<i>Dipterix</i>	0,95	1,00	0,98	20
<i>Dipteryx odorata</i>	0,69	1,00	0,82	20
<i>Eperua oleifera</i>	0,75	0,90	0,82	20
Garapeira	0,62	0,75	0,68	20
<i>Manilkara huberi</i>	0,74	0,85	0,79	20
Maracatiara	0,80	0,80	0,80	20
Peroba	0,78	0,35	0,48	20
Roxinho	1,00	0,90	0,95	20
<i>Sextoniarubra</i>	0,94	0,85	0,89	20
Tajuba	1,00	0,75	0,86	20
<i>Vochysia</i>	0,78	0,90	0,84	20
Média	0,85	0,85	0,83	

Fonte: Autoria própria (2025).

umentar o desempenho do modelo, pois ao utilizar as imagens segmentadas, para algumas imagens, a região segmentada ficou com tamanho reduzido.

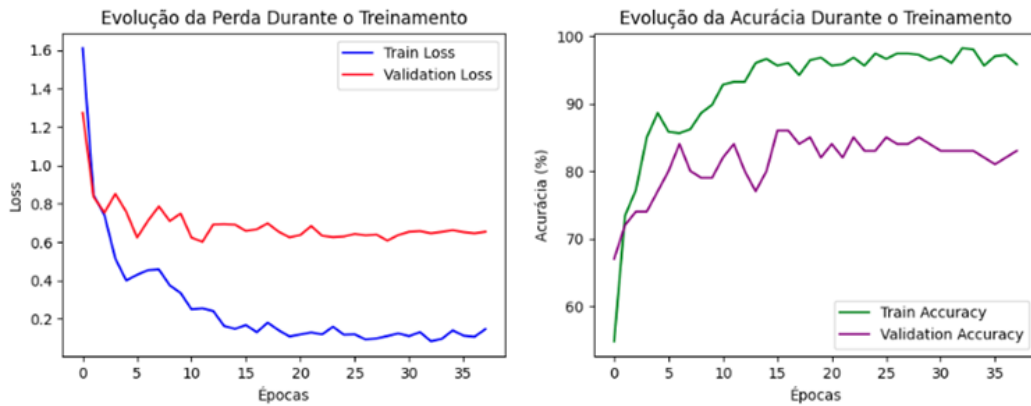
Figura 23 – Crop



Fonte: Autoria própria (2025).

O modelo foi treinado por 38 épocas, sendo interrompido por critério de *early stopping*. Nas primeiras épocas, houve uma melhora constante no desempenho do modelo com a perda (*loss*) inicial de 1,61 caindo para 0,10 no decorrer do treinamento e a acurácia passando de 54,80% para 97,20%. A Figura 24 apresenta a evolução da perda e da acurácia durante as 38 épocas de treinamento. O modelo apresentou bom desempenho durante o treinamento, oscilando a acurácia e a perda em épocas intermediárias.

Figura 24 – Evolução das métricas durante o treinamento



Fonte: Autoria própria (2025).

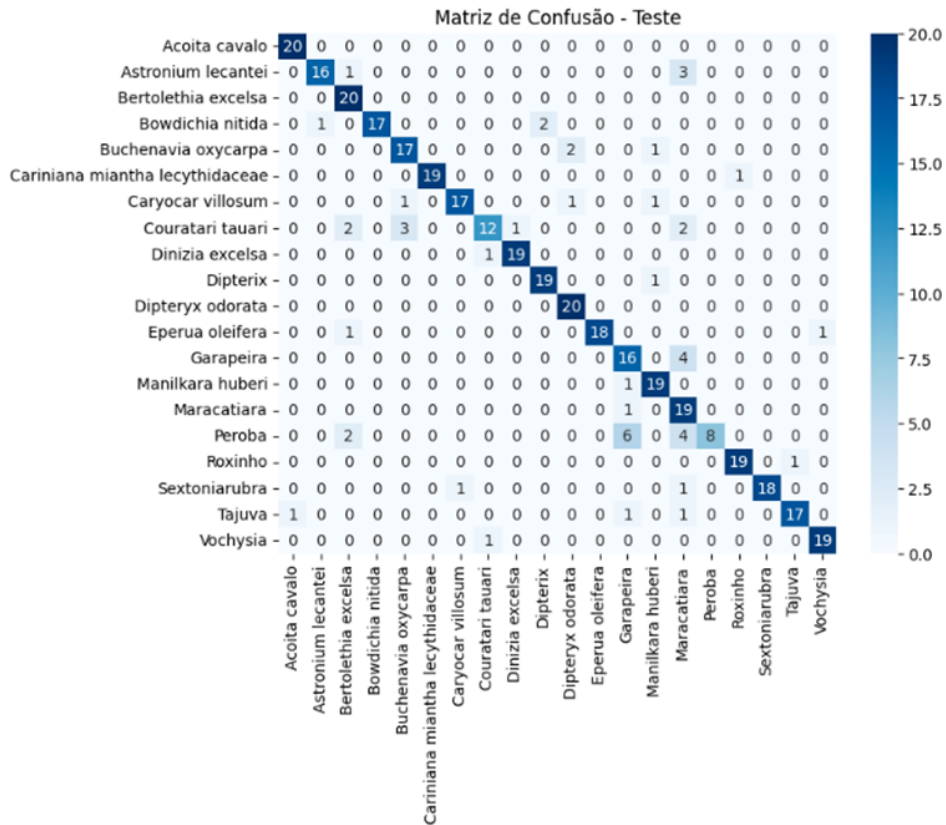
A Tabela 9 apresenta os resultados de *precision*, *recall* e *f1-score* para cada classe no conjunto de teste, e a acurácia global foi de 87%. O relatório demonstra bons resultados no geral, com a precisão alta para muitas classes, porém, algumas espécies apresentam *recall* baixo, como *Peroba* com 0,40, indicando que o modelo teve dificuldades de identificar todas as ocorrências dessa classe. A matriz de confusão do modelo, apresentada na Figura 25, fornece uma visão detalhada do desempenho do modelo em cada classe, indicando o número de predições corretas e incorretas.

Tabela 9 – Teste - Relatório de classificação

Espécie	Precisão	Recall	F1-score	Imagens
Açoita cavalo	0,95	1,00	0,98	20
<i>Astronium lecantei</i>	0,94	0,80	0,86	20
<i>Bertolethia excelsa</i>	0,77	1,00	0,87	20
<i>Bowdichia nitida</i>	1,00	0,85	0,92	20
<i>Buchenavia oxycarpa</i>	0,81	0,85	0,83	20
<i>Cariniana miantha</i>	1,00	0,95	0,97	20
<i>Caryocar villosum</i>	0,94	0,85	0,89	20
<i>Couratari tauari</i>	0,86	0,60	0,71	20
<i>Dinizia excelsa</i>	0,95	0,95	0,95	20
<i>Dipteryx</i>	0,90	0,95	0,93	20
<i>Dipteryx odorata</i>	0,87	1,00	0,93	20
<i>Eperua oleifera</i>	1,00	0,90	0,95	20
Garapeira	0,64	0,80	0,71	20
<i>Manilkara huberi</i>	0,86	0,95	0,90	20
Maracatiara	0,56	0,95	0,70	20
Peroba	1,00	0,40	0,57	20
Roxinho	0,95	0,95	0,95	20
<i>Sextonia rubra</i>	1,00	0,90	0,95	20
Tajuva	0,94	0,85	0,89	20
<i>Vochysia</i>	0,95	0,95	0,95	20
Média	0,89	0,87	0,87	

Fonte: Autoria própria (2025).

Figura 25 – Matriz de confusão - Teste



Fonte: Autoria própria (2025).

4.5 Avaliação

Conforme a Tabela 10, ao analisar o desempenho sem e com segmentação, observa-se que em geral as métricas tiveram uma queda após a segmentação. A precisão média sem a segmentação foi de 0,91, enquanto com segmentação diminuiu para 0,85. Indicando que o modelo de sem a segmentação foi mais preciso. Já a o *recall* saiu de uma média de 0,90 sem a segmentação para 0,85 com a segmentação. No entanto, ao usar a técnica de *Crop*, houve uma melhora, com a precisão média aumentando para 0,89 e o *recall* de para 0,87. Embora a segmentação tenha melhorado o desempenho para algumas espécies como *Dinizia excelsa*, com aumento na precisão e *recall*, passando de 77% para 95%. No geral, a segmentação causou queda no desempenho da acurácia geral dos modelos, as acurácias finais dos modelos foram, para imagens sem segmentação de 91%, imagens segmentadas 83%, e para imagens segmentadas com o *Crop* de 87%.

Tabela 10 – Comparação entre resultados com e sem segmentação

Especie	Sem Segmentação			Com Segmentação			Segmentadas (Crop)			Imagens
	Prec	Recall	F1	Prec	Recall	F1	Prec	Recall	F1	
<i>Acoita cavalo</i>	1,00	1,00	1,00	1,00	0,95	0,97	0,95	1,00	0,98	20
<i>Astronium lecantei</i>	1,00	0,80	0,89	0,94	0,85	0,89	0,94	0,80	0,86	20
<i>Bertolethia excelsa</i>	1,00	0,75	0,86	0,75	0,90	0,82	0,77	1,00	0,87	20
<i>Bowdichia nitida</i>	0,86	0,90	0,88	1,00	0,90	0,95	1,00	0,85	0,92	20
<i>Buchenavia oxycarpa</i>	0,89	0,85	0,87	0,76	0,80	0,78	0,81	0,85	0,83	20
<i>Cariniana miantha</i>	1,00	1,00	1,00	0,87	1,00	0,93	1,00	0,95	0,97	20
<i>Caryocar villosum</i>	0,89	0,85	0,87	0,77	0,85	0,81	0,94	0,85	0,89	20
<i>Couratari tauari</i>	0,78	0,90	0,84	0,79	0,55	0,65	0,86	0,60	0,71	20
<i>Dinizia excelsa</i>	0,77	0,85	0,81	0,88	0,75	0,81	0,95	0,95	0,95	20
<i>Dipterix</i>	0,87	1,00	0,93	0,95	1,00	0,98	0,90	0,95	0,93	20
<i>Dipteryx odorata</i>	0,87	1,00	0,93	0,69	1,00	0,82	0,87	1,00	0,93	20
<i>Eperua oleífera</i>	0,94	0,80	0,86	0,75	0,90	0,82	1,00	0,90	0,95	20
Garapeira	0,76	0,95	0,84	0,62	0,75	0,68	0,64	0,80	0,71	20
<i>Manilkara huberi</i>	1,00	1,00	1,00	0,74	0,85	0,79	0,86	0,95	0,90	20
Maracatiara	0,74	1,00	0,85	0,80	0,80	0,80	0,56	0,95	0,70	20
Peroba	0,95	0,90	0,92	0,78	0,35	0,48	1,00	0,40	0,57	20
Roxinho	1,00	0,75	0,86	1,00	0,90	0,95	0,95	0,95	0,95	20
<i>Sextoniarubra</i>	1,00	0,90	0,95	0,94	0,85	0,89	1,00	0,90	0,95	20
Tajuva	0,95	1,00	0,98	1,00	0,75	0,86	0,94	0,85	0,89	20
<i>Vochysia</i>	0,94	0,80	0,86	0,78	0,90	0,84	0,95	0,95	0,95	20
Média	0,91	0,90	0,90	0,85	0,85	0,83	0,89	0,87	0,87	400

Fonte: Autoria própria (2025).

5 CONCLUSÃO

Este estudo teve como objetivo analisar a aplicação do modelo de segmentação YOLO11 para melhorar a acurácia do modelo de classificação de imagens *Resnet-50*. Através da comparação entre classificações realizadas com e sem segmentação, assim foi possível analisar os impactos da segmentação na tarefa de classificação.

Foi construído um conjunto de imagens de madeira tiradas a partir de um celular comum, no qual aparecem objetos como lupa, partes do corpo humano, entre outros, simulando a operação em campo por parte de agentes fiscalizadores. Tal conjunto de dados é composto por 20 espécies com 50 amostras cada, totalizando 1000 imagens. Todo esse conjunto de imagens foi rotulado, e a partir disso foram geradas as respectivas máscaras de segmentação.

Para o objetivo de segmentar as imagens foi usado o modelo YOLO11 para realizar a segmentação com base nas máscaras, no qual diferentes parâmetros foram testados, gerando um coeficiente dice com 81%, demonstrando a capacidade do modelo em realizar a segmentação com base nas caixas anotadas.

Com o intuito de classificar as espécies florestais, três abordagens foram propostas, a primeira utilizando a imagem completa sem a segmentação, que tinha como objetivo ter uma análise inicial do comportamento do modelo *Resnet-50*, foi obtida uma precisão de 91%. Já a segunda abordagem, foi realizada com as imagens segmentadas com fundo preto, com o tamanho original de 640x640, o que resultou em uma precisão de 83%, eliminando o fundo, garantindo que não atrapalhariam. E a terceira abordagem, com a região de interesse redimensionando as imagens, retirando os *pixels* pretos, verificando se o fundo preto iria ou não interferir na classificação, com resultado de 87%.

A partir desses experimentos, pode-se perceber que a segmentação não teve impacto significativo na acurácia da classificação das espécies. Adicionando o custo computacional requerido para a etapa de segmentação, que inclui o treinamento, a geração das máscaras, e a aplicação no conjunto de dados, não se justificou a utilização da segmentação. Assim, o modelo de classificação treinado com imagens originais, sem a segmentação, se mostrou mais eficaz para a tarefa proposta, com uma acurácia geral de 91%, em comparação com o modelo que utilizou segmentação, que obteve 87%.

Apesar disso, observou-se que o modelo YOLO se comportou bem, porém, considerando que a segmentação não teve um impacto significativo na acurácia de classificação, poderiam ser avaliadas outras abordagens para tarefa de segmentação. Como alternativa, poderiam ser utilizados modelos alternativos, como o SAM, validando se a segmentação pode ter uma maior contribuição na identificação das espécies.

5.1 Trabalhos futuros

Os estudos apresentados neste trabalho podem ser aprofundados testando outros modelos de segmentação de última geração, tais como o SAM, a fim de avaliar o impacto com abordagens distintas. Além disso, testar diferentes combinações de parâmetros e configurações podem contribuir para um melhor desempenho. Para tornar a solução aplicável em cenários reais, como o uso em campo por parte dos agentes fiscalizadores, pode-se buscar o desenvolvimento de um modelo enxuto para execução em dispositivos móveis, utilizando técnicas com abordagens como *Vision Transformer (ViT)* e *MobileViT*.

Outra melhoria, seria o aumento do conjunto de dados, aumentando o número de espécies, como também o número de amostras por espécie. Também, o conjunto de dados pode ser utilizado em outros problemas relacionados à segmentação e classificação de espécies. Por fim, a utilização da técnica *Gradient-weighted Class Activation Mapping (Grad-CAM)*, que pode ajudar na interpretação das decisões do modelo de classificação, indicando quais foram as características da madeira que tiveram maior influência na classificação da espécie.

REFERÊNCIAS

- ACHATZ, J. *et al.* Convolutional neural networks for quality and species sorting of roundwood with image and numerical data. **Expert Systems with Applications**, v. 246, p. 123117, 2024. Disponível em: https://www.researchgate.net/publication/377167038_Convolutional_neural_networks_for_quality_and_species_sorting_of_roundwood_with_image_and_numerical_data. Acesso em: 28 mai. 2024.
- AMJOURD, A. B.; AMROUCH, M. Object detection using deep learning, cnns and vision transformers: A review. **IEEE Access**, v. 11, p. 35479–35516, 2023. Disponível em: <https://ieeexplore.ieee.org/document/10098596>. Acesso em: 04 jun. 2024.
- ANDREZA, M. N. d. S. *et al.* Interferência do desmatamento na qualidade ambiental e social. **Brazilian Journal of Development**, v. 10, p. 1420–1436, 2024. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/66410>. Acesso em: 01 jun. 2024.
- BOTOSSO, P. C. Identificação macroscópica de madeiras: guia prático e noções básicas para o seu reconhecimento. **Embrapa**, Embrapa, maio 2011. Disponível em: <https://www.infoteca.cnptia.embrapa.br/handle/doc/894373>. Acesso em: 09 mai. 2024.
- BRAGAGNOLO, L.; DA SILVA, R.; GRZYBOWSKI, J. Amazon forest cover change mapping based on semantic segmentation by u-nets. **Ecological Informatics**, v. 62, p. 101279, 2021. ISSN 1574-9541. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574954121000704>. Acesso em: 29 mai. 2024.
- CARVALHO, W. *et al.* Uma visão sobre a estrutura, composição e biodegradação da madeira. **Química Nova**, v. 32, p. 2191–2195, 2009. ISSN 0100-4042. Disponível em: <https://doi.org/10.1590/S0100-40422009000800033>. Acesso em: 02 jun. 2024.
- CASTRO, A.; ANDRADE, D. O custo econômico do desmatamento da floresta amazônica brasileira (1988-2014). **Perspectiva Econômica**, v. 12, p. 21, 2016. Disponível em: https://www.researchgate.net/publication/305313317_O_custo_economico_do_desmatamento_da_Floresta_Amazonica_brasileira_1988-2014. Acesso em: 01 jun. 2024.
- CHEN, Y.-P. *et al.* A deep learning-based segmentation strategy for diabetic foot ulcers: Combining the strengths of hardnet-mseg and sam models. *In: 2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)*. [s.n.], 2023. p. 378–381. Disponível em: https://www.researchgate.net/publication/373355559_A_Deep_Learning-Based_Segmentation_Strategy_for_Diabetic_Foot_Ulcers_Combining_the_Strengths_of_HarDNet-MSEG_and_SAM_Models. Acesso em: 09 fev. 2024.
- COLAB, G. **Google Colaboratory**. 2024. Disponível em: <https://colab.google/>. Acesso em: 11 mai. 2024.
- CVAT. **Open Data Annotation Platform**. 2024. Disponível em: <https://www.cvat.ai/>. Acesso em: 12 mai. 2024.
- DICKMANN, I.; LIOTTI, L. Educação ambiental crítica: Mudanças climáticas. *In: ____*. [s.n.], 2024. p. 189. ISBN 9786580329663. Disponível em: <https://www.researchgate.net/publication/3803>. Acesso em: 20 mai. 2024.

DING, B.; QIAN, H.; ZHOU, J. Activation functions and their characteristics in deep neural networks. In: **2018 Chinese Control And Decision Conference (CCDC)**. [s.n.], 2018. p. 1836–1841. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8407425>. Acesso em: 09 mai. 2024.

DINIZ, M. B. **Desmatamento e ausência de riqueza na Amazônia**. Belém : Editora Paka-Tatu, 2017. <https://www.gbv.de/dms/zbw/1668938774.pdf>. Disponível em: <https://www.econbiz.de/Record/desmatamento-e-aus%C3%AAncia-de-riqueza-na-amaz%C3%B4nia-diniz-marcelo-bentes/10012025755>. Acesso em: 30 mai. 2024.

FABIJAŃSKA, A.; DANEK, M.; BARNIAK, J. Wood species automatic identification from wood core images with a residual convolutional neural network. **Computers and Electronics in Agriculture**, v. 181, p. 105941, 2021. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S016816992033146X>. Acesso em: 21 dez. 2023.

FEARNSIDE, P. Desmatamento na amazônia brasileira: História, índices e consequências. In: _____. [s.n.], 2020. ISBN 978-85-211-0193-2. Disponível em: https://www.researchgate.net/publication/340923558_Desmatamento_na_Amazonia_brasileira_Historia_indices_e_consequencias. Acesso em: 29 mai. 2024.

GEUS, A. *et al.* Amazon wood species classification: a comparison between deep learning and pre-designed features. **Wood Science and Technology**, v. 55, 2021. Disponível em: https://www.researchgate.net/publication/350244797_Amazon_wood_species_classification_a_comparison_between_deep_learning_and_pre-designed_features. Acesso em: 28 mai. 2024.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. MIT Press, 2016. Disponível em: <http://www.deeplearningbook.org>. Acesso em: 20 mai. 2024.

GUO, B. *et al.* Detection of the grassland weed phlomis umbrosa using multi-source imagery and an improved yolov8 network. **Agronomy**, v. 13, n. 12, 2023. ISSN 2073-4395. Disponível em: <https://www.mdpi.com/2073-4395/13/12/3001>. Acesso em: 05 jan. 2024.

HAFIZ, A. M.; BHAT, G. M. A survey on instance segmentation: state of the art. **International Journal of Multimedia Information Retrieval**, v. 9, n. 3, p. 171–189, 2020. ISSN 2192-662X. Disponível em: <https://doi.org/10.1007/s13735-020-00195-x>. Acesso em: 02 jun. 2024.

HE, K. *et al.* Deep residual learning for image recognition. **ArXiv**, abs/1512.03385, p. 12, 2015. Disponível em: <https://arxiv.org/abs/1512.03385>. Acesso em: 18 maio. 2024.

HEATON, J. **Artificial Intelligence for Humans: Deep learning and neural networks**. Heaton Research, Incorporated., 2015. (Artificial Intelligence for Humans Series). ISBN 9781505714340. Disponível em: <https://books.google.com.br/books?id=q9mijgEACAAJ>. Acesso em: 29 mai. 2024.

HWANG, S.-W.; SUGIYAMA, J. Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review. **Plant Methods**, v. 17, 2021. Disponível em: https://www.researchgate.net/publication/351137832_Computer_vision-based_wood_identification_and_its_expansion_and_contribution_potentials_in_wood_science_A_review. Acesso em: 02 jun. 2024.

IBRAHIM, I. *et al.* Tree species recognition system based on macroscopic image analysis. **Wood Science and Technology**, v. 51, 2017. Disponível em: <https://link.springer.com/article/10.1007/s00226-016-0859-4#citeas>. Acesso em: 02 jun. 2024.

Jardim Botânico do Rio de Janeiro. **REFLORA - Plantas do Brasil: Resgate Histórico e Herbário Virtual para o Conhecimento e Conservação da Flora Brasileira**. 2024. Disponível em: <http://floradobrasil.jbrj.gov.br/>. Acesso em: 28 mai. 2024.

JOCHER, G.; CHAURASIA, A.; QIU, J. **Ultralytics YOLOv8**. 2024. Disponível em: <https://github.com/ultralytics/ultralytics>. Acesso em: 11 mai. 2024.

JUNIOR, S. L. de A. **IDENTIFICAÇÃO DE PÓLIPOS ATRAVÉS DE TÉCNICAS DE SEGMENTAÇÃO DE IMAGENS**. 2023. 61 p. Monografia — Universidade Tecnológica Federal do Paraná, Medianeira, PR, Brasil, 2023.

KING, R. **Brief summary of YOLOv8 model structure**. 2023. Acesso em: 1 jun. 2024. Disponível em: <https://github.com/ultralytics/ultralytics/issues/189>.

KIRILLOV, A. *et al.* Panoptic segmentation. **ArXiv**, p. 10, 2019. Disponível em: <https://arxiv.org/abs/1801.00868>. Acesso em: 15 jan. 2024.

KOCH, G.; HAAG, V. Control of internationally traded timber - the role of macroscopic and microscopic wood identification against illegal logging. **Journal of Forensic Research**, v. 06, 2015. Disponível em: https://www.researchgate.net/publication/305184738_Control_of_Internationally_Traded_Timber_-_The_Role_of_Macroscopic_and_Microscopic_Wood_Identification_against_Illegal_Logging. Acesso em: 30 mai. 2024.

KIRBAŞ İsmail; ÇIFCI, A. An effective and fast solution for classification of wood species: A deep transfer learning approach. **Ecological Informatics**, v. 69, p. 101633, 2022. ISSN 1574-9541. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574954122000826>. Acesso em: 15 dez. 2023.

LEI, L. *et al.* Deep learning implementation of image segmentation in agricultural applications: a comprehensive review. **Artificial Intelligence Review**, v. 57, n. 6, p. 149, 2024. ISSN 1573-7462. Disponível em: <https://doi.org/10.1007/s10462-024-10775-6>. Acesso em: 25 mai. 2024.

LENZ, I.; LEE, H.; SAXENA, A. Deep learning for detecting robotic grasps. **The International Journal of Robotics Research**, v. 34, p. 705–724, 2015. Disponível em: <https://doi.org/10.1177/0278364914549607>. Acesso em: 04 jun. 2024.

LI, Y.; HU, M.; YANG, X. Polyp-sam: Transfer sam for polyp segmentation. **ArXiv**, p. 11, 2023. Disponível em: <https://arxiv.org/abs/2305.00293>. Acesso em: 10 dez. 2023.

LIU, W.; HASAN, I.; LIAO, S. Center and scale prediction: Anchor-free approach for pedestrian and face detection. **Pattern Recognition**, v. 135, p. 109071, 2023. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320322005519>. Acesso em: 04 jun. 2024.

MAGHANAKI, M. *et al.* Improving early diagnosis: The intersection of lean healthcare and computer vision in cancer detection. **Springer Nature Switzerland**, v. 956, p. 404–413, 2024. ISSN 978-3-031-56950-0. Disponível em: https://link.springer.com/chapter/10.1007/978-3-031-56950-0_34#citeas. Acesso em: 02 jun. 2024.

MATPLOTLIB. **Matplotlib: Visualization with Python**. 2024. Disponível em: <https://matplotlib.org/>. Acesso em: 12 mai. 2024.

MINAEE, S. *et al.* Image segmentation using deep learning: A survey. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, abs/2305.06422, p. 44, 2022. ISSN 1939-3539.

Disponível em: <https://www.computer.org/csdl/journal/tp/2022/07/09356353/1rigXK0s5Ak>. Acesso em: 15 mai. 2024.

Ministério do Meio Ambiente. **Prevenção e controle do desmatamento**. 2024. Disponível em: <https://www.gov.br/mma/pt-br/assuntos/combate-ao-desmatamento>. Acesso em: 29 mai. 2024.

MITCHELL, T. **Machine Learning**. McGraw-Hill, 1997. (McGraw-Hill International Editions). ISBN 9780071154673. Disponível em: <https://books.google.com.br/books?id=EoYBngEACAAJ>. Acesso em: 29 mai. 2024.

NAFEA, A. A. *et al.* A short review on supervised machine learning and deep learning techniques in computer vision. **Babylonian Journal of Machine Learning**, v. 2024, p. 48–55, 2024. ISSN 0031-3203. Disponível em: <https://mesopotamian.press/journals/index.php/BJML/article/view/261>. Acesso em: 04 jun. 2024.

NUMPY. **NumPy: The fundamental package for scientific computing with Python**. 2024. Disponível em: <https://numpy.org/>. Acesso em: 12 mai. 2024.

OLIVEIRA, W. de. **Software para reconhecimento de espécies florestais a partir de imagens digitais de madeiras utilizando deep learning**. 2018. 103 p. Monografia (Mestrado) — Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná, Medianeira, PR, Brasil, 2018. Disponível em: <https://repositorio.utfpr.edu.br/jspui/handle/1/3189>. Acesso em: 21 dez. 2023.

OPENCV. **OpenCV - Open Computer Vision Library**. 2024. Disponível em: <https://opencv.org/>. Acesso em: 12 mai. 2024.

PAULA FILHO, P. L. de. **Reconhecimento de Espécies Florestais Através de Imagens Macroscópicas**. 2012. 105 p. Tese (Doutorado) — Programa de Pós-Graduação em Informática, Universidade Federal do Paraná, Curitiba, PR, Brasil, 2012.

PAULA, J. de; ALVES, J. de H. **Madeiras nativas: anatomia, dendrologia, dendrometria, produção, uso**. Fundação Mokiti Okada, 1997. Disponível em: <https://books.google.com.br/books?id=vpvBXwAACAAJ>. Acesso em: 02 jun. 2024.

POYSER, M.; BRECKON, T. P. Neural architecture search: A contemporary literature review for computer vision applications. **Pattern Recognition**, v. 147, p. 110052, 2024. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320323007495>. Acesso em: 02 jun. 2024.

PYTHON DOCUMENTATION. **Python Tutorial**. 2024. Acesso em: 11 de maio de 2024. Disponível em: <https://docs.python.org/3/tutorial/index.html>.

PYTORCH. **PyTorch: Get Started**. 2024. Disponível em: <https://pytorch.org/>. Acesso em: 11 mai. 2024.

REDMON, J. *et al.* You only look once: Unified, real-time object detection. **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, p. 779–788, 2016. Disponível em: <https://ieeexplore.ieee.org/document/7780460>. Acesso em: 15 jun. 2024.

Rizwan I Haque, I.; NEUBERT, J. Deep learning approaches to biomedical image segmentation. **Informatics in Medicine Unlocked**, v. 18, p. 100297, 2020. ISSN 2352-9148. Disponível em: <https://www.sciencedirect.com/science/article/pii/S235291481930214X>. Acesso em: 10 mai. 2024.

ROBOFLOW. **Roboflow: Computer vision tools for developers and enterprises**. 2024. Disponível em: <https://roboflow.com/>. Acesso em: 12 mai. 2024.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. Pearson, 2020. (Pearson series in artificial intelligence). ISBN 9780134610993. Disponível em: <https://books.google.com.br/books?id=koFptAEACAAJ>. Acesso em: 20 mai. 2024.

SA, R. J. da Silva de *et al.* A importância da biodiversidade amazônica. **Multidisciplinary Reviews**, v. 2, p. 4, 2019. ISSN 2595-3982. Disponível em: <https://malque.pub/ojs/index.php/mr/article/view/107>. Acesso em: 02 jun. 2024.

SANTOS, S. N. C. dos. Análise estrutural e de superfície de madeiras submetidas a tratamento térmico e degradadas pelo fungo *pycnoporus sanguineus*. Escola de Engenharia de São Carlos, mar. 2015. Disponível em: <https://www.teses.usp.br/teses/disponiveis/18/18158/tde-22092015-105627/pt-br.php>. Acesso em: 22 ago. 2023.

SCHEEREN, M. H. **APLICAÇÃO DE TÉCNICAS DE APRENDIZADO PROFUNDO PARA SEGMENTAÇÃO: UM ESTUDO DE CASO EM FOLHAS DE CAFÉ**. 2022. 66 p. Monografia — Universidade Tecnológica Federal do Paraná, Medianeira, PR, Brasil, 2022.

SCHMARJE, L. *et al.* A survey on semi-, self- and unsupervised learning for image classification. **IEEE Access**, v. 09, p. 82146–82168, 2021. ISSN 2169-3536. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9442775>. Acesso em: 05 jun. 2024.

SCHUBERT, M.; PANZARASA, G.; BURGERT, I. Sustainability in wood products: A new perspective for handling natural diversity. **Chemical Reviews**, v. 123, 2022. Disponível em: https://www.researchgate.net/publication/366414988_Sustainability_in_Wood_Products_A_New_Perspective_for_Handling_Natural_Diversity. Acesso em: 29 mai. 2024.

SCIKIT-LEARN. **Scikit-learn: Machine Learning in Python**. 2024. Disponível em: <https://scikit-learn.org/stable/>. Acesso em: 11 mai. 2024.

SHARMA, A.; GARG, S.; SHARMA, V. Atr-ftir spectroscopy and machine learning for sustainable wood sourcing and species identification: Applications to wood forensics. **Microchemical Journal**, v. 200, p. 110467, 2024. ISSN 0026-265X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0026265X24005794>. Acesso em: 26 mai. 2024.

SILVA, G. *et al.* Reconhecimento de espécies florestais utilizando visão computacional. In: **Anais do XII Seminário de Extensão e Inovação & XXVII Seminário de Iniciação Científica e Tecnológica da UTFPR**. Santa Helena (PR): UTFPR Santa Helena, 2022. Disponível em: <https://www.even3.com.br/anais/seisicite2022/549155-reconhecimento-de-especies-florestais-utilizando-visao-computacional>. Acesso em: 24 maio. 2024.

SINGH, A.; SINGH, P. Image classification: A survey. **Journal of Informatics Electrical and Electronics Engineering (JIEEE)**, v. 1, p. 1–9, 2020. Disponível em: <https://jieee.a2zjournals.com/index.php/ieee/article/view/2>. Acesso em: 05 jun. 2024.

SOUZA, D. *et al.* An automatic recognition system of brazilian flora species based on textural features of macroscopic images of wood. **Wood Science and Technology**, v. 54, p. 28, 2020. Disponível em: https://www.researchgate.net/publication/342422935_An_automatic_recognition_system_of_Brazilian_flora_species_based_on_textural_features_of_macroscopic_images_of_wood. Acesso em: 26 mai. 2024.

- SOUZA, V. *et al.* Análise comparativa de redes neurais convolucionais no reconhecimento de cenas. 2020. Disponível em: https://www.researchgate.net/publication/346593579_Analise_Comparativa_de_Redес_Neурais_Convolucionais_no_Reconhecimento_de_Cenas. Acesso em: 10 mai. 2024.
- SUN, Z. *et al.* Human action recognition from various data modalities: A review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 45, p. 3200–3225, 2023. Disponível em: <https://ieeexplore.ieee.org/document/9795869>. Acesso em: 04 jun. 2024.
- TAN, K. *et al.* Integrating advanced computer vision and ai algorithms for autonomous driving systems. **Journal of Theory and Practice of Engineering Science**, v. 4, p. 41–48, 2024. Disponível em: <https://centuryscipub.com/index.php/jtpes/article/view/427>. Acesso em: 02 jun. 2024.
- TERVEN, J.; ESPARZA, D.-M. C.; ROMERO-GONZALEZ, J.-A. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. **Machine Learning and Knowledge Extraction**, MDPI AG, v. 5, n. 4, p. 1680–1716, nov 2023. ISSN 2504-4990. Disponível em: <http://dx.doi.org/10.3390/make5040083>. Acesso em: 03 jun. 2024.
- VASILEV, I. *et al.* Python deep learning: Exploring deep learning techniques and neural network architectures with pytorch, keras, and tensorflow. *In: _____*. 2. ed. Birmingham: Packt Publishing, 2019. p. 386. ISBN 978-1-78934-846-0. Disponível em: <https://books.google.com.br/books?id=0LWVzgEACAAJ>. Acesso em: 29 mai. 2024.
- VIEIRA, G. de S. *et al.* Identification of wood from the Amazon by characteristics of Haralick and Neural Network: image segmentation and polishing of the surface. **iForest - Biogeosciences and Forestry**, n. 4, p. 234–239, 2022. Disponível em: <https://iforest.sisef.org/contents/?id=ifor3906-015>. Acesso em: 09 dez. 2023.
- WANG, M. *et al.* Tsw-yolo-v8n: Optimization of detection algorithms for surface defects on sawn timber. **North Carolina State University**, p. 27, 2023. ISSN 19302126. Disponível em: <https://www.proquest.com/docview/2893447982?fromopenview=true&pq-origsite=gscholar&sourcetype=Scholarly%20Journals>. Acesso em: 21 jan. 2024.
- WANG, Y.; ZHAO, Y.; PETZOLD, L. An empirical study on the robustness of the segment anything model (sam). **ArXiv**, abs/2305.06422, p. 27, 2023. Disponível em: <https://arxiv.org/abs/2305.06422>. Acesso em: 15 jan. 2024.
- YU, C. *et al.* Bisenet: Bilateral segmentation network for real-time semantic segmentation. *In: FERRARI, V. et al. (Ed.). Computer Vision – ECCV 2018*. Cham: Springer International Publishing, 2018. p. 334–349. ISBN 978-3-030-01261-8. Disponível em: https://www.researchgate.net/publication/373355559_A_Deep_Learning-Based_Segmentation_Strategy_for_Diabetic_Foot_Ulcers_Combining_the_Strengths_of_HarDNet-MSEG_and_SAM_Models. Acesso em: 09 fev. 2024.
- ZHANG, J. *et al.* Research on the application of computer vision based on deep learning in autonomous driving technology. **ArXiv**, p. 5, 2024. ISSN 2406.00490. Disponível em: <https://arxiv.org/abs/2406.00490v1>. Acesso em: 02 maio. 2024.