

RECONHECIMENTO DE ESPÉCIES ARBÓREAS COM BASE EM IMAGENS DE FOLHAS

Marcos Vinícius Ribeiro Da Silva¹, Pedro Luiz De Paula Filho², Mariana Cabral Hetka Bczuska²

¹Universidade Tecnológica Federal do Paraná (UTFPR), Medianeira - PR, Brasil,
marcosviniciussilva@alunos.utfpr.edu.br.

²Universidade Tecnológica Federal do Paraná (UTFPR), Medianeira - PR, Brasil

Resumo: Uma possível alternativa para o reconhecimento de espécies arbóreas, que geralmente é feita via inspeção visual, é o uso de um sistema automático baseado em redes neurais. O objetivo deste trabalho foi treinar e avaliar uma arquitetura de rede neural na tarefa de reconhecer espécies nativas do Oeste do Paraná e exóticas do Campus Medianeira da UTFPR usando imagens de folhas. A arquitetura escolhida foi a YOLO12, atingindo 76,06% de acurácia no subconjunto de teste usado após o treino.

Palavras-chave: classificação; rede neural; YOLO12.

INTRODUÇÃO

As florestas são essenciais para a manutenção da biodiversidade, da ciclagem de água, do estoque de carbono, entre vários outros benefícios ambientais (Fearnside, 2004). Além de medidas como o combate ao desmatamento, também é necessário um manejo florestal para garantir esta manutenção e, para que tal manejo seja sustentável, é fundamental a identificação das espécies, tendo como intuito certificar que os vários exemplares da biodiversidade de determinada região possam ser preservados de modo apropriado.

Como Raven et al (2014) descrevem, este reconhecimento deve-se iniciar com a observação de estruturas vegetativas presentes nos espécimes, como folhas e cascas, o que permite distinção entre grupos taxonômicos através das variações morfológicas presentes nestas estruturas.

Normalmente, essa identificação é realizada por especialistas que demoraram um tempo considerável até que conseguissem um grau de competência adequado e, por conta dessa dificuldade de formação, estão em falta (Paula Filho, 2012). Como Paula Filho et al. (2014) afirmaram, essa é a razão pela qual a indústria e as agências de supervisão também optam por uma alternativa ao trabalho exercido por especialistas.

Tendo esse cenário em mente, a utilização de um sistema de reconhecimento automático com base em redes neurais é uma das possíveis alternativas, como, por exemplo, as abordagens que utilizam características como folhas para a tarefa do reconhecimento das espécies.

Já existem diversos outros trabalhos nessa área que tratam de diversas abordagens diferentes para contextos diferentes, como o de Gama et al. (2025) que é voltado especificamente para o reconhecimento de árvores comerciais presentes no Bioma Amazônico.

O objetivo deste trabalho foi treinar modelos de redes neurais para a tarefa de reconhecer espécies nativas do Bioma da Mata Atlântica, além de algumas espécies exóticas, que podem ser encontradas no Campus Medianeira da Universidade Tecnológica Federal do Paraná (UTFPR), usando como base imagens digitais de folhas.

MATERIAL E MÉTODOS

Para a realização deste presente trabalho, o conjunto de dados utilizado no treinamento, na validação e na avaliação dos modelos foi fornecido Grupo de Inteligência Computacional (GIC) da UTFPR, sendo composto por 704 imagens de 15 espécies diferentes, divididos em subconjuntos de treinamento, validação e teste, seguindo uma proporção de 70/20/10.

As imagens digitais são fotografias das folhas das árvores, que correspondem aos órgãos laterais encontrados no caule principal e nos ramos das plantas (Marchiori, 2004; Lemos e Andrade, 2022).

É importante notar que existe um desbalanceamento no número de imagens de cada espécie presente no conjunto, pois, por exemplo, há 83 e 79 imagens de Araçá (*Psidium cattleianum*) e de Brinco-de-índio (*Cojoba arborea*) respectivamente, enquanto quase todas as demais possuem por volta de 30 imagens.

As 15 espécies arbóreas presentes no conjunto de dados e a quantidade de imagens de cada uma estão demonstradas na Tabela 1.

Tabela 1. Espécies arbóreas do conjunto de dados.

Nome Popular	Nome Científico	Quantidade
Abacateiro	<i>Persea americana</i>	42
Araçá	<i>Psidium cattleianum</i>	83
Brinco-de-índio	<i>Cojoba arborea</i>	79
Cajueiro	<i>Anacardium occidentale</i>	76
Carvalho	<i>Quercus robur</i>	30
Cateretê	<i>Machaerium paraguariense</i>	30
Cerejeira	<i>Prunus serrulata</i>	30
Coité	<i>Crescentia cujete</i>	33
Fruta-do-Conde	<i>Annona squamosa</i>	54
Grevilha	<i>Grevillea robusta</i>	47
Jambolão	<i>Syzygium cumini</i>	30
Laranja Champanhe	<i>Citrus sinensis</i>	31
Louro-pardo	<i>Cordia trichotoma</i>	31
Pau Brasil	<i>Paubrasilia echinata</i>	78
Peroba-rosa	<i>Aspidosperma polyneuron</i>	30

Além do problema do desbalanceamento, também é relevante que as imagens não são das folhas isoladas em um fundo simples, mas sim possuem fundos variados, como algumas que estão em cima de calçadas, de gramado ou até mesmo nos galhos. Essa diferença dos fundos das imagens pode ser vista na Figura 1.

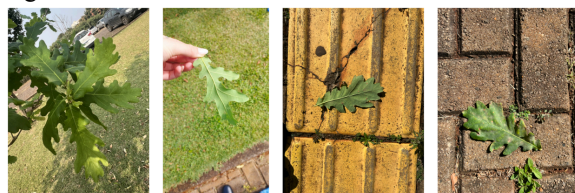


Figura 1. Instâncias de folhas de carvalho com fundos diferentes.

A arquitetura de rede neural pré-treinada escolhida foi a YOLO12, lançada em 18 de fevereiro de 2025, sendo parte da família YOLO (*You Only Look Once*), que trata-se de uma família de sistemas de detecção de objetos cujo primeiro modelo foi originalmente proposto por Redmon et al. (2016).

O treinamento envolveu a técnica de *Grid Search*, na qual todas as combinações possíveis de um conjunto de alguns hiperparâmetros foram testadas até que se encontre a melhor combinação de acordo com uma função de avaliação. Alguns dos hiperparâmetros que foram testados são a Taxa de Aprendizagem (ou chamada de *Learning Rate*), que define o tamanho dos ajustes realizados nos pesos internos do modelo durante cada etapa do treinamento, e a Época, que é responsável por definir o número de iterações em que o modelo verá o conjunto de treinamento e consequentemente irá ter os seus pesos ajustados (Ultralytics, 2024).

Adicionalmente, foi aplicada a técnica de *Transfer Learning*, no qual apenas as camadas finais dos modelos pré-treinados foram treinadas novamente, pois as camadas iniciais desses modelos já aprenderam a reconhecer características universais, como bordas, texturas e formas (Ultralytics, 2024). Por fim, ainda no treinamento, foi também aplicada a técnica de *Early Stopping*. Esta técnica consiste em parar precocemente o treinamento de um modelo assim que a performance no conjunto de validação para de crescer, de modo a prevenir que o modelo apenas memorize as imagens do treinamento e se torne incapaz de reconhecer imagens novas (Ultralytics, 2024).

Após cada treinamento de um modelo, seu desempenho com a melhor combinação de hiperparâmetros do conjunto atual foi avaliado utilizando métricas, neste caso, a Acurácia e o *F1-Score*, com o intuito de quantificar a performance do modelo de classificação (Grandini; Bagli; Visani, 2020).

Caso um modelo fosse julgado inadequado, ou seja, se não fosse capaz de diferenciar as espécies na grande maioria dos casos, o treinamento seria refeito com um novo conjunto modificado de hiperparâmetros, até que obtivesse um desempenho adequado.

RESULTADOS E DISCUSSÃO

Os hiperparâmetros, testados por meio da técnica de *Grid Search*, foram as épocas (com os valores 10, 20 e 30) e a Taxa de Aprendizagem (com os valores 0,0005; 0,0001; 0,00005 e 0,00001). Os valores da Acurácia e do *F1-Score* obtidos pelo modelo YOLO12 no subconjunto de teste após o treinamento foram, respectivamente, 0,7606 e 0,7198.

A Figura 2 demonstra o desempenho do modelo treinado na classificação do subconjunto de teste, para melhor visualização do desempenho do modelo. Isto é exibido por meio de uma matriz de confusão, que demonstra as espécies que as imagens do subconjunto pertencem na realidade e as espécies que o modelo previu para as imagens.

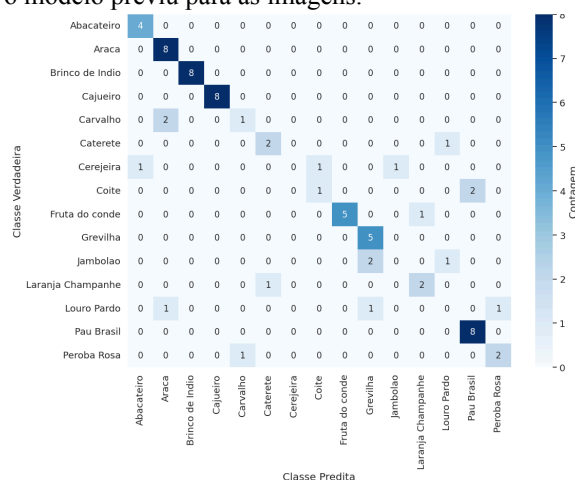


Figura 2. Matriz de Confusão demonstrando a performance de classificação do modelo no subconjunto de treinamento.

A combinação de hiperparâmetros que apresentou o melhor desempenho para este modelo, com este conjunto de dados, foi a Taxa de Aprendizado de valor 0,0001 com 20 épocas. Como visto na Figura 1, as espécies que apresentaram a pior acurácia foi a Cerejeira (*Prunus serrulata*), o Jambolão (*Syzygium cumini*) e o Louro-pardo (*Cordia trichotoma*), pois o modelo previu todas as instâncias dessas espécies erroneamente.

Além disso, as espécies Araça (*Psidium cattleianum*) e Grevilha (*Grevillea robusta*) foram as que mais causaram confusão no modelo, já que foram as espécies que mais tiveram instâncias de outras espécies classificadas como pertencentes a elas.

A Figura 2 demonstra uma instância de Araça no conjunto de dados comparada com instâncias das outras duas espécies que causaram confusão no modelo.

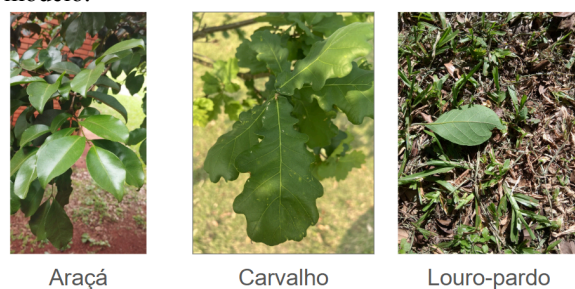


Figura 2. Comparação de uma instância de Araça com instâncias de outras espécies.

CONCLUSÃO

A arquitetura YOLO12, que não é originalmente um classificador, apresentou valores razoavelmente bons de acurácia e *F1-score* no conjunto de testes, apesar do tamanho pequeno do conjunto de dados comparado aos que geralmente são utilizados em trabalhos similares, no entanto, as espécies que apresentaram o pior desempenho foram a Cerejeira, o Jambolão e o Louro-pardo.

A principal razão desse baixo desempenho nessas espécies se deve que essas espécies são as que possuem o menor número de instâncias no conjunto de dados. Além disso, o baixo desempenho para as espécies Araça e Grevilha também demonstra um problema que vários sistemas de reconhecimento automático similares também possuem, sendo este o fato que as folhas destas espécies apresentam semelhanças visuais consideráveis com outras espécies do conjunto de dados, como a forma de sua lâmina, o que causa confusão na rede neural. Por fim, é importante ressaltar que é possível que parte dos erros do modelo sejam causados pelo mesmo ter aprendido a utilizar os fundos das imagens para reconhecer as espécies invés das folhas em si, como visto na Figura 2, no qual a folha de Araça não é similar a uma folha de Carvalho.

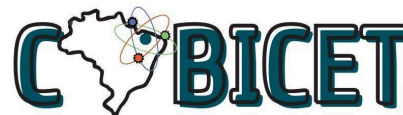
Em trabalhos futuros, o conjunto de dados utilizado deverá ser expandido com mais imagens de modo a corrigir os problemas de desbalanceamento, além de testar outras arquiteturas de rede neural para servirem de comparativo e treinar um modelo de segmentação para separar os fundos dos objetos de interesse antes da classificação.

AGRADECIMENTOS

Agradecemos ao CNPq pela bolsa de Produtividade em Desen. Tec. e Extensão Inovadora (DT-II - Processo: 303279/2023-0).

REFERÊNCIAS

- FEARNSIDE, P. A. A água de São Paulo e a Floresta Amazônica. *Ciência Hoje*, v. 34, p. 63–65, 2004. Disponível em: https://www.researchgate.net/publication/287878871_A_agua_de_Sao_Paulo_e_a_floresta_amazonica. Acesso em: 16 set. 2025.
- GAMA, N. C. et al. Automatic recognition of commercial tree species from the Amazon flora using bark images and transfer learning. *Forests*, v. 16, n. 9, p. 1374, 2025. Disponível em: <https://www.mdpi.com/1999-4907/16/9/1374>. Acesso em: 4 set. 2025.
- GRANDINI, M.; BAGLI, E.; VISANI, G. Metrics for multi-class classification: an overview. *arXiv*, 2020. Disponível em:



<https://arxiv.org/abs/2008.05756>. Acesso em: 2 out. 2025.

LEMOS, J.; ANDRADE, I. Glossário ilustrado de Botânica. Oficina de Textos, [S.l.], 2022.

MARCHIORI, J. N. C. Elementos de Dendrologia. 2. ed. Editora UFSM, Santa Maria, 2004.

PAULA FILHO, P. L. Sistema de reconhecimento de espécies florestais por meio de imagens digitais de madeira utilizando aprendizado de máquina. Dissertação (Mestrado em Ciências Florestais) — Universidade Federal do Paraná, Curitiba, 2012.

PAULA FILHO, P. L. et al. Forest species recognition using macroscopic images. Machine Vision and Applications, v. 25, 2014. Disponível em: <https://link.springer.com/article/10.1007/s00138-014-0592-7>. Acesso em: 13 set. 2025.

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. Biologia Vegetal. Guanabara Koogan, Rio de Janeiro, 2014.

REDMON, J. et al. You only look once: unified, real-time object detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, [S.l.], 2016. p. 779–788. Disponível em: <https://doi.org/10.1109/CVPR.2016.91>. Acesso em: 3 out. 2025.

ULTRALYTICS. Ultralytics glossary. 2024. Disponível em: <https://www.ultralytics.com/glossary>. Acesso em: 18 out. 2025.