



IDENTIFICAÇÃO DE SEIS ESPÉCIES DE PALMEIRAS NATIVAS NA AMAZÔNIA UTILIZANDO REDES NEURAIAS CONVOLUCIONAIS E IMAGENS DE VANT

Airton Gaio Júnior¹; Rodrigo Pinheiro Ribas²

¹ Universidade Estadual de Santa Catarina - UDESC, Centro de Ciências Humanas e da Educação - FAED, Florianópolis, Brasil. airton.gaio@edu.udesc.br, <https://orcid.org/0000-0002-4694-5769>

² Universidade Estadual de Santa Catarina - UDESC, Centro de Ciências Humanas e da Educação - FAED, Florianópolis, Brasil. rodrigo.ribas@udesc.br, <https://orcid.org/0000-0002-0054-3123>

RESUMO. As florestas tropicais, especialmente a Amazônia, são vitais para a biodiversidade e o equilíbrio climático global. As palmeiras, abundantes nesse ecossistema, possuem grande importância ecológica e socioeconômica, sendo utilizadas para alimentação, medicina, construção e artesanato. No entanto, a identificação e mapeamento dessas espécies ainda é um desafio, devido aos métodos tradicionais de inventário florestal que são dispendiosos e imprecisos. Este estudo propõe o uso de veículos aéreos não tripulados (VANTs) e redes neurais convolucionais (CNNs) para a identificação de seis espécies de palmeiras nativas da Amazônia (açai, cocão, jaci, patauá, paxiúba e tucumã) a partir de imagens de alta resolução adquiridas por VANTs. Os resultados indicam que o modelo alcançou alta precisão para algumas classes, como "Açaí" (91,16%), "Jací" (89,56%) e "Floresta" (95,32%). Porém, houve dificuldades significativas com outras classes, especialmente "Patauá", a qual não teve previsões corretas, e "Tucumã" (22,6%). A hipótese de que uma rede CNN pode identificar padrões de diferentes palmeiras foi confirmada, mas melhorias significativas são necessárias para algumas classes. Esses resultados têm potencial para transformar o manejo florestal, proporcionando métodos mais eficientes e precisos para a conservação e uso sustentável das florestas tropicais. O uso de CNNs demonstrou ser uma solução promissora para a classificação de espécies vegetais, contribuindo para o manejo e monitoramento florestal na Amazônia, em comparação com as abordagens tradicionais.

Palavras-chave: Identificação de Espécies; Veículos Aéreos Não Tripulados (VANT); Palmeiras; Redes Neurais Convolucionais (CNN); Floresta Tropical.

ABSTRACT. Tropical forests, especially the Amazon, are vital for biodiversity and global climate balance. Palms, abundant in this ecosystem, have significant ecological and socioeconomic importance, being used for food, medicine, construction, and handicrafts. However, the identification and mapping of these species remain challenging due to the traditional forest inventory methods, which are costly and inaccurate. This study proposes the use of unmanned aerial vehicles (UAVs) and convolutional neural networks (CNNs) to identify six native palm species of the Amazon (açai, cocão, jaci, patauá, paxiúba, and tucumã) from high-resolution images acquired by UAVs. The results indicate that the model achieved high accuracy for some classes, such as "Açaí" (91.16%), "Jací" (89.56%), and "Forest" (95.32%). However, there were significant difficulties with other classes, especially "Patauá," which had no correct predictions, and "Tucumã" (22.6%). The hypothesis that a CNN can identify patterns of different palm species was confirmed, but significant improvements are necessary for some classes. These results have the potential to transform forest management, providing more efficient and accurate methods for the conservation and sustainable use of tropical forests. The use of CNNs has proven to be a promising solution for plant species classification, contributing to forest management and monitoring in the Amazon, compared to traditional approaches.

Keywords: Species Identification; Unmanned Aerial Vehicles (UAV); Palms; Convolutional Neural Networks (CNN); Tropical Forest.



1. INTRODUÇÃO

As florestas tropicais desempenham um papel vital na manutenção da biodiversidade global e no equilíbrio climático. A Amazônia, em particular, abriga uma enorme diversidade de espécies, sendo considerada um dos ecossistemas mais ricos do planeta. As palmeiras, da família Arecaceae, são essenciais não apenas pela sua importância ecológica, mas também pelo seu valor econômico e cultural para as comunidades locais (Ferreira, 2006). Espécies como o açaí (*Euterpe oleracea*), o cocão (*Attalea speciosa*), o jaci (*Attalea maripa*), a patauí (*Oenocarpus bataua*), a paxiuba (*Socratea exorrhiza*) e o tucumã (*Astrocaryum aculeatum*) são amplamente exploradas para alimentação, medicina, construção e artesanato, entre outros usos (Lorenzi; Noblick, 2010). Apesar de sua importância, a identificação e o mapeamento dessas espécies ainda enfrentam desafios significativos. Métodos tradicionais de inventário florestal que envolvem trabalho de campo intensivo são caros, demorados e frequentemente imprecisos devido à dificuldade de acesso e à alta diversidade de espécies nas florestas tropicais. Nesse contexto, a aplicação de novas tecnologias, tais quais o imageamento por meio de veículos aéreos não tripulados (VANTs) e a classificação de imagens por meio de Redes Neurais Convolucionais (CNNs), apresentam-se como soluções promissoras. Os VANTs, aliados às CNNs, permitem a captura de imagens de alta resolução e a análise automatizada de grandes volumes de dados, oferecendo uma alternativa eficaz e de baixo custo para a identificação de espécies. Estudos recentes demonstraram o potencial das CNNs na classificação de espécies vegetais a partir de imagens aéreas. Por exemplo, Li *et al.* (2019) desenvolveram um método de rede neural convolucional para detectar palmeiras de óleo na Malásia com 96,5% de precisão. Culman; Delalieux e Tricht (2020) usaram aprendizado profundo para detectar palmeiras individuais na Espanha, atingindo 86% de precisão. Gibril *et al.* (2021) aplicaram uma arquitetura U-Net para mapear palmeiras datileiras, obtendo a métrica IoU de 0,93. Htet e Sein (2021) identificaram palmeiras de Toddy em Mianmar com o modelo Mask R-CNN, alcançando alta precisão. Wahed *et al.* (2022) desenvolveram o modelo CraunNet para detectar a maturidade das palmeiras de sagu com 85,7% de precisão. Letsoin *et al.* (2022) usaram técnicas de transferência de conhecimento para reconhecer palmeiras de sagu, com a ResNet-50 atingindo 98,5% de precisão. Marin; Mondragon e Colorado (2022) identificaram palmeiras Moriche na Amazônia com 98% de precisão usando Mask R-CNN. Gibril *et al.* (2023) utilizaram transformadores de visão profunda para mapear palmeiras-datas, mostrando alta precisão e capacidade de generalização. Aliado aos estudos, o uso de dados adquiridos por VANTs, combinado com aprendizado profundo, revela-se altamente eficaz para o mapeamento e monitoramento de espécies em florestas tropicais (Anderson e Gaston, 2013). Diferenciando-se de outros estudos que classificam apenas uma espécie de palmeira, a hipótese deste é que uma Rede Neural Convolucional (CNN) pode alcançar altas taxas de precisão na classificação de múltiplas espécies de palmeiras em um único modelo. Portanto, este estudo propõe o uso de uma CNN para identificar até seis espécies de palmeiras nativas - açaí, cocão, jaci, patauí, paxiuba e tucumã - em ambientes complexos, como as áreas de floresta tropical amazônica. Utilizando imagens de alta resolução adquiridas por VANTs, o objetivo é classificar essas espécies de palmeiras, contribuindo para o manejo e monitoramento florestal na Amazônia. A identificação precisa dessas espécies não só facilita a conservação e o uso sustentável dos recursos florestais, mas também fortalece a preservação da biodiversidade e apoia as comunidades locais que dependem dessas palmeiras para diversas finalidades.

2. MATERIAL E MÉTODOS

A pesquisa foi realizada em uma área de floresta tropical no Projeto de Assentamento Dirigido Santa Luzia (PAD), em Cruzeiro do Sul, Acre, Brasil. A área de estudo, com cerca de 100 hectares, é caracterizada por uma alta incidência de palmeiras (Figura 1). A região possui clima equatorial chuvoso, com precipitação anual média de 1.950 mm e temperatura média de 24,8°C (INMET, 2020).



Figura 1. Mapa localização da pesquisado.

Fonte: Elaborado pelo autor com base nos dados (ACRE, 2006).

2.1. Esquema metodológico

Esta seção descreve o método utilizado, que começa com a aquisição de imagens por um VANT, resultando em uma série de imagens individuais. Após isso, a calibração e a geração de mosaicos são feitas automaticamente. Como a rede neural aceita apenas matrizes de 512x512 pixels com 3 dimensões (RGB), a banda Alpha, que representa a altimetria, é excluída, mantendo apenas as bandas RGB. Em seguida, as imagens são recortadas conforme o limite da propriedade e exportadas no formato GeoTiff (uint8). Nesta etapa, as regiões de interesse das imagens são selecionadas, removendo informações desnecessárias para reduzir a complexidade dos dados. Após o recorte, aplica-se o algoritmo *Mean Shift* para segmentação, gerando segmentos em formato vetorial, que são rotulados de acordo com as classes definidas. A rotulação é realizada com base em uma inspeção visual em escala 1:250. Os segmentos rotulados são então convertidos para o formato raster (byte), tornando-os compatíveis com algoritmos de *machine learning*. Por fim, a imagem e os rótulos são recortados em blocos uniformes de 512x512 pixels. O recorte uniforme das imagens garante consistência no tamanho e resolução dos dados de entrada para o treinamento da rede neural. Para lidar com o desequilíbrio de classes, foi utilizada a técnica de sobreamostragem, gerando novos exemplos artificiais das classes minoritárias. As imagens originais foram espelhadas horizontal e verticalmente (*Flip e Flop*), e combinadas (*Flip/Flop*), triplicando a área de pesquisa e o número de amostras. Com os dados preparados, iniciou-se o processo de treinamento, teste e validação com redes neurais convolucionais, eficazes no reconhecimento de padrões e classificação de imagens. Os resultados do modelo treinado são avaliados com métricas apropriadas para medir sua precisão e eficácia, além de identificar áreas de melhoria. Se os resultados não forem

satisfatórios, ajustes nos parâmetros da rede são feitos, e o processo de treino e teste é repetido para otimizar o desempenho.

2.2. Imagens de VANT

A expansão dos VANTs para aplicações civis, como mapeamento e agricultura de precisão, começou no início dos anos 2000. Desde então, avanços em autonomia, sensores e carga útil tornaram esses equipamentos mais acessíveis e versáteis. Anderson e Gaston (2013) destacam que VANTs leves revolucionaram a pesquisa ecológica, fornecendo imagens de alta resolução espacial, especialmente para o mapeamento da copa das árvores. No entanto, desafios como a padronização de protocolos de aquisição e o desenvolvimento de algoritmos mais precisos ainda precisam ser superados. De modo geral, os VANTs têm trazido avanços importantes em pesquisa ecológica, agricultura e monitoramento de recursos naturais. As imagens deste estudo foram capturadas em dezembro de 2020, utilizando um VANT Phantom 4 Professional equipado com uma câmera RGB de 20 megapixels. O voo foi realizado a 120 metros de altura, com uma velocidade constante de 13 m/s, cobrindo uma área significativa. As capturas ocorreram entre 09h00 e 12h00 (UTC -05:00), em condições ideais de iluminação. Foram obtidas 672 imagens com resolução de 4,3 cm/pixel (Figura 2). As imagens foram calibradas e um ortomosaico foi gerado com o software Pix4Dmapper®, no sistema de coordenadas WGS 84/UTM zona 18S.

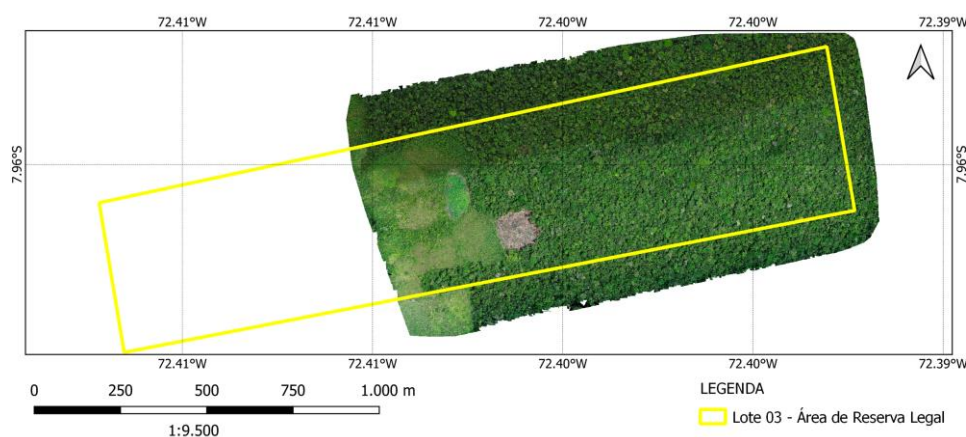


Figura 2. Mosaico de imagens aéreas obtidas através do sobrevoio do VANT.

Fonte: Elaborado pelo autor com base nos dados (ACRE, 2006).

2.3. Espécies de palmeiras

Os dados de inventário das espécies fornecido pela Fundação de Tecnologia do Acre (FUNTAC) revelaram uma grande diversidade de palmeiras na área estudada, com destaque para 6 (seis) espécies dominantes de grande importância socioeconômica para a população local (Lorenzi; Noblick, 2010): - **Açaí** (*Euterpe precatoria*): Fornece palmito comestível e o fruto é usado para preparar o "vinho de açaí". Os caules são utilizados na construção de casas e as raízes têm aplicação medicinal; - **Jaci** (*Attalea butyraceae*): Produz óleo comestível a partir do endosperma e as folhas são usadas para cobertura de casas; - **Patauá** (*Oenocarpus bataua*): Os frutos são utilizados para preparar suco e extrair óleo semelhante ao de oliva. As folhas são empregadas no artesanato trançado; - **Cocão** (*Attalea tessmannii*): As folhas são usadas para cobertura de casas rústicas e as amêndoas são comestíveis. Os caroços são queimados para esfumaçar látex da borracha; - **Paxiúba** (*Socratea exorrhiza*): Os caules rachados são usados

em construções como assoalhos e paredes. A planta também possui potencial ornamental para paisagismo; e - **Tucumã** (*Astrocaryum aculeatum*): O mesocarpo é comestível e os frutos são vendidos nos mercados. As fibras são usadas na confecção de redes.

A Tabela 1 fornece um panorama das características morfológicas dessas espécies (Lorenzi; Noblick, 2010), enriquecendo nossa compreensão sobre cada uma delas.










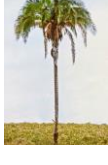


Espécie	Características	Imagem VANT	Vista do Solo
Açaí	Possui caule simples e ereto de 3 a 20 m, com raízes visíveis na base e palmito liso no topo. Folhas pinadas e planas, de 10 a 20, com bainha fechada de 0,7 a 1,6 m, formando um tubo verde ou verde com listas amarelas. Frutos globosos de 1,0 a 1,3 cm, de cor púrpura-negra.		
Patauí	Possui caule solitário de 5 a 25 m de altura e 20 a 45 cm de diâmetro, com raízes fasciculadas visíveis na base e sem palmito no topo. Tem 10 a 20 folhas eretas e divergentes. Produz frutos oblongos de 2,7 a 4,5 cm, de cor roxo-escuro e endosperma ruminado.		
Cocão	Possui caule solitário de 12 a 24 m de altura e 25 a 46 cm de diâmetro, coberto por restos de folhas. As folhas pinadas e eretas (10 a 12) atingem até 10 m de comprimento. Seus frutos oblongos medem 11 a 15 cm e contêm 1 a 3 sementes.		
Paxiuba	Possui caule solitário, ereto e liso, de 10 a 20 m de altura e 10 a 18 cm de diâmetro, com palmito liso no topo, sustentado por cerca de 25 raízes aéreas espinescentes na base. Suas folhas são pinadas e plumosas, em pequeno número (cerca de 7).		
Jaci	Possui caule solitário de 8 a 18 m de altura e 45 a 90 cm de diâmetro, com folhas pinadas e arqueadas, raque de 5 a 8 m e 180 a 230 pinas. Os frutos medem de 5 a 8 cm, com mesocarpo doce e endocarpo duro contendo 1 a 2 sementes.		
Tucumã	Possui caule simples de até 25 m de altura e 40 cm de diâmetro, com folhas pinadas e ascendentes (5 a 15) e espinhos de até 10 cm. As pinas lineares (100 a 120) por lado estão em grupos de 2 a 5, com pinas medianas medindo 1 m de comprimento e 4 a 6 cm de largura.		

Tabela 1 - Descrição das espécies de palmeiras nativas presentes na região da pesquisa.

2.4. Preparação dos dados e rotulação das imagens

Para aumentar o conjunto de dados de imagens, utilizou-se uma técnica comum de espelhamento horizontal e vertical (*Flip* e *Flop*), além de combinações (*Flip/Flop*), triplicando a área de pesquisa e o número de amostras. O processo foi automatizado com o software ImageMagick® para imagens e rótulos. Na fase inicial de rotulação das imagens, o algoritmo de *Mean Shift Clustering* foi adotado para criar segmentos homogêneos que serão posteriormente etiquetados com as categorias correspondentes às espécies de palmeiras. Esse



algoritmo, fundamentado na densidade, identifica regiões com alta concentração de pontos em um espaço de características. Cada ponto é iterativamente deslocado em direção à média dos pontos vizinhos até que sua posição se estabilize, formando clusters (Cheng, 1995 & Comaniciu; Meer, 2002). Uma vantagem desse método é que os resultados são gerados em formato vetorial. A área de estudo abrange um total de 1.230.190 segmentos criados pelo algoritmo. A atribuição das categorias de espécies de palmeiras aos segmentos foi feita por meio de uma análise visual conduzida por um especialista, utilizando uma paleta de cores naturais no ortomosaico em escala 1:250, com suporte do Instituto de Pesquisas da Amazônia (INPA). A plataforma QGIS® foi usada para classificar os segmentos, atribuindo os seguintes valores: 1 para Açaí, 2 para Cocão, 3 para Jaci, 4 para Patauí, 5 para Paxiúba, 6 para Tucumã, e 9 para Floresta. Ao final, a distribuição ficou assim: Floresta com 1.125.328 segmentos, Açaí com 13.556, Cocão com 1.901, Jaci com 70.983, Patauí com 1.088, Paxiúba com 12.554 e Tucumã com 4.780 segmentos. A distribuição das classes na área de estudo revela que a classe "Floresta" é a mais dominante, com 3.987.159,41 pixels, representando 93,68% do total. "Jaci" ocupa 4,25% com 18.094.788 pixels. "Açaí" e "Paxiúba" possuem valores semelhantes, com 0,81% (3.437.494 pixels) e 0,80% (3.392.452 pixels), respectivamente. "Tucumã" tem 0,30% (1.279.702 pixels), "Cocão" representa 0,11% (457.790 pixels), e "Patauí" é a menos representada, com 0,06% (248.975 pixels). Essas porcentagens destacam a predominância de certas espécies na região. Desconsiderando a classe "Floresta", a classe "Jaci" é a mais prevalente entre as palmeiras, com 67,24% dos rótulos, destacando sua dominância na área de estudo. A classe "Patauí" é a menos representada, com apenas 0,93%, sugerindo raridade ou dificuldades na identificação. "Açaí" é a segunda maior classe, com 12,77%, seguida por "Paxiúba" com 12,61%. "Tucumã" tem 4,76%, e "Cocão" apresenta a menor prevalência, com 1,70%. Não foram atribuídos rótulos aos pixels sem imagem ("*no-data*"), tarefa realizada com segurança na inspeção visual, graças à alta resolução das imagens (4,3 cm/pixel).

2.5. Arquitetura da CNN

O número de camadas, ou a profundidade da rede, é diretamente relacionado à capacidade de aprender características dos dados, o que ajuda a reduzir erros de classificação (Shimodaira, 2000 & Simonyan e Zisserman, 2014). Com o aumento da profundidade, mais camadas foram adicionadas para melhorar os resultados. No entanto, alguns estudos demonstraram um aumento nos erros de treinamento com redes mais profundas (Simonyan e Zisserman, 2014). Para mitigar esse problema, foi proposta a Rede Residual (ResNet) por He et al. (2016). A ResNet é composta por blocos residuais que realizam conexões de salto, permitindo que as ativações de uma camada sejam encaminhadas para uma camada mais profunda. As variações comuns da ResNet incluem ResNet-18, ResNet-50 e ResNet-101, que diferem no número de camadas residuais. Neste estudo, a arquitetura ResNet-18 foi escolhida, baseada no trabalho de Júnior e Ribas (2023), devido ao equilíbrio entre tempo de processamento e precisão observado em testes preliminares. Conforme He et al. (2016), a ResNet-18 oferece desempenho comparável ou superior a outras arquiteturas profundas, mesmo com menos dados de treinamento, graças ao módulo residual que permite um aprendizado mais profundo e robusto, e à normalização, que ajuda a evitar o problema de gradientes desaparecendo, acelerando o treinamento. A ResNet-18 é composta por 18 camadas, divididas em dois tipos: camadas residuais, formadas por duas camadas convolucionais seguidas de uma soma com a entrada, e camadas de transformação, compostas por uma camada convolucional e normalização. Essa arquitetura foi incorporada ao modelo DeepLabv3+, reconhecido por sua eficiência em segmentação semântica de imagens. O modelo foi treinado e testado utilizando um total de 3.560 imagens



que foram rotacionadas e transformadas aleatoriamente para evitar *overfitting*, garantindo que o modelo pudesse generalizar melhor em novos dados.

2.6. Configuração, Treino e validação da CNN

Treinamos o modelo com sete classes: Açai, Cocão, Jaci, Patauá, Paxiúba, Tucumã e Floresta, sendo a inclusão da classe "Floresta" crucial para que a rede aprendesse a diferenciar palmeiras de outras árvores. Os pesos iniciais foram calculados a partir dos dados de treinamento, e durante o processo, as imagens de entrada foram rotacionadas e transformadas aleatoriamente ao longo dos eixos. A taxa de aprendizado inicial foi 0,05, sendo reduzida pela metade a cada duas épocas completas. Os dados foram particionados em 80% (2.848 arquivos de 512×512 pixels) para treinamento e 20% (712 arquivos) para validação e teste, selecionados aleatoriamente. O modelo foi treinado em mini-batches de 16, por até 10 épocas, usando o algoritmo *Stochastic Gradient Descent with Momentum* (SGDM). A cada época, foram extraídos 512 patches aleatórios, e a validação foi realizada 50 vezes durante o treinamento para avaliar a capacidade de generalização do modelo. O treinamento foi executado em um laptop com CPU Intel i7-12700H, 32 GB de RAM e GPU NVIDIA RTX 3070Ti, utilizando Malab versão 2023b.

3. RESULTADOS E DISCUSSÃO

A Matriz de Confusão (MC) apresentada na Tabela 2, fornece uma visão detalhada do desempenho do modelo na classificação de diferentes classes. Cada linha da matriz representa as instâncias reais das classes ("Açai", "Cocão", "Jaci", "Patauá", "Paxiúba", "Tucumã" e "Floresta"), e cada coluna representa as previsões do modelo. Valores diagonais indicam previsões corretas, enquanto os valores fora da diagonal refletem erros. Esta ferramenta é essencial para identificar onde o modelo está cometendo erros e ajustar parâmetros para melhor performance (Wilkinson e Friendly, 2009). O modelo obteve uma precisão de 91,16% para "Açai", mas cometeu 113.231 erros para "Jaci" e 180.741 para "Floresta". Para "Cocão", a precisão foi de 77,42%, com erros principalmente em "Jaci" (50.952) e "Floresta" (31.712). "Jaci" teve uma precisão de 89,56%, mas mostrou erros consideráveis para "Floresta" (1.116.136) e "Paxiúba" (394.334). "Patauá" não teve previsões corretas, com 242.584 erros para "Jaci". A classe "Paxiúba" teve 80,78% de precisão, mas apresentou 207.034 erros para "Jaci" e 250.355 para "Floresta". "Tucumã" obteve 22,6% de precisão, com a maioria dos erros em "Paxiúba" (894.928) e "Floresta" (39.329). "Floresta" foi a classe com melhor desempenho, alcançando 95,32% de precisão, embora tenha cometido 10.769.946 erros para "Jaci" e 4.716.234 para "Açai". A análise revela variações significativas de desempenho entre as classes. "Açai" e "Jaci" apresentaram alta precisão, enquanto "Patauá" não teve previsões corretas, possivelmente devido ao desequilíbrio de dados ou dificuldades do modelo em identificar suas características. Erros frequentes entre "Jaci" e "Floresta" sugerem que o modelo está incluindo uma área maior de cobertura das espécies, levando a confusões entre essas classes. A classe "Patauá", em particular, apresentou uma alta quantidade de erros, principalmente com "Jaci", o que indica a necessidade de mais dados de treinamento ou uma revisão nas características utilizadas para identificar esta classe. A classe "Floresta", que apresentou a maior precisão, parece ser bem identificada pelo modelo devido à sua grande quantidade de dados e características visuais mais distintas. Para melhorar o desempenho geral, recomenda-se aumentar a quantidade de dados de treinamento para classes com baixa precisão, como "Patauá" e "Tucumã", revisar as características usadas no modelo e ajustar seus parâmetros. Realizar



mais ciclos de treinamento também pode ajudar a melhorar o aprendizado das classes menos representadas. Em resumo, o modelo teve um desempenho sólido para várias classes, mas ainda há um espaço considerável para melhorias, especialmente nas classes com precisão mais baixa ou inexistente, como "Patauá".

Matriz de Confusão (MC)

Classes Verdadeiras	Açaí	3133560	108	113231	0	9848	6	180741
	Cocão	575	354405	50952	0	13175	6971	31712
	Jaci	276260	79242	1620609	0	394334	22719	1116136
	Patauá	903	0	242584	0	0	0	5488
	Paxiúba	19522	12291	207034	0	2740365	162885	250355
	Tucumã	5868	5690	45229	0	894928	288658	39329
	Floresta	4716234	298092	1076994	0	2705568	165736	3800603
		Açaí	Cocão	Jaci	Patauá	Paxiúba	Tucumã	Floresta
Classes Preditas								

Tabela 2. Resultados da matriz de confusão gerada pelo modelo.

Conforme Congalton e Green (1999), as métricas de Erro de Comissão (EC) e Erro de Omissão (EO) fornecem uma visão detalhada dos erros de classificação. O erro de comissão ocorre quando uma classe é atribuída incorretamente, e o erro de omissão quando uma classe não é identificada corretamente. As fórmulas aplicadas foram: $EC = (FP/(TP+FP)) * 100$ e $EO = (FN/(TP+FN)) * 100$, onde FP são Falsos Positivos, TP Verdadeiros Positivos e FN Falsos Negativos. A Tabela 3 apresenta os cálculos de EC e EO para cada classe.

Classe	Nr. de Pixels	EC	EO	Classe	Nr. de Pixels	EC	EO	
Açaí	TP: 3133560	61.59%	8.84%	Paxiúba	TP: 2740365	59.45%	19.22%	
	FN: 303934				FN: 652087			
	FP: 5019362				FP: 4017853			
Cocão	TP: 354,405	52.70%	22.58%	Tucumã	TP: 288658	55.38%	77.44%	
	FN: 103385				FN: 991044			
	FP: 395423				FP: 358317			
Jaci	TP: 16206097	41.36%	10.44%	Floresta	TP: 380060365	0,43%	4.68%	
	FN: 1888691				FN: 18655576			
	FP: 11428976				FP: 1623761			
Patauá	TP: 0	0.00%	100.00%					
	FN: 248975							
	FP: 0							

Tabela 3. Cálculo dos EC e EO.

Os cálculos de Erro de Omissão (EO) e Erro de Comissão (EC) fornecem informações importante sobre o desempenho do modelo. Para a classe "Açaí", o EO foi de 8,84%, indicando poucas omissões, mas um EC de 61,59% revela significativa confusão com outras classes. "Cocão" apresentou um EO de 22,58% e um EC de 52,70%, indicando dificuldades moderadas na identificação. "Jaci" teve um EO de 10,44%, com bom desempenho, mas um EC de 41,36% mostra confusões. A classe "Patauá" foi crítica, com um EO de 100%, pois o modelo não reconheceu nenhuma instância, e um EC de 0%. "Paxiúba" teve um EO de 19,22% e EC de 59,45%, sugerindo confusão com outras classes. "Tucumã" mostrou baixos resultados, com EO de 77,44% e EC de 55,38%. Já a classe "Floresta" teve excelente desempenho, com EO de

4,68% e EC de 0,43%, destacando sua alta eficácia na identificação correta. Os elevados valores de Erro de Comissão (EC) observados podem ser explicados pela imagem de diferença (Figuras 3c e 3d), que mostra que o modelo incluiu mais pixels nas bordas do que a rotulação do analista (Figura 3b). O modelo parece tentar garantir que toda a espécie seja englobada, incorporando pixels extras que, em sua maioria, pertencem à classe "Floresta". Para comparar os resultados da segmentação gerada pelo modelo com a rotulação feita pelo analista, foi utilizada a operação de diferença de imagem. Essa técnica, descrita por Gonzalez e Woods (2000), calcula a diferença pixel a pixel entre duas imagens. A fórmula utilizada foi: $D(x, y) = I1(x, y) - I2(x, y)$, onde $D(x, y)$ representa o valor do pixel na imagem de diferença. A Figura 3 ilustra a comparação entre a classificação visual feita pelo analista e o resultado obtido pelo modelo, além da imagem de diferença entre ambos. A figura contém quatro imagens alinhadas horizontalmente, representando diferentes estágios do processo de segmentação: (a) a imagem aérea original da floresta capturada pelo VANT, (b) a rotulação do analista, onde áreas em verde/roxo são palmeiras e o amarelo é a floresta, (c) o resultado da classificação do modelo e (d) a imagem de diferença, mostrando erros de classificação, com pixels pretos indicando acertos e brancos indicando erros.

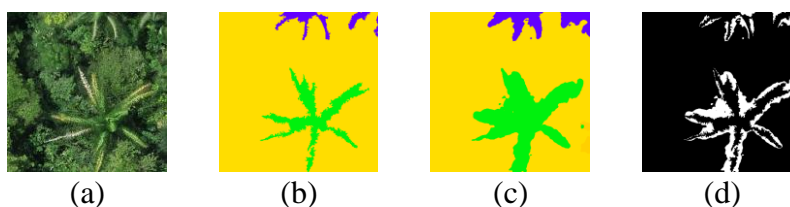


Figura 3 – Exemplificação da predição de uma imagem pelo modelo comparada ao rótulo do analista e o resultado da diferença entre os dois rótulos. (a) Detalhe da imagem na escala 1:250. (b) Rótulo do analista. (c) Rótulo do modelo. (d) Diferença entre a segmentação do modelo e os rótulos de dados.

A Tabela 4 apresenta os resultados da imagem de diferença entre os rótulos de dados da área de estudo produzidos pelo analista através de inspeção visual, a segmentação e identificação da área de estudo realizada pelo modelo proposto.

Resultado	Nr. Pixels	%
Correto	1169923117	98,08
Erros	22846115	1,92

Tabela 4. Resultado do processamento da diferença.

A análise das diferenças entre as imagens mostrou que 1.169.923.117 pixels, representando 98,08% do total, foram classificados corretamente, enquanto 22.846.115 pixels (1,92%) foram classificados incorretamente. Esses resultados indicam uma alta precisão do modelo de segmentação, com a maioria das classificações corretas. No entanto, a classe "Floresta", que representa 93,68% das amostras, domina o conjunto de dados, o que pode influenciar a precisão geral, já que o modelo tem mais dados dessa classe para aprender. Embora a precisão seja elevada, o desequilíbrio nas amostras deve ser considerado ao avaliar a eficácia do modelo em classes menos representadas, e há espaço para melhorias na redução dos erros de classificação.

4. CONCLUSÕES



Os resultados da pesquisa mostram que o objetivo foi parcialmente alcançado, com sucesso na identificação de algumas das seis espécies de palmeiras. Classes como "Açaí", "Jaci" e "Floresta" tiveram alta precisão, enquanto "Patauá" não teve previsões corretas. A hipótese de que uma rede CNN poderia identificar padrões visuais de palmeiras foi confirmada, demonstrando sua eficácia, embora sejam necessárias melhorias para certas classes. O estudo destacou as capacidades e limitações do modelo na classificação das diferentes classes. A classe "Açaí" alcançou 91,16% de precisão, enquanto "Cocão" obteve 77,42%. "Jaci" apresentou um bom desempenho com 89,56%, mas "Patauá" foi a mais problemática, sem previsões corretas (0,00% de precisão). "Paxiúba" teve 80,78% de precisão, "Tucumã" alcançou 22,6%, e a classe "Floresta" foi a mais precisa, com 95,32%. Os resultados indicam que o modelo conseguiu identificar com alta precisão as classes "Açaí", "Jaci" e "Floresta", com "Floresta" alcançando 95,32% de precisão devido à grande quantidade de dados disponíveis. No entanto, houve desafios com a classe "Patauá", que não teve previsões corretas, e com as classes "Cocão", "Tucumã" e "Paxiúba", que apresentaram precisões mais baixas. Esses problemas podem ser atribuídos à insuficiência de exemplos no conjunto de treinamento. Os altos erros de EC e EO podem ser explicados pela análise da imagem de diferença, que revela a tendência do modelo em expandir a área de classificação, incorporando pixels de outras classes para garantir a cobertura completa das espécies de palmeiras. Os resultados deste estudo têm importantes implicações práticas para o manejo florestal e políticas ambientais na Amazônia. O uso de CNNs para identificar espécies pode aumentar a eficiência e precisão dos inventários florestais, reduzindo custos e acelerando a coleta de dados. Isso pode facilitar políticas mais eficazes de conservação e uso sustentável dos recursos. A identificação precisa das palmeiras é vital para a preservação da biodiversidade e o suporte às comunidades locais que dependem dessas espécies. A melhoria dos modelos de CNN e o aumento dos dados de treinamento são essenciais para maximizar o impacto positivo dessas tecnologias no manejo florestal. Para melhorias futuras, sugere-se coletar mais dados de treinamento, especialmente para classes com baixa precisão, como "Patauá" e "Tucumã". A análise da matriz de confusão mostra que, apesar do bom desempenho geral do modelo, há espaço para aprimoramento, especialmente em classes com baixa precisão. Com a continuidade da pesquisa, espera-se avançar no uso de aprendizado de máquina para a classificação de espécies em ambientes complexos como a Floresta Amazônica.

REFERÊNCIAS

ACRE, Governo do Estado. Programa estadual de zoneamento ecológico-econômico do estado do Acre. Zoneamento ecológico-econômico do Acre Fase II: documento síntese – Escala, v. 1, n. 250.000, 2006.

ANDERSON, Karen; GASTON, Kevin J. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, Wiley Online Library, v. 11, n. 3, p. 138–146, 2013.

CHENG, Y. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 17, n. 8, p. 790-799, 1995. DOI: 10.1109/34.400568.

COMANICIU, D.; MEER, P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 5, p. 603-619, 2002. DOI: 10.1109/34.1000236.



- CONGALTON, R. G.; GREEN, K. Assessing the accuracy of remotely sensed data: principles and practices. Boca Raton: CRC Press, 1999.
- CULMAN, M.; DELALIEUX, S.; VAN TRICHT, K. Individual palm tree detection using deep learning on RGB imagery to support tree inventory. *Remote Sensing*, v. 12, n. 21, p. 3476, 2020. Disponível em: <https://doi.org/10.3390/rs12213476>.
- FERREIRA, E. J. L. Manual das palmeiras do Acre, Brasil. Rio Branco: Instituto Nacional de Pesquisas da Amazônia, 2006. Disponível em: https://www.nybg.org/bsci/acre/www1/manual_palmeiras.htm.
- GIBRIL, M. B. A. *et al.* Deep convolutional neural network for large-scale date palm tree mapping from UAV-based images. *Remote Sensing*, v. 13, n. 14, p. 2787, 2021. Disponível em: <https://doi.org/10.3390/rs13142787>.
- GIBRIL, Mohamed Barakat A. *et al.* Large-scale date palm tree segmentation from multiscale UAV-based and aerial images using deep vision transformers. *Drones*, v. 7, n. 2, 2023. ISSN 2504-446X. Disponível em: <https://www.mdpi.com/2504-446X/7/2/93>.
- GONZALEZ, R. C.; WOODS, R. E. Processamento de imagens digitais. São Paulo: Editora Blucher, 2000.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. *IEEE*, 2016. DOI: 10.1109/CVPR.2016.90.
- HTET, Khaing Suu; SEIN, Myint Myint. Toddy palm trees classification and counting using drone video: Retuning hyperparameter mask-rcnn. p. 196–200, 2021.
- INMET, Instituto Nacional de Meteorologia. Normais climatológicas Brasil - período: 1991-2020. Ministério da Agricultura, 2020.
- JÚNIOR, A. G.; RIBAS, R. P. Identificação de Palmeiras (Arecaceae) Nativas em Áreas de Floresta Tropical baseado em Rede Neural Convolutiva com imagens de VANT. *Revista Brasileira de Geografia Física*, v. 16, n. 5, p. 2360-2374, 2023.
- LI, W.; DONG, R.; FU, H.; YU, L. Large-Scale Oil Palm Tree Detection from High-Resolution Satellite Images Using Two-Stage Convolutional Neural Networks. *Remote Sensing*, v. 11, n. 1, p. 11, 2019. Disponível em: <https://doi.org/10.3390/rs11010011>.
- LETSOIN, Sri Murniani Angelina *et al.* Recognition of sago palm trees based on transfer learning. *Remote Sensing*, v. 14, n. 19, 2022. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/14/19/4932>.
- LORENZI, H.; NOBLICK, L. R.; KAHN, F.; FERREIRA, E. Flora Brasileira Lorenzi: Arecaceae (Palmeiras). Nova Odessa, 2010.
- MARIN, Willintong; MONDRAGON, Ivan F.; COLORADO, Julian D. Aerial identification of Amazonian palms in high-density forest using deep learning. *Forests*, v. 13, n. 5, 2022. ISSN 1999-4907. Disponível em: <https://www.mdpi.com/1999-4907/13/5/655>.
- MURPHY, K. P. Machine learning: a probabilistic perspective. MIT press, 2012.
- SHIMODAIRA, H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, v. 90, n. 2, p. 227–244, 2000. DOI: 10.1016/S0378-3758(00)00115-4.



SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. arXiv, 2014. Disponível em: <https://arxiv.org/abs/1409.1556>.

WAHED, Zulhakim et al. Sago palm detection and its maturity identification based on improved convolution neural network. *Pertanika Journal of Science & Technology*, v. 30, n. 2, p. 1219–1236, 2022.

WILKINSON, L.; FRIENDLY, M. The history of the cluster heat map. *The American Statistician*, v. 63, n. 2, p. 179-184, 2009.