ReGeo ISSN: 2177-3246

AVALIAÇÃO COMPARATIVA DAS ARQUITETURAS EFFICIENTNET-B0 E MOBILENETV2 PARA A IDENTIFICAÇÃO AUTOMÁTICA DE ESPÉCIES DE TUBARÕES

COMPARATIVE EVALUATION OF EFFICIENTNET-B0 AND MOBILENETV2 ARCHITECTURES FOR THE AUTOMATIC IDENTIFICATION OF SHARK SPECIES

EVALUACIÓN COMPARATIVA DE LAS ARQUITECTURAS EFFICIENTNET-B0 Y MOBILENETV2 PARA LA IDENTIFICACIÓN AUTOMÁTICA DE ESPECIES DE TIBURONES

https://doi.org/10.56238/revgeov17n11-083

Data de submissão: 16/10/2026 **Data de publicação:** 16/11/2026

João Antonio Assis de Santana Batista
Mestre em Modelagem e Simulação de Biossistemas
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: jantonioa2810@gmail.com
Lattes: http://lattes.cnpq.br/8346245065149508

Jéssica Figuera Oliveira Miranda
Doutoranda em Ciências Ambientais
Instituição: Universidade Estadual de Feira de Santana (UEFS)
E-mail: ma.jessicafiguera@gmail.com
Lattes: http://lattes.cnpq.br/8985717830628943

Andressa Morais Amâncio
Graduada em Licenciatura em Ciências Biológicas
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: dressamancio@gmail.com
Lattes: http://lattes.cnpq.br/7754580261539226

Beatriz Souza de Assis
Mestranda em Modelagem e Simulação de Biossistemas
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: beatrizassis@outlook.com
Lattes: http://lattes.cnpq.br/7095774571990585

Hannah Daiahnara da Assunção Cruz
Mestranda em Modelagem e Simulação de Biossistemas
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: hannahcruuz@gmail.com
Lattes: http://lattes.cnpq.br/0768203789308016

Ísis Maria Teles de Oliveira
Bacharelanda em Engenharia de Software
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: isiselet@gmail.com
Lattes: http://lattes.cnpq.br/4857281988131506

ReGeo

ISSN: 2177-3246

Neide Sales dos Santos
Especialista em Educação Científica e Popularização das Ciências
Instituição: Instituto Federal Baiano (IF Catu)
E-mail: nsalesbio@gmail.com
Lattes: http://lattes.cnpq.br/1132800455861776

Gildasio Costa Menezes Neto
Graduando em Licenciatura em Ciências Biológicas
Instituição: Universidade do Estado da Bahia (UNEB)
E-mail: biologildabz@gmail.com
Lattes: http://lattes.cnpq.br/5427687704202978

RESUMO
O artigo compara as arquiteturas de Redes Neurais Convolucionais EfficientNet-B0 e MobileNetV2 na classificação automática de 14 espécies de tubarões, visando apoiar o monitoramento e a conservação da biodiversidade marinha. Utilizou-se o conjunto Shark Species Identification, com 1.546 imagens distribuídas entre treino (70%), validação (20%) e teste (10%). Ambas as redes empregaram Transfer Learning, com pesos pré-treinados do ImageNet, sob condições idênticas durante 40 épocas. A MobileNetV2 destacou-se pela eficiência computacional, concluindo o treinamento em 17,45 minutos, sendo 40,8% mais rápida e demandando menos parâmetros e memória. Entretanto, apresentou acurácia de 74,84% e F1-Score ponderado de 69,17%, com desempenho inferior para Bull, Blacktip e Mako. A EfficientNet-B0 alcançou maior robustez, com acurácia de 79,35% e F1-Score ponderado de 79,31%, mantendo F1-Score superior a 0,5 em todas as espécies e melhor controle de overfitting. Os erros foram associados principalmente às semelhanças morfológicas entre espécies e às condições adversas das imagens subaquáticas. Assim, a EfficientNet-B0 mostrou-se mais adequada à bioidentificação de elasmobrânquios, pois sua maior precisão taxonômica compensou o custo computacional superior.

Palavras-chave: Bioidentificação. Classificação de Imagens. Elasmobrânquios. Redes Neurais Convolucionais.

ABSTRACT
This article compares the EfficientNet-B0 and MobileNetV2 Convolutional Neural Network architectures for the automatic classification of 14 shark species, aiming to support marine biodiversity monitoring and conservation. The study utilized the Shark Species Identification dataset, comprising 1,546 images distributed across training (70%), validation (20%), and testing (10%) sets. Both networks employed transfer learning with pre-trained ImageNet weights under identical conditions over 40 epochs. MobileNetV2 stood out for its computational efficiency, completing training in 17.45 minutes making it 40.8% faster and requiring fewer parameters and less memory. However, it achieved an accuracy of 74.84% and a weighted F1-score of 69.17%, showing lower performance for Bull, Blacktip, and Mako sharks. EfficientNet-B0 demonstrated greater robustness, achieving an accuracy of 79.35% and a weighted F1-score of 79.31%, while maintaining an F1-score above 0.5 for all species and better controlling for overfitting. Errors were primarily attributed to morphological similarities between species and the challenging conditions of underwater imagery. Consequently, EfficientNet-B0 proved more suitable for the bio-identification of elasmobranchs, as its superior taxonomic precision outweighed its higher computational cost.

Keywords: Bio-identification. Image Classification. Elasmobranchs. Convolutional Neural Networks.

RESUMEN
El artículo compara las arquitecturas de redes neuronales convolucionales EfficientNet-B0 y MobileNetV2 para la clasificación automática de 14 especies de tiburones, con el objetivo de apoyar el monitoreo y la conservación de la biodiversidad marina. Se utilizó el conjunto de datos Shark Species Identification, compuesto por 1.546 imágenes distribuidas en entrenamiento (70 %), validación (20 %) y prueba (10 %). Ambas redes emplearon aprendizaje por transferencia (Transfer Learning) con pesos preentrenados de ImageNet, bajo condiciones idénticas durante 40 épocas. MobileNetV2 destacó por su eficiencia computacional, completando el entrenamiento en 17,45 minutos; fue un 40,8 % más rápida y requirió menos parámetros y memoria. Sin embargo, presentó una precisión del 74,84 % y un F1-Score ponderado del 69,17 %, con un rendimiento inferior para las especies Bull, Blacktip y Mako. EfficientNet-B0 logró mayor robustez, con una precisión del 79,35 % y un F1-Score ponderado del 79,31 %, manteniendo un F1-Score superior a 0,5 en todas las especies y un mejor control del sobreajuste (overfitting). Los errores se asociaron principalmente a las similitudes morfológicas entre especies y a las condiciones adversas de las imágenes subacuáticas. Así, EfficientNet-B0 resultó más adecuada para la bioidentificación de elasmobranquios, ya que su mayor precisión taxonómica compensó el costo computacional más elevado.

Palabras clave: Bioidentificación. Clasificación de Imágenes. Elasmobranquios. Redes Neuronales Convolucionales.

1 INTRODUÇÃO

A biologia marinha é fundamental para a compreensão, conservação e manejo dos ecossistemas oceânicos. Nesse contexto, os peixes cartilaginosos, como tubarões, raias e quimeras, desempenham funções ecológicas essenciais ao regularem populações e contribuírem para o equilíbrio das cadeias tróficas. Contudo, a pesca excessiva e a degradação ambiental têm provocado o declínio dessas espécies, sendo que aproximadamente um terço encontra-se ameaçado de extinção (Dulvy *et al.*, 2017). Esse cenário evidencia a necessidade de ferramentas que apoiem o monitoramento, a pesquisa e a conservação da biodiversidade marinha.

Nesse contexto, a identificação taxonômica correta das espécies constitui uma etapa fundamental para inventários da biodiversidade, monitoramento pesqueiro, fiscalização ambiental e desenvolvimento de estratégias de conservação. Na prática, contudo, esse processo ainda depende fortemente da experiência de especialistas e da consulta a guias taxonômicos, chaves dicotômicas e catálogos científicos, tornando a identificação demorada, complexa e suscetível a erros, especialmente para profissionais em formação ou em situações de campo. No Brasil, onde a diversidade de elasmobrânquios é elevada e a pressão sobre esses organismos permanece significativa, ferramentas computacionais capazes de agilizar e padronizar esse processo podem contribuir diretamente para pesquisadores, estudantes, órgãos ambientais e instituições de conservação (Chelotti; Santos, 2020).

Sob a perspectiva teórica, a aplicação de tecnologias computacionais à identificação biológica fundamenta-se na integração entre a taxonomia, e sistemática e a inteligência artificial. Nos últimos anos, a Visão Computacional e o Aprendizado Profundo têm demonstrado elevado potencial para automatizar tarefas de classificação biológica por meio da extração automática de características visuais presentes nas imagens. Entre essas abordagens, as Redes Neurais Convolucionais (*Convolutional Neural Networks* conhecidas como CNNs) destacam-se por sua elevada capacidade de aprender padrões relacionados à forma, textura, coloração e outras características morfológicas discriminativas, reduzindo a necessidade de extração manual de atributos e alcançando resultados expressivos em diferentes problemas de reconhecimento de imagens (LeCun; Bengio; Hinton, 2015).

Diversos estudos têm evidenciado o potencial dessas arquiteturas na identificação automática de organismos aquáticos e na conservação da biodiversidade (Miranda *et al.*, 2026; Ivanoff *et al*., 2019; Zhou *et al.*, 2023; Sami; Ahmed, 2026). Modelos como *MobileNetV2* e *EfficientNet-B0* vêm sendo amplamente empregados em aplicações que exigem alta acurácia aliada à eficiência computacional, permitindo sua utilização em dispositivos com recursos limitados e sistemas *web*. Enquanto a *MobileNetV2* foi projetada para oferecer rapidez e baixo custo computacional por meio de convoluções separáveis em profundidade (Sandler *et al*., 2018), a *EfficientNet-B0* utiliza um método de

escalonamento composto que otimiza simultaneamente profundidade, largura e resolução da rede, proporcionando melhor equilíbrio entre desempenho e complexidade computacional. Esses modelos já apresentaram resultados promissores em tarefas de classificação de espécies e reconhecimento de padrões biológicos, indicando seu potencial para aplicações em biologia marinha (Tan; Le, 2019).

Apesar desses avanços, o conhecimento disponível ainda apresenta limitações importantes quando aplicado ao contexto da identificação de tubarões, raias e quimeras. Estudos recentes demonstram o potencial de técnicas de inteligência artificial para a identificação de elasmobrânquios, embora parte das aplicações permaneça concentrada em conjuntos de dados específicos, número limitado de espécies ou contextos particulares de reconhecimento (Beviá-Ballesteros *et al*., 2026). Em paralelo, pesquisas voltadas à classificação automática de peixes têm demonstrado que diferentes arquiteturas de redes neurais apresentam compromissos distintos entre desempenho preditivo e eficiência computacional, tornando necessária a avaliação conjunta de métricas de classificação, tempo de inferência e demanda de recursos computacionais.

Nesse cenário, ainda são necessárias investigações comparativas especificamente direcionadas aos peixes cartilaginosos que permitam avaliar, sob condições experimentais equivalentes, arquiteturas leves e eficientes e sua adequação a aplicações práticas (Mohammadisabet *et al*., 2025). Diante desse cenário, o presente estudo busca preencher essa lacuna por meio da comparação entre as arquiteturas *EfficientNet-B0* e *MobileNetV2* na tarefa de classificação automática de espécies de tubarões. A pesquisa avaliará métricas de desempenho, como acurácia, precisão, revocação e F1-*score*, além da eficiência computacional e do potencial de aplicação prática em um sistema *web* de identificação de espécies. Espera-se, assim, identificar a arquitetura que oferece o melhor equilíbrio entre desempenho e custo computacional.

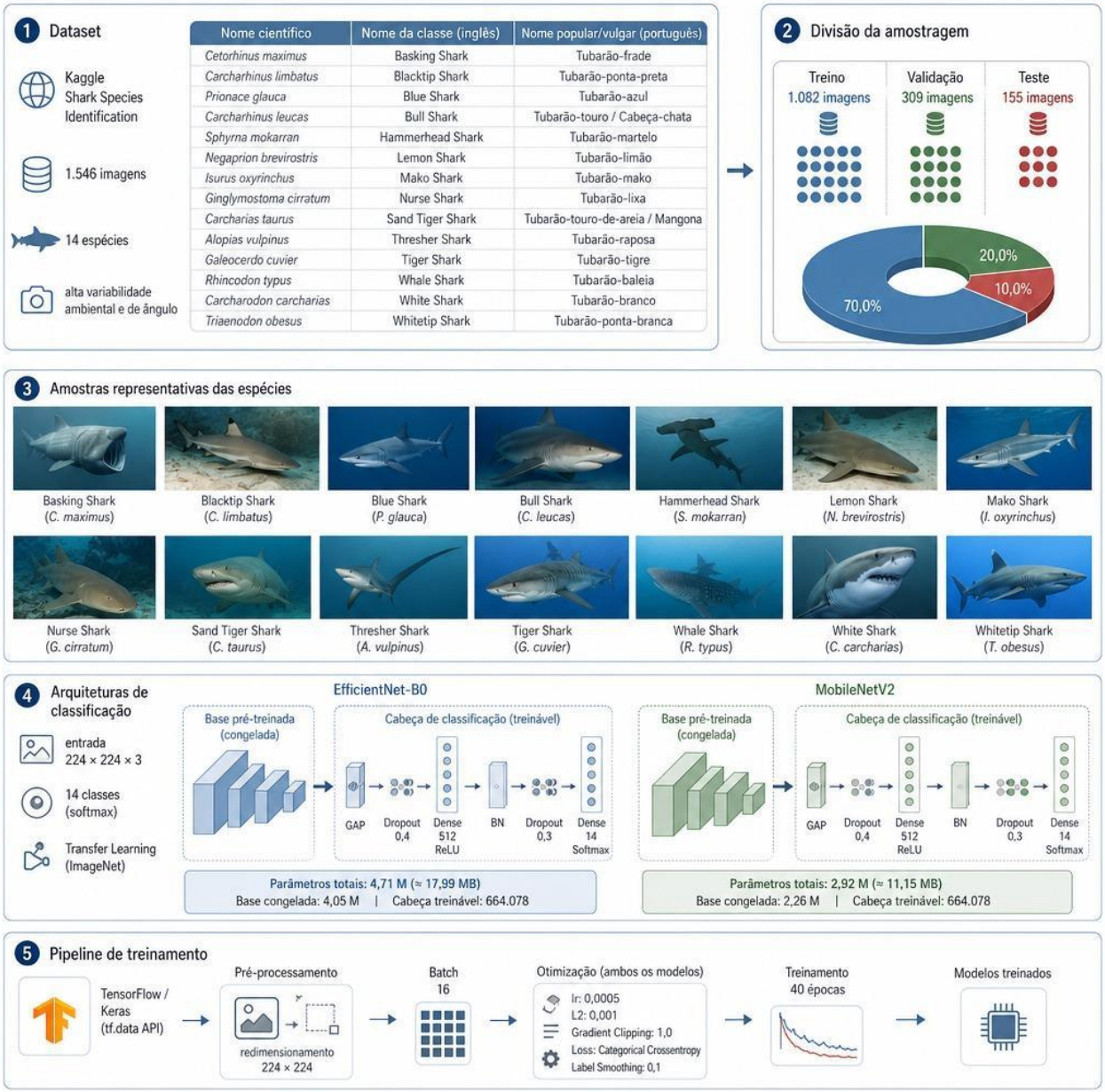
3 METODOLOGIA

A metodologia deste estudo, voltada à classificação automatizada de espécies da fauna marinha, foi estruturada a partir do *pipeline* metodológico apresentado na Figura 1, que sistematiza as principais etapas desenvolvidas ao longo da pesquisa. Para a construção e avaliação dos modelos de classificação, utilizou-se o conjunto de dados públicos *Shark Species Identification*, disponibilizado por Sayak Samanta na plataforma *Kaggle*. A escolha desse conjunto de dados fundamentou-se em sua diversidade taxonômica e na disponibilidade de imagens representativas de diferentes espécies de tubarões.

O conjunto de dados é composto por 1.546 imagens, distribuídas entre 14 espécies: Basking Shark (*Cetorhinus maximus*), Blacktip Shark (*Carcharhinus limbatus*), Blue Shark (*Prionace glauca*), Bull Shark (*Carcharhinus leucas*), Hammerhead Shark (*Sphyrna mokarran*), Lemon Shark (*Negaprion*

brevirostris), Mako Shark (*Isurus oxyrinchus*), Nurse Shark (*Ginglymostoma cirratum*), Sand Tiger Shark (*Carcharias taurus*), Thresher Shark (*Alopias vulpinus*), Tiger Shark (*Galeocerdo cuvier*), Whale Shark (*Rhincodon typus*), White Shark (*Carcharodon carcharias*) e Whitetip Shark (*Triaenodon obesus*). Essa composição permitiu estabelecer um problema de classificação multiclasse, no qual as diferentes espécies constituíram as classes utilizadas no treinamento, validação e teste dos modelos de aprendizado profundo.

Figura 1. Pipeline metodológico aplicado ao estudo



Fonte: elaboração própria (2026).

Com o objetivo primário de garantir a correta avaliação da capacidade de generalização do modelo e prevenir qualquer tipo de vazamento de dados durante as etapas de ajuste, a amostragem foi dividida rigorosamente em três subconjuntos independentes.

A distribuição quantitativa da base de dados mostrada na Figura 1, evidencia que o conjunto de treino reuniu 1.082 imagens destinadas ao aprendizado direto dos parâmetros da rede, o conjunto de validação conteve 309 imagens reservadas para o monitoramento da evolução e ajuste fino de hiperparâmetros, e o conjunto de teste final foi estruturado com 155 imagens mantidas sob sigilo total em relação ao treinamento para mensurar o desempenho em dados não vistos. As imagens que compõem o repositório exibem alta complexidade visual e variabilidade ambiental, visto que foram capturadas sob diferentes condições de iluminação subaquática, níveis de turbidez da água e ângulos de tomada, englobando desde planos gerais do animal em nado livre até enquadramentos focados em detalhes anatômicos específicos. Foram selecionadas duas arquiteturas de redes neurais convolucionais consagradas na literatura para tarefas de visão computacional: a *EfficientNet-B0* e a *MobileNetV2*. Ambas foram configuradas a partir de pesos pré-treinados no *dataset ImageNet* via aprendizado por transferência (*Transfer Learning*), visando classificar 14 espécies de tubarões extraídas de um *dataset* composto por 1.546 imagens, distribuídas em 1.082 amostras para treino, 309 para validação e 155 para teste. A justificativa para a escolha de ambas reside na busca pelo equilíbrio ideal entre capacidade de representação abstrata e custo computacional.

A *EfficientNet-B0* apresenta aproximadamente 4,71 milhões de parâmetros totais, equivalentes a cerca de 17,99 *megabytes* em memória, dos quais 4,05 milhões pertencem à base congelada. Sua cabeça de classificação customizada possui 664.078 parâmetros treináveis, composta por uma camada *GlobalAveragePooling2D*, *Dropout* de 0,4, uma camada densa intermediária de 512 neurônios com ativação ReLU e *Batch Normalization*, um segundo *Dropout* de 0,3 e uma camada final *Dense* com 14 neurônios e ativação *Softmax*.

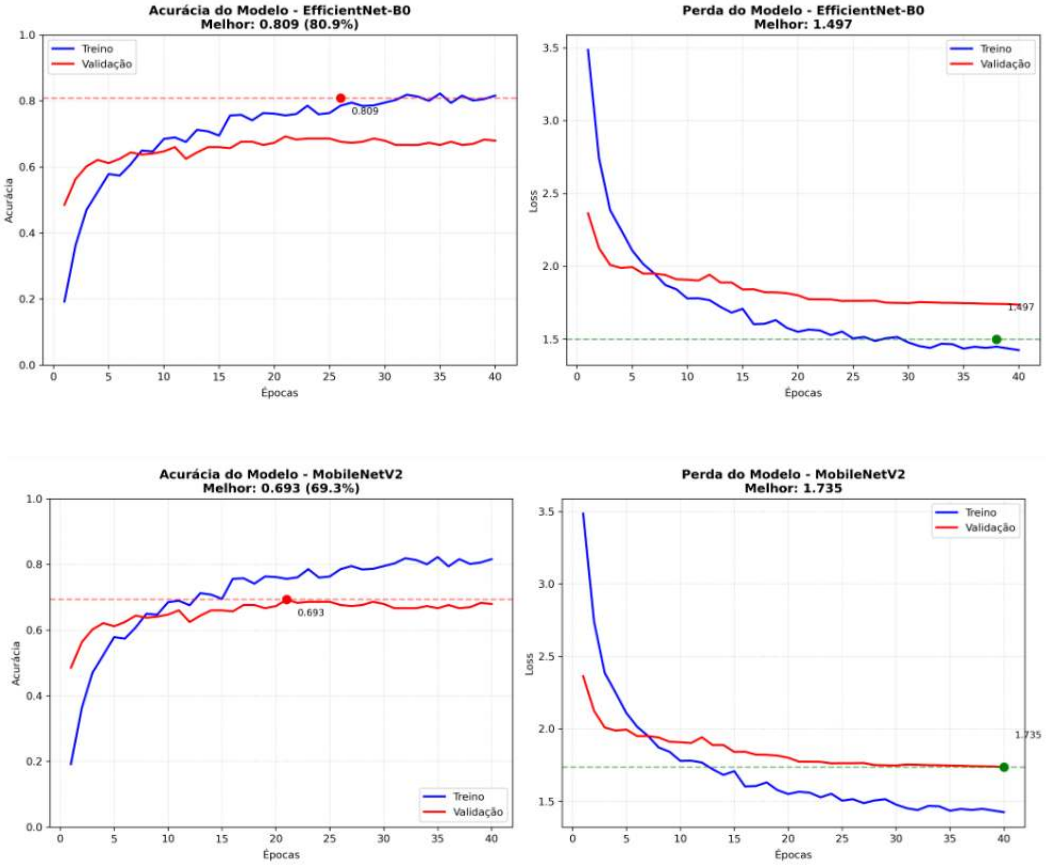
Por sua vez, a *MobileNetV2* apresenta uma estrutura ainda mais compacta, somando 2,92 milhões de parâmetros totais, correspondentes a 11,15 *megabytes* em memória ou cerca de 14 *megabytes* de pesos salvos em disco, contando com 2,26 milhões de parâmetros congelados no *backbone*. A cabeça de classificação manteve a mesma estrutura sequencial do primeiro modelo, contendo rigorosamente os mesmos 664.078 parâmetros treináveis. Enquanto a *EfficientNet-B0* exige um consumo estimado de memória total entre 1,4 e 1,6 *gigabytes* durante o processamento de lotes de 16 imagens na resolução de 224 por 224 *pixels*, a *MobileNetV2* opera de forma mais leve, demandando entre 1,0 e 1,2 *gigabytes*, tornando-a uma alternativa voltada ao empacotamento e execução em dispositivos com maior restrição de *hardware*.

A infraestrutura e o *pipeline* de execução do treinamento de ambas as redes foram padronizados no ecossistema *TensorFlow* e *Keras* com suporte à API *tf.data*. As imagens foram redimensionadas para a resolução padrão de 224 por 224 *pixels* em lotes de 16 amostras. Ambos os modelos foram submetidos a condições idênticas de otimização, que incluíram taxa de aprendizado inicial de 0,0005, regularização L2 de 0,001, *Gradient Clipping* fixado em 1,0 e função de perda *Categorical Crossentropy* com *Label Smoothing* de 0,1. O ciclo completo de ajuste foi conduzido ao longo de 40 épocas contínuas.

4 RESULTADOS E DISCUSSÕES

O diagnóstico comparativo de desempenho revelou comportamentos nitidamente distintos entre as duas arquiteturas ao longo do ciclo de otimização, bem como nas etapas subsequentes de validação e teste. No que concerne à eficiência computacional e velocidade de processamento, a *MobileNetV2* demonstrou expressiva superioridade temporal, concluindo o ciclo completo de 40 épocas em 17,45 minutos, o que corresponde a uma cadência média de 26,18 segundos por época. Em contrapartida, a arquitetura *EfficientNet-B0* demandou 24,58 minutos para finalizar o mesmo ciclo, um tempo computacional cerca de 40,8% superior, registrando um custo médio de 36,87 segundos por época. Tal disparidade decorre da maior complexidade das conexões e operações de escalonamento composto (*compound scaling*) inerentes ao *backbone* da *EfficientNet-B0*.

Figura 2. Curvas de aprendizado das arquiteturas EfficientNet-B0 e MobileNetV2 nos conjuntos de treino e validação.



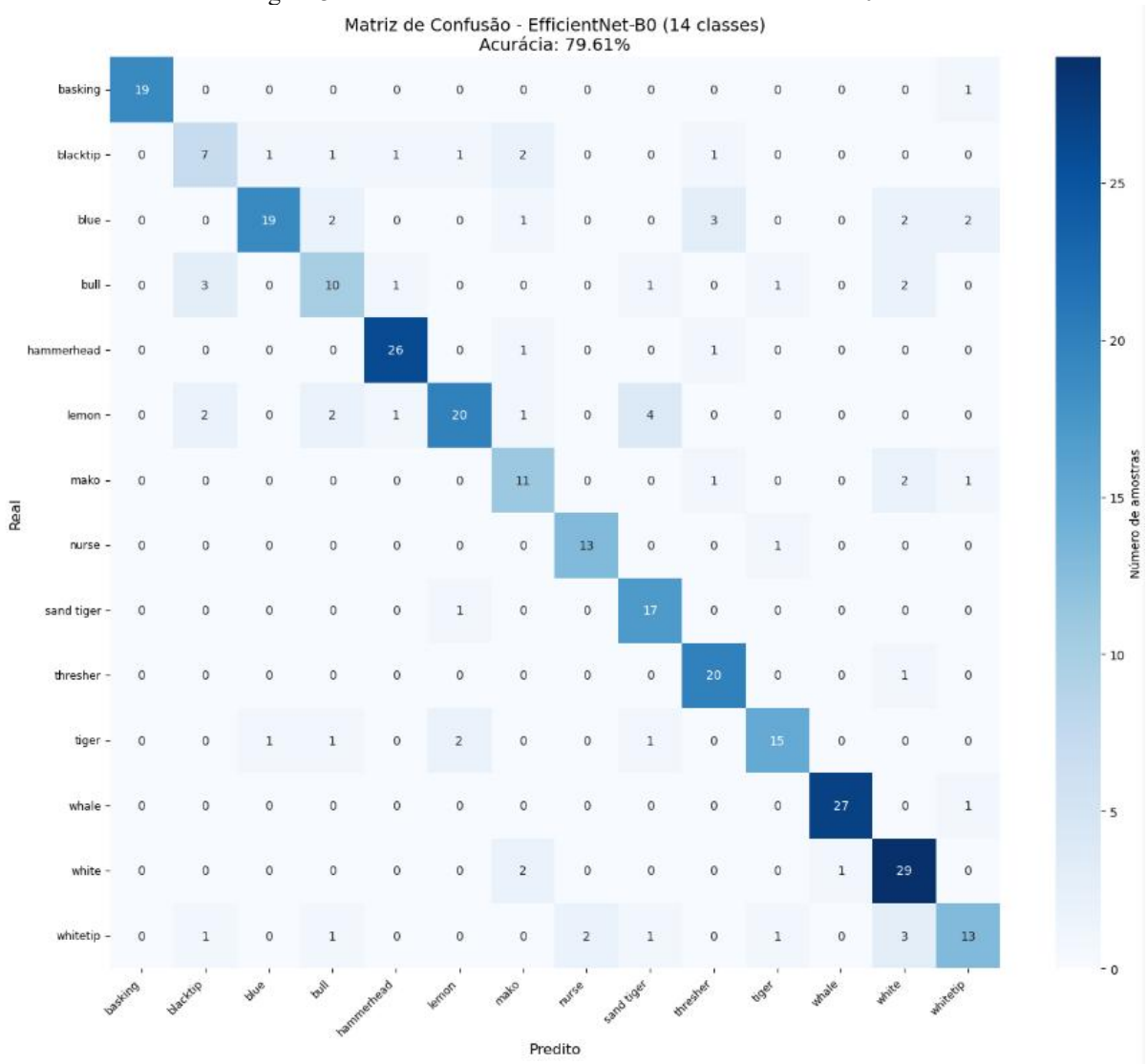
Fonte: elaboração própria (2026).

Analisando a capacidade de generalização e o comportamento dinâmico das redes ilustrados na Figura 02, a *EfficientNet-B0* exibiu maior estabilidade e robustez preditiva. Durante a fase de validação, essa rede atingiu seu pico de acurácia em 80,91% na época 26 e alcançou o menor valor da função de perda (*loss*) fixado em 1,497 na época 38. Esse comportamento reflete um controle rigoroso do fenômeno de *overfitting*, evidenciado por uma discrepância estreita de apenas 3,38% entre a acurácia de treino e validação. Quando submetida ao conjunto de teste independente (composto por 155 imagens inéditas), a EfficientNet-B0 manteve a consistência de seus resultados, registrando uma acurácia global de 79,35%, *loss* de 1,4778 e F1-*Score* ponderado de 79,31%. Notavelmente, a rede não apresentou nenhuma classe com F1-*Score* inferior ao limite crítico de 0,5.

Por outro lado, a *MobileNetV2* manifestou limitações em sua capacidade de convergência e abstração de feições complexas. A acurácia máxima de validação do modelo estabilizou-se em 69,26% na época 21, com a menor perda registrada em 1,735 na 40 época. A divergência entre a acurácia de treino (75,60%) e a de validação assinalou um *overfitting* moderado de 6,35%. No conjunto de teste, embora tenha obtido uma acurácia nominal de 74,84% e *loss* de 1,6470, o F1-*Score* ponderado da

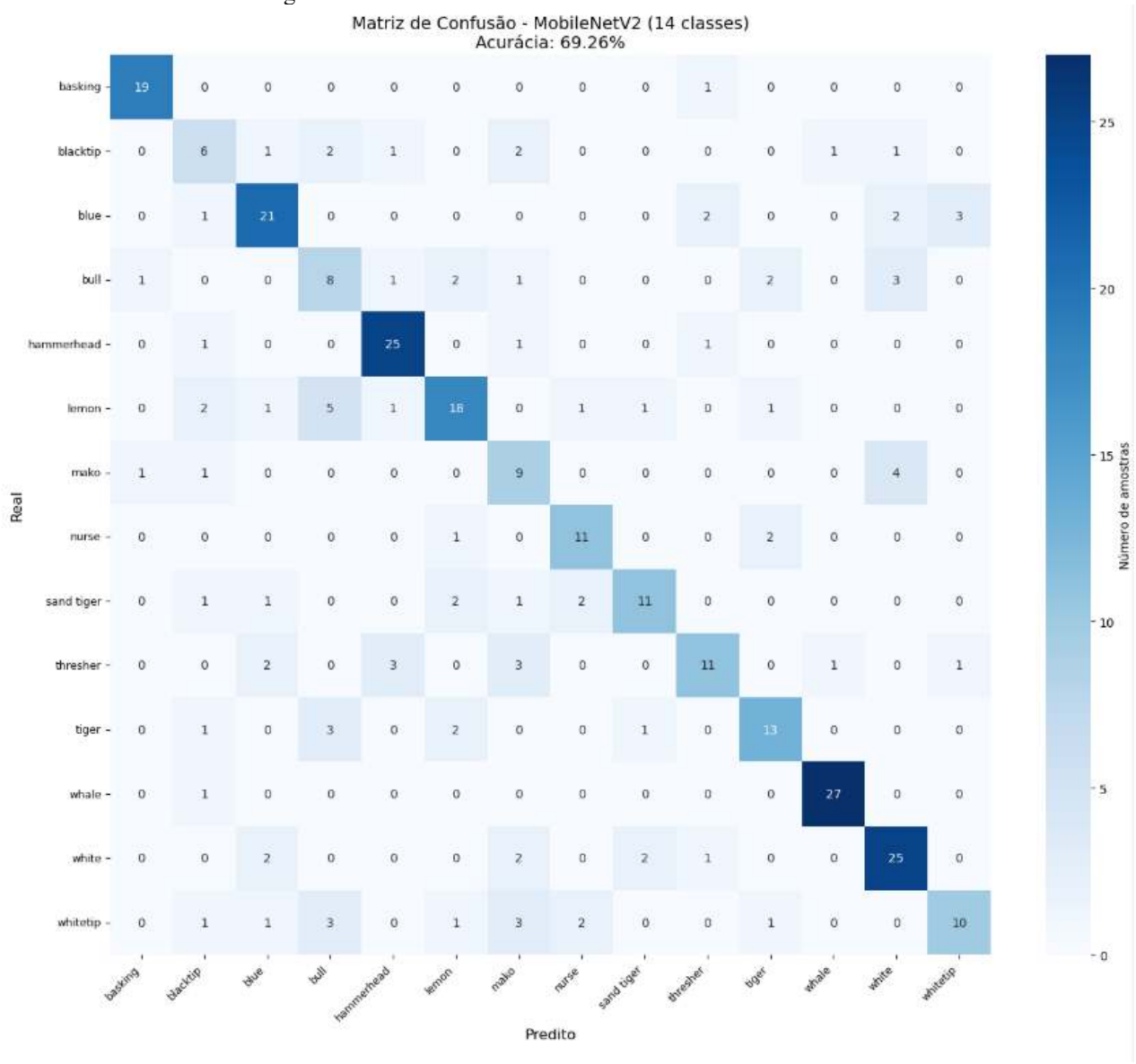
MobileNetV2 sofreu uma queda acentuada para 69,17%, sugerindo um desequilíbrio na taxa de acertos entre as diferentes categorias.

Figura 3. Matrizes de confusão normalizadas EfficientNet-B0.



Fonte: elaboração própria (2026).

Figura 4. Matrizes de confusão normalizadas MobileNetV2.



Fonte: elaboração própria (2026).

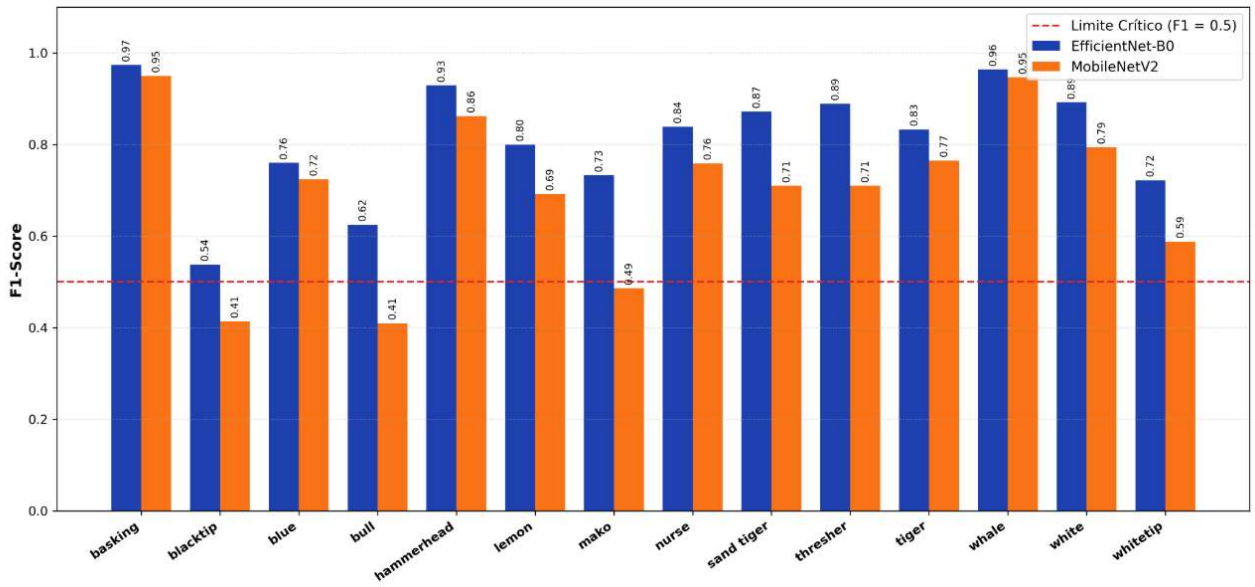
A inspeção detalhada das matrizes de confusão apresentadas na Figura 03 e na Figura 04 permite diagnosticar visualmente os padrões de erro e as ambiguidades preditivas de cada arquitetura. Enquanto a *EfficientNet-B0* concentrou a esmagadora maioria de suas predições ao longo da diagonal principal, obtendo acertos expressivos e consistentes em classes como *Whale* com 27 acertos, *White* com 29 acertos, *Hammerhead* com 26 acertos, *Thresher* com 20 acertos e *Lemon* com 20 acertos, a MobileNetV2 exibiu uma dispersão acentuada fora da diagonal, revelando fragilidades significativas na discriminação interclasses.

Na *MobileNetV2*, notam-se confusões severas em espécies como *Mako*, em que 4 amostras reais foram incorretamente classificadas como *White*, e em *Lemon*, onde 5 amostras foram marcadas

como *Bull*, além da classe *Whitetip*, que sofreu intensa pulverização preditiva ao ter suas amostras erroneamente distribuídas entre *Mako* (3), *Bull* (3), *Nurse* (2) e *Lemon* (1).

Em contrapartida, a *EfficientNet-B0* manteve erros pontuais e estritamente contidos, onde espécies de alta complexidade morfológica, como *Sand Tiger* com 17 acertos de 18 amostras e *Nurse* com 13 acertos de 14 amostras, demonstraram baixíssima taxa de dispersão, de modo que suas raras ambiguidades limitaram-se a sobreposições visuais clássicas, a exemplo de 3 amostras de *Blue* classificadas como *Thresher* e 3 amostras de *Bull* classificadas como *Blacktip*. Essa análise comprova a superioridade da *EfficientNet-B0* em extrair representações de alto nível e feições taxonômicas sutis, enquanto a *MobileNetV2* tende a fundir padrões morfológicos de táxons semelhantes devido à menor capacidade expressiva de seu *backbone*.

Figura 5. Gráficos de barras comparativos do F1-Score.



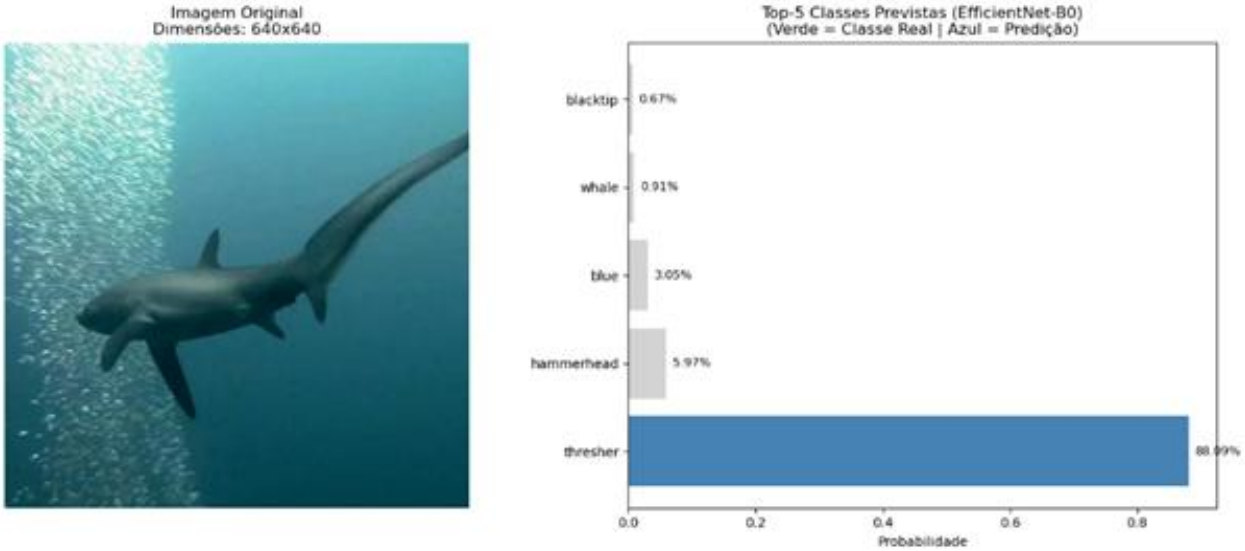
Fonte: elaboração própria (2026).

Como pode ser observado na análise do *F1-Score* por espécie Figura 05, a *MobileNetV2* sofreu uma degradação severa de desempenho em classes caracterizadas por alta variabilidade intra-específica ou sobreposição visual. Três táxons críticos registraram F1-Scores situados abaixo do patamar de 0,5: a espécie *Bull* (*Carcharhinus leucas*, F1-*Score* de 0,410), *Blacktip* (*Carcharhinus limbatus*, F1-*Score* de 0,414) e *Mako* (*Isurus oxyrinchus*, F1-*Score* de 0,486).

Tais fragilidades indicam que, a despeito do seu menor custo computacional e maior velocidade, a arquitetura *MobileNetV2* compromete a precisão taxonômica em cenários de alta complexidade fenotípica. Em contrapartida, a *EfficientNet-B0* provou ser substancialmente mais apta

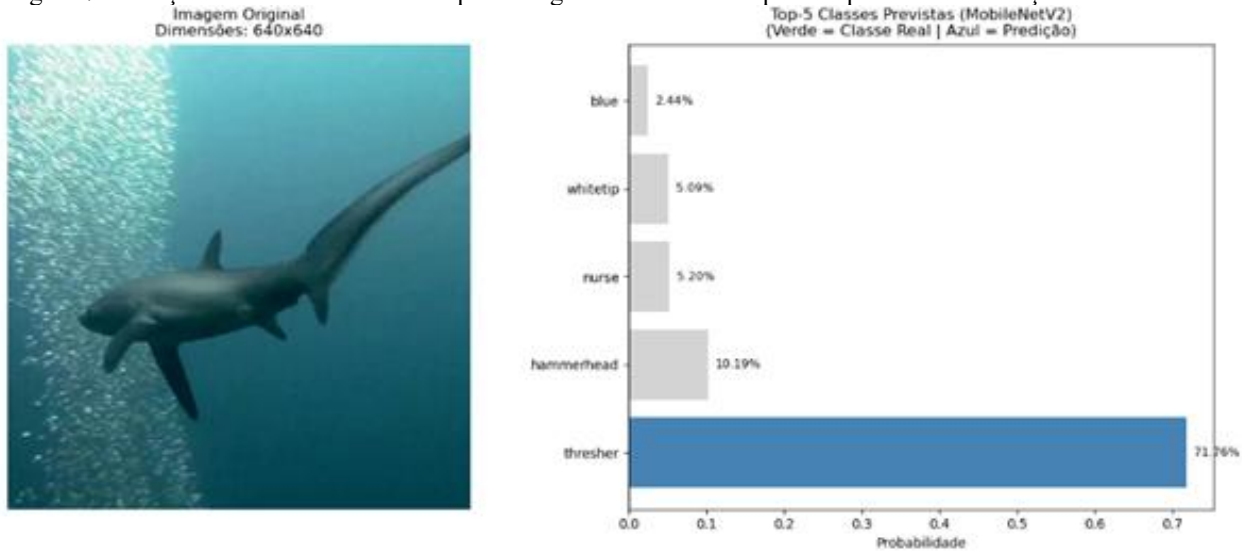
para a tarefa de bioidentificação marinha automatizada, compensando o custo computacional superior com uma distribuição de acertos equilibrada e acurada em todas as 14 espécies analisadas.

Figura 6. Predições usando EfficientNet-B0 para imagens de teste de Alopias vulpinus com indicação da classe real.



Fonte: elaboração própria (2026).

Figura 7. Predições usando MobileNetV2 para imagens de teste de Alopias vulpinus com indicação da classe real.



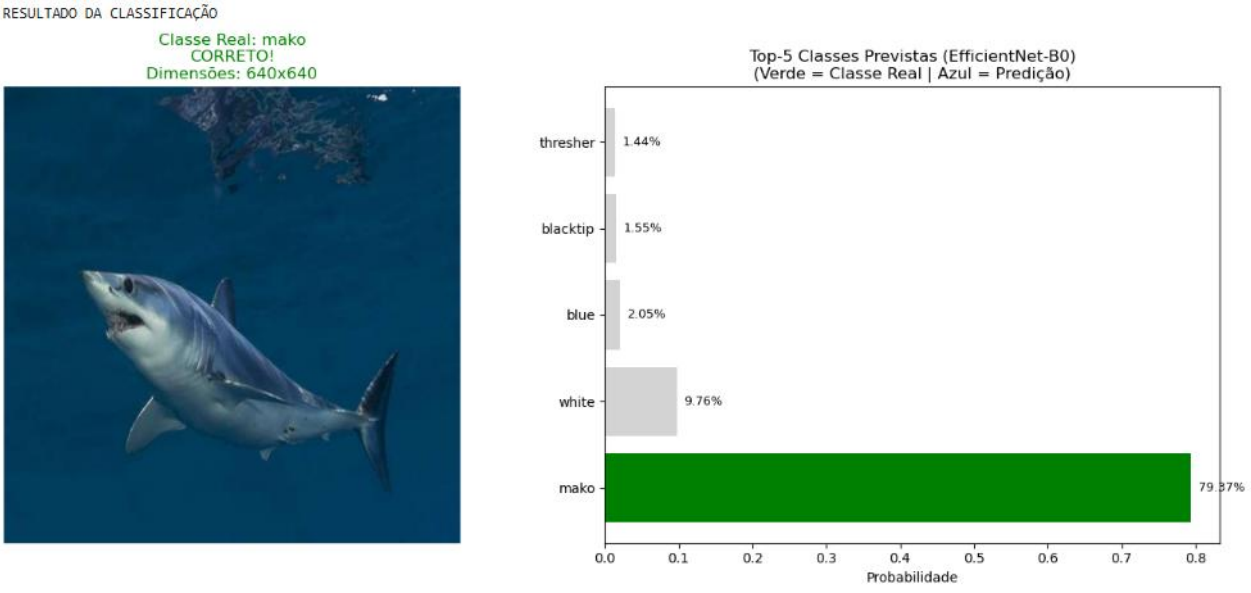
Fonte: elaboração própria (2026).

A avaliação comparativa de predição direta sobre a mesma amostra visual do *Thresher* (*Alopias vulpinus*), apresentada nas Figuras 06 e 07, evidencia o impacto das diferentes arquiteturas de extração de características no grau de certeza da predição. O resultado do modelo *EfficientNet-B0*, presente na Figura 06, demonstrou um desempenho superior de convergência discriminativa ao classificar o espécime como *Thresher* (*Alopias vulpinus*) com 88,09% de probabilidade, mantendo uma dispersão de incerteza muito reduzida entre as demais classes do ranking, tais como *Hammerhead* (*Sphyrna*

mokarran) (5,97%), *Blue* (*Prionace glauca*) (3,05%), *Whale* (*Rhincodon typus*) (0,91%) e *Blacktip* (*Carcharhinus limbatus*) (0,67%). Esse expressivo nível de confiança confirma a eficácia do mecanismo de *compound scaling* da rede em mapear com alta precisão o padrão morfológico contínuo e a extensão da nadadeira caudal.

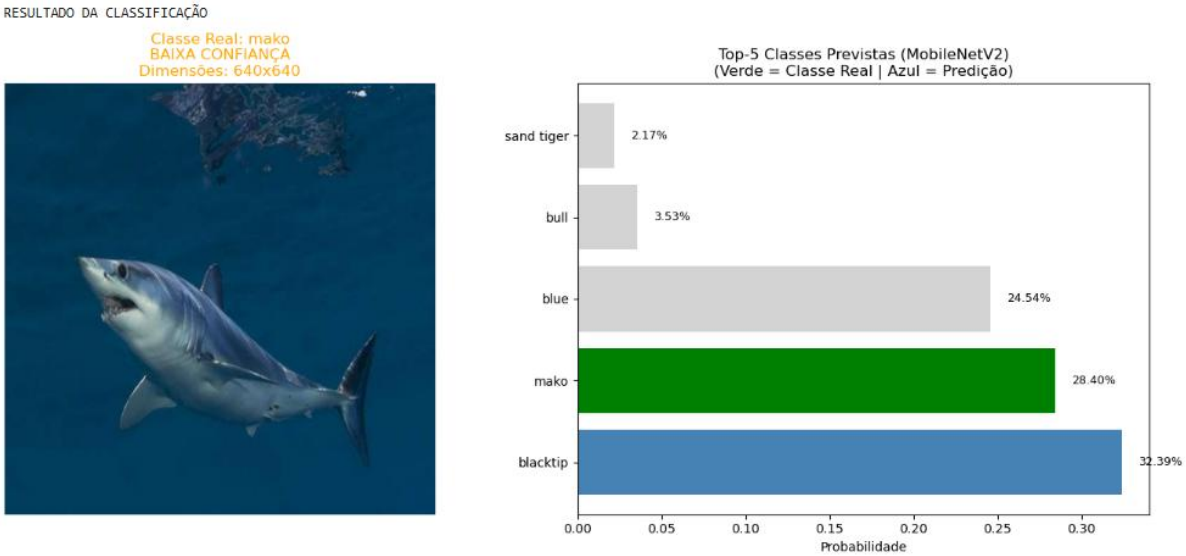
Por sua vez, a *MobileNetV2*, apresentada na Figura 07, embora tenha classificado corretamente a espécie como *Thresher* (*Alopias vulpinus*) com 71,76% de probabilidade, apresentou uma distribuição de massa probabilística mais ruidosa. Em sua predição, a segunda classe mais provável foi *Hammerhead* (*Sphyrna mokarran*) (10,19%), seguida por *Nurse* (*Ginglymostoma cirratum*) (5,20%), *Whitetip* (*Carcharhinus longimanus*) (5,09%) e *Blue* (*Prionace glauca*) (2,44%). Essa diferença de mais de 16 pontos percentuais em favor da *EfficientNet-B0* reforça que, perante ruídos de fundo marinho e colunas de bolhas de água presentes na cena, arquiteturas mais robustas conseguem filtrar artefatos contextuais e alocar maior probabilidade à classe verdadeira, garantindo inferências mais assertivas para a taxonomia automatizada de tubarões.

Figura 8. Predições usando EfficientNet-B0 para imagens de teste de Isurus oxyrinchus com indicação da classe real.



Fonte: elaboração própria (2026).

Figura 9. Predições usando MobileNetV2 para imagens de teste de Isurus oxyrinchus com indicação da classe real.



Fonte: elaboração própria (2026).

A confusão ocorrida na figura 9 entre o tubarão-ponta-preta (*Carcharhinus limbatus*), denominado *Blacktip*, e o tubarão-mako (*Isurus oxyrinchus*) pode ser explicada pela semelhança de algumas características morfológicas externas e pelas condições da imagem utilizada na classificação. Embora pertençam a famílias diferentes, Carcharhinidae e Lamnidae, respectivamente, ambas as espécies apresentam corpo fusiforme e hidrodinâmico, região dorsal acinzentada ou azulada, ventre claro, focinho relativamente pontiagudo e nadadeira dorsal triangular. Essas semelhanças estão relacionadas à adaptação à natação ativa e podem produzir silhuetas visualmente próximas, principalmente em fotografias subaquáticas nas quais o animal aparece distante, inclinado ou parcialmente encoberto. Nesse sentido, a semelhança não indica necessariamente uma relação evolutiva próxima entre as espécies, mas pode ser interpretada como resultado de adaptações funcionais semelhantes ao deslocamento eficiente no ambiente aquático (Carpenter, De Angelis, 2016).

Apesar da semelhança geral, o *Mako* apresenta características morfológicas que permitem diferenciá-lo do *Blacktip*, como o focinho acentuadamente cônico, as nadadeiras peitorais relativamente estreitas, a segunda nadadeira dorsal e a nadadeira anal reduzidas, além da nadadeira caudal em formato de crescente e das quilhas laterais desenvolvidas no pedúnculo caudal. O *Blacktip*, por sua vez, apresenta marcas escuras características nas extremidades de diferentes nadadeiras, constituindo um atributo visual importante para sua identificação. Entretanto, características diagnósticas dessa natureza podem se tornar menos evidentes em fotografias obtidas em condições subaquáticas. Fatores como variações de iluminação, absorção e dispersão da luz pela água, turbidez, desfoque, baixa resolução e diferentes perspectivas de captura podem alterar a aparência das estruturas e reduzir a qualidade das informações visuais disponíveis para a classificação (Xiao *et al*., 2024; Zhang

et al., 2024). Nessas condições, diferenças morfológicas relevantes entre espécies visualmente semelhantes podem se tornar menos evidentes, aumentando a complexidade do processo de identificação.

Os resultados apresentados demonstram essa dificuldade de diferenciação. A predição ocorrida do *EfficientNet-B0* apresentada na figura 08 reconheceu corretamente o exemplar como mako, atribuindo aproximadamente 79,87% de probabilidade à classe verdadeira. A *MobileNetV2* (figura 09), entretanto, classificou o mesmo indivíduo como *Blacktip*, com aproximadamente 32,39%, enquanto atribuiu 28,40% à classe *Mako* e 24,54% à classe *Blue*. A proximidade entre essas probabilidades indica baixa segurança do modelo e revela que os padrões visuais presentes na imagem também foram associados a outras espécies de tubarões com formato corporal e coloração semelhantes. Portanto, não se trata de uma identificação taxonômica segura, mas de uma situação em que o modelo ficou dividido entre diferentes classes visualmente próximas.

Outro fator que pode ter contribuído para o erro é o aprendizado de elementos secundários presentes no conjunto de treinamento. Caso muitas imagens da classe *Blacktip* apresentem tubarões fotografados lateralmente, em águas costeiras azuladas ou sob determinadas condições de luminosidade, a rede neural pode estabelecer associações entre esses elementos contextuais e a classe, utilizando informações do fundo, da posição do animal e do contraste da imagem como referências para a classificação, em vez de se concentrar exclusivamente nos caracteres anatômicos relevantes (Yan *et al*., 2024). Esse aspecto torna-se particularmente importante na classificação de organismos visualmente semelhantes, uma vez que variações ambientais e condições adversas de aquisição das imagens podem comprometer a identificação de características discriminativas e aumentar a ocorrência de classificações equivocadas (Mohammadisabet *et al*., 2025). No caso dos tubarões, essas características podem envolver diferenças relativamente sutis no formato do focinho, nas proporções e posições das nadadeiras, na estrutura da cauda e nos padrões de coloração (Compagno, 2002).

A diferença observada entre os modelos também pode estar relacionada à capacidade de cada arquitetura de extrair e preservar informações visuais discriminativas. Na imagem analisada, a *EfficientNet-B0* aparentemente conseguiu reconhecer melhor os elementos associados ao *Mako*, enquanto a *MobileNetV2* apresentou maior dificuldade na identificação dos caracteres diagnósticos da espécie. Entretanto, esse resultado individual não permite estabelecer a superioridade de uma arquitetura sobre a outra, uma vez que diferentes modelos podem apresentar compromissos distintos entre capacidade preditiva, extração de características e eficiência computacional (Mohammadisabet *et al.*, 2025). Uma comparação mais consistente deve considerar o desempenho global no conjunto de

teste e métricas como acurácia, precisão, revocação, F1-*score* e matriz de confusão, incluindo a análise do comportamento dos modelos em cada classe.

A classificação equivocada do *Mako* como *Blacktip* pode, portanto, estar relacionada à combinação entre semelhança morfológica, baixa visibilidade dos caracteres diagnósticos, perspectiva da fotografia, interferência das condições do ambiente subaquático e limitações do modelo na identificação de características discriminativas. Mohammadisabet *et al.* (2025) destacam que variações nas condições de aquisição das imagens e a elevada similaridade visual entre espécies constituem desafios relevantes para sistemas automáticos de classificação de peixes. Para reduzir esses erros, estratégias como ampliação e balanceamento do conjunto de dados, inclusão de imagens obtidas sob diferentes ângulos e condições ambientais e técnicas de aumento de dados podem favorecer a capacidade de generalização dos modelos. Adicionalmente, os autores demonstram a aplicabilidade de métodos de inteligência artificial explicável, como o *Grad-CAM*, para identificar as regiões da imagem que mais contribuíram para a decisão da rede, possibilitando verificar se a classificação está fundamentada em características relevantes do organismo ou em elementos secundários do cenário.

5 CONCLUSÃO

O presente estudo avaliou comparativamente as arquiteturas de Redes Neurais Convolucionais *EfficientNet-B0* e *MobileNetV2* para a identificação automática de 14 espécies de tubarões, utilizando o conjunto de dados *Shark Species Identification*. A análise dos resultados evidenciou que ambas as arquiteturas apresentam *trade-offs* distintos no tocante ao custo de processamento e à capacidade discriminativa visual.

No que tange ao desempenho preditivo e à capacidade de generalização, a arquitetura *EfficientNet-B0* destacou-se como a opção superior em termos de robustez e precisão taxonômica. Atingindo uma acurácia global de 79,35% no conjunto de teste independente e um F1-*Score* ponderado de 79,31%, o modelo demonstrou rigoroso controle de *overfitting*, apresentando uma variação de apenas 3,38% entre o treino e a validação. Ao avaliar o F1-Scores por classe, foi evidenciado que todas as classes estavam com a métrica acima do limite crítico de 0,5, essa rede deixou claro sua alta competência na extração de feições anatômicas sutis e na filtragem de artefatos visuais típicos do ambiente subaquático, tais como variações de iluminação e colunas de bolhas.

Por outro lado, a *MobileNetV2* confirmou sua vocação para ambientes de hardware restrito ao finalizar o treinamento em 17,45 minutos, sendo cerca de 40,8% mais rápida do que a *EfficientNet-B0*, além de demandar menor quantidade de parâmetros e consumo de memória entre 1,0 e 1,2 GB de RAM. Contudo, essa maior leveza estrutural comprometeu diretamente sua capacidade de

diferenciação. O modelo registrou um F1-*Score* ponderado de 69,17% e sofreu uma queda severa de desempenho, obtendo métricas F1 abaixo de 0,5 em espécies caracterizadas por alta complexidade fenotípica ou expressiva sobreposição visual, a exemplo dos tubarões *Bull*, *Blacktip* e *Mako*.

Essa análise detalhada dos erros revelou que as confusões entre espécimes visualmente semelhantes decorrem da convergência adaptativa do formato corporal fusiforme e de restrições inerentes às fotografias subaquáticas, como limitações de ângulo e perda de contraste de caracteres diagnósticos. Diante desses cenários desafiadores, o mecanismo de *compound scaling* da *EfficientNet-B0* garantiu maior grau de certeza e menor dispersão probabilística nas inferências. Dessa forma, conclui-se que a *EfficientNet-B0* é a arquitetura mais recomendada para a classificação automatizada de elasmobrânquios, uma vez que o ganho expressivo na confiabilidade taxonômica compensa o custo computacional moderadamente superior.

REFERÊNCIAS

BEVIÁ-BALLESTEROS, I. *et al.* eLasmobranc Dataset: An Image Dataset for Elasmobranch Species Recognition and Biodiversity Monitoring. arXiv, 2026. DOI: 10.48550/arXiv.2603.10724.

CARPENTER, K. E.; DE ANGELIS, N. The living marine resources of the Eastern Central Atlantic. Volume 2: Bivalves, gastropods, hagfishes, sharks, batoid fishes, and chimaeras. FAO, 2016.

CHELOTTI, L. D.; SANTOS, P. R. S. dos. Biodiversity and conservation of marine elasmobranchs in the extreme south of Brazil, Southwestern Atlantic. Regional Studies in Marine Science, v. 38, p. 101371, 2020.

COMPAGNO, L. J. V. Sharks of the world: an annotated and illustrated catalogue of shark species known to date. Volume 2: Bullhead, mackerel and carpet sharks (Heterodontiformes, Lamniformes and Orectolobiformes). Rome: Food and Agriculture Organization of the United Nations (FAO), 2002.

DULVY, N. K. *et al.* Challenges and priorities in shark and ray conservation. Current Biology, v. 27, n. 11, p. R565-R572, 2017.

IVANOFF, R.; PENNINO, M. G.; RUFENER, M. C.; VOOREN, C. M.; KINAS, P. G. Modelagem espacial bayesiana para riqueza de elasmobrânquios do extremo sul do Brasil. Revista CEPSUL – Biodiversidade e Conservação Marinha, v. 8, e2019002, 2019.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, v. 521, n. 7553, p. 436–444, 2015.

MIRANDA, J. F. O.; BATISTA, J. A. A. de S.; SANTOS, N. S. dos; ASSIS, B. S. de; CAMPOS, Y. T. A.; CRUZ, H. D. da A. Aplicação de Deep Learning na identificação de espécies de sapos utilizando ResNet-18. Revista Projeção e Docência, [S. l.], v. 17, n. 3, 2026.

MOHAMMADISABET, A.; HASAN, R.; DATTANA, V.; MAHMOOD, S.; HUSSAIN, S. CNN-Based Optimization for Fish Species Classification: Tackling Environmental Variability, Class Imbalance, and Real-Time Constraints. Information, v. 16, n. 2, art. 154, 2025.

SAMI, M.; AHMED, F. Automated seaweed species classification using deep learning and large language models. Scientific Reports, v. 16, art. 63136, 2026.

SANDLER, M. *et al.* Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 4510-4520, 2018.

TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR, p. 6105-6114, 2019.

XIAO, S.; SHEN, X.; ZHANG, Z.; WEN, J.; XI, M.; YANG, J. Underwater image classification based on image enhancement and information quality evaluation. Displays, v. 82, p. 102635, 2024.

YANG, C. *et al.* FishAI: Automated hierarchical marine fish image classification with vision transformer. Engineering Reports, v. 6, n. 12, p. e12992, 2024.

ZHANG, W.; CHEN, G.; ZHUANG, P.; ZHAO, W.; ZHOU, L. CATNet: Cascaded attention transformer network for marine species image classification. Expert Systems with Applications, v. 256, p. 124932, 2024.

ZHOU, Z.; FU, G.-Y.; FANG, Y.; YUAN, Y.; SHEN, H.-B.; WANG, C.-S.; XU, X.-W.; ZHOU, P.; PAN, X. EchoAI: a deep-learning based model for classification of echinoderms in global oceans. Frontiers in Marine Science, v. 10, art. 1147690, 2023. DOI: 10.3389/fmars.2023.1147690.