

## 14<sup>o</sup> Congresso de Inovação, Ciência e Tecnologia do IFSP - 2023

### Avaliação de redes neurais compactas para detecção de incêndios florestais

WESLEY P. TERRÃO<sup>1</sup>, GABRIEL M. ALVES<sup>2</sup>

<sup>1</sup>Graduando em Bacharelado em Ciência da Computação, Bolsista PIBIFSP, IFSP, Câmpus São João da Boa Vista, wesley.terrao@aluno.ifsp.edu.br.

<sup>2</sup>Professor de Ciência da Computação, Câmpus São João da Boa Vista

Área de conhecimento (Tabela CNPq): 1.03.03.05-7 Processamento Gráfico (Graphics).

**RESUMO:** Os incêndios florestais podem se transformar em uma ameaça causando impactos catastróficos. O avanço da tecnologia, sobretudo da Internet das Coisas (IoT) e da Visão Computacional, possibilitou o surgimento de novas abordagens de monitoramento e detecção precoce de incêndios. No entanto, atualmente também tem aumentado o interesse em desenvolver sistemas em tempo real. Tais sistemas utilizam redes neurais profundas para realizar a detecção de incêndios em imagens digitais. Portanto, um novo desafio tem se apresentado e que consiste em utilizar redes neurais leves que possam ser embarcadas em dispositivos com menores capacidades computacionais e que ainda mantenham boa acurácia na classificação das imagens. Neste sentido, o objetivo deste trabalho foi identificar a rede neural compacta apropriada para a detecção de incêndios florestais. Para isso, foram avaliadas as redes compactas MobileNetV3 e EfficientNetB0 em um conjunto de 24.426 imagens de incêndios florestais. Pôde-se observar que as arquiteturas compactas MobileNetV3 e EfficientNetB0 alcançaram acurácias consideradas satisfatórias, pois se aproximaram dos resultados obtidos pelas arquiteturas convencionais VGG19 e ResNet50V2 ao mesmo tempo que ocorreu redução significativa no número de parâmetros e no uso de memória para armazenamento.

**PALAVRAS-CHAVE:** redes neurais convolucionais; aprendizado profundo; mobilenetv3; efficient-netb0.

#### Compact neural network evaluation for forest fire detection

**ABSTRACT:** Forest fires can become a threat, causing catastrophic impacts. The advancement of technology, particularly in the Internet of Things (IoT) and Computer Vision, has enabled the emergence of new approaches for monitoring and early detection of fires. However, there has also been a growing interest in developing real-time systems. These systems use deep neural networks to detect fires in digital images. Therefore, a new challenge has arisen, which is to use lightweight neural networks that can be embedded in devices with lower computational capabilities while still maintaining good accuracy in image classification. In this regard, the aim of this work is to identify the appropriate compact neural network for forest fire detection. To achieve this, the compact architectures MobileNetV3 and EfficientNetB0 were evaluated on a dataset of 24,426 forest fire images. It was observed that the compact architectures MobileNetV3 and EfficientNetB0 achieved satisfactory accuracies, as

they approached the results obtained by conventional architectures VGG19 and ResNet50V2, while significantly reducing the number of parameters and memory usage for storage.

**KEYWORDS:** convolutional neural networks; deep learning; mobilenetv3; efficientnetb0.

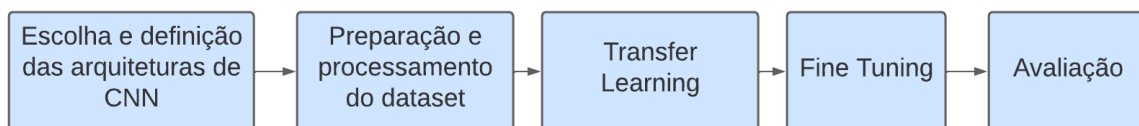
## INTRODUÇÃO

Os incêndios florestais podem se transformar em uma ameaça global, causando impactos catastróficos na natureza, nos animais, nos seres humanos, na economia (Li et al. 2022). Além disso, a contenção pode trazer sérios riscos para os profissionais que atuam no enfrentamento dos incêndios. Outro aspecto é que o tempo gasto entre a detecção do fogo e a notificação aos órgãos competentes é um fator importante na redução dos incêndios. Neste sentido, com o avanço da tecnologia, sobretudo da Internet das Coisas (IoT) e da Visão Computacional, surgiram novas abordagens de monitoramento e detecção precoce de incêndios a exemplo dos sistemas baseados em veículos aéreos não tripulados, também conhecidos por *drones*. No entanto, atualmente também tem aumentado o interesse em desenvolver sistemas em tempo real. Tais sistemas utilizam redes neurais profundas para realizar a detecção de incêndios em imagens digitais. Portanto, um novo desafio tem se apresentado e que consiste em utilizar redes neurais leves que possam ser embarcadas em dispositivos com menores capacidades computacionais e que ainda mantenham boa acurácia na classificação das imagens. Neste sentido, o objetivo deste trabalho é identificar rede neural compacta apropriada para a detecção de incêndios florestais. Para isso, foram avaliadas as redes compactas MobileNetV3 e EfficientNetB0 em um conjunto de 24.426 imagens de incêndios florestais.

## MATERIAIS E MÉTODOS

O desenvolvimento deste trabalho foi estruturado em cinco etapas, conforme apresentado pela Figura 1, que considerou: i) a escolha e definição das arquiteturas de Redes Neurais Convolucionais (CNN); ii) a preparação e processamento do *dataset* de imagens; iii) a transferência de aprendizagem; iv) o *fine tuning*, e; v) a avaliação das redes. Além disso, foi utilizada a linguagem de programação Python juntamente com a biblioteca Tensorflow (Abadi et al. 2015) e o ambiente de desenvolvimento colaborativo Google Colab<sup>TM</sup>.

Figura 1: Etapas para a avaliação de redes neurais compactas para detecção de incêndios florestais.



Fonte: Elaborado pelo autores.

A escolha e definição das arquiteturas CNN a serem avaliadas considerou o uso das redes compactas MobileNetV3 (Howard et al. 2019) e EfficientNetB0 (Tan e Le 2019), a fim de identificar uma arquitetura viável de ser utilizada em dispositivos móveis, ou até mesmo em dispositivos com capacidades reduzidas de processamento. Para fins de comparação da acurácia na detecção de incêndios florestais, as arquiteturas MobileNetV3 e EfficientNetB0 utilizou-se como referências as redes VGG19 (Simonyan e Zisserman 2014) e ResNet50V2 (He et al. 2015), que podem ser consideradas não-compactas uma vez que possuem número elevado de parâmetros e alto uso de memória.

A MobileNetV3 utiliza o conceito de *blocos* que constroem a rede e são projetados para serem eficientes computacionalmente. A rede também utiliza uma combinação de camadas convolucionais e camadas de ativação (como *ReLU6*) para otimizar a arquitetura que conta com a função de ativação não-linear *hard-swish*, que é uma aproximação da função ReLU6, e contribui para a redução da carga computacional e mantém eficiência satisfatória. A MobileNetV3 utiliza de camadas do tipo *Bottleneck*, onde se é aplicada uma sequência de convoluções de diferentes tamanhos para reduzir o número de canais intermediários, ajudando a reduzir a complexidade computacional e o número de parâmetros. Além disso, a rede inclui camadas de detecção de bordas que melhora a precisão do modelo e o bloco *Squeeze-and-Excitation*, que tem como objetivo adaptar a importância de cada canal intermediário em cada camada, melhorando a representação de características relevantes. Por fim, é necessário citar que a MobileNetV3 conta com duas variantes, *small* e *large*, sendo que neste trabalho foi utilizado o modelo *small*, pois é mais adequado para dispositivos com recursos computacionais limitados.

Já a EfficientNetB0 utiliza de escalonamento composto, permitindo realizar um ajuste entre o desempenho e a eficiência do modelo com a variação de três fatores: número de canais, número de camadas e o tamanho da entrada. Utiliza também, blocos de construção *Mobile Inverted Bottleneck MBConv* que são projetados para otimizar a relação entre eficiência e desempenho utilizando convoluções  $1 \times 1$  e  $3 \times 3$ . Além disso, conta a função de ativação não-linear *swish*, semelhante à ReLU com desempenho melhorado em tarefas de *deep learning*. Foi organizada uma base de imagens que consistiu em 24.426 imagens a partir do agrupamento dos *datasets* Fire Dataset (Gamaleldin et al. 2020), DFireDataset (Venâncio et al. 2022) e Forest Fire Dataset (Khan et al. 2022). Ambos datasets rotulam as imagens como *fogo* e *não-fogo* e o intuito do agrupamento foi aumentar a robustez da avaliação, considerando que são bases com imagens de mesmas características e mesmos rótulos. É importante mencionar que 80% da base foi destinada para o treinamento (19.554 imagens) e 20% destinada para o teste e validação (4872 imagens). Antes da etapa de treino a base de imagens foi redimensionada para  $224 \times 224$  pixels.

O treino se deu pelas etapas de *transfer learning* e *fine tuning*, no qual, inicialmente, foi realizado o treino e avaliação por dez épocas das arquiteturas MobileNetV3 *small* e EfficientNetB0 que foram previamente treinadas no *dataset* ImageNet (Deng et al. 2009) . Após isso, todas as camadas foram descongeladas para que pudessem ser treinadas totalmente no *dataset* organizado neste trabalho. Este treino foi realizado juntamente com funções que auxiliam o modelo a treinar somente enquanto estiver obtendo resultados, sem a necessidade de treinar desnecessariamente por muitas épocas, como também foi empregado o uso de funções que auxiliam o modelo com parâmetros de treino que se adaptam durante a sessão de treinamento, entre elas estão *EarlyStopping* e *ReduceLROnPlateau*.

A avaliação consistiu, além de medir a precisão do modelo com métricas de acurácia, precisão e *recall*, realizar o mesmo treinamento com arquiteturas CNN's, ResNet50V2 e VGG19. Além disso, foi avaliada a performance de cada arquitetura, considerando o número de parâmetros e a memória necessária para o armazenamento das mesmas.

## RESULTADOS E DISCUSSÃO

A avaliação das arquiteturas considerou o número total de parâmetros, que são valores ajustáveis aprendidos pela rede durante o processo de treinamento, a acurácia obtida e a quantidade de memória necessária para armazenamento da arquitetura em análise. A Tabela 1 apresenta os resultados obtidos durante a avaliação das arquiteturas.

Pode-se observar pela Tabela 1 que a rede MobileNetV3 apresentou uma redução de 94,20% no

Tabela 1: Resultados obtidos durante a avaliação das arquiteturas para detecção de incêndios florestais.

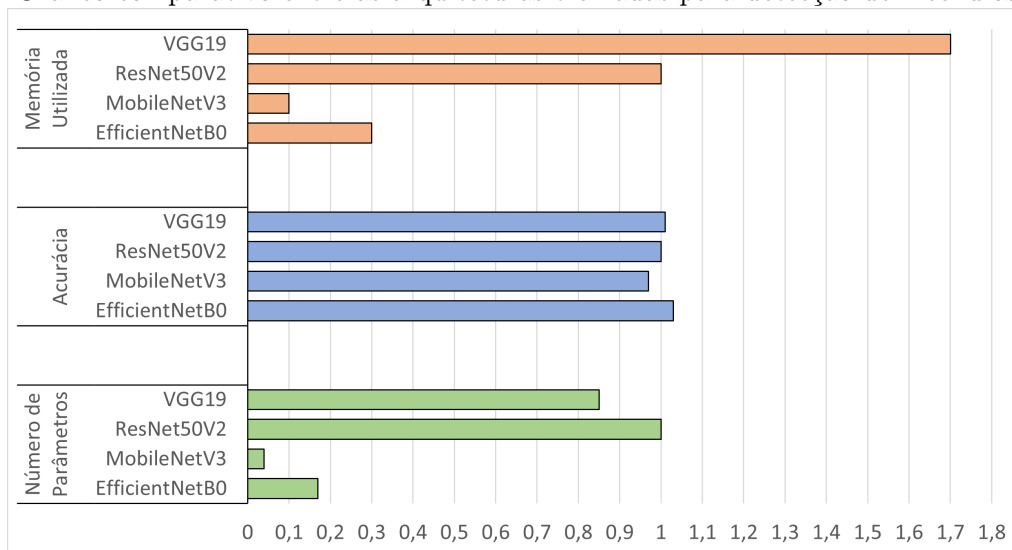
Arquitetura	Número de Parâmetros	Acurácia Obtida(%)	Memória Necessária(Mb)
VGG19	20024897	96,91	459,00
ResNet50V2	23566849	95,21	269,00
MobileNetV3	939697	92,61	26,60
EfficientNetB0	4050852	98,14	86,20

Fonte: Elaborado pelos autores.

uso de memória em relação a rede VGG19 e uma redução de 90,11% no uso de memória em relação a rede ResNet50V2. Ao mesmo tempo que a acurácia foi 4,43% menor do que a VGG19 e 2,73% menor que a ResNet50V2. Já a rede EfficientNetB0 apresentou uma redução de 81,22% no uso de memória em relação a rede VGG19 e 67,95% em relação a rede ResNet50V2. Enquanto que, em termos de acurácia, a EfficientNetB0 apresentou um aumento de 1,27% em relação a rede VGG19 e 3,07% em relação a rede ResNet50V2.

A Figura 2 apresenta um gráfico comparativo entre os modelos treinados e avaliados que tomou como referência os valores obtidos pela arquitetura ResNet50V2. É possível observar que as arquiteturas compactas MobileNetV3 e EfficientNetB0 conseguem atingir acurácias próximas das arquiteturas convencionais, VGG19 e ResNet50V2, mesmo utilizando uma quantidade menor de parâmetros e uma quantidade menor de memória para o armazenamento.

Figura 2: Gráfico comparativo entre as arquiteturas treinadas para detecção de incêndios florestais.



Fonte: Elaborado pelos autores.

As Figuras 3, 4, 5 e Figura 6 apresentam os resultados visuais de detecção de incêndios florestais obtidos nas redes VGG19, ResNet50V2, MobileNetV3 e EfficientNetB0. Para cada rede, foram selecionadas 20 imagens aleatoriamente do conjunto de teste definido no *dataset* organizado neste trabalho e que é composto por 4872 imagens. Para cada imagem apresentada nas Figuras 3, 4, 5 e 6 é possível observar a notação *real* que significa a classe real da imagem, ou seja se é *fogo* ou *não-fogo*, e a notação *pred* que indica a previsão do modelo, e as cores das notações indicam se o modelo acertou (verde) ou erro (vermelho).

Figura 3: Resultados obtidos em 20 imagens pela VGG19 na detecção de incêndios florestais.



Fonte: Elaborado pelos autores.

Figura 4: Resultados obtidos em 20 imagens pela ResNet50V2 na detecção de incêndios florestais.



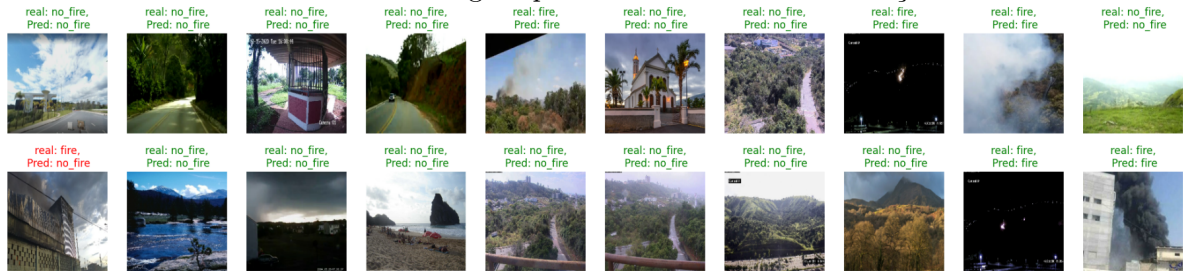
Fonte: Elaborado pelos autores.

Figura 5: Resultados obtidos em 20 imagens pela MobileNetV3 na detecção de incêndios florestais.



Fonte: Elaborado pelos autores.

Figura 6: Resultados obtidos em 20 imagens pela EfficientNetB0 na detecção de incêndios florestais.



Fonte: Elaborado pelos autores.

## CONCLUSÕES

Este trabalho realizou o estudo, a estruturação e a avaliação de duas arquiteturas CNN, consideradas compactas, para a detecção de incêndios. Pôde-se observar que as arquiteturas compactas MobileNetV3 e EfficientNetB0 alcançaram acurácias consideradas satisfatórias, pois se aproximaram dos resultados obtidos pelas arquiteturas convencionais VGG19 e ResNet50V2 ao mesmo tempo que ocorreu redução significativa no número de parâmetros e no uso de memória para armazenamento.

Entre as redes compactas verificou-se a MobileNetV3 é a mais apropriada para a utilização em dispositivos com recursos computacionais limitados uma vez que apresentou número menor de parâmetros, bem como uso de memória menor que a rede EfficientNetB0. Por outro lado, a rede EfficientNetB0 apresentou acurácia 5,63% maior que a MobileNetV3. Os resultados indicam que o uso de redes compactas é viável para a detecção de incêndios e que possuem potencial para serem utilizadas em dispositivos móveis e, até mesmo, em dispositivos com recursos limitados. Como trabalhos futuros, pretende-se investigar a abordagem TinyML para reduzir ainda mais a rede CNN sem comprometer a acurácia, com o apoio de ferramentas como o TensorFlow Lite™.

## CONTRIBUIÇÕES DOS AUTORES

Wesley P. Terrão: concepção, coleta de dados, análise de dados, elaboração do manuscrito, discussão dos resultados. Gabriel M. Alves: concepção, coleta de dados, análise de dados, elaboração do manuscrito, discussão dos resultados. Todos os autores contribuíram com a revisão do trabalho e aprovaram a versão submetida.

## AGRADECIMENTOS

Ao Programa Institucional de Bolsas de Iniciação Científica do IFSP, à Pro Reitoria de Pesquisa (PRE) do IFSP e também à Coordenação de Pesquisa e Inovação (CPI) do Campus.

## REFERÊNCIAS

- Abadi, M. et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.
- Deng, J. et al. ImageNet: A Large-Scale Hierarchical Image Database. In: *CVPR09*. [S.l.: s.n.], 2009.
- Gamaleldin, A.; Atef, A.; Saker, H.; Shaheen, A. *Fire Dataset*. 2020. <https://www.kaggle.com/datasets/phylake1337/fire-dataset>.
- He, K.; Zhang, X.; Ren, S.; Sun, J. *Deep Residual Learning for Image Recognition*. arXiv, 2015. Disponível em: <<https://arxiv.org/abs/1512.03385>>.
- Howard, A. et al. *Searching for MobileNetV3*. arXiv, 2019. Disponível em: <<https://arxiv.org/abs/1905.02244>>.
- Khan, A.; Hassan, B.; Khan, S.; Ahmed, R.; Adnan, A. Deepfire: A novel dataset and deep transfer learning benchmark for forest fire detection. *Mobile Information System*, v. 2022, p. 5358359, 2022.
- Li, X.; Vázquez-López, A.; Sáez, J. S. del R.; Wang, D.-Y. Recent advances on early-stage fire-warning systems: Mechanism, performance, and perspective. *Nano-Micro Letters*, Springer Science and Business Media LLC, v. 14, n. 1, out. 2022. Disponível em: <<https://doi.org/10.1007/s40820-022-00938-x>>.
- Simonyan, K.; Zisserman, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. arXiv, 2014. Disponível em: <<https://arxiv.org/abs/1409.1556>>.
- Tan, M.; Le, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv, 2019. Disponível em: <<https://arxiv.org/abs/1905.11946>>.
- Venâncio, P. V. A. Borges de; Lisboa, A. C.; Barbosa, A. V. An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices. *Neural Computing and Applications*, 2022.