

Rede Neural Convolucional EfficientNet: Uma Proposta de Classificação de Fogo Usando Transferência de Aprendizado

Ítalo José da Silva Melo¹, José Jailton Henrique Ferreira Junior¹

¹Faculdade de Computação – Universidade Federal do Pará (UFPA)
Caixa Postal 479, CEP: 66075-110, Castanhal – PA – Brasil

italosilva1411@gmail.com, jjj@ufpa.br

Abstract. *The alarming increase in wildfires in Brazil in 2024 reinforces the urgency of real-time automatic fire detection systems. This work proposes a binary classifier based on the EfficientNet architecture, pre-trained on ImageNet and fine-tuned using transfer learning and data augmentation techniques. The model was trained and evaluated on a public dataset, achieving 98% accuracy and 100% recall for the fire class, ensuring that no fire cases were missed. The results confirm the robustness and efficiency of the solution, which is suitable for integration into embedded devices and early warning applications. Future work includes expanding datasets, reducing false positives, and conducting field tests for practical validation.*

Resumo. *O crescimento alarmante das queimadas no Brasil em 2024 reforça a urgência de sistemas automáticos de detecção de incêndios em tempo real. Este trabalho propõe um classificador binário baseado na arquitetura EfficientNet, pré-treinada no ImageNet e ajustada por fine-tuning, utilizando técnicas de transfer learning e data augmentation. O modelo foi treinado e avaliado em um conjunto de dados público, obtendo acurácia de 98% e revocação de 100% para a classe fogo, garantindo que nenhum caso de incêndio passasse despercebido. Os resultados confirmam a robustez e a leveza da solução, que se mostra adequada para integração em dispositivos embarcados e aplicações de alerta precoce. Aponta-se como trabalhos futuros a expansão de bases de dados, redução de falsos positivos e testes em cenários de campo para validação prática.*

1. Introdução

O Brasil viveu em 2024 um ano recorde de incêndios, com mais de 30,8 milhões de hectares queimados, um aumento de 79 % em relação a 2023, representando um território maior que o da Itália [do Fogo 2025]. Na Amazônia, principal componente da região Norte, foram consumidos 17,9 milhões de hectares, mais da metade do total nacional [do Fogo 2025]. No Pará, estado-chave da região, foi declarado estado de emergência ambiental devido à intensidade e à extensão das queimadas [Reuters 2024]. Esses números evidenciam a gravidade da situação e reforçam a necessidade de ferramentas para monitoramento e detecção de incêndios, capazes de atuar em tempo real e reduzir o impacto socioambiental.

Neste contexto, torna-se essencial o desenvolvimento de sistemas automáticos de detecção de incêndios que conciliem desempenho, baixo custo computacional e aplicabilidade em dispositivos de campo. As redes neurais convolucionais (CNNs) destacam-se

por sua capacidade de aprender padrões visuais complexos e apresentar resultados expressivos em tarefas de classificação de imagens [Diaconu 2023]. Entre essas arquiteturas, a EfficientNet se sobressai por sua leveza e eficiência, viabilizando aplicações em dispositivos com recursos computacionais limitados, sem comprometer a acurácia [Khan and colleagues 2022].

Assim, este estudo propõe o uso da EfficientNet pré-treinada no ImageNet e ajustada por meio de *fine-tuning* para a tarefa de classificação binária (fogo/não-fogo). Foram empregadas técnicas de *transfer learning* e *data augmentation* para ampliar a capacidade de generalização do modelo e mitigar riscos de sobreajuste. O conjunto de dados FIRE, amplamente utilizado na literatura, foi dividido em treino, validação e teste, garantindo avaliação justa e representativa do desempenho do modelo.

O treinamento envolveu a experimentação de hiperparâmetros como taxa de aprendizado, uso de *callbacks* para ajuste dinâmico e critério de *early stopping*, assegurando convergência eficiente. A avaliação foi conduzida com métricas clássicas de aprendizado de máquina, tais como: acurácia, precisão, *recall* e *F1-score*, complementadas pela análise da matriz de confusão para investigar falsos positivos e negativos.

Os resultados obtidos demonstraram excelente desempenho, com acurácia de 98% no conjunto de teste e *recall* de 100% para a classe fogo, garantindo que nenhum caso de incêndio passasse despercebido. A ocorrência de poucos falsos positivos mostra-se aceitável no contexto de sistemas de alerta, em que a prioridade é minimizar falsos negativos. Assim, esta pesquisa fornece uma base para o uso de modelos leves em sistemas de monitoramento automático de incêndios, com potencial de aplicação em tempo real e integração futura a dispositivos embarcados.

2. Fundamentação Teórica

2.1. Redes Neurais Convolucionais (CNNs)

As redes neurais convolucionais são um dos pilares do aprendizado profundo para dados visuais. Seu funcionamento baseia-se na aplicação de filtros convolucionais que extraem características locais em diferentes escalas, criando mapas de ativação hierárquicos. Esse mecanismo permite detectar bordas, texturas e formas mais complexas à medida que as camadas se aprofundam [He et al. 2016]. Diferente das redes densas, as CNNs reduzem drasticamente o número de parâmetros devido ao compartilhamento de pesos, o que as torna mais eficientes e menos propensas a sobreajuste.

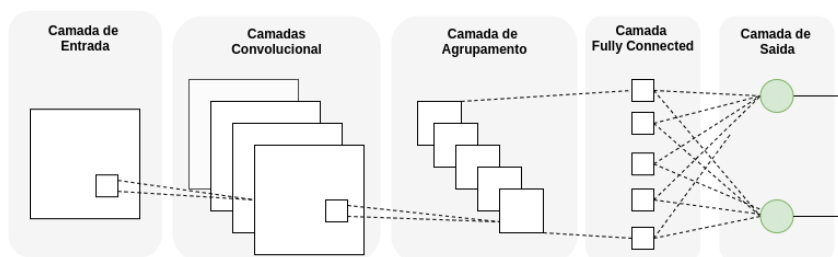


Figure 1. Fluxo de processamento de uma CNN típica, com camadas de convolução, pooling e classificação totalmente conectada. Fonte: [Barbosa et al. 2021]

Arquiteturas clássicas como AlexNet, VGG e ResNet impulsionaram a popularização das CNNs ao bater recordes no ImageNet, mostrando que redes mais profundas aumentam a capacidade de aprendizado [Krizhevsky et al. 2012, He et al. 2016].

2.2. EfficientNet

A *EfficientNet* é uma família de CNNs proposta para otimizar a relação entre profundidade, largura e resolução através de um método chamado *compound scaling*, que expande uniformemente essas três dimensões segundo um fator de escala [Tan and Le 2019]. O resultado é uma rede com alta acurácia e menor custo computacional em comparação com modelos clássicos como ResNet ou Inception. Sua estrutura está descrita na Figura 2.

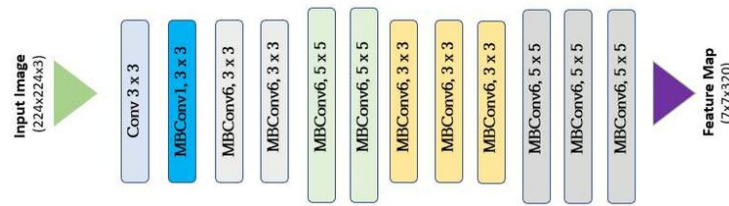


Figure 2. Estrutura da rede EfficientNet. Fonte: [Hisaria et al. 2024].

Sua estrutura é baseada em blocos MBConv com convoluções separáveis em profundidade, reduzindo FLOPs e mantendo desempenho robusto. Essa característica é importante em aplicações em tempo real e dispositivos *edge*, como drones ou câmeras de vigilância.

2.3. Transfer Learning

O *transfer learning* consiste em reaproveitar pesos de uma rede previamente treinada em uma tarefa genérica, como classificação no *ImageNet*, para uma tarefa específica. Essa técnica é especialmente útil quando há limitação de dados, pois diminui a necessidade de grandes bases de treinamento e acelera a convergência [He et al. 2016].

Formalmente, se $\mathcal{D}_S$ é o domínio de origem com tarefa $\mathcal{T}_S$ e $\mathcal{D}_T$ o domínio de destino com tarefa $\mathcal{T}_T$, o aprendizado por transferência busca melhorar $P(y_T|x_T)$ aproveitando $P(y_S|x_S)$ onde $\mathcal{D}_S \neq \mathcal{D}_T$ ou $\mathcal{T}_S \neq \mathcal{T}_T$.

2.4. Fine-Tuning

Após o carregamento dos pesos pré-treinados, realiza-se o *fine-tuning*, destravando camadas superiores da rede e ajustando seus pesos ao novo domínio [Yosinski et al. 2014]. Esse processo melhora a especialização do modelo, permitindo que ele capture peculiaridades do conjunto de dados alvo sem perder a robustez aprendida.

2.5. Data Augmentation On-the-Fly

O *data augmentation on-the-fly* consiste em aplicar transformações aleatórias às imagens durante o treinamento, sem gerar novas amostras físicas no disco. Operações como rotações, inversões, mudanças de brilho e zoom aumentam a diversidade das amostras, reduzindo o risco de sobreajuste [Shorten and Khoshgoftaar 2019].

2.6. Métricas de Avaliação

A análise do desempenho de modelos de classificação binária exige métricas que avaliem não apenas a taxa global de acerto, mas também o equilíbrio entre erros de cada classe. A acurácia é definida por:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

A precisão mede a proporção de predições corretas da classe positiva:

$$Precision = \frac{TP}{TP + FP}$$

O *recall* indica a sensibilidade do modelo à classe positiva:

$$Recall = \frac{TP}{TP + FN}$$

Por fim, o F1-score é a média harmônica de precisão e *recall*:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

A matriz de confusão apresenta essas relações de forma visual, permitindo identificar se há predominância de falsos positivos ou falsos negativos.

3. Trabalhos correlatos

Diversos estudos têm explorado o uso de arquiteturas leves de redes neurais convolucionais para a detecção automática de fogo, com destaque para a EfficientNet, dada sua eficiência computacional e aplicabilidade em sistemas em tempo real.

Nesse sentido, o estudo de Fernandes et al. (2024) propõe um sistema automático de detecção de incêndios florestais com EfficientNets (B0–B2) e saída em mosaico para localizar plumas de fumaça em longas distâncias, usando um dos maiores conjuntos públicos reportados para fumaça em torres de vigilância (26,204 imagens com fumaça e 51,075 sem). As redes superam *InceptionV3* e *MobileNetV2* no mesmo conjunto, alcançando aproximadamente 89,2% de verdadeiros positivos e 0,306% de falsos positivos no teste; o foco é vigilância de fumaça ao ar livre, com ênfase em pAUROC e mitigação de desbalanceamento [Fernandes et al. 2024].

Já os autores Farman et al. (2023) apresentam o *E-EFNet*, um modelo leve para detecção de fogo baseado na rede EfficientNet (B0/B3) como *backbone*, acoplado a codificadores densamente conectados, visando reduzir falsos alarmes e viabilizar inferência em *edge*. Em *benchmarks* clássicos (Foggia e Yar), reportam melhor acurácia e FPR/FNR que SOTA e demonstram *throughput* em CPU e *Raspberry Pi*. O foco é reconhecimento em vídeos/quadros de vigilância, não um protocolo padronizado no FIRE [Farman et al. 2023].

Wang et al. (2024), por sua vez, propõe-se um método híbrido *CBAM–EfficientNetV2* com *label smoothing* para melhorar a atenção a alvos pequenos (chamas/fumaça). Os autores relatam até 98,9% de acurácia em reconhecimento de fogo (e 98,5% para fumaça), com ganhos em precisão/*recall*/F1 frente a *baselines*; a avaliação é realizada em *datasets* de pesquisa (não FIRE) e o objetivo é robustez a alvos diminutos [Wang et al. 2024].

Outrossim, Sreenath et al. (2022), apresentam uma abordagem de *ensemble* para detecção de incêndios aéreos combinando *EfficientNet-B5* e *DenseNet-201* para classificação (fogo/não-fogo), além de *EfficientSeg/TransUNet/TransFire* para segmentação de fogo. Utilizando o *FLAME* (imagens RGB/IR embarcadas), os autores reportam superioridade do *ensemble* sobre *EfficientNet-B1–B5* isoladas e outras CNNs, com desempenho forte em cenários de UAV; o foco é aéreo/multimodal, não o FIRE [Ghali and Akhloufi 2024].

Nesse sentido, os quatro trabalhos confirmam a efetividade da família *EfficientNet* em fogo/fumaça, mas diferem quanto à tarefa e aos dados. Para sintetizar os principais achados dos trabalhos que utilizam a *MobileNetV2* para detecção de fogo, a Tabela 1 resume objetivo, lacunas e dataset utilizado. Assim, este trabalho destaca-se por realizar a classificação binária fogo/não-fogo com a rede *EfficientNet* pré-treinada e fine-tuning diretamente. Além disso, o trabalho posiciona-se como uma solução leve e reproduzível.

Table 1. Comparativo dos trabalhos com EfficientNet

Referência	Objetivo	Lacunas (relação a este trabalho)	Dataset
Fernandes et al. (2024)	Detectar fumaça em longa distância com <i>EfficientNet</i> e saída em mosaico, reduzindo FPs com pAUROC/ajuste de classe.	Foco em fumaça (não chamas) e vigilância de torres; maior complexidade (mosaico).	Conjunto próprio de torres (26k fumaça / 51k não-fumaça)
Farman et al. (2023)	<i>E-EFNet: EfficientNet-B0/B3</i> + codificadores densos para reduzir falsos alarmes e operar em <i>edge</i> .	Mistura cenários (Foggia/Yar) e vídeos; arquitetura mais complexa (autoencoders densos) que pode elevar custo de implantação.	Foggia; Yar
Wang et al. (2024)	<i>CBAM–EfficientNetV2</i> com <i>label smoothing</i> para alvos pequenos e maior generalização.	Ênfase em atenção (overhead) e <i>datasets</i> de pesquisa não reproduzíveis; foco parcial em fumaça.	Bases de pesquisa
Sreenath et al. (2022)	<i>Ensemble</i> com <i>EfficientNet-B5</i> + <i>DenseNet-201</i> para classificação (fogo/não-fogo) e segmentação com <i>EfficientSeg/Transformers</i> .	Cenário aéreo/multimodal (RGB/IR); uso de <i>ensemble</i> aumenta complexidade e custo computacional.	FLAME (UAV; RGB/IR)

4. Materiais e Métodos

4.1. Arquitetura de CNN

Para a etapa de classificação, foi escolhida a arquitetura *EfficientNet*, uma rede neural convolucional (CNN) leve e eficiente, amplamente utilizada em tarefas de visão com-

putacional. A EfficientNet foi projetada para ser computacionalmente econômica, apresentando baixo número de parâmetros e reduzido custo de inferência, o que a torna adequada para aplicações em tempo real e em dispositivos com recursos limitados [Fernandes et al. 2024].

A escolha dessa arquitetura justifica-se por seu bom equilíbrio entre acurácia e eficiência, além de ser amplamente validada na literatura para tarefas de classificação de imagens de fogo. Sua estrutura baseada em *depthwise separable convolutions* reduz significativamente o custo computacional sem comprometer o desempenho, tornando-a apropriada para o problema de detecção binária de fogo e não-fogo.

Optou-se por utilizar essa rede pré-treinada no conjunto de dados ImageNet, aproveitando a técnica de *transfer learning*. Essa abordagem permite inicializar a rede com pesos já otimizados para reconhecimento de características visuais genéricas, acelerando o processo de convergência, reduzindo a necessidade de grandes quantidades de dados e mitigando o risco de *overfitting*. A rede foi posteriormente ajustada (*fine-tuning*) para a tarefa específica de classificação binária proposta neste trabalho.

4.2. Base de Dados

Para o treinamento e avaliação da classificação binária (fogo / não-fogo), utilizou-se o banco de dados público *FIRE Dataset*, disponível na plataforma Kaggle [Saied 2018]. Esse conjunto de dados é composto por 999 imagens que representam duas classes distintas: uma classe com cenas contendo incêndios (“fire”) e outra com cenas que não apresentam fogo (“non-fire”). As imagens abrangem diferentes condições de iluminação, cenários naturais e artificiais.

Estando divididas em 2 pastas, a primeira com 755 imagens classificadas como fogo contendo fotos de fogueiras ao ar livre, algumas delas com fumaça pesada, fogo em ambientes fechados, florestas, etc e a segunda com fotos que remetem a espaços de não fogo, contendo 244 imagens da natureza, tais como: floresta, árvore, grama, rio, pessoas, floresta com neblina, lago, animal, estrada e cachoeira.

A escolha do *FIRE Dataset* justifica-se por sua disponibilidade pública, facilidade de acesso e uso consolidado na literatura para tarefas de detecção de fogo. Além disso, a diversidade visual do conjunto favorece o aprendizado de representações generalizáveis, aumentando a capacidade de o modelo detectar incêndios em diferentes contextos.

Apesar disso, é importante reconhecer que seu tamanho relativamente pequeno pode limitar a capacidade de generalização do modelo treinado. Conjuntos de dados reduzidos podem não capturar toda a variabilidade presente em cenários reais, como diferentes intensidades de chama, condições climáticas, fontes de iluminação ou tipos de vegetação. Ainda assim, a escolha de uma base de dados menor foi deliberada, considerando as restrições de poder computacional disponíveis para o desenvolvimento deste trabalho. O processamento foi realizado em um Notebook Samsung Essentials E30 (NP350), equipado com processador Intel Core i5, 4 GB de memória RAM, HD de 1 TB e SSD adicional, o que impôs limitações de desempenho, especialmente em tarefas de treinamento de redes neurais convolucionais.

Dessa forma, optou-se por não utilizar bases mais extensas nem realizar a execução em plataformas de computação em nuvem, priorizando a viabilidade experi-

mental e o controle do ambiente de desenvolvimento. No entanto, para mitigar o impacto dessa problemática (tamanho reduzido do conjunto de dados e evitar o sobreajuste (overfitting)), foram aplicadas estratégias de aumento de dados (data augmentation) durante o treinamento.

O banco de dados foi dividido em três subconjuntos: treino, validação e teste. Seguiu-se a estratégia de reservar aproximadamente 20% das imagens para teste, garantindo que este conjunto permanecesse completamente separado do processo de treinamento e servisse apenas para avaliação final do modelo. Dessa forma, o conjunto de teste ficou composto por 200 imagens, sendo 158 da classe *fire* e 42 da classe *non-fire*.

As imagens restantes foram utilizadas na etapa de treinamento e subdivididas em treino (aproximadamente 80%) e validação (aproximadamente 20%), mantendo o balanceamento entre as classes.

A Fig. 3 apresenta exemplos visuais de cada classe, extraídos do próprio banco: à esquerda uma imagem da classe “*fire*” mostrando chamas visíveis, e à direita uma imagem da classe “*non-fire*”, sem presença de fogo.

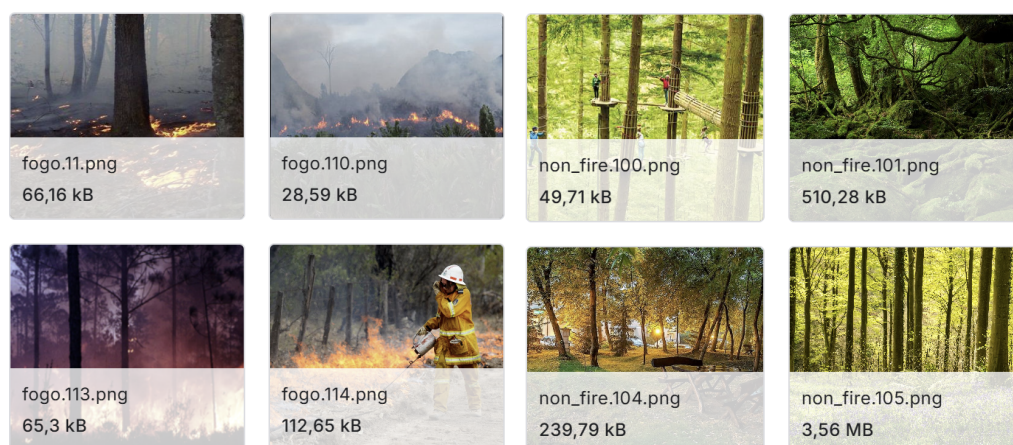


Figure 3. Exemplos de imagens do *FIRE Dataset*: (a) classe *fire*, (b) classe *non-fire*.

4.3. Pré-processamento das Imagens

As imagens do banco de dados FIRE foram inicialmente redimensionadas para 224×224 pixels, dimensão compatível com a entrada esperada pela *MobileNetV2*. Em seguida, os valores dos pixels foram normalizados para o intervalo $[0, 1]$, de modo a facilitar o processo de otimização e acelerar a convergência do modelo.

Posteriormente, foram aplicadas técnicas de *data augmentation* para aumentar a variabilidade das amostras durante o treinamento. Foram incluídas transformações como rotações aleatórias, inversões horizontais, pequenas translações, *zoom* e ajustes de brilho. Essas transformações foram implementadas de forma dinâmica (*on-the-fly*), utilizando as funcionalidades de geração de lotes de imagens do *Keras*. Dessa forma, a cada época de treinamento, novas variações são geradas a partir das imagens originais, o que torna o conjunto de treino virtualmente infinito em termos de combinações possíveis.

É importante destacar que, por essa razão, o número total de imagens não foi

alterado, isso pois não houve criação de novas amostras físicas no disco. As imagens originais permanecem em número constante nos subconjuntos de treino, validação e teste, sendo apenas transformadas em tempo de execução. Essa estratégia é fundamental para reduzir o *overfitting* sem a necessidade de ampliar manualmente o banco de dados.

4.4. Hiperparâmetros e Configuração de Treinamento

Para o treinamento do modelo, utilizou-se a arquitetura *EfficientNet* pré-treinada no conjunto *ImageNet*, com a camada de classificação original removida e substituída por uma camada de *Global Average Pooling* seguida de uma camada totalmente conectada com ativação sigmoide para produzir a saída binária. A função de perda empregada foi a entropia cruzada binária, e a métrica monitorada foi a acurácia no conjunto de validação. O otimizador utilizado foi o *Adam*, com taxa de aprendizado inicial escolhida em escala logarítmica e ajustada automaticamente durante o treinamento por meio de um *callback* do tipo *ReduceLROnPlateau*, reduzindo o valor sempre que a métrica de validação se mantinha estagnada. O treinamento foi realizado com mini-lotes embaralhados a cada época e com aplicação de *data augmentation* dinâmica (*on-the-fly*), incluindo rotações, translações, inversões horizontais, variação de brilho e *zoom*, de modo a gerar maior diversidade no conjunto de treino sem alterar o número total de imagens disponíveis. O número de épocas foi definido como 100, porém foi empregado *early stopping* para interromper o treinamento ao detectar ausência de melhora na perda de validação por um número consecutivo de épocas, preservando o melhor modelo obtido por meio de um *ModelCheckpoint*. Ao final do processo, a convergência foi alcançada por volta da época 40, com perda de validação em torno de 0,04 e acurácia de validação aproximada de 98,1%, indicando que o modelo atingiu um bom ponto de generalização sem apresentar sinais de sobreajuste.

4.5. Métodos de Avaliação do Modelo

Por fim, o modelo foi avaliado utilizando métricas amplamente aceitas em problemas de classificação binária, como acurácia, precisão, revocação e F1-score. A utilização dessas métricas, em conjunto com a matriz de confusão, permitiu analisar não apenas o desempenho geral, mas também o comportamento do classificador diante de falsos positivos e falsos negativos.

5. Resultados e Discussão

O modelo *EfficientNet* ajustado para a tarefa binária de detecção de fogo apresentou excelente desempenho tanto na fase de treinamento quanto na avaliação em dados não vistos. A Figura 4 apresenta as curvas de perda e acurácia ao longo das épocas. Observa-se uma rápida convergência já nas primeiras dez épocas, com estabilização da acurácia de validação próxima de 98% e perda de validação em torno de 0,04, indicando boa capacidade de generalização e ausência de sobreajuste.

A Tabela 2 resume os resultados obtidos no conjunto de teste. Nota-se que o modelo alcançou 98% de acurácia, com perda de teste de aproximadamente 0,0777. Os valores de precisão, revocação e *F1-score* são elevados para ambas as classes, com destaque para a revocação de 100% para a classe *fire*, fundamental para aplicações de detecção de incêndios, uma vez que reduz a probabilidade de deixar casos positivos sem detecção.

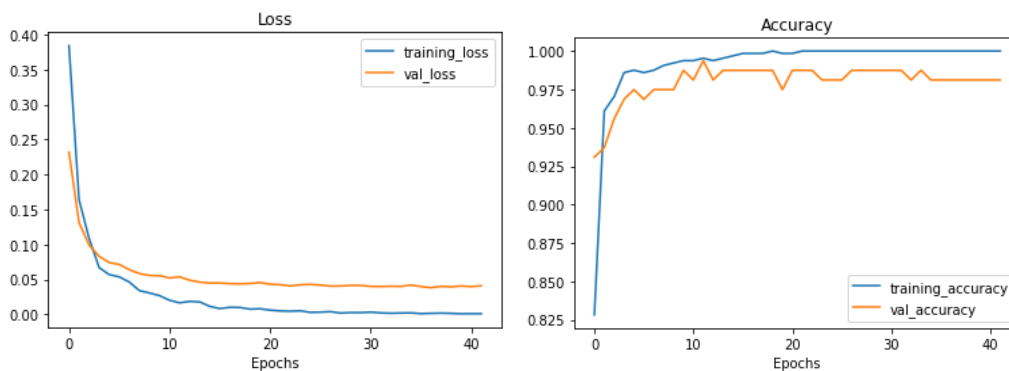


Figure 4. Evolução da perda (à esquerda) e da acurácia (à direita) durante o treinamento e validação.

Table 2. Métricas de desempenho no conjunto de teste.

Classe	Precisão	Revocação	F1-score	Suporte
fire_images	0,975	1,000	0,988	158
non_fire_images	1,000	0,905	0,950	42
Acurácia		0,980		200
Média Macro	0,988	0,952	0,969	200
Média Ponderada	0,980	0,980	0,980	200

A Figura 5 apresenta a matriz de confusão obtida. Nota-se que o modelo classificou corretamente todas as imagens de fogo, não apresentando falsos negativos nessa classe. Houve quatro falsos positivos, em que imagens da classe *non-fire* foram preditas como fogo. Esse comportamento, embora ligeiramente reduza a precisão geral, é aceitável em um contexto de sistemas de alerta, em que é preferível detectar todos os casos de fogo, ainda que com alguns alarmes falsos.

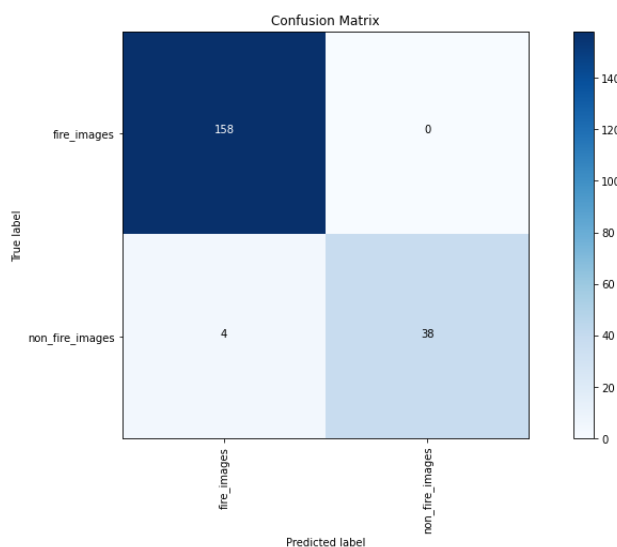


Figure 5. Matriz de confusão para o conjunto de validação.

Sobre as predições realizadas pelo modelo, a maioria dos casos foi classificada

corretamente, incluindo cenas de incêndios em diferentes condições de iluminação, perspectiva e intensidade. As poucas classificações incorretas ocorreram em imagens de vegetação sem fogo, que possuíam cores ou padrões semelhantes aos de incêndios florestais, indicando que o modelo pode ter sensibilidade aumentada a tonalidades alaranjadas ou áreas com contraste intenso.

De modo geral, os resultados confirmam a efetividade da abordagem de *transfer learning* com *EfficientNet* para a detecção de fogo em imagens. A alta revocação na classe *fire* é especialmente desejável para sistemas de monitoramento, visto que a prioridade é minimizar falsos negativos. A ligeira perda de precisão para a classe *non-fire* pode ser mitigada em aplicações práticas combinando este modelo com filtros adicionais ou etapas de validação manual para reduzir alarmes falsos.

6. Conclusão

O presente trabalho teve como objetivo desenvolver e avaliar um modelo de classificação binária para detecção automática de fogo em imagens, utilizando a arquitetura *EfficientNet* com técnicas de *fine-tuning*, *transfer learning* e *data augmentation*. O propósito central foi investigar se uma rede leve e eficiente, treinada a partir de um conjunto de dados público e de tamanho reduzido, poderia alcançar desempenho competitivo e aplicabilidade em sistemas de monitoramento em tempo real. Os resultados obtidos confirmaram o alcance desse objetivo, demonstrando que a abordagem proposta é capaz de detectar incêndios de forma precisa, rápida e com baixo custo computacional.

O modelo atingiu 98% de acurácia global e 100% de revocação para a classe fogo, garantindo que nenhum caso de incêndio fosse deixado de ser identificado, característica essencial em sistemas de alerta precoce. Ainda que tenha ocorrido um pequeno número de falsos positivos (quatro casos), esse comportamento é aceitável no contexto da vigilância ambiental, em que a prioridade é evitar falsos negativos, ou seja, impedir que incêndios reais passem despercebidos.

Além disso, a *EfficientNet* confirmou sua adequação como arquitetura de alto desempenho e baixo custo, validando a escolha por um modelo leve e eficiente que pode ser adaptado a dispositivos de campo ou *edge computing*. O uso de técnicas de *data augmentation on-the-fly* e *early stopping* mostrou-se fundamental para evitar o sobreajuste e garantir a generalização do modelo. O trabalho também demonstrou que é possível obter resultados expressivos mesmo com restrições de hardware.

Durante a análise qualitativa, verificou-se que as imagens com vegetação densa e tonalidades alaranjadas foram as principais responsáveis pelos falsos positivos. Isso ocorreu porque o modelo, ao aprender padrões visuais característicos do fogo, tais como brilho intenso, coloração avermelhada e texturas irregulares, acabou associando erroneamente certas combinações de cor e contraste típicas do pôr do sol, folhagens secas ou reflexos de luz em superfícies ao comportamento visual do fogo. Esse tipo de confusão é comum em classificadores baseados em visão computacional, pois tais padrões cromáticos e texturais se sobrepõem parcialmente aos das chamas reais.

De modo geral, o trabalho comprovou a efetividade da combinação entre *transfer learning*, *fine-tuning* e *data augmentation* para o problema de detecção de fogo, mostrando que redes pré-treinadas podem ser ajustadas com êxito mesmo em conjuntos

pequenos e heterogêneos. A abordagem proposta se destaca por ser reprodutível, leve e de fácil implementação, atendendo aos requisitos de desempenho e eficiência necessários para aplicações em tempo real.

Assim, este trabalho contribui para o avanço da pesquisa em detecção automática de incêndios, oferecendo uma solução leve, reprodutível e com resultados de alto impacto para monitoramento em tempo real. Com a continuidade das pesquisas e o desenvolvimento das melhorias sugeridas, este estudo pode servir como base para sistemas maiores de prevenção e resposta a incêndios em alta escala, auxiliando na mitigação de impactos socioambientais, especialmente no norte onde este estudo está sendo desenvolvido.

6.1. Trabalhos Futuros

Apesar dos bons resultados, o estudo apresenta algumas limitações. O conjunto de dados utilizado contém poucas imagens, o que pode limitar a diversidade de cenários representados e impactar a generalização do modelo em ambientes distintos. Além disso, o modelo apresentou quatro falsos positivos, sugerindo que imagens com tonalidades alaranjadas ou padrões semelhantes a fogo podem induzir erros. Outro ponto é que o estudo concentrou-se exclusivamente em classificação binária, sem abordar localização espacial de incêndios ou detecção em fluxos de vídeo em tempo real.

Para trabalhos futuros, pretende-se tratar a expansão do conjunto de dados, incorporando outras bases públicas para verificar o modelo em diferentes condições climáticas, horários e regiões. Também será investigado a integração do modelo em dispositivos *edge*, como drones ou câmeras de vigilância, de modo a validar sua viabilidade em cenários de campo.

Para trabalhos futuros, pretende-se tratar a expansão do conjunto de dados, incorporando outras bases públicas para verificar o modelo em diferentes condições climáticas, horários e regiões. Também será investigada a integração do modelo em dispositivos *edge*, como drones ou câmeras de vigilância, de modo a validar sua viabilidade em cenários de campo.

Além disso, propõe-se a realização de testes em fluxos de vídeo em tempo real, avaliando o desempenho do modelo em situações dinâmicas e contínuas, aproximando-se de aplicações reais de monitoramento. Pretende-se também implementar protótipos em hardware embarcado, empregando plataformas de baixo custo e ampla disponibilidade, como o Raspberry Pi 4 Model B e a NVIDIA Jetson Nano, que oferecem suporte a processamento paralelo via GPU e execução de modelos em frameworks como TensorFlow Lite ou PyTorch Mobile. Esses dispositivos podem ser integrados a módulos de câmeras compatíveis (por exemplo, Raspberry Pi Camera Module v2 ou Arducam) para captura contínua de vídeo e acoplados a sensores auxiliares, como sensores de temperatura e gases (MQ-2, DHT22), de modo a enriquecer o processo de detecção com dados multimodais. O objetivo é construir um sistema embarcado autônomo, capaz de detectar fogo em tempo real com baixo consumo energético e operação em áreas remotas, aproximando o modelo de uma aplicação prática de vigilância ambiental e alerta precoce.

Outro aspecto relevante para trabalhos futuros é a comparação sistemática entre diferentes arquiteturas de redes neurais, como ResNet, MobileNet, VGG e DenseNet, a fim de evidenciar de forma quantitativa as vantagens da EfficientNet em termos de desempenho, custo computacional e capacidade de generalização. Essa análise comparativa

poderá reforçar a escolha do modelo adotado neste estudo e indicar eventuais oportunidades de aprimoramento.

Por fim, sugere-se investigar estratégias para redução de falsos positivos, como o uso de filtros temporais para análise de sequência de quadros em vídeo, integração de informações de sensores térmicos ou ambientais, e aplicação de técnicas de *post-processing* baseadas em aprendizado por consenso entre múltiplos modelos. Tais medidas podem aumentar a confiabilidade do sistema e reduzir alarmes incorretos em ambientes com iluminação variável ou objetos com tonalidades semelhantes ao fogo.

References

- Barbosa, G., Bezerra, G., Medeiros, D., Andreoni, M., and Menezes, D. (2021). *Segurança em Redes 5G: Oportunidades e Desafios em Detecção de Anomalias e Predição de Tráfego Baseadas em Aprendizado de Máquina*, pages 145–189.
- Diaconu, B. M. (2023). Recent advances and emerging directions in fire detection with machine learning. *Fire*.
- do Fogo, M. . M. (2025). Área queimada no brasil cresce 79% em 2024 e supera os 30,8 milhões de hectares. <https://brasil.mapbiomas.org/2025/01/22/area-queimada-no-brasil-cresce-79-em-2024-e-supera-os-30-milhoes-de-hectares/>.
- Farman, H., Nasralla, M. M., Khattak, S. B. A., and Jan, B. (2023). Efficient fire detection with e-efnet: A lightweight deep learning-based approach for edge devices. *Applied Sciences*, 13(23):12941.
- Fernandes, A., Utkin, A. B., and Chaves, P. (2024). Enhanced automatic wildfire detection system using big data and efficientnets. *Fire*, 7(8):286.
- Ghali, R. and Akhloufi, M. A. (2024). Deep learning approach for wildland fire recognition using rgb and thermal infrared aerial image. *Fire*, 7(10):343.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Hisaria, S., Sharma, P., Gupta, R., and Konatham, S. (2024). An analysis of multi-criteria performance in deep learning-based medical image classification: A comprehensive review.
- Khan, S. and colleagues (2022). Ffirenet: Deep learning based forest fire classification using mobilenetv2. *Symmetry*.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, volume 25, pages 1097–1105.
- Reuters (2024). Pará decreta estado de emergência ambiental por queimadas. <https://www.reuters.com/fact-check/portugues/FLIY5GGKGVJOTI7ZVN56SYHV2Y-2024-09-25/>.
- Saied, A. (2018). Fire dataset. <https://www.kaggle.com/datasets/phylake1337/fire-dataset>. Acesso em: set. 2025.
- Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):1–48.
- Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114.
- Wang, B., Huang, G., Li, H., Chen, X., Zhang, L., and Gao, X. (2024). Hybrid cbam-efficientnetv2 fire image recognition method with label smoothing in detecting tiny targets. *Machine Intelligence Research*, 21(6):1145–1161.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328.