

Detecção e rastreamento de fogo em tempo real com YOLOv5

Utilizando a linguagem de programação *Python* para implementar um algoritmo de detecção e rastreamento de fogo por meio do modelo YOLOv5.

Daniel Ribeiro da Silva
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás
daniel.ribeiro@discente.ufg.br

Fabrycio Leite Nakano Almada
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás
fabrycio@discente.ufg.br

Maykon Adriell Dutra
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás
maykonadriell@discente.ufg.br

Victor Emanuel da Silva Monteiro
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás
victor_emanuel@discente.ufg.br

Abstract—Incêndios representam um desafio complexo devido aos diversos fatores que podem desencadeá-los e às suas consequências devastadoras em termos materiais, ecológicos e humanos. No entanto, é possível mitigar esses efeitos por meio da detecção precoce, o que abre caminho para seu controle e extinção. Este artigo apresenta um estudo sobre a detecção de incêndios em tempo real usando a arquitetura YOLOv5. Com alta precisão e velocidade, o modelo foi criado com foco em oferecer integração em câmeras de segurança e drones, permitindo uma detecção rápida e confiável de chamas. A combinação de tecnologias culmina em resultados consistentes e confiáveis na identificação da presença de fogo, contribuindo para a prevenção.

Index Terms—incêndios, detecção e rastreamento, YOLOv5, visão computacional, drones.

I. INTRODUÇÃO

Incêndios são fenômenos que podem ocorrer quando diferentes riscos - como presença de materiais combustíveis, negligência em prevenção de acidentes ou atos de vandalismo - são combinados resultando em uma reação química, motivo gerador de prejuízos materiais, ecológicos, financeiros e até mesmo perda de vidas [1]. Contudo, através da detecção desses eventos quando o fogo ainda está em estágio inicial, é possível reduzir a incidência de incêndios, bem como mitigar seus impactos, permitindo que ele possa ser controlado e então extinguido completamente. Nesse sentido, é possível utilizar algoritmos de visão computacional e integrá-los a câmeras de segurança, drones e quaisquer outros dispositivos que permitam a visualização da presença de chamas, principalmente se essa detecção ocorrer em tempo real.

Entre as abordagens promissoras, destaca-se o YOLOv5: uma evolução da família YOLO (You Only Look Once) que tem ganhado popularidade devido à sua precisão e velocidade de detecção [2]. Esse é um modelo de aprendizado profundo que utiliza uma arquitetura baseada em redes neurais

convolucionais, tornando-o ideal para aplicações em tempo real, como a detecção de fogo em imagens capturadas por câmeras de segurança ou drones. Sua capacidade de identificar múltiplos objetos em uma única passada pela imagem, aliada ao uso de técnicas avançadas de reconhecimento, permite uma detecção rápida e precisa das chamas, tornando-o um candidato promissor para esse cenário específico.

Além disso, os modelos baseados em YOLO são altamente personalizáveis e podem ser treinados para reconhecer classes específicas de objetos, incluindo incêndios, com um conjunto de dados apropriado [3]. Isso o torna uma ferramenta versátil para atender às necessidades específicas de diferentes ambientes e cenários de detecção de fogo.

Aliado a implementação do modelo, a captação dos dados com câmeras e sensores adequados são de suma importância para a detecção de incêndios de forma confiável [4]. Além disso, a utilização de robôs para captura das imagens nos permite a garantia de maior segurança para as pessoas, de tal forma, surgem os drones como possibilidade, uma vez que permitem uma grande abrangência em área de detecção e bom controle para evitar prejuízos maiores [5]. Dessa forma, entende-se que a utilização de sensores com boa qualidade de imagem aliados com a utilização de drone e do modelo YOLOv5 nos resultará em resultados consistentes na identificação da presença de fogo.

II. FUNDAMENTOS TEÓRICOS

Para a dissertação do tema é de suma importância conceituar os fundamentos teóricos da arquitetura YOLOv5 envolvidos na detecção de fogo em tempo real em imagens. Esse artigo foi desenvolvido com base em diversos trabalhos prévios que fazem uso do supracitado modelo para detecção de objetos em tempo real, em especial, aqueles que sofrem alguma

deformação ao longo do tempo, haja vista as propriedades dinâmicas do fogo.

Nesse contexto, Zhang, Yu et al. [2] ilustram a aplicabilidade do YOLOv5 em contextos dinâmicos ao aplicar uma versão aprimorada desse modelo para detecção de veículos em tempo real, validando assim a sua eficiência em identificar formas variáveis ao longo do tempo, uma característica crucial que se assemelha com algumas dificuldades no processo da detecção de chamas. Adicionalmente, Dos Santos et al. [4] desenvolvem um sistema de monitoramento de incêndios ambientais utilizando o módulo ESP32-CAM e alguns sensores para o monitoramento e envio de dados de variáveis como temperatura, umidade e concentração de gás carbônico, validando, a viabilidade e importância do monitoramento de incêndios por meio de técnicas de Inteligência Artificial (IA). Por fim, Kinaneva [5] enfatiza a eficiência na combinação do uso de drones com modelos de visão computacional para a detecção precoce de incêndios florestais, demonstrando a importância da detecção em tempo real, ainda que o YOLOv5 não tenha sido utilizado. Coletivamente, esses trabalhos oferecem uma base sólida para a aplicação do YOLOv5 na detecção de fogo, corroborando a viabilidade e pertinência deste estudo.

A. YOLOv5

O YOLOv5 é uma evolução da arquitetura YOLO, introduzindo melhorias significativas em termos de desempenho e precisão. Nesse sentido, adota blocos de convolução residuais, inspirados no ResNet, para extrair características mais ricas e complexas das imagens, contribuindo para o aumento do desempenho geral. A eficiência computacional é aprimorada, tornando-o adequado para implantação em dispositivos com recursos limitados, como drones. No que diz respeito à aplicação prevista, ainda, pode-se sugerir que a capacidade de detectar múltiplos objetos em tempo real é uma característica essencial para a detecção ágil de incêndios em cenas dinâmicas, quando devidamente treinado para essa finalidade.

B. Treinamento

O treinamento do modelo YOLOv5 é realizado com um conjunto de dados anotados contendo imagens com regiões delimitadas por *bounding box* balizado ao redor do fogo. O processo de treinamento emprega algoritmos de otimização para ajustar os pesos da rede com o objetivo de minimizar a função de perda de detecção (YOLO detection loss). Essa função de perda leva em consideração a precisão das detecções e das localizações das *bounding boxes*. No que diz respeito a isso, o treinamento intensivo fundamenta-se essencial para o aprendizado adequado do modelo em termos da tarefa proposta.

C. Integração com drone

A integração da rede treinada com o drone DJI Tello permite a implementação prática do sistema em cenários reais. O aparelho é equipado com uma câmera de resolução HD, responsável por capturar e transmitir a área sob monitoramento. Os frames são então transmitidos ao computador central,

responsável pelo processamento das imagens e comandos de voo do drone, onde o modelo realiza a detecção de fogo em tempo real. A partir disso, os comandos necessários são calculados e enviados de volta ao Veículo Aéreo Não Tripulado (VANT), que passa a acompanhar e seguir o rasto do fogo sem perdê-lo, enquanto mantém a distância correta do foco de incêndio. Esse cálculo é realizado considerando a área e a posição da *bounding box*.

D. Bibliotecas e frameworks utilizados

- **Djittelopy**: biblioteca Python que fornece uma interface para o controle do drone DJI Tello. Essa biblioteca possibilita o estabelecimento da conexão com o drone, a transmissão de vídeo em tempo real e o envio de comandos de voo;
- **OpenCV**: Utilizada para realizar o pré-processamento das imagens capturadas pela câmera do drone. Assim, inclui a conversão das imagens para escala de cinza, o redimensionamento para o tamanho desejado e a aplicação de outras técnicas pertinentes com vista o aumento da capacidade de generalização do modelo YOLOv5.
- **Pytorch**: Para o treinamento do modelo - processo que envolve a atualização iterativa dos pesos da rede neural para ajustá-los aos dados anotados - de modo a maximizar a precisão da detecção de fogo. Sua escolha ocorre, pois permite uma integração suave com a arquitetura YOLOv5 e outras redes neurais.

E. Apoio à tomada de decisão

Posterior a detecção do fogo, o sistema pode adotar diferentes ações, dependendo da aplicação específica e do contexto de emprego. Pode, por exemplo, enviar informações imediatas para centros de controle, equipes de emergência ou sistemas de monitoramento, permitindo uma resposta rápida e coordenada. Adicionalmente, pode ser acrescido ao sistema um modelo preditivo capaz de estimar a trajetória do incêndio, prever possíveis áreas de risco e auxiliar na tomada de decisões estratégicas. No que tange a esta provável aplicação, surge a possibilidade de empregar autômatos celulares para simulação da dinâmica do fogo em áreas urbanas ou rurais.

F. Conclusão teórica

Os fundamentos teóricos aqui descritos, delineiam os principais mecanismos, técnicas e algoritmos envolvidos na solução proposta para a detecção de fogo em tempo real utilizando imagens e sua integração com um drone. A combinação dessas abordagens pode representar um avanço significativo na combinação de técnicas próprias de visão computacional e sistemas autônomos, com aplicações promissoras em monitoramento e combate a incêndios, e contribuições importantes para a segurança pública, proteção ambiental e infraestrutura crítica.

III. METODOLOGIA

No início do processo de implementação, foi imperativo empregar conjuntos de dados contendo representações visuais de chamas, essenciais para a subsequente avaliação do desempenho dos modelos desenvolvidos. Após uma investigação preliminar minuciosa, dois conjuntos de dados destacaram-se: o Fire and Smoke Dataset [6], abrangendo imagens de fogo e fumaça, e o fire and gun dataset [7], contendo imagens de fogo e armas. Todavia, alinhado ao escopo específico deste estudo, optou-se por direcionar o foco exclusivamente às imagens que retratam chamas, as quais foram selecionadas para compor o conjunto de treinamento dos modelos em desenvolvimento. Ao todo, foram utilizadas 3527 imagens contendo pelo menos uma chama, seja em áreas rurais ou urbanas e em diferentes contextos. Além disso, é relevante destacar que, devido ao pré-processamento prévio realizado pelo autor do dataset, não foi necessário executar qualquer etapa adicional de pré-processamento nos dados antes do treinamento.

Em posse dos dados, o treinamento do modelo YOLOv5 foi realizado no ambiente Google Collaboratory, uma vez que ele oferece acesso a unidades de processamento GPU T4, o que aumenta significativamente o desempenho no treinamento de modelos de inteligência artificial. As unidades de processamento foram utilizadas por cerca de 3 horas para o treinamento de 10 épocas, resultando em um modelo com 157 camadas com um total de 7.012.822 parâmetros.

Após o treinamento do modelo YOLOv5 para detecção de focos de incêndio, foi desenvolvida uma etapa de implementação dedicada à integração do drone DJI Tello com a capacidade de predição de fogo em ambientes reais. Essa integração foi concretizada com o uso da biblioteca Djitellopy, desempenhando um papel crucial ao viabilizar a comunicação fluida entre o drone e o ambiente de programação Python.

A fim de assegurar o funcionamento preciso e coordenado do sistema, um script elaborado foi responsável pela importação e incorporação do modelo previamente treinado no ambiente do drone, bem como pela definição de estratégias de resposta do drone a cenários diversos. Entre as estratégias de reação estão a realização de movimentos específicos, como recuar quando a detecção de foco de incêndio é próxima e avançar quando distante, além de ajustar a orientação da câmera horizontalmente para enquadrar corretamente a detecção com base na área da bounding box. Essa implementação unificada culminou em um sistema que, por meio da fusão entre visão computacional avançada e controle autônomo, se mostra promissor para a detecção e acompanhamento automatizado de focos de incêndio em cenários do mundo real, capturando as imagens do drone, enviando ao computador para processar os dados e retornando os comandos para o drone.

No que tange ao apoio à tomada de decisão, foi realizada a implementação de um autômato celular para prever o espalhamento de chamas em um determinado ambiente, consideramos como informações relevantes para análise a direção de ventos e os componentes com altas chances de incêndio, como

vegetações secas. Com essas informações, é possível realizar o cálculo das probabilidades para onde o fogo se espalhará em diferentes direções com base em um modelo pré-treinado. O autômato celular, então, realiza o espalhamento das chamas com base nas probabilidades passadas e, assim, é possível determinar as regiões que podem sofrer ação das chamas. Nesse sentido, foram consideradas as seguintes características do autômato:

- **Estado 0:** Representa uma célula vazia ou espaço não afetado pelo fogo.
- **Estado 1:** Representa uma célula com vegetação.
- **Estado 2:** Representa uma célula em chamas.

As regras de transição do autômato celular são definidas na função `atualizar_grade`, onde a propagação do fogo ocorre de acordo com as probabilidades de espalhamento definidas no dicionário `prob_propagacao_fogo`. Para cada célula em chamas (Estado 2), a função verifica as células vizinhas nas direções definidas pelas chaves do dicionário ((0, 1), (0, -1), (1, 0), -1, 0)) e verifica se as células vizinhas são de vegetação (Estado 1). Se sim, a propagação do fogo para a célula vizinha ocorre com base na probabilidade associada àquela direção.

Pode-se, ainda, resumir o comportamento da máquina de estados através do conjunto de três regras:

- Se uma célula está em chamas (Estado 2), ela tentará propagar o fogo para suas células vizinhas de vegetação (Estado 1) de acordo com as probabilidades definidas no dicionário `prob_propagacao_fogo`.
- As células vazias ou não afetadas pelo fogo (Estado 0) permanecem inalteradas.
- As células em chamas (Estado 2) continuam queimando em iterações subsequentes até que a propagação do fogo pare.

Além disso, outra metodologia, implementada inicialmente como um protótipo, trata da criação de uma rede neural artificial capaz de classificar a vegetação em solo dentro de duas categorias pré-determinadas. Sendo elas:

- **1. Vegetação Seca.**
- **2. Vegetação não Seca**

A fim de cumprir tal objetivo, foi escolhida a biblioteca TensorFlow a fim de determinar a melhor arquitetura possível para a classificação de vegetação. A biblioteca Skimage foi utilizada para criar um dataset de imagens sintéticas que tentasse replicar algumas das características usuais do ambiente florestal e, por se tratar de um protótipo, julga-se satisfatória o comportamento da arquitetura escolhida. Sendo esta última também utilizada em datasets de contexto próximo como o Fruits and Vegetables Image Recognition Dataset [8].

IV. RESULTADOS

Para avaliação dos resultados obtidos no trabalho, primeiro avaliamos o desempenho no treinamento do modelo YOLOv5 e, assim, podemos observar que as losses para detecção de objetos e para geração da bounding box diminuíram com o passar das épocas. Assim, mostrando que o modelo conseguiu aprender a presença de fogo em diferentes cenários.



Fig. 1. Diminuição das losses durante o treinamento do modelo YOLOv5.

Além disso, pode-se verificar a detecção de fogo do modelo com base em fotos e vídeos utilizados para inferência.

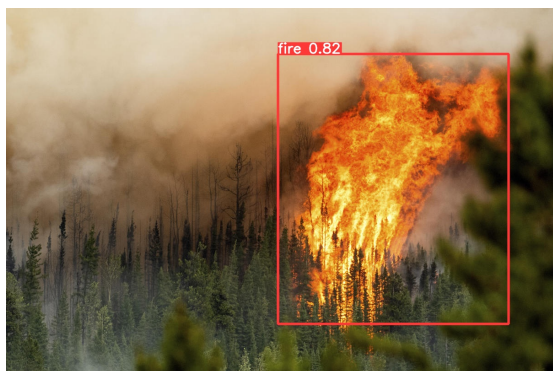


Fig. 2. Detecção de fogo em uma imagem real com a utilização do modelo YOLOv5.

Podemos observar, também, através da aplicação desenvolvida para utilização no drone, que a detecção de fogo ocorreu de forma robusta e diversos testes realizados para o movimento do drone performaram conforme o planejado, um vídeo mostrando testes realizados está anexado junto a este artigo [9].

No que diz respeito à predição do avanço do fogo em áreas florestais, ambas as metodologias complementares obtiveram resultados satisfatórios. O autômato celular consegue, de maneira heurística, simular o espalhamento do fogo com base em um dicionário de probabilidades direcionais. Na geração de tal dicionário, utilizando-se de uma rede neural artificial, pode-se denotar a viabilidade da classificação vegetal em solo com base em suas características morfológicas. O

modelo, entretanto, por se tratar de um protótipo carece de novos dados de entrada que representem a realidade ambiental com todas as suas nuances e transições temporais.

O projeto, portanto, segue para etapas futuras de desenvolvimento onde pretende-se criar um conjunto de dados mais adequado, com imagens reais, além de definir como outras variáveis, a citar; direção do vento, umidade relativa e topografia, vão influenciar no dicionário de probabilidades utilizadas pelo autômato.

V. CONCLUSÕES

Por fim, uma vez que o modelo conseguiu detectar de forma robusta a presença de chamas e os algoritmos para utilização de drone e predição da propagação de fogo performaram de forma correta, pode-se concluir que a aplicação desenvolvida neste trabalho é eficiente na detecção e predição de chamas em diferentes locais.

Com isso, constata-se que tais métodos podem ser utilizados em contextos reais a fim de prevenir problemas oriundos do fogo, como por exemplo acidentes domésticos e propagação de chamas em florestas. Com esse contexto, cria-se uma ferramenta capaz de gerar ganhos sociais, comerciais, financeiros e até salvar vidas.

VI. REFERÊNCIAS

- 1) SEITO, Alexandre Itiu. Fundamentos de fogo e incêndio. In: A Segurança contra Incêndio no Brasil. São Paulo: Projeto Editora. Acesso em: 03 ago 2023.
- 2) Zhang, Yu. Et al. Real-Time Vehicle Detection Based on Improved YOLO v5. Tongji University, Shanghai, 27 set 2022. Disponível em: <<https://www.mdpi.com/2071-1050/14/19/12274>>. Acesso em: 03 ago 2023.
- 3) Zhao, Lei. Fire-YOLO: A Small Target Object Detection Method for Fire Inspection. Taiyuan University of Technology, Taiyuan, 20 abr 2022. Disponível em: <<https://www.mdpi.com/2071-1050/14/9/4930>>. Acesso em: 03 ago 2023.
- 4) Dos SANTOS Et. al. Proposta de um sistema de monitoramento de incêndios ambientais utilizando o ESP32-CAM. In: ESCOLA REGIONAL DE INFORMÁTICA DE MATO GROSSO (ERI-MT), 21., 2021. Disponível em: <<https://sol.sbc.org.br/index.php/erimt/article/view/18229/18063>>. Acesso em: 03 ago 2023.
- 5) Kinaneva, D. Early Forest Fire Detection Using Drones and Artificial Intelligence. Institute of Electrical and Electronics Engineers, 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8756696>>. Acesso em: 03 ago 2023.
- 6) DataCluster Labs. Fire and Smoke Dataset, 2023. Disponível em: <<https://www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset>>. Acesso em: 17 ago. 2023.
- 7) KUMAR, A. Fire and gun dataset, 2020. Disponível em: <<https://www.kaggle.com/datasets/atulyakumar98/fire-and-gun-dataset>>. Acesso em: 17 ago. 2023.

- 8) KUMAR, A. Fruits and Vegetables Image Recognition Dataset, 2022. Disponível em: <<https://www.kaggle.com/datasets/kritikseth/fruit-and-vegetable-image-recognition>>. Acesso em: 18 ago. 2023.
- 9) SILVA, D. R. da; ALMADA, F. L. N.; DUTRA, M. A.; MONTEIRO, V. E. da S. Github. 2023-1-INF0417-VisaoComputacional. Disponível em: <bit.ly/3KNnyoS>. Acesso em: 18 ago. 2023.