

UNIVERSIDADE FEDERAL DE SÃO CARLOS - CAMPUS SOROCABA

CCGT - CIÊNCIA DA COMPUTAÇÃO

Processamento de Imagens

Prof. Dr. Jurandy Almeida

PROJETO DE PROCESSAMENTO DE IMAGENS

Proposta de Projeto

Felipe Bonadia de Oliveira Bravo - 813908

João Vitor Naves Mesa - 814149

28/04/2025

Proposta : Detecção automática de fumaça e fogo com Fine-Tuning de Yolo5 no D-Fire Dataset

A proposta deste projeto é desenvolver, em até dois meses e com dedicação parcial de dois alunos iniciantes, um sistema básico de detecção automática de fumaça e fogo em imagens, voltado para aplicações de vigilância em pequenos ambientes industriais ou residenciais. Para embasar o trabalho, revisaremos o artigo de Venâncio et al. (2022), que apresenta o D-Fire Dataset, e um tutorial prático de fine-tuning de modelos YOLO em PyTorch.

Dessa maneira, utilizaremos o D-Fire Dataset do Kaggle, com suas 21000 imagens já divididas em treino, validação e teste e anotadas em formato YOLO para as classes “Smoke” e “Fire”, sem necessidade de coleta adicional de dados. A abordagem consistirá em baixar uma versão pré-treinada do YOLOv5 e realizar fine-tuning usando parâmetros padrão, ao mesmo tempo em que incorporamos data augmentation on-the-fly na pipeline de treinamento por meio da implementação Ultralytics ou da biblioteca Albumentations.

Assim, quando aplicarmos flips horizontais, o framework recalcula automaticamente as coordenadas das bounding boxes (por exemplo, em uma imagem de largura W , cada caixa com x_{min} e x_{max} passa a ter $x_{min}' = W - x_{max}$ e $x_{max}' = W - x_{min}$), enquanto ajustes de brilho e contraste, por serem fotométricos, não alteram as anotações. Para rotações — 90° , 180° e 270° — as próprias funções prontas realizam a “rotação” das caixas junto com a imagem, usando fórmulas como, no caso de 90° para a direita, $(x, y) \rightarrow (y, W - x)$ e similares para os demais ângulos, garantindo consistência sem trabalho manual. Manteremos como baseline o modelo pré-treinado sem data augmentation e avaliaremos os resultados de forma qualitativa — exibindo imagens de teste com caixas delimitadoras e exemplos de falsos positivos e negativos — e quantitativa, por meio das métricas $mAP@0.5$, precisão e recall fornecidas pela biblioteca Ultralytics, além de medir a velocidade de inferência (FPS) em CPU comum.

Diante disso, espera-se entregar um relatório e um demonstrador simples que atinja pelo menos 70 % de $mAP@0.5$, servindo como ponto de partida para projetos futuros em monitoramento visual.

Referências

Pedro Vinícius Almeida Borges de Venâncio, Adriano Chaves Lisboa, Adriano Vilela Barbosa. "An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices." *Neural Computing and Applications*, vol. 34, no. 18, 2022, pp. 15349–15368. DOI: 10.1007/s00521-022-07467-z.

J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.