

Detecção de Carros com Análise de Movimento

- Murilo Lima Lirani de Carvalho (11234673)
 - Enzo Castelo Branco Biondi (14568142)
 - Vinicio Yusuke Hayashibara (13642797)
 - Julia Pravato (14615054)
 - Marcos Paulo Rocha Baltazar (13782373)
 - Laura Nannetti Moreira (12557504)
-

1. Introdução

Este projeto visa aprimorar um modelo de detecção de veículos, tornando-o mais preciso e eficiente. Embora o objetivo final seja a contagem e o rastreamento de veículos em vídeo, identificamos que a detecção é o alicerce fundamental e mais complexo dessa tarefa. Nossa abordagem consiste em otimizar o detector ao fornecer-lhe informações contextuais sobre o movimento na cena, criando uma base sólida para futuros trabalhos de rastreamento.

2. Metodologia

A metodologia partiu da aquisição de dados e culminou na avaliação comparativa de dois modelos. O processo envolveu a criação de um dataset, a extração de informações de movimento e o uso dessas informações para pré-processar as imagens de treino, realçando seletivamente as áreas de interesse.

2.1. Preparação do Dataset e Extração de Movimento

Utilizamos um vídeo de um cruzamento de ruas, do qual extraímos frames a cada 3 segundos. Para gerar os rótulos de referência (*ground truth*), aplicamos o modelo YOLOv5 aos frames, agrupando todas as classes de veículos (carro, caminhão) em uma única classe: "automóvel". Para isolar o movimento, o fundo estático da cena foi gerado pelo cálculo da moda de cada píxel em um segmento de 5 minutos do vídeo. Como a câmera é estática, a cor mais frequente em qualquer ponto corresponde ao cenário, efetivamente "apagando" os objetos móveis.

2.2. Pré-processamento com Realce e Treinamento

Como a arquitetura YOLOv8 (usada para o treino) opera com 3 canais de cor (RGB), descartamos a ideia de adicionar o movimento como um 4º canal. A solução foi usar a informação de movimento para pré-processar as imagens:

1. **Máscara de Movimento:** Calculamos a diferença absoluta entre cada frame e a imagem de fundo para gerar uma máscara binária indicando as regiões de movimento. Foram ignoradas regiões muito ao fundo, onde os ruídos eram tão presentes que praticamente todas as regiões seriam realçadas
2. **Realce da Imagem:** A máscara guiou um realce seletivo nas imagens originais, aumentando a saturação dos pixels nas áreas de movimento para destacar visualmente os veículos.
3. **Treinamento Comparativo:** Para validar a abordagem, treinamos e comparamos dois modelos YOLOv8: um modelo de base com as imagens originais e um modelo otimizado com as imagens realçadas. O processo visa utilizar um processo de fine-tuning para tentar replicar os labels feitos pelo modelo inicial. A comparação não se dá pela modelo v5 e v8, mas sim,

entre os dois modelos v8 que utilizaram os mesmos conjuntos de imagens de treino e teste, apenas diferenciando-se pelo realce dado nas regiões de movimento.

3. Resultados e Análise

Para facilitar a visualização comparativa e a organização do trabalho, segue o acesso aos gráficos resultantes do processo de fine-tuning:  [GRAFICOS_NEURAIIS.pdf](#)

O modelo treinado com as imagens realçadas apresentou uma convergência mais rápida durante o treinamento, necessitando de menos épocas para replicar os rótulos do ground truth. Embora os ganhos de desempenho tenham sido modestos, essa diferença indica uma leve vantagem da abordagem de realce de movimento. Um possível fator que limitou os resultados foi o valor do hiperparâmetro alfa, que regula a intensidade do realce e pode ter sido pequeno demais para causar alterações significativas nas imagens. Além disso, não foram aplicadas máscaras sobre os rótulos dos carros no ground truth, o que significa que alguns veículos ao fundo (que estavam corretamente anotados) não receberam o realce. Isso pode ter impedido o modelo otimizado de aprender de forma mais eficaz em comparação com o modelo base.

Além disso, um fator importante a ser considerado é a própria qualidade do modelo escolhido para o fine-tuning. O modelo YOLOv8 utilizado já apresentou um desempenho bastante elevado mesmo antes das alterações, o que pode ter limitado o impacto perceptível do realce. Ou seja, a baixa diferença observada entre os modelos pode estar mais relacionada à robustez do modelo base do que à ineficácia da técnica de realce. É possível que, ao aplicar a mesma abordagem em um modelo com desempenho inicial inferior, os efeitos do realce se tornem mais evidentes, resultando em melhorias mais expressivas.

4. Conclusão e Trabalhos Futuros

O projeto sugere que o pré-processamento de imagens para realçar áreas de movimento é uma abordagem válida, contribuindo para uma melhora modesta na precisão final e na velocidade de treinamento do modelo de detecção.

Balanco da Abordagem:

- **Vantagens:** Observou-se um ganho incremental nas métricas de detecção e uma leve aceleração na convergência do modelo durante o treinamento.
- **Desafios:** A principal contrapartida é a adição de custo computacional na etapa de pré-processamento e a necessidade de otimizar o hiperparâmetro de intensidade do realce (alfa).

Apesar dos ganhos modestos, provavelmente pelos problemas citados, como a alta acurácia base do modelo v8 e a escolha do hiperparâmetro, a metodologia foi validada e sugere uma melhora no treinamento de modelos de imagens, servindo como um ponto de partida interessante para as possíveis próximas etapas lógicas do projeto. Estas incluem a implementação de algoritmos de rastreamento (*tracking*) para associar detecções do mesmo veículo ao longo do tempo e, futuramente, o desenvolvimento da funcionalidade de contagem, buscando cumprir o objetivo original.