

Relatório de Desenvolvimento

Resumo

Medir a densidade de tráfego convencionalmente requer um esforço significativo para gerenciar pessoas e gastar dinheiro. Estudos recentes mostram que a aplicação de técnicas de Aprendizado Profundo e Visão Computacional transforma o problema em algo mais acessível e prático (1, 2, 3). Assim, meu projeto tem como objetivo empregar essas técnicas, unindo detecção e rastreamento de objetos para contar o número de veículos que cruzam uma estrada em um vídeo pré-gravado. Para isso, apliquei o algoritmo YOLOV8 para detectar os objetos em um quadro, e um rastreador (codificado por mim) baseado em medições de IoU (Interseção sobre União) para rastrear e contar o número de veículos. Embora o número correto de veículos no vídeo não seja exatamente conhecido, com uma contagem manual, pode-se observar que o algoritmo apresenta resultados esperados.

Metodologia

Conjunto de Dados

O conjunto de dados está disponível no repositório. Ele consiste em 67 vídeos da mesma estrada, cada um com 30 segundos (25 fps). A câmera e a posição são as mesmas para todos os vídeos. No repositório, você pode ver um exemplo de um quadro retirado do primeiro vídeo.

Algoritmo

You Only Look Once (YOLO) é um algoritmo de aprendizado profundo projetado para detecção de objetos em tempo real em imagens e vídeos. Desenvolvido por Joseph Redmon e Santosh Divvala, o YOLO revolucionou a detecção de objetos dividindo a imagem em uma grade e prevendo caixas delimitadoras e probabilidades de classe para objetos dentro de cada célula da grade em uma única passagem. Ao contrário de métodos tradicionais que exigem passagens múltiplas, a abordagem de única passagem do YOLO permite uma detecção em tempo real mais rápida e eficiente. O YOLO equilibra velocidade e precisão, sendo amplamente utilizado em aplicações como vigilância, veículos autônomos e robótica, onde o reconhecimento rápido e preciso de objetos é crucial. Neste projeto, a 8ª versão pré-treinada do algoritmo é usada para detectar os veículos presentes no vídeo. As caixas delimitadoras retornadas por ele são, então, processadas por um rastreador que as compara com os veículos já previstos para decidir se é um veículo anteriormente visto ou não. A comparação é baseada em um valor positivo (> 0.0) na medição de IoU de ambas as caixas delimitadoras (previamente prevista e nova). Além disso, o rastreador considera apenas os veículos que cruzam um determinado valor de altura y entre 400 e 500 pixels.

Validação do Modelo e Resultados

Até agora, o modelo não foi rigorosamente validado; os resultados esperados foram baseados em uma contagem manual limitada de veículos nos vídeos, que ficou em torno de 70-90 veículos. Mas a medida manual não foi feita para todos os vídeos (apenas os primeiros). E o algoritmo obteve um valor de resultado entre o esperado para esses.

Conclusões

Com um modelo de detecção bem definido (YOLO) e um rastreador simples, podemos ver o incrível potencial da aplicação de técnicas de Aprendizado Profundo e Visão Computacional na contagem de tráfego (uma ótima maneira de aprender mais sobre essas técnicas). No entanto, este projeto ainda carece de alguns detalhes que poderiam aprimorar sua aplicabilidade. Em primeiro lugar, precisamos de uma abordagem mais científica para validar melhor o modelo com métricas bem definidas e resultados de referência. Em segundo lugar, é muito importante compará-lo com outros modelos existentes (incluindo alguns rastreadores já bem conhecidos). Portanto, há muito espaço para explorar e mais variáveis para testar neste projeto. Além disso, reforça-se a intenção educacional deste projeto como uma forma de entender como lidar com um problema da vida real como pesquisador em Visão Computacional e IA.