

Problema e Solução para a Detecção de Semáforos e seus Estados

Problema:

Detectar semáforos e identificar seus estados (vermelho, amarelo, verde) em imagens de tráfego urbano apresenta diversos desafios. Variáveis como diferentes condições de iluminação, ângulos de visão variados, condições climáticas adversas e diversidade nos designs de semáforos complicam a detecção precisa. Reflexos e ofuscamentos frequentes, bem como obstruções causadas por árvores, edifícios e veículos, agravam ainda mais o problema. Diferentes posicionamentos de câmeras, em termos de altura e distância, também influenciam na visibilidade e aparência dos semáforos, dificultando a tarefa de detecção.

Primeira Solução: Visão Computacional Clássica

A abordagem inicial utilizou técnicas clássicas de visão computacional com a biblioteca OpenCV para detectar semáforos com base em suas características visuais, como forma redonda e cores específicas (vermelho e verde).

Pipeline de Manipulação de Imagem:

Detecção de Círculos de Hough: Utilizou-se a Transformada de Hough para identificar formas circulares correspondentes aos semáforos.

HOG (Histogram of Oriented Gradients): Implementou-se o HOG para identificar e classificar os semáforos nas imagens.

Apesar de eficazes em condições ideais, essas técnicas enfrentaram dificuldades significativas:

Alta Variabilidade: Semáforos com diferentes posições e formatos não foram corretamente identificados, levando a muitos falsos positivos e negativos.

Condições de Iluminação: Alterações na iluminação afetaram drasticamente a precisão da detecção.

Segunda Solução: Faster R-CNN

Devido às limitações da abordagem clássica, adotou-se o modelo Faster R-CNN, uma técnica avançada de detecção de objetos que integra a geração de propostas de regiões e a classificação em uma única rede neural.

Arquitetura do Modelo:

Backbone (ResNet-50): Extrai características de alto nível das imagens.

Region Proposal Network (RPN): Gera propostas de regiões de interesse (Rois) onde os objetos podem estar localizados.

RoI Pooling: Ajusta as regiões de interesse para um tamanho fixo.

Classificação e Refinamento: Classifica cada região e ajusta as caixas delimitadoras.

Treinamento e Avaliação:

Dataset: Utilizou-se um dataset anotado com 2400 imagens para treino e 400 para teste, cobrindo diferentes horários, condições de iluminação e tipos de semáforo.

Avaliação: O modelo foi testado em cenários variados, garantindo validação em condições reais.

Comparação e Melhorias Futuras:

Visão Computacional Clássica: A abordagem inicial foi limitada pela incapacidade de lidar com a variabilidade nas condições de iluminação e formatos dos semáforos.

Faster R-CNN: Proporcionou melhorias significativas na precisão e robustez da detecção, superando as limitações da abordagem clássica.

Melhorias Propostas:

Aumento do Dataset: Incluir exemplos variados cobrindo diferentes condições.

Pré-processamento: Aplicar técnicas como normalização de iluminação e redução de ruídos.

Modelos Especializados: Desenvolver e combinar modelos especializados para diferentes condições.

Regularização e Ajuste Fino: Aplicar técnicas de regularização e realizar ajuste fino com dados específicos.

Comparação com Outros Modelos: Implementar e comparar com modelos como Cascade R-CNN para avaliação de desempenho.

Com essas melhorias, espera-se aumentar ainda mais a eficácia do modelo em cenários reais de tráfego urbano, tornando-o uma ferramenta robusta para a detecção de semáforos e seus estados.