

Revista Científica Di Fatto, ISSN 2966-4527. Edição 4. Ano: 2025.

Submissão em: 25/12/2024

Aprovação em: 01/01/2025

Publicado em: 01/01/2025

Disponível em: <https://revistadifatto.com.br/artigos/deteccao-de-objetos-em-imagens-de-transito-urbano/>

Detecção de objetos em imagens de trânsito urbano

Juan Carlos Quevedo Weimar

Mestrando em Computação, Tecnólogo em Sistemas para Internet.

Resumo

O aumento da urbanização e a complexidade dos centros urbanos têm intensificado os desafios na gestão do tráfego veicular, levantando preocupações sobre segurança viária e eficiência do transporte urbano. Nesse contexto, a detecção de objetos em trânsito urbano surge como uma solução promissora para reduzir riscos e melhorar a mobilidade. Com características dinâmicas e diversificadas, os ambientes urbanos demandam tecnologias que identifiquem, rastreiem e antecipem a movimentação de veículos, pedestres e ciclistas em tempo real. Este artigo foca na aplicação prática da detecção de objetos em cenários urbanos utilizando a arquitetura YOLO (You Only Look Once). A pesquisa detalha os métodos de configuração do modelo, a execução do projeto e os resultados alcançados, além de apresentar uma análise crítica sobre os impactos da tecnologia no contexto urbano. Ao explorar a detecção em tempo real como ferramenta essencial para melhorar a segurança e a fluidez do tráfego, o estudo busca contribuir para o avanço das soluções tecnológicas voltadas à mobilidade e à segurança viária.

Palavras-Chave: Detecção de objetos. Trânsito urbano. Arquitetura YOLO. Segurança viária. Mobilidade urbana.

Abstract

Increased urbanization and the complexity of urban centers have intensified the challenges in managing vehicular traffic, raising concerns about road safety and urban transport efficiency. In this context, object detection in urban traffic emerges as a promising solution to reduce risks and improve mobility. With dynamic and diverse characteristics, urban environments demand technologies that identify, track and anticipate the movement of vehicles, pedestrians and cyclists in real time. This article focuses on the practical application of object detection in urban scenarios using the YOLO (You Only Look Once) architecture. The research details the model configuration methods, project execution and results achieved, in addition to presenting a critical analysis of the impacts of technology in the urban context. By exploring real-time detection as an essential tool for improving traffic safety and fluidity, the study seeks to contribute to the advancement of technological solutions aimed at mobility and road safety.

Keywords: *Object detection. Urban traffic. YOLO Architecture. Road safety. Urban mobility.*

1. INTRODUÇÃO

O acelerado processo de urbanização e a consolidação dos centros urbanos têm provocado profundas mudanças no cenário das cidades, trazendo consigo desafios significativos na gestão do tráfego veicular e na promoção de um transporte eficiente e seguro. Esse fenômeno, amplamente documentado, tem gerado preocupações crescentes sobre a segurança viária e a eficiência no deslocamento, especialmente em metrópoles que enfrentam níveis elevados de congestionamento e riscos de acidentes (De Vasconcellos, 2000). Com a densidade populacional cada vez maior e a multiplicidade de elementos interagindo nas vias urbanas, as soluções convencionais de gestão de tráfego mostram-se insuficientes para lidar com a complexidade e o dinamismo das cidades modernas. Nesse contexto, a detecção de objetos em trânsito urbano apresenta-se como uma ferramenta tecnológica promissora, fornecendo insights essenciais para a tomada de decisão em tempo real. Tecnologias de visão computacional e aprendizado de máquina têm permitido avanços significativos na identificação e rastreamento de veículos, pedestres, ciclistas e outros elementos presentes no trânsito urbano.

As cidades contemporâneas demandam tecnologias capazes de integrar diferentes sistemas de mobilidade, promover segurança e otimizar o fluxo veicular. A detecção de objetos em tempo real, sustentada por arquiteturas avançadas como o YOLO (You Only Look Once), destaca-se nesse cenário por sua eficiência, permitindo identificar múltiplos objetos simultaneamente com alta precisão e velocidade (Redmon et al., 2016). Essa abordagem é especialmente relevante em ambientes urbanos, onde a variabilidade de situações, como diferentes condições climáticas, iluminação e padrões de tráfego, requerem soluções robustas e adaptáveis.

O objetivo deste artigo é explorar em detalhes a aplicação prática da detecção de objetos em cenários urbanos utilizando a arquitetura YOLO, abordando desde a configuração técnica do modelo até a análise dos resultados alcançados. Serão apresentados estudos de caso que demonstram o impacto dessa tecnologia na redução de acidentes, no planejamento viário e na gestão de tráfego em tempo real. Além disso, será discutida a importância de integrar essas tecnologias a sistemas mais amplos de cidades inteligentes, promovendo maior conectividade e eficiência nas operações urbanas (TACHIZAWA, 2024).

A adoção de tecnologias de detecção de objetos em tempo real não apenas melhora a segurança viária, mas também contribui para a sustentabilidade urbana, reduzindo os tempos de deslocamento e o consumo de combustível. A integração de tais soluções com sistemas de transporte público e

dispositivos inteligentes pode potencializar ainda mais seus benefícios, promovendo um ambiente mais seguro e eficiente para todos os usuários das vias urbanas.

Por fim, o artigo busca contribuir para o avanço da implementação dessas tecnologias, oferecendo uma análise crítica sobre suas aplicações e discutindo os desafios e oportunidades que ainda precisam ser enfrentados. Este trabalho visa não apenas ampliar a compreensão sobre o potencial da detecção de objetos no trânsito urbano, mas também fornecer subsídios para o desenvolvimento de políticas públicas e estratégias que fomentem sua adoção em larga escala.

2. METODOLOGIA

Para implementar a detecção de objetos em trânsito urbano, foi escolhida a arquitetura YOLOv4 (You Only Look Once), reconhecida por sua alta eficiência e precisão em tarefas de detecção em tempo real. Desenvolvida por Bochkovskiy et al. (2020), a YOLOv4 combina inovação tecnológica com uma abordagem simplificada para a predição de objetos, dividindo a imagem em uma grade e realizando a detecção em uma única passada. Esse método, além de aumentar a velocidade do processamento, torna a arquitetura altamente adequada para aplicações em ambientes complexos e dinâmicos, como o tráfego urbano.

Uma das principais vantagens da YOLOv4 é sua capacidade de realizar inferências em tempo real sem comprometer a precisão, mesmo em cenários que envolvem variabilidade de iluminação, movimento rápido e presença de múltiplos objetos (REDMON e FARHADI, 2018). Esses atributos tornam a arquitetura uma escolha ideal para enfrentar os desafios da detecção de objetos em áreas urbanas, onde a velocidade de processamento e a precisão são fundamentais para garantir a segurança e a eficiência viária.

A formação do dataset foi uma etapa essencial neste projeto, envolvendo a combinação de imagens devidamente rotuladas do Open Images Dataset e de imagens extraídas da web. Estas últimas passaram por um rigoroso processo de rotulação manual por meio da ferramenta labelImg, garantindo a conformidade com os padrões exigidos pela YOLOv4. No total, foram utilizadas 3.000 imagens para treinamento e 600 imagens para validação, compondo um conjunto de dados robusto e diversificado. Segundo Everingham et al. (2010), a qualidade e a diversidade do dataset são determinantes para o desempenho dos modelos de visão computacional, especialmente em aplicações que demandam generalização em diferentes contextos.

A configuração do modelo envolveu a adaptação de parâmetros essenciais para maximizar sua eficiência no ambiente urbano. O número de classes foi ajustado para incluir objetos de alta relevância, como veículos, placas de sinalização, pedestres, ciclistas e faixas de segurança. Além disso, o tamanho de entrada da imagem foi otimizado para atender às demandas de processamento

em tempo real, assegurando um equilíbrio entre precisão e velocidade. Camadas adicionais e filtros foram configurados para aprimorar a detecção em ambientes dinâmicos, alinhando-se às especificidades do tráfego urbano. De acordo com Liu et al. (2016), a adaptação de parâmetros técnicos é um passo crucial para ajustar os modelos às características únicas do cenário de aplicação, garantindo resultados superiores.

Por fim, a execução do treinamento foi realizada utilizando recursos computacionais de alta performance, com o auxílio de GPUs, para garantir que o modelo alcançasse os melhores resultados em termos de precisão e eficiência. A validação foi conduzida com rigor, utilizando métricas amplamente reconhecidas, como mAP (mean Average Precision) e FPS (Frames Per Second), que permitiram avaliar o desempenho do modelo em diferentes condições de trânsito. Como afirmado por Huang et al. (2017), métricas objetivas e padronizadas são indispensáveis para validar a eficácia de modelos de visão computacional em aplicações reais.

3. TREINAMENTO

Durante o treinamento, foi utilizado o framework Darknet, amplamente reconhecido por sua eficiência na execução de modelos baseados na arquitetura YOLO. A configuração específica empregou os parâmetros da YOLOv4, ajustados para atender às necessidades do cenário de detecção de objetos em trânsito urbano. Inicialmente, foi necessária a criação de um arquivo de configuração nomeado como `obj.data`, responsável por armazenar informações sobre classes, caminhos dos dados e diretórios de backup. Essa etapa, segundo Redmon e Farhadi (2018), é fundamental para garantir que o modelo opere de forma otimizada e específica para a tarefa proposta.

O arquivo padrão de configuração da YOLO foi adaptado para incluir a quantidade de classes, o número máximo de batches e os valores de steps definidos como 80% e 90% do total de batches. O arquivo customizado, nomeado como `yolov4_custom.cfg`, foi preparado para maximizar a performance do modelo. Além disso, utilizou-se um conjunto de pesos pré-treinados (`yolov4.conv.137`), que, conforme Bochkovskiy et al. (2020), serve como um ponto de partida eficiente para acelerar o processo de treinamento, aproveitando conhecimentos previamente adquiridos.

O treinamento foi realizado ao longo de 6.000 épocas, seguindo a métrica de (2.000 épocas x número de classes). A cada 100 épocas, um arquivo de pesos atualizado era salvo, permitindo interromper e retomar o treinamento sem perda de progresso. Esse método, descrito por Liu et al. (2016), é essencial em cenários que demandam longos períodos de processamento, oferecendo flexibilidade e segurança durante o processo de ajuste do modelo.

A fase de treinamento foi monitorada rigorosamente para garantir que o modelo convergisse adequadamente. O treinamento foi interrompido ao atingir as 6.000 épocas e os pesos obtidos nessa etapa foram utilizados para avaliação em um conjunto de dados de teste independente. Segundo Everingham et al. (2010), a validação do modelo com dados não vistos é indispensável para garantir a capacidade de generalização do modelo em cenários reais.

4. RESULTADOS

Para verificar os resultados obtidos durante o treinamento, foi realizada uma análise criteriosa dos pesos gerados, com o objetivo de determinar aquele que apresentasse o melhor desempenho em termos de detecção de objetos em imagens reais. Os testes foram conduzidos em diferentes pesos ao longo do treinamento, sendo identificado o peso da época 5.000 como o ideal, dado que este foi o último a apresentar um aumento significativo nas métricas de avaliação. Segundo Huang et al. (2017), o

monitoramento das métricas ao longo das iterações é essencial para identificar o momento ideal de interrupção do treinamento, evitando tanto o subajuste quanto o sobreajuste do modelo.

O modelo treinado demonstrou capacidade significativa para identificar e delimitar com precisão objetos de interesse, como placas de sinalização, pedestres e faixas de segurança, mesmo sob condições desafiadoras. Essas condições incluíram variações de iluminação, diferentes densidades de tráfego e presença de múltiplos elementos na cena. A robustez do modelo em lidar com tais variações reflete a eficácia do treinamento realizado e a qualidade do conjunto de dados utilizado. Conforme

Redmon e Farhadi (2018), a capacidade de generalização é um indicador crítico do desempenho de modelos de detecção em cenários reais.

As métricas de desempenho avaliadas incluíram precisão, recall e F1-score, apresentando valores promissores, mesmo considerando que o conjunto de dados utilizado não era extremamente robusto. A precisão elevada indica que o modelo foi capaz de minimizar falsas detecções, enquanto o recall alto reflete a capacidade de

identificar corretamente a maioria dos objetos presentes. O F1-score, como métrica harmônica entre precisão e recall, apresentou resultados consistentes, evidenciando o equilíbrio alcançado pelo modelo. Essas métricas, de acordo com Everingham et al. (2010), são amplamente utilizadas na avaliação de modelos de visão computacional, fornecendo uma visão abrangente sobre seu desempenho em diferentes aspectos.

Por fim, os resultados obtidos destacam o potencial da arquitetura YOLOv4 aplicada ao trânsito urbano. Mesmo com um dataset limitado, o modelo demonstrou alta eficácia na tarefa proposta, sendo capaz de atender às demandas de aplicações em tempo real. Os avanços alcançados

corroboram os estudos prévios que apontam a YOLO como uma das arquiteturas mais eficientes para detecção de objetos em cenários dinâmicos e desafiadores, como enfatizado por Bochkovski et al. (2020).

5. CONSIDERAÇÕES FINAIS

A detecção de objetos em trânsito urbano utilizando a arquitetura YOLOv4 mostrou-se não apenas eficaz, mas também uma solução promissora para os desafios encontrados em ambientes urbanos complexos e dinâmicos. Este estudo evidenciou que a implementação prática desse modelo, com as devidas adaptações, é capaz de capturar com precisão nuances específicas do tráfego urbano, contribuindo significativamente para a segurança e eficiência nas vias públicas.

Os resultados obtidos durante a pesquisa demonstram que o modelo é altamente eficiente na identificação e rastreamento de uma ampla gama de objetos relevantes para o contexto viário, como placas de sinalização, pedestres e faixas de segurança. A capacidade de realizar essas detecções em tempo real, mesmo sob condições adversas, como variações de iluminação e diferentes densidades de tráfego, reforça a robustez e a aplicabilidade da solução desenvolvida.

Além disso, a abordagem utilizada revelou-se adequada para otimizar recursos computacionais e maximizar o desempenho, destacando o equilíbrio entre precisão e eficiência. Essa característica torna a aplicação do modelo viável em cenários reais, onde a necessidade de respostas rápidas e precisas é fundamental.

Com base nos avanços obtidos, abre-se a possibilidade de futuras aplicações e aprimoramentos. O modelo pode ser expandido para incluir outros tipos de objetos, integrar-se a sistemas mais amplos de gestão de tráfego urbano ou até mesmo ser utilizado como base para soluções autônomas em veículos inteligentes. Assim, o presente estudo não apenas cumpre os objetivos propostos, mas também pavimentamos o caminho para novas pesquisas e inovações no campo da visão computacional aplicada ao trânsito urbano.

Portanto as contribuições deste trabalho representam um passo importante na busca por soluções tecnológicas que atendam às demandas de mobilidade urbana, contribuindo para a construção de cidades mais seguras, organizadas e eficientes.

REFERÊNCIAS

ALVES, G. Detecção de Objetos com YOLO – Uma abordagem moderna, 2020. Disponível em: https://iaexpert.academy/2020/10/13/deteccao-deobjetos-com-yolo-uma-abordagem-moderna/doing_wp_cron=1703638729.1622610092163085937500. Acesso em: 29 nov. 2023.

BOCHKOVSKIY, A.; WANG, C.; LIAO, H. YOLOv4: Optimal Speed and Accuracy of Object Detection. ArXiv, 2020.

DE VASCONCELLOS, Eduardo Alcântara. Transporte urbano nos países em desenvolvimento: reflexões e propostas. Annablume, 2000.

EVERINGHAM, M.; VAN GOOL, L.; WILLIAMS, C. K. I.; WINN, J.; ZISSERMAN, A. The Pascal Visual Object Classes (VOC) Challenge. International Journal of Computer Vision, v. 88, n. 2, p. 303-338, 2010.

HUANG, J.; RATHOD, V.; SUNDERMANN, M. A. et al. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. CVPR, 2017.

LIU, W.; ANGUELOV, D.; ERHAN, D.; SZEGEDY, C.; REED, S.; FU, C.; BERG, A. SSD: Single Shot MultiBox Detector. European Conference on Computer Vision, 2016.

REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You Only Look Once: Unified, RealTime Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 779-788, 2016.

REDMON, J.; FARHADI, A. YOLOv3: An Incremental Improvement. ArXiv, 2018.

TACHIZAWA, Takeshy. Cidades inteligentes e sustentáveis e o avanço das tecnologias urbana. Revista Conecta, v. 7, n. 1, p. 35-52, 2024.