

Sistema de Detecção de Usuários Vulneráveis da Via Robusto à Baixa Iluminação

Vinicius Avena, Rodrigo S. Couto, Miguel Elias M. Campista,
Luís Henrique M. K. Costa *

¹Universidade Federal do Rio de Janeiro - PEE/COPPE/GTA

{avena, rodrigo, miguel, luish}@gta.ufrj.br

Resumo. Neste trabalho, é proposto um sistema ADAS baseado em visão computacional para aumentar a segurança de Usuários Vulneráveis das Vias (VRUs) em regiões de ponto cego, utilizando exclusivamente imagens de câmeras traseiras de veículos. O método desenvolvido integra detecção e rastreamento multi-objetos para emitir alertas de proximidade de forma eficiente. Além disso, o sistema incorpora um mecanismo auxiliar para condições de baixa luminosidade, baseado na detecção de faróis. Para apoiar a pesquisa nessa área, é apresentado o FilterLane-VRU, um novo conjunto de dados com cenas reais de tráfego urbano capturadas por câmeras traseiras e anotadas com rótulos de alerta. O pipeline proposto oferece uma solução confiável e de baixo custo para aumentar a segurança dos VRUs em ambientes urbanos.

Abstract. This work proposes a computer vision-based ADAS system designed to enhance the safety of Vulnerable Road Users (VRUs) in blind spot regions using only rear-view camera input. The method combines object detection and multi-object tracking to generate proximity alerts in real time. The system also integrates a fallback mechanism for low-light conditions, based on headlight detection. To support research in this domain, we introduce FilterLane-VRU, a new dataset composed of real-world rear-view urban traffic scenarios annotated with alert labels. The proposed pipeline offers a cost-effective and reliable solution for VRU protection in urban contexts.

1. Introdução

O desenvolvimento de Sistemas Avançados de Assistência ao Condutor (*Advanced Driver Assistance System* - ADAS) tem evoluído rapidamente nos últimos anos, impulsionado por melhorias tanto no hardware como em soluções de software baseadas em aprendizado profundo. No entanto, a adoção em larga escala desses sistemas ainda enfrenta desafios significativos, especialmente no que se refere à percepção do ambiente em tempo real [Mannion, 2019]. A capacidade de identificar, rastrear e prever com precisão o movimento de outros usuários da via é fundamental para que os ADAS desempenhem funções críticas de segurança, como evitar colisões e realizar manobras com segurança. Entre esses

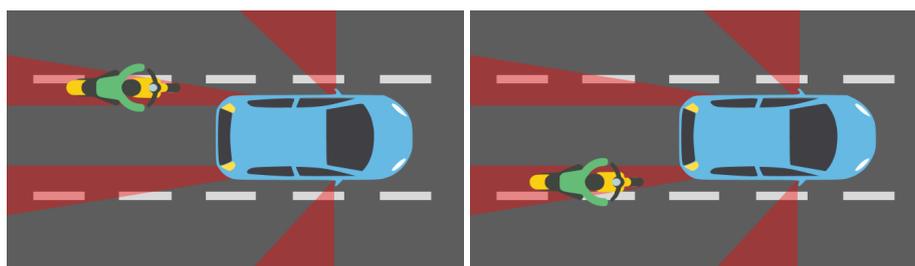
*O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) - Código de Financiamento 001. O trabalho também foi financiado pela Fundação de Desenvolvimento da Pesquisa - Fundep - Rota 2030, CNPq (SEI:01300.010794/2023-1, 408255/2023-4 e 309304/2021-0), FAPERJ (E-26/204.562/2024 e E-26/204.122/2024) e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), auxílios no. 23/00673-7 e 23/00811-0.

usuários, a proteção dos Usuários Vulneráveis das Vias (*Vulnerable Road User - VRU*), como pedestres, ciclistas e motociclistas, impõe desafios adicionais à percepção devido à sua maior exposição, alta mobilidade e comportamento imprevisível [Mannion, 2019].

Os Usuários Vulneráveis das Vias representam uma parcela expressiva dos acidentes de trânsito graves em todo o mundo. Em países em desenvolvimento, a ampla dependência de motocicletas e bicicletas tem agravado esse problema: no Brasil, por exemplo, as mortes de motociclistas quase dobraram na última década, passando de 16,8% para 30,2% [de Carvalho e Guedes, 2023]. Assim, garantir a segurança dos VRUs é um desafio crítico nos sistemas de transporte inteligentes modernos.

Um dos principais fatores que contribuem para acidentes envolvendo automóveis e VRUs é a presença de pontos cegos, áreas ao redor do veículo onde o campo de visão do condutor é obstruído. Embora espelhos retrovisores e sensores ultrassônicos auxiliem o condutor, esses sistemas têm alcance e campo de visão limitados, o que dificulta a antecipação de riscos em tempo hábil. Essa questão é particularmente preocupante em “corredores”, espaços estreitos entre veículos de faixa adjacentes frequentemente utilizados por motociclistas e ciclistas em tráfego urbano congestionado. Essas áreas são particularmente perigosas por estarem fora do campo de visão do condutor, dificultando a percepção de VRUs que trafegam entre faixas em cenários congestionados, situação comum em mudanças de faixa ou curvas repentinas.

As imagens da Figura 1 ilustram duas situações de risco envolvendo motociclistas se aproximando pelos corredores adjacentes do veículo observador. As áreas em vermelho indicam zonas de ponto cego típicas, onde o condutor não possui visibilidade direta. Em ambos os casos, o VRU encontra-se em uma posição crítica: fora do campo de visão dos espelhos retrovisores e em trajetória potencial de colisão. Esses cenários evidenciam a vulnerabilidade dos VRUs e reforçam a necessidade de sistemas proativos de detecção e alerta.



(a) Cenário de VRU no corredor à esquerda. (b) Cenário de conversão pela direita.

Figura 1. Cenários comuns de ponto cego envolvendo VRUs no tráfego urbano.

Embora métodos atuais detectem e classifiquem VRUs, falham em não incorporar a dinâmica do movimento para avaliar risco. Detectar a presença de um VRU, por si só, é insuficiente; um sistema eficaz deve distinguir se esse está se aproximando ou se afastando, adaptando a resposta conforme o contexto para evitar colisões. Além disso, a maioria das soluções foca na visão frontal, ignorando o alerta de pontos cegos a partir do uso de câmeras traseiras, um cenário ainda não explorado na literatura. Este trabalho avança ao propor um sistema que estima o risco com base na trajetória dos objetos, mesmo

em condições adversas, como oclusões e baixa iluminação.

Dessa forma, é proposto um sistema ADAS baseado em visão computacional, projetado para reforçar a segurança de VRUs em regiões de ponto cego traseiras, especialmente nos corredores entre faixas. A proposta busca complementar soluções já existentes, como sensores ultrassônicos ou radares de ponto cego, oferecendo uma percepção mais detalhada em ambientes urbanos densos, onde a movimentação de VRUs é frequente e imprevisível. Como contribuição, este trabalho apresenta o FilterLane-VRU, um novo conjunto de dados composto por gravações reais de câmera traseira com anotações temporais de alertas de proximidade, voltado para esse cenário ainda pouco explorado na literatura.

O sistema desenvolvido combina um detector de objetos estado-da-arte com estratégias complementares que ampliam sua robustez: um mecanismo auxiliar para condições de baixa iluminação baseado na detecção de faróis, rastreamento multiobjetos com filtro temporal, e análise de trajetória para identificar aproximações laterais consistentes. Essa abordagem viabiliza um pipeline embarcado, eficiente e sensível ao contexto, capaz de emitir alertas com alta confiabilidade mesmo em situações desafiadoras como oclusões ou iluminação reduzida. Todos os experimentos foram conduzidos com o conjunto de dados apresentado neste trabalho, e demonstraram que as técnicas propostas melhoram significativamente a precisão e a sensibilidade do sistema, evidenciando seu potencial para aplicações práticas em segurança veicular.

2. Trabalhos Relacionados

A segurança de Usuários Vulneráveis das Vias, com base em técnicas de detecção, tem sido amplamente estudada nos últimos anos, impulsionada pela demanda por sistemas de percepção mais robustos em ADAS e Veículos Autônomos (*Autonomous Vehicles - AV*). Embora o número de pesquisas sobre detecção e classificação de VRUs tenha aumentado significativamente [Silva et al., 2024], questões como oclusões, variações de iluminação e avaliação de risco em tempo real ainda persistem como desafios abertos, em especial nas soluções baseadas em visão computacional.

2.1. Abordagens Baseadas em Infraestrutura

Uma das abordagens para aumentar a segurança de VRUs envolve o uso da infraestrutura viária como fonte complementar de percepção. Nessa estratégia, sensores fixos, como câmeras, radares e LiDARs, são instalados ao longo das vias para monitorar o tráfego de forma distribuída e fornecer dados relevantes aos veículos. Esse modelo amplia o campo de visão e permite antecipar riscos, especialmente em cenários urbanos complexos.

[Borba et al., 2025] propõem uma arquitetura colaborativa entre veículos automatizados e infraestrutura inteligente, capaz de detectar VRUs mesmo em situações de oclusão. Seguindo direção semelhante, [Malinverno et al., 2018] exploram um sistema de prevenção de colisões baseado em comunicação veicular com a infraestrutura (*Vehicle-to-Infrastructure - V2I*), enquanto [Al-Qassab et al., 2018] propõem um sistema C-ADAS cooperativo no qual a infraestrutura fornece informações visuais para reforçar a percepção embarcada. Apesar dos avanços, tais soluções exigem conectividade estável e infraestrutura dedicada, o que dificulta sua adoção em larga escala, especialmente em países em desenvolvimento.

2.2. Detecção com Sensores Embarcados

Apesar do potencial das abordagens baseadas em infraestrutura, os sistemas embarcados continuam sendo a principal estratégia adotada para a detecção de VRUs, especialmente por sua independência de conectividade e infraestrutura externa. Essa abordagem fundamenta-se em sensores instalados diretamente nos veículos, como câmeras, radares e LiDAR [Ortiz et al., 2023].

Soluções baseadas em radar vêm sendo exploradas por sua robustez em condições climáticas adversas. Métodos desse tipo [Alaqueel et al., 2023, Dubey et al., 2020] demonstram potencial em cenários críticos, utilizando respostas em frequência ou dados sintéticos para distinguir pedestres e ciclistas. No entanto, sofrem com resolução limitada e desempenho reduzido em altas velocidades. Já métodos com sensores LiDAR, como [Wu et al., 2018], segmentam pedestres e ciclistas diretamente na nuvem de pontos 3D. Apesar da alta precisão, essas abordagens exigem sensores caros e processamento especializado, limitando sua adoção em sistemas ADAS embarcados de baixo custo em veículos comerciais.

Por outro lado, abordagens com câmeras tratam a detecção como uma tarefa de classificação de objetos e se beneficiam dos avanços recentes em aprendizado profundo. Ainda assim, enfrentam desafios em ambientes urbanos complexos, como oclusões e variações de iluminação. Para aumentar a robustez, métodos de fusão multissensorial [Liu et al., 2023, Wang et al., 2024] integram dados de múltiplos sensores em arquiteturas unificadas. Esses sistemas tendem a apresentar maior precisão e robustez, mas demandam hardware embarcado mais sofisticado e caro, o que restringe sua adoção em soluções ADAS de baixo custo.

2.3. Detecção de VRUs em Ponto Cego

Apesar da relevância dos pontos cegos na segurança viária, poucos trabalhos investigam especificamente a detecção de VRUs nessas regiões, seja em veículos autônomos ou conduzidos por humanos. A literatura sobre o tema ainda é escassa. [Van Beeck e Goedemé, 2016] propõem um sistema de alerta baseado em visão para caminhões e veículos de grande porte, que emite avisos quando um VRU entra em zonas de perigo pré-definidas. No entanto, o uso de câmeras fixas montadas em posições elevadas restringe sua aplicação a veículos convencionais e compromete a adaptabilidade do sistema a cenários urbanos dinâmicos.

Em outro estudo, [Han et al., 2022] apresentam um método baseado em detecção de objetos com estimativa monocular de distância, voltado para pontos cegos em conversões à direita. Embora o modelo apresente baixa complexidade computacional, sua dependência de parâmetros fixos da câmera compromete a capacidade de generalização para diferentes cenários de tráfego e configurações veiculares. Além disso, o método é limitado a situações específicas de curvas à direita, não abrangendo corredores laterais ou outros comportamentos frequentes no trânsito urbano.

2.4. Posicionamento do Trabalho

Grande parte das soluções existentes para detecção de VRUs concentra-se na visão frontal dos veículos, explorando cenários como travessias de pedestres ou colisões frontais

iminentes. No entanto, a visão traseira também representa uma região crítica, especialmente em tráfego urbano denso, onde motociclistas e ciclistas avançam entre faixas pelos corredores laterais. Essas áreas, fora do campo de visão do condutor, estão sujeitas a manobras de risco como mudanças de faixa e freadas bruscas, com alto potencial de colisão com VRUs advindos dos pontos cegos traseiros.

Este trabalho propõe uma solução ADAS de baixo custo baseada apenas em câmera traseira, que opera embarcada e em tempo real, sem dependência de infraestrutura externa ou configurações multissensoriais. Para apoiar o desenvolvimento e avaliação do método, foi construído o FilterLane-VRU, um conjunto de dados com vídeos reais de tráfego urbano e anotações de alertas de aproximação de VRUs às zonas de risco. A proposta integra detecção de objetos, rastreamento multiobjetos e filtragem temporal em um fluxo unificado, voltado à proteção de VRUs em regiões de ponto cego pouco exploradas na literatura.

3. Conjunto de Dados FilterLane-VRU

Apesar da ampla disponibilidade de conjuntos de dados voltados para percepção visual em ambientes veiculares, a maioria das bases públicas com sinais de câmera concentra-se na visão frontal, negligenciando os desafios associados à percepção traseira e aos corredores, que são regiões críticas onde VRUs frequentemente circulam em meio ao tráfego urbano denso. Embora algumas bases amplamente utilizadas, como o *Waymo Open Dataset* [Sun et al., 2020] e o *nuScenes* [Caesar et al., 2020], incluam cobertura de 360 graus, suas anotações não contemplam explicitamente interações entre VRUs e veículos em faixas adjacentes a partir da perspectiva traseira.

O FilterLane-VRU foi construído para suprir essa lacuna, fornecendo dados voltados especificamente à percepção traseira de VRUs em contextos urbanos. O conjunto é composto por 95 vídeos reais capturados por uma câmera veicular Intelbras Duo DC 3201 2K+, posicionada de forma a simular a visão de uma câmera de ré convencional, com lente angular suficiente para cobrir as regiões de ponto cego nas laterais traseiras do veículo. A Figura 1 ilustra essas zonas de risco, onde o condutor não possui visibilidade direta e ocorrem interações frequentes com motociclistas e ciclistas em movimento. Cada sequência possui 10 segundos de duração, gravada a 30 quadros por segundo (FPS), totalizando 28.200 quadros anotados.

As gravações foram realizadas em variados contextos urbanos da cidade do Rio de Janeiro, abrangendo vias expressas, cruzamentos, áreas residenciais e trechos congestionados, sob diferentes condições de iluminação: diurna, noturna e ao entardecer. Foram consideradas motocicletas, bicicletas e seus respectivos condutores, abrangendo VRUs motorizados e não motorizados que trafegam entre faixas ou em regiões com visibilidade reduzida. A Figura 2 apresenta exemplos visuais desses diferentes cenários, que refletem a complexidade do tráfego urbano dinâmico e de alto risco.

Cada sequência foi anotada manualmente com a presença ou ausência de alerta de risco, indicando o instante exato de início do alerta e o lado (esquerdo ou direito) da aproximação do VRU em relação ao campo de visão traseiro do veículo observador. A Figura 3 apresenta as zonas de risco consideradas nas anotações, localizadas nas laterais traseiras do veículo — áreas tipicamente fora do campo de visão do condutor. Ao todo, o conjunto inclui 83 eventos anotados como alerta (60 do lado direito e 23 do lado es-

querdo), além de 26 eventos de “quase alerta”, nos quais há movimentação próxima às zonas de risco, mas sem condições suficientes para justificar a emissão do aviso, permitindo a avaliação da robustez do método proposto frente a falsos positivos.

Dessa forma, o FilterLane-VRU oferece suporte ao desenvolvimento e à avaliação de métodos de detecção sensíveis à trajetória em cenários de ponto cego. Para incentivar a reprodutibilidade e fomentar novos trabalhos, o conjunto está disponível para fins de avaliação mediante solicitação aos autores.

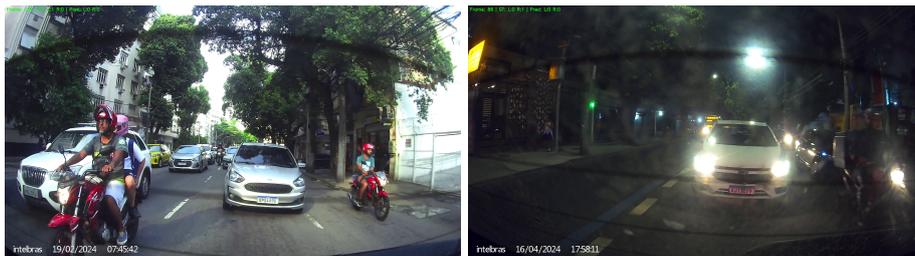


Figura 2. Exemplos de imagens do conjunto de dados FilterLane-VRU.



Figura 3. Definição dos alertas laterais (esquerdo e direito) no conjunto de dados FilterLane-VRU.

4. Sistema de Detecção de VRUs em Pontos Cegos

O sistema proposto visa aprimorar a detecção de VRUs em pontos cegos traseiros e emitir alertas em tempo real com base na dinâmica de aproximação lateral. Esse sistema processa fluxos de vídeo da câmera traseira para detectar e rastrear VRUs, avaliando seu movimento e proximidade em relação ao veículo observador. O sistema é composto por três módulos principais: (i) detecção de objetos, com mecanismo auxiliar para condições de baixa iluminação; (ii) rastreamento multiobjetos; e (iii) análise de trajetória para geração de alertas. A Figura 4 apresenta uma visão geral da arquitetura proposta. Cada componente é detalhado nas subseções a seguir.

4.1. Detecção de Objetos com Mecanismo Auxiliar em Baixa Iluminação

O sistema proposto utiliza a versão *X-large* do YOLOv8 [Jocher et al., 2023], um detector de objetos amplamente utilizado em aplicações em tempo real. O *You Only Look Once* (YOLO) [Redmon et al., 2016] é uma arquitetura de detecção de estágio único que realiza a previsão simultânea das caixas delimitadoras e das classes dos objetos diretamente a partir da imagem de entrada, sem a necessidade de etapas intermediárias de proposição de regiões, como ocorre em arquiteturas de dois estágios, a exemplo da Faster

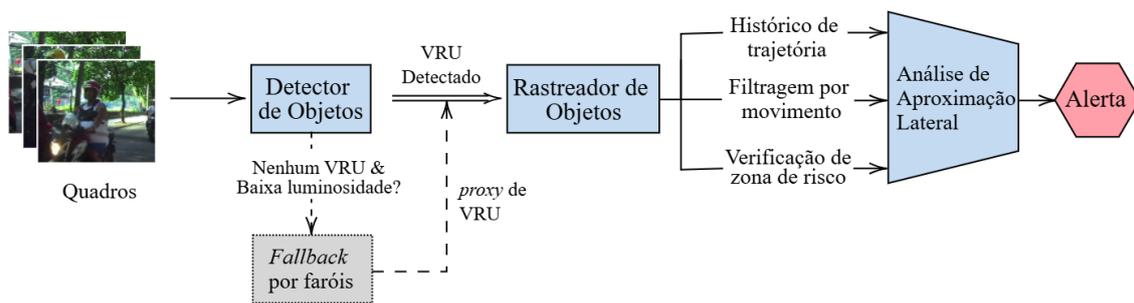


Figura 4. Pipeline do sistema proposto: da entrada do vídeo até a geração do alerta de risco de VRU.

R-CNN [Ren et al., 2015]. Essa abordagem direta é particularmente vantajosa para tarefas embarcadas em visão computacional, onde a latência de inferência é um fator crítico.

Entretanto, em ambientes com baixa iluminação, mesmo detectores robustos como o YOLO podem falhar ao identificar VRUs, principalmente quando esses estão subexpostos. Para mitigar esse problema, foi implementado um mecanismo auxiliar baseado na detecção de faróis, que atua como uma alternativa de reforço ao detector principal. Esse módulo é ativado apenas quando duas condições são simultaneamente atendidas: (i) o brilho médio da imagem está abaixo de um limiar mínimo e (ii) nenhuma detecção de VRU próximo foi realizada pelo modelo principal.

O funcionamento do módulo auxiliar (*fallback*) segue uma lógica baseada no processamento de imagem. Primeiramente, o quadro é convertido para escala de cinza e segmentado com base em intensidade luminosa, de modo a isolar regiões de alta luminosidade, típicas de faróis. Em seguida, aplica-se a extração de contornos e filtros espaciais para rejeitar reflexos irrelevantes. Apenas regiões localizadas na porção inferior da imagem são consideradas, assumindo que objetos relevantes devem estar próximos ao plano da via e dentro do campo de visão traseiro da câmera.

Cada região luminosa detectada é tratada como uma *região proxy*, ou seja, uma estimativa indireta da possível presença de um VRU. Embora essa técnica não identifique explicitamente as classes dos objetos, ela contribui para reduzir omissões ao explorar indícios visuais ainda perceptíveis em ambientes escuros. Vale destacar que, nesta etapa, o sistema não diferencia faróis de motocicletas daqueles de veículos maiores. Essa distinção é realizada apenas nas fases posteriores, por meio da análise da trajetória e da permanência da detecção nas regiões de risco.

4.2. Rastreamento de Múltiplos VRUs

O segundo módulo do sistema é responsável pelo rastreamento multiobjetos (*Multi-Object Tracking – MOT*), ou seja, por manter a identidade de cada VRU ao longo do tempo. Isso permite acompanhar suas trajetórias em quadros consecutivos, o que é fundamental para a análise de aproximação dos VRUs.

Este trabalho utiliza o algoritmo *ByteTrack* [Zhang et al., 2022], que apresenta bom desempenho em tempo real e baixa complexidade computacional, características importantes para aplicações embarcadas. Diferentemente de abordagens como o *DeepSORT* [Wojke et al., 2017], que utilizam descritores visuais, o *ByteTrack* baseia-se exclusi-



Figura 5. Amostras noturnas em que nenhum VRU foi detectado pelo YOLO, mas o mecanismo de apoio por detecção de farol inferiu com sucesso a presença de possíveis veículos.

vamente em sobreposição espacial entre caixas de detecção, sem depender de redes neurais adicionais.

A cada quadro, o algoritmo processa todas as detecções recebidas, sejam elas provenientes do YOLO ou dos *proxies* luminosos do mecanismo de apoio descrito na Seção 4.1. As detecções de alta confiança, geradas pelo YOLO, são priorizadas na associação com rastros existentes, com base na sobreposição entre suas posições na imagem. Em seguida, detecções de menor confiança, como os *proxies* de farol, são usadas para manter rastros ativos que não foram atualizados, aumentando a robustez do sistema em situações de oclusão ou falha temporária de detecção. Rastros que permanecem inativos por um número determinado de quadros são descartados, enquanto novas detecções não associadas iniciam rastreamentos com identificadores únicos.

A saída do módulo consiste em objetos rastreados com histórico recente de posições, que serão analisados no módulo seguinte para verificar padrões de aproximação de VRUs e emitir alertas.

4.3. Análise de Trajetórias para Geração de Alertas

O terceiro módulo do sistema avalia se os objetos rastreados estão se aproximando do veículo observador ao longo dos corredores laterais, correspondentes às regiões de ponto cego. Para isso, mantém-se, para cada objeto, um histórico das posições das caixas delimitadoras nos últimos quadros. Com base nesse histórico, calcula-se um vetor médio de deslocamento, que indica a direção predominante do movimento recente.

Considera-se que um VRU está em processo de aproximação quando seu vetor médio aponta de forma consistente no sentido longitudinal, em direção às laterais esquerda ou direita, do veículo observador durante uma janela temporal pré-definida. Essa janela atua como um filtro para suavizar variações ruidosas e garantir que apenas movimentos contínuos sejam interpretados como aproximação real. Essa análise permite descartar objetos que estão se afastando, atravessando perpendicularmente ou permanecendo praticamente parados, reduzindo, assim, o número de alertas falsos.

A mesma lógica de filtragem por trajetória é aplicada aos objetos detectados pelo mecanismo auxiliar baseado em faróis, descrito na Seção 4.1. Para reduzir os falsos positivos causados por faróis de veículos não-VRUs, o sistema aplica a análise de trajetória também aos *proxies* luminosos. Quando dois faróis se mantêm alinhados e com deslocamento coordenado, são descartados como veículos maiores. Já detecções isoladas e consistentes nas zonas de risco são mantidas como possíveis VRUs. Como o rastreamento

unifica detecções do YOLO e do fallback ao longo do tempo, é comum que o mesmo objeto seja alternadamente detectado pelos dois módulos em diferentes quadros do vídeo, aumentando a robustez em condições de oclusão ou baixa iluminação.

Por fim, o sistema verifica se o objeto identificado como VRU em aproximação está nas regiões de risco definidas, correspondentes aos corredores laterais. Quando essa condição é atendida, o alerta correspondente, esquerdo ou direito, é registrado. Esse mecanismo garante que os alertas sejam acionados apenas quando há risco real de aproximação, considerando não só a presença de um objeto, mas também seu movimento consistente em direção aos pontos cegos, aumentando a precisão e a utilidade prática do sistema.

5. Avaliação Experimental

A eficácia do sistema proposto para detecção de VRUs em regiões de ponto cego foi avaliada com experimentos que abrangem a utilização do conjunto de dados FilterLane-VRU e os diferentes módulos implementados. Nesta seção, é descrito o método de avaliação, a definição da configuração de referência (*baseline*), a apresentação de resultados quantitativos e a análise qualitativa de casos representativos.

5.1. Protocolo de Avaliação

A tarefa é formulada como dois problemas de classificação binária independentes: a cada quadro, o sistema decide se deve acionar um alerta de risco para o lado esquerdo ou direito do veículo, com base na detecção de VRUs se aproximando pelas zonas de risco laterais (Figura 3). Como cada lado é avaliado separadamente, alertas podem ser emitidos de forma simultânea ou independente.

A avaliação é feita em nível de evento, correspondente a um intervalo de quadros anotado como situação de risco, considerando um alerta correto se emitido dentro desse intervalo. Isso evita penalizações por pequenas variações temporais e reflete melhor o objetivo prático do sistema: reagir a aproximações reais em momentos críticos. Utilizam-se métricas padrão de classificação binária: precisão, *recall*, taxas de falsos positivos e falsos negativos — esta última especialmente relevante em contextos de segurança veicular.

Os modelos (YOLOv8 e variantes) foram usados apenas em fase de inferência, sem qualquer ajuste fino com os dados do FilterLane-VRU. O conjunto foi utilizado exclusivamente para avaliação, visando mensurar o impacto das técnicas propostas. A adaptação dos modelos ao domínio específico do FilterLane-VRU, por meio de *fine-tuning* supervisionado, é considerada uma possibilidade para trabalhos futuros.

Como configuração de referência, adotou-se a detecção com YOLOv8X pré-treinado, seguida de uma verificação geométrica que testa a interseção das caixas com regiões laterais de risco. Embora simples, essa abordagem oferece uma base eficaz para monitoração em tempo real, mas não considera aspectos temporais e é vulnerável a falsos positivos e negativos causados por oclusões, baixa iluminação ou aparições momentâneas.

5.2. Comparativo de Detectores: Precisão vs. Velocidade

Para justificar a escolha do detector, foi realizada uma avaliação comparativa de três modelos pré-treinados, YOLOv8n, YOLOv8x e RT-DETR [Zhao et al., 2024], utilizando a configuração de referência para detecção de VRUs. A Tabela 1 apresenta os

resultados de precisão, *recall* e taxa de inferência de quadros por segundo média (FPS), obtidos em uma GPU NVIDIA RTX 2080 Ti, plataforma com desempenho comparável a soluções embarcadas de alto desempenho encontradas em veículos, como a série NVIDIA Drive AGX.

Tabela 1. Comparação de desempenho dos detectores no FilterLane-VRU.

Detector	Precisão	Recall	FPS
YOLOv8n	50,4	69,8	84,2
YOLOv8x	60,8	84,3	41,1
RT-DETR	59,5	78,3	14,8

Como mostrado na Tabela 1, o YOLOv8n apresentou maior velocidade, mas desempenho insuficiente em precisão e *recall*. O RT-DETR teve bons resultados, mas sua baixa taxa de FPS inviabiliza aplicações embarcadas. O YOLOv8x conciliou alta precisão e operação em tempo real (41 FPS), o que o torna adequado para sistemas ADAS e será adotado como detector principal nos experimentos subsequentes. Avaliações futuras incluirão testes em plataformas embarcadas reais para validar a viabilidade prática da proposta.

5.3. Impacto da Detecção com Módulo de Apoio em Condições de Baixa Iluminação

Para lidar com a limitação do modelo de referência em ambientes noturnos, foi introduzido o mecanismo de detecção complementar baseado na análise de faróis (Seção 4.1). Essa abordagem atua apenas quando o brilho do quadro é insuficiente e o detector principal falha, buscando capturar possíveis VRUs com base em padrões luminosos característicos.

A Tabela 2 mostra os resultados da avaliação em um subconjunto noturno do conjunto de dados. Neste experimento, o efeito do mecanismo de apoio por faróis foi avaliado isoladamente, ainda sem a aplicação do módulo de rastreamento e filtragem por trajetória, de forma a avaliar sua capacidade bruta de recuperar detecções em baixa iluminação.

Tabela 2. Impacto do mecanismo de apoio em cenas noturnas.

Método	Precisão	Recall	Taxa de FN	Taxa de FP
Referência	63,6%	41,2%	58,8%	36,4%
Com mecanismo de apoio	37,5%	70,6%	29,4%	62,5%

Sem o mecanismo de apoio, o sistema apresentou *recall* de 41,2%, precisão de 63,6% e taxa de falsos negativos de 58,8%. Com o módulo adicional, a taxa de falsos negativos foi reduzida para 29,4%, e o *recall* aumentou para 70,6%. Entretanto, a taxa de falsos positivos aumentou, reduzindo a precisão para 37,5%. Apesar do aumento, a expressiva redução na taxa de falsos negativos motiva o uso do mecanismo auxiliar em contextos de baixa iluminação, priorizando a sensibilidade em cenários de risco.

5.4. Avaliação Final do Sistema com Rastreamento e Análise de Trajetórias

Para avaliar os benefícios do rastreamento e da análise de trajetória, analisou-se o impacto desses módulos sobre as versões prévias do sistema. O objetivo é eliminar

detecções instáveis que não persistem ao longo de múltiplos quadros e distinguir, com base no movimento, os VRUs que de fato representam risco iminente. Essa combinação visa reduzir tanto falsos positivos quanto falsos negativos, melhorando a precisão e a robustez dos alertas emitidos.

A Tabela 3 resume os resultados obtidos para quatro configurações: a versão de referência com detecção isolada, o uso do mecanismo de apoio por faróis, a introdução do rastreamento com análise de trajetória no sistema de detecção simples e, por fim, o sistema completo, ou seja a detecção com mecanismo de apoio e rastreamento das trajetórias.

Tabela 3. Desempenho do sistema sob diferentes configurações.

Configuração	Precisão	Recall	Taxa FP	Taxa FN
Detecção isolada (referência)	60,9%	84,3%	39,1%	15,7%
Com mecanismo de apoio	40,4%	91,6%	59,6%	8,4%
Rastreamento + trajetória	80,1%	81,9%	19,9%	18,1%
Sistema completo	79,4%	92,8%	20,6%	7,2%

A versão de referência, com detecção isolada, já apresenta boa cobertura dos eventos reais, mas sofre com muitos falsos positivos, resultado da ausência de qualquer forma de filtragem por rastreamento. A introdução do mecanismo de apoio baseado em faróis aumenta o *recall*, mas à custa de uma queda significativa na precisão, devido à inclusão de objetos irrelevantes em condições de baixa iluminação.

Na configuração com rastreamento e análise de trajetória, observa-se uma melhora expressiva na precisão. Isso mostra que a análise da trajetória é eficaz em rejeitar detecções breves ou inconsistentes e que a verificação da direção de movimento contribui para suprimir alertas irrelevantes, especialmente de objetos estáticos ou que se afastam.

Finalmente, o sistema completo combina os benefícios dos dois caminhos: recupera detecções perdidas em baixa luz por meio do mecanismo de apoio e valida os alertas com base em consistência temporal e direção de movimento. Essa integração resulta em um desempenho mais balanceado, com *recall* de 92,8% e precisão de 79,4%, mantendo baixas taxas de falsos negativos e falsos positivos.

Apesar do alto índice de falsos positivos quando usado isoladamente (59,6%), o mecanismo de apoio tem seu impacto controlado no sistema completo. O rastreamento e a análise de trajetória permitem descartar padrões compatíveis com veículos maiores, evitando alertas indevidos. Como resultado, a taxa de falsos positivos praticamente não se altera — de 19,9% para 20,6% — evidenciando que o módulo auxiliar é eficaz quando integrado aos demais componentes do pipeline.

5.5. Resultados Qualitativos

Para complementar a avaliação quantitativa, a Figura 6 apresenta exemplos visuais representativos de diferentes cenários enfrentados pelo sistema. As imagens ilustram a capacidade do modelo em lidar com múltiplos VRUs, variações de iluminação e complexidade do tráfego urbano.

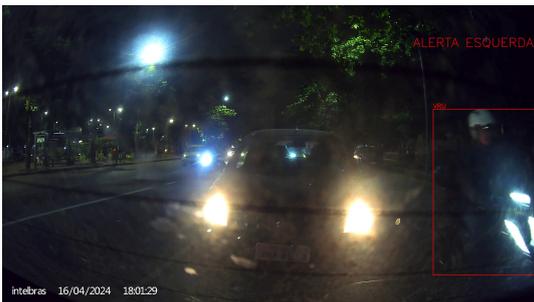
Esses exemplos reforçam a robustez do sistema proposto ao lidar com situações realistas. Os alertas são acionados de forma precisa tanto em ambientes bem iluminados



(a) Alerta unilateral e rastreamento de múltiplos VRUs se aproximando.



(b) Detecção simultânea via YOLO e mecanismo de apoio.



(c) VRU detectado com sucesso em baixa iluminação.



(d) VRU detectado com sucesso sob oclusão parcial.

Figura 6. Exemplos de geração de alertas.

quanto em cenas noturnas, e o mecanismo de rastreamento evita falsos positivos causados por faróis ou detecções momentâneas inconsistentes.

6. Conclusão

Este trabalho apresentou um sistema ADAS baseado exclusivamente em visão traseira, voltado para a detecção de aproximação de VRUs ao longo de corredores entre faixas adjacentes, regiões críticas e pouco exploradas na literatura. O método proposto integra detecção de objetos, rastreamento multiobjetos e análise de trajetória para acionar alertas sensíveis ao contexto de movimentação, filtrando objetos irrelevantes e reduzindo falsos positivos. Também introduziu-se um mecanismo auxiliar de detecção em baixa iluminação, baseado na identificação de faróis. Além disso, foi construído o FilterLaneVRU, um conjunto de dados voltado para esse cenário desafiador e ainda pouco explorado.

Os resultados experimentais demonstram que cada módulo contribui de forma complementar para o desempenho final. O rastreamento e a análise de trajetória aumentaram a precisão ao filtrar detecções instáveis, enquanto o mecanismo de apoio aumentou a sensibilidade do sistema em cenas noturnas. Embora o mecanismo de apoio, usado isoladamente, gere falsos positivos, esse efeito é amplamente mitigado na versão completa do sistema, que mantém alto *recall* sem comprometer a precisão — evidenciando que o módulo auxiliar é eficaz apenas quando combinado com mecanismos temporais e espaciais mais robustos. Como limitações, destaca-se que o sistema ainda depende de modelos pré-treinados, que podem não generalizar bem para aparências atípicas de VRUs ou condições extremas de iluminação. Além disso, o módulo auxiliar utiliza critérios visuais simples,

que podem ser refinados futuramente.

Para trabalhos futuros, propõe-se ampliar o conjunto de dados com novos exemplos de aproximações traseiras em condições desafiadoras, o que permitirá investigar o treinamento de modelos personalizados mais sensíveis aos padrões visuais de VRUs e melhor adaptados à detecção com base na visão traseira. Além disso, pretende-se implementar o sistema em plataformas embarcadas de baixo consumo, como unidades de processamento dedicadas em veículos, avaliando seus requisitos mínimos de hardware e o tempo de resposta para emissão de alertas. Isso permitirá validar sua viabilidade prática em cenários veiculares reais, conforme exigido em aplicações comerciais de ADAS.

Referências

- Al-Qassab, H., Pang, S., Al-Qizwini, M. e Radha, H. (2018). Visual sensor fusion and data sharing across connected vehicles for active safety. Technical Report 2018-01-0026, SAE International, Warrendale, PA, USA.
- Alaqeel, A. A., Alburadi, A., Nashashibi, A. Y., Sarabandi, K. e Shaman, H. (2023). Detection and identification of pedestrians and bicyclists using j-band automotive radars. Em *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, p. 6157–6159.
- Borba, T. d., Vaculín, O., Marzbani, H. e Jazar, R. N. (2025). Increasing safety of vulnerable road users in scenarios with occlusion: A collaborative approach for smart infrastructures and automated vehicles. *IEEE Access*, 13:8851–8885.
- Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G. e Beijbom, O. (2020). nuscenes: A multimodal dataset for autonomous driving. Em *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, p. 11621–11631.
- de Carvalho, C. H. R. e Guedes, E. P. (2023). Balanço da primeira década de ação pela segurança no trânsito no brasil e perspectivas para a segunda década. *Administração Pública. Governo. Estado*.
- Dubey, A., Fuchs, J., Reissland, T., Weigel, R. e Lurz, F. (2020). Uncertainty analysis of deep neural network for classification of vulnerable road users using micro-doppler. Em *2020 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT)*, p. 23–26.
- Han, L., Zheng, P., Li, H., Chen, J., Hua, Z. e Zhang, Z. (2022). A novel early warning strategy for right-turning blind zone based on vulnerable road users detection. *Neural Computing and Applications*, p. 1–20.
- Jocher, G., Qiu, J. e Chaurasia, A. (2023). Ultralytics YOLO.
- Liu, Z., Tang, H., Amini, A., Yang, X., Mao, H., Rus, D. L. e Han, S. (2023). Bevfusion: Multi-task multi-sensor fusion with unified bird's-eye view representation. Em *2023 IEEE international conference on robotics and automation (ICRA)*, p. 2774–2781.
- Malinverno, M., Avino, G., Casetti, C., Chiasserini, C.-F., Malandrino, F. e Scarpina, S. (2018). Performance analysis of c-v2i-based automotive collision avoidance. Em *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, p. 1–9.

- Mannion, P. (2019). Vulnerable road user detection: state-of-the-art and open challenges. *arXiv preprint arXiv:1902.03601*.
- Ortiz, F. M., Sammarco, M., Costa, L. H. M. K. e Detyniecki, M. (2023). Applications and services using vehicular exteroceptive sensors: A survey. *IEEE Transactions on Intelligent Vehicles*, 8(1):949–969.
- Redmon, J., Divvala, S., Girshick, R. e Farhadi, A. (2016). You only look once: Unified, real-time object detection. Em *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 779–788.
- Ren, S., He, K., Girshick, R. e Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- Silva, R. M., Azevedo, G. F., Berto, M. V. V., Rocha, J. R., Fidelis, E. C., Nogueira, M. V., Lisboa, P. H. e Almeida, T. A. (2024). Vulnerable road user detection and safety enhancement: A comprehensive survey.
- Sun, P., Kretschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B. et al. (2020). Scalability in perception for autonomous driving: Waymo open dataset. Em *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, p. 2446–2454.
- Van Beeck, K. e Goedemé, T. (2016). The automatic blind spot camera: A vision-based active alarm system. Em *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part I 14*, p. 122–135. Springer.
- Wang, Z., Huang, Z., Gao, Y., Wang, N. e Liu, S. (2024). Mv2dfusion: Leveraging modality-specific object semantics for multi-modal 3d detection. *arXiv preprint arXiv:2408.05945*.
- Wojke, N., Bewley, A. e Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. Em *2017 IEEE international conference on image processing (ICIP)*, p. 3645–3649.
- Wu, B., Wan, A., Yue, X. e Keutzer, K. (2018). Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud. Em *2018 IEEE international conference on robotics and automation (ICRA)*, p. 1887–1893.
- Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W. e Wang, X. (2022). Bytetrack: Multi-object tracking by associating every detection box. Em *European conference on computer vision*, p. 1–21.
- Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., Liu, Y. e Chen, J. (2024). Detrs beat yolos on real-time object detection. Em *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 16965–16974.