



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA

Detecção de Presença de Pedestres em Áreas de Interesse Usando Múltiplas Câmeras

Kenedy Felipe dos Santos da Silva

Recife - PE, Fevereiro 2024

Kenedy Felipe dos Santos da Silva

Detecção de Presença de Pedestres em Áreas de Interesse Usando Múltiplas Câmeras

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Informática Aplicada da Pró-Reitoria de Pós-Graduação da Universidade Federal Rural de Pernambuco, como parte dos requisitos necessários para obtenção do grau de Mestre.

Universidade Federal Rural de Pernambuco

Pró-Reitoria de Pós-Graduação

Programa de Pós-Graduação em Informática Aplicada

Orientador: Prof. Dr. João Paulo Silva do Monte Lima

Recife - PE

Fevereiro 2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

S586d Silva, Kenedy Felipe dos Santos da
Detecção de Presença de Pedestres em Áreas de Interesse Usando Múltiplas Câmeras / Kenedy Felipe dos Santos da
Silva. - 2024.
66 f. : il.

Orientador: Joao Paulo Silva do Monte Lima.
Inclui referências.

Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Informática
Aplicada, Recife, 2024.

1. Detecção. 2. Rastreamento. 3. Presença de Pedestres. 4. Áreas de Interesse. 5. Múltiplas Câmeras. I. Lima, Joao
Paulo Silva do Monte, orient. II. Título

CDD 004

Kenedy Felipe dos Santos da Silva

Detecção de Presença de Pedestres em Áreas de Interesse Usando Múltiplas Câmeras

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Informática Aplicada da Pró-Reitoria de Pós-Graduação da Universidade Federal Rural de Pernambuco, como parte dos requisitos necessários para obtenção do grau de Mestre.

Aprovada em Recife - PE, Fevereiro 22, 2024:

**Prof. Dr. João Paulo Silva do Monte
Lima** (Orientador)
Universidade Federal Rural de Pernambuco

Prof. Dr. Filipe Rolim Cordeiro
Universidade Federal Rural de Pernambuco

**Prof. Dr. João Marcelo Xavier Natário
Teixeira**
Universidade Federal de Pernambuco

Recife - PE
Fevereiro 2024

Dedico a Harlen Duarte Freire, grande irmão e amigo. (in memoriam)

Agradecimentos

Agradeço a Deus primeiramente, que sempre me deu sabedoria, indicou as melhores ações e reações em toda a minha jornada para aquisição de conhecimento durante o mestrado e toda a minha vida acadêmica.

Agradeço imensamente a minha querida e companheira de todos os momentos, Rayane Gomes, que me ajuda e me dá força a todo tempo, me encorajando a seguir em frente, nas dificuldades e nas adversidades. Amo você, muito obrigado por fazer parte de mais uma conquista, uma conquista nossa.

Agradeço a toda a minha família, principalmente meus pais, que sempre me apoiam. Meu pai, que sempre fomentou a minha carreira para a área de tecnologia e a minha mãe, que sempre me deu toda educação e oportunidades de crescimento. Minha irmã Jéssica, que sempre me mostrou que ser mais nova não quer dizer que será apenas ensinada e que aprendi muito com você. Ao meu irmão Harlen (In Memoriam), que por ser mais velho, me mostrou muito o quanto eu poderia aprender e conhecer mais da área de tecnologia. Aos meus irmãos do Rio de Janeiro e meus familiares que moram lá também. Meus avós, que sempre dizem que sou um orgulho e agradeço por todo apoio. Minhas tias que desejam muito serem minhas mães.

Agradeço ao meu orientador João Paulo Lima, que sempre me motivou e mostrou os melhores cenários e por ser um exemplo de pesquisador, compartilhando muito conhecimento e conquistas. Agradeço ao time do Voxar Labs, em especial a Isabella Andrade por compartilhar ideias e soluções nas aplicações de detecção, a Rafael Roberto, Kelvin Cunha, Francisco Simões (Chico), Maria Euzébio, Veronica Teichrieb (VT) e a todo o time Voxar que participaram e me ajudaram em todas as etapas do mestrado, sou eternamente grato por tudo que fizeram e estão fazendo por mim.

Agradecimentos ao time Neurotech, em especial a minha equipe Squad Crédito 3 - 220V, que sempre me apoiou e ajudou em todos os momentos e compartilhamento de conhecimento, agradecimento especial ao meu coordenador João Carlos Almeida, que sempre encorajou o crescimento, a busca por conhecimento e a conquistar o mundo. Ao time atual da Squad 3 e aqueles que passaram pelo nosso time, todos têm a sua cota de contribuição nesta minha jornada.

“Porque o Senhor dá a sabedoria, e da sua boca vem o conhecimento e o entendimento.”
(Provérbios 2:6)

Resumo

O monitoramento de pedestres em cenários de segurança é um problema que abrange diversos ambientes da sociedade. A pesquisa sobre detecção de pessoas em áreas de interesse é motivada pela necessidade de abordar a segurança em indústria observando todas as movimentações e compartilhamento do ambiente com objetos e veículos. A complexidade inerente ao monitoramento de pedestres em tais cenários impulsiona a busca por soluções eficazes, focando em técnicas de detecção e rastreamento. O presente estudo tem como objetivo não apenas compreender, mas também aprimorar a detecção de pedestres em áreas específicas, visando promover uma maior segurança para os transeuntes. Dada a constante demanda por vigilância e a complexidade do problema, a nossa pesquisa busca abordagens que minimizem os custos associados a treinamentos extensivos, proporcionando resultados precisos e eficientes na identificação e monitoramento de pedestres em áreas críticas. O foco está na problemática da detecção e rastreamento de pedestres em ambientes urbanos, com uma ênfase especial na avaliação e comparação de técnicas utilizando múltiplas câmeras. O estudo propõe soluções para identificar a presença de pedestres em locais específicos, considerando a complexidade do mundo real e a evolução da visão computacional. A implementação de um algoritmo de validação da presença de pedestres, utilizando técnicas avançadas de visão computacional e aprendizado de máquina, é realizada para alcançar resultados mais precisos. A avaliação dos métodos é conduzida através de métricas de desempenho, comparando as previsões com a informação real de cada exemplo. Os resultados evidenciam a eficácia dos métodos desenvolvidos, contribuindo para um entendimento dos padrões de movimento dos pedestres em espaços urbanos. A introdução de um conjunto de dados valida a integração da tecnologia proposta. A análise comparativa entre a técnica de detecção e a técnica de rastreamento de pedestres revela resultados notáveis com o dataset Wildtrack, com uma acurácia de 0,952 e F-score de 0,970 para a detecção, e acurácia de 0,948 e F-score de 0,967 para o rastreamento. A precisão foi de 0,977 e 0,984 para detecção e rastreamento, respectivamente. Resultados médios para a presença de pedestres nas áreas de interesse são apresentados, destacando a acurácia, F-score e precisão de ambas as técnicas.

Palavras-chaves: Detecção, Rastreamento, Presença de Pedestres, Áreas de Interesse, Múltiplas Câmeras.

Abstract

Pedestrian monitoring in security scenarios is a problem that encompasses various societal environments. Research on people detection in areas of interest is motivated by the need to address safety in industries by observing all movements and sharing the environment with objects and vehicles. The inherent complexity of pedestrian monitoring in such scenarios drives the search for practical solutions, focusing on detection and tracking techniques. The present study aims to understand and enhance pedestrian detection in specific areas, promoting excellent pedestrian safety. Given the constant demand for surveillance and the complexity of the problem, our research seeks approaches that minimize the costs associated with extensive training, providing accurate and efficient results in pedestrian identification and monitoring in critical areas. The focus is on the problem of pedestrian detection and tracking in urban environments, with a particular emphasis on evaluating and comparing techniques using multiple cameras. The study proposes solutions to identify the presence of pedestrians in specific locations, considering the complexity of the natural world and the evolution of computer vision. Implementing a pedestrian presence validation algorithm, using advanced computer vision and machine learning techniques, is conducted to achieve more accurate results. The methods are evaluated through performance metrics, comparing predictions with the actual information of each example. The results demonstrate the effectiveness of the developed methods, contributing to an understanding of pedestrian movement patterns in urban spaces. The introduction of a dataset validates the integration of the proposed technology. Comparative analysis between pedestrian detection and tracking techniques reveals notable results with the Wildtrack dataset, with an accuracy of 0.952 and an F-score of 0.970 for detection and an accuracy of 0.948 and an F-score of 0.967 for tracking. The precision for detection and tracking was 0.977 and 0.984, respectively. Average results for pedestrian presence in areas of interest are presented, highlighting both techniques' accuracy, F-score, and precision.

Keywords: Detection, Tracking, Pedestrian Presence, Areas of Interest, Multiple Cameras.

Lista de ilustrações

Figura 1 – Circulação de pessoas em um ambiente industrial (MECALUX, 2024)	2
Figura 2 – Simulação de acidente em indústria (ONFLEX, 2024)	3
Figura 3 – Detecções de Poses - <i>Alphapose</i> (FANG et al., 2022)	6
Figura 4 – Tecnologia <i>Deepsort</i> (WOJKE; BEWLEY, 2018)	8
Figura 5 – Metodologia de calibração de câmeras (ZHANG, 2000)	10
Figura 6 – Re-identificação de pessoas em múltiplas câmeras (ZHENG et al., 2015)	11
Figura 7 – <i>ETH Pedestrian Detection Dataset</i> (GEIGER; LENZ; URTASUN, 2012)	13
Figura 8 – Rastreamento 3D de 7 câmeras do conjunto de dados Wildtrack. (VO et al., 2020)	17
Figura 9 – Multi-Target Multi-Camera Tracking by Tracklet-to-Target Assignment. (HE et al., 2020)	18
Figura 10 – Detecção com modelo SSD. (HE et al., 2020)	20
Figura 11 – Identificação de limitações. (SHAHZAD; JALAL, 2021)	21
Figura 12 – Rastreamento online de múltiplos objetos(ZHANG; WANG; GU, 2021)	22
Figura 13 – Resultados de detecção de pedestres em 3D. (LIMA et al., 2022)	25
Figura 14 – Regularização <i>DropView</i> . (VORA et al., 2023)	26
Figura 15 – Projeções de características e seu desempenho. (HOU; ZHENG; GOULD, 2020)	27
Figura 16 – Abordagem da técnica com multi visualização não rotulada. (LIMA et al., 2023)	27
Figura 17 – Rastreamento de pedestres em 3D. (LYRA et al., 2022)	28
Figura 18 – Detecção de anomalias por Pustokhina (PUSTOKHINA et al., 2021)	29
Figura 19 – Combinação implementada por Asarnia et al. (ANSARNIA et al., 2022)	30
Figura 20 – Dataset <i>Wildtrack</i>	32
Figura 21 – Capturas do Dataset Si.U.	33
Figura 22 – Momentos Dataset Si.U.	33
Figura 23 – Detecção de pedestres ao mesmo tempo instantânea e visualização do plano de solo de cada pedestre. (LIMA et al., 2022)	34
Figura 24 – Detecção de pedestres ao mesmo tempo instantânea e visualização do plano de solo de cada pedestre. (HOU; ZHENG; GOULD, 2020)	35
Figura 25 – Resultados de rastreamento de pedestres vistos em múltiplas câmeras ao mesmo tempo instantaneamente t e sua trilha no plano terrestre. (LYRA et al., 2022)	36
Figura 26 – Áreas de interesse (SILVA; LIMA; TEICHRIEB, 2023)	37

Lista de tabelas

Tabela 1 – Grupos de áreas de interesse 1 - Wildtrack	39
Tabela 2 – Grupos de áreas de interesse 2 - Wildtrack	39
Tabela 3 – Métricas de avaliação dos métodos de Detecção e Rastreamento de pedestres dentro das áreas de interesse para dataset <i>Wildtrack</i>	40
Tabela 4 – Métricas de avaliação de frames com detecção de pedestres entre métodos de Detecção e Rastreamento	41
Tabela 5 – Comparação de frames com detecção de pedestres entre métodos de Detecção e Rastreamento	41
Tabela 6 – Métricas de avaliação dos métodos de Detecção de pedestres dentro das áreas de interesse para os últimos 40 frames do dataset <i>Wildtrack</i>	42
Tabela 7 – Grupos de áreas de interesse 1 - Si.U	43
Tabela 8 – Grupos de áreas de interesse 2 - Si.U	43
Tabela 9 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - LA01	43
Tabela 10 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - LA02	44
Tabela 11 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - NI01	44

Lista de abreviaturas e siglas

3D	Tridimensional
CMU	<i>Carnegie Mellon University</i>
CNN	<i>Convolutional Neural Network</i>
DeepSort	<i>Deep Tracking System</i>
DukeMTMC	<i>Duke Multi-Target, Multi-Camera Tracking</i>
GCN	<i>Graph Convolutional Network</i>
GloVT	<i>Global Vision Transformer</i>
GNN	<i>Graph Neural Networks</i>
HCI	<i>Internacional Conference On Human-Computer Interaction</i>
HD	<i>High Definition</i>
HOG	<i>Histogram of oriented gradients</i>
IoU	<i>Intersection Over Union</i>
KITTI	<i>Karlsruhe Institute of Technology and Toyota Technological Institute</i>
LA	<i>Last Afternoon</i>
LiDAR	<i>Light Detection and Ranging</i>
mAP	<i>Mean Average Precision</i>
MODA	<i>Multi-Objetc Detection Accuracy</i>
MODP	<i>Multi-Object Detection Precision</i>
MOT	<i>Multiple Object Tracking</i>
NI	<i>Night</i>
PETS	<i>Performance Evaluation of Tracking and Surverllance</i>
PREC	<i>Precisão (Precision)</i>
R-CNN	<i>Region Based Convolutional Neural Networks</i>

RCLL	<i>Recall</i>
ReID	Re-Identificação
RGB	<i>Red Green Blue</i>
RNN	<i>Recurrent Neural Network</i>
RSSF	Redes de Sensores Sem Fio
Si.U	<i>Spatial Intelligence for Urban Intelligence</i>
SIFT	<i>Scale Invariant Feature Transform</i>
SSD	<i>Single Shot Multibox Detector</i>
SURF	<i>Speeded-Up Robust Features</i>
UFRPE	Universidade Federal Rural de Pernambuco
VIRAT	<i>Video and Image Retrieval and Analysis Tool</i>
YOLO	<i>You Only Look Once</i>

Sumário

1	INTRODUÇÃO	1
1.1	Motivação	1
1.2	Justificativa	1
1.3	Objetivos	2
1.4	Organização do trabalho	3
2	REFERENCIAL TEÓRICO	4
2.1	Redes Neurais	4
2.2	Detecção de Objetos	5
2.3	Detecção de Pedestres	5
2.4	Rastreamento de Pedestres	7
2.5	Múltiplas Câmeras	9
2.6	Integração de Dados de Detecção para o Rastreamento	11
2.7	Conjuntos de Dados (Datasets)	11
2.8	Paradigmas e Abordagens	13
2.9	Desafios Atuais e Futuros	14
3	TRABALHOS RELACIONADOS	16
3.1	Abordagens de Detecção e Rastreamento	16
3.2	Avanços em Rastreamento de Pedestres	19
3.3	Detecção e Rastreamento de Pedestres	21
3.4	Estudos Integrados sobre Detecção e Rastreamento de Pedestres	23
3.5	Detecção de Presença de Pedestres em Áreas de Interesse	28
3.6	Análise e Tendências	30
4	MATERIAIS E MÉTODOS	31
4.1	Datasets Utilizados	31
4.2	Metodologia de Detecção em Áreas de Interesse	33
4.3	Definição e Verificação das Áreas de Interesse	36
4.4	Avaliação e Métricas de Desempenho	37
5	RESULTADOS	39
5.1	Dataset <i>Wildtrack</i>	39
5.2	Dataset Si.U	42
6	CONCLUSÕES	45
6.1	Contribuições	46

6.2	Trabalhos futuros	46
	REFERÊNCIAS	47

1 Introdução

A detecção e o rastreamento de pedestres em ambientes urbanos têm se tornado uma área de pesquisa crucial, especialmente com o advento de sistemas de vigilância e segurança baseados em câmeras. Este estudo aborda a problemática da detecção de pedestres em áreas de interesse, utilizando métodos de detecção (LIMA et al., 2022) (HOU; ZHENG; GOULD, 2020) (LIMA et al., 2023) (VORA et al., 2023) e de rastreamento (LYRA et al., 2022). A pesquisa foca na análise comparativa destes métodos, buscando compreender suas eficácias e limitações em diferentes cenários.

1.1 Motivação

Em um cenário urbano dinâmico, a segurança e a eficiência dos espaços públicos dependem, em grande parte, da capacidade de monitorar e compreender a movimentação de pedestres (Figura 1). Em análise ao mundo real, indústrias se mostram preocupadas com a possibilidade de acidentes junto aos seus colaboradores (Figura 2). A segurança em ambientes urbanos é uma preocupação crescente, e a avaliação da presença de pedestres desempenha um papel fundamental nesse contexto. Esta dissertação propõe uma análise da presença de pedestres em áreas de interesse determinadas pelo usuário, empregando métodos avançados de visão computacional. O foco principal é avaliar a eficácia de técnicas de detecção e rastreamento de pedestres, bem como investigar a integração de dados e tecnologias que visem para possibilitar a aplicação em novos conjuntos de dados, observando cenários diferentes entre si.

1.2 Justificativa

A necessidade de sistemas de monitoramento urbano mais adaptáveis e personalizáveis tem se tornado cada vez mais necessário (JABBAR et al., 2019), parceiros na indústria citam que o compartilhamento do ambiente entre veículos e pedestres é um grande risco e monitorar as ações de cada um no ambiente se faz necessário para melhorar a segurança de ambos. Ao centrar-se na avaliação da presença de pedestres em áreas de interesse determinadas pelo usuário, áreas estas identificadas como locais perigosos, que não podem ter passagem de pedestres, locais que poderão levar insegurança para os pedestres. Esta pesquisa contribui para a identificação de pedestres em áreas que ofereçam riscos, melhorando assim a segurança e a gestão do espaço das indústrias.



Figura 1 – Circulação de pessoas em um ambiente industrial (MECALUX, 2024)

1.3 Objetivos

O principal objetivo desta dissertação é avaliar, de maneira abrangente, a presença de pedestres em áreas de interesse personalizadas. A pesquisa não visa apenas identificar a eficácia dos métodos de visão computacional aplicados a detecção e rastreamento de pedestres, mas também compreender como essas técnicas podem ser adaptadas e personalizadas para atender às necessidades específicas de diferentes usuários.

Como objetivos específicos, pretende-se:

- **Avaliar Métodos de Detecção e Rastreamento de Pedestres:** Comparar e analisar criticamente métodos de visão computacional para detecção e rastreamento de pedestres. Examinar a precisão, eficiência e robustez desses métodos em ambientes urbanos, com foco nas áreas de interesse definidas pelo usuário.
- **Investigar a Integração de Novos Datasets:** Explorar a viabilidade e os desafios associados à integração de novos dados obtidos pelas câmeras para permitir a detecção de pedestres em áreas de interesse com sucesso. Analisar como os métodos podem ser combinados e integrados de forma a ter soluções com variações nas condições de iluminação, configurações de câmera e padrões de movimento dos pedestres.



Figura 2 – Simulação de acidente em indústria (ONFLEX, 2024)

1.4 Organização do trabalho

A dissertação está estruturada de forma a abordar diferentes aspectos relacionados à detecção e rastreamento de pedestres em áreas urbanas personalizadas. A seção de Referencial Teórico (Capítulo 2) fornece uma base conceitual, abordando tópicos como detecção de pedestres, rastreamento, utilização de múltiplas câmeras, integração de dados, conjuntos de dados relevantes, e desafios atuais e futuros. A seção Trabalhos Relacionados (Capítulo 3) apresenta uma revisão da literatura, destacando abordagens de detecção e rastreamento, avanços recentes, estudos integrados e paradigmas relevantes para o tema. Na seção Materiais e Métodos (Capítulo 4), a metodologia adotada é detalhada. Inclui a descrição dos datasets utilizados, processamento prévio dos dados, inclusão de novos datasets, algoritmos de detecção e rastreamento, definição e verificação de áreas de interesse, e métricas de avaliação de desempenho. Os Resultados e Discussão (Capítulo 5) apresentam e analisam os resultados obtidos com os datasets Wildtrack e Si.U, incluindo a definição de áreas de interesse, avaliação das métricas e discussão dos resultados. A seção de Conclusões (Capítulo 6) destaca as contribuições do trabalho, fornecendo uma visão geral das descobertas e propondo possíveis direções para trabalhos futuros.

2 Referencial Teórico

A detecção e rastreamento de pedestres em áreas de interesse utilizando múltiplas câmeras é uma tarefa essencial em aplicações de vigilância, monitoramento de tráfego, segurança pública e até mesmo na era dos veículos autônomos. A capacidade de identificar, localizar e rastrear pedestres de forma eficaz desempenha um papel fundamental na prevenção de acidentes, na otimização do tráfego e na segurança geral. O desenvolvimento de técnicas precisas e eficientes para detectar a presença de pedestres em ambientes dinâmicos e complexos é crucial para aprimorar a segurança e otimizar o gerenciamento de espaços urbanos.

Este referencial teórico busca proporcionar uma visão abrangente das técnicas, abordagens e desafios envolvidos na detecção e rastreamento de pedestres em ambientes monitorados por múltiplas câmeras.

2.1 Redes Neurais

As redes neurais têm se destacado como uma abordagem promissora para a detecção de pedestres em cenas complexas. Inspiradas pelo funcionamento do cérebro humano, essas redes são compostas por camadas de neurônios artificiais interconectados, capazes de aprender e extrair padrões complexos dos dados de entrada (LECUN; BENGIO; HINTON, 2015).

Em particular, as redes neurais convolucionais (CNNs) têm sido amplamente empregadas com sucesso na detecção de pedestres. As CNNs são capazes de aprender características discriminativas diretamente dos dados brutos de imagem, eliminando a necessidade de extração manual de características (KRIZHEVSKY; SUTSKEVER; HINTON, 2017).

Isso permite que as CNNs se adaptem automaticamente a diferentes condições de iluminação, variações de pose e fundos complexos, tornando-as altamente eficazes para tarefas de detecção de pedestres em tempo real (REN et al., 2016).

Além disso, o surgimento de arquiteturas avançadas, como redes neurais de detecção de objetos em uma única etapa (Single-Shot Detectors - SSD) e redes neurais de detecção de regiões (Region-Based Convolutional Neural Networks - R-CNNs), tem impulsionado ainda mais o desempenho e a precisão da detecção de pedestres" (LIU et al., 2016).

"Essas arquiteturas foram projetadas para lidar com desafios específicos da detecção de objetos, como a detecção de múltiplos objetos em uma única imagem e a redução do tempo de processamento" (GIRSHICK, 2015).

As redes neurais oferecem uma abordagem poderosa e flexível para a detecção de pedestres, permitindo a criação de sistemas de visão computacional robustos e eficazes para uma variedade de aplicações, desde monitoramento de segurança até veículos autônomos (LONG; SHELHAMER; DARRELL, 2015).

2.2 Detecção de Objetos

A detecção de objetos é uma tarefa fundamental em visão computacional, com aplicações que vão desde a segurança e vigilância até a automação industrial e veículos autônomos. Essa área visa identificar e localizar objetos de interesse em imagens ou vídeos, fornecendo informações essenciais para sistemas inteligentes interagirem com o ambiente circundante.

As técnicas tradicionais de detecção de objetos costumavam depender de métodos de extração de características manuais, como descritores de características específicas ou filtros de convolução projetados manualmente. No entanto, com o avanço das CNNs, a detecção de objetos deu um salto significativo em desempenho e precisão (LECUN; BENGIO; HINTON, 2015).

Uma das arquiteturas mais influentes para detecção de objetos é a *Faster R-CNN*, que combina uma rede neural convolucional para extração de características com um módulo de proposta de região para gerar regiões candidatas para detecção. Esse modelo introduz um mecanismo eficiente para a geração de propostas de região, permitindo uma detecção rápida e precisa de objetos em imagens (REN et al., 2016).

Além disso, abordagens como YOLO (*You Only Look Once*) tem ganhado destaque por sua capacidade de realizar detecções em tempo real com uma única passagem pela rede, tornando-as ideais para aplicações que exigem baixa latência, como sistemas de assistência ao motorista e vigilância em tempo real (ADARSH; RATHI; KUMAR, 2020).

Em resumo, a detecção de objetos baseada em redes neurais oferece uma abordagem poderosa e flexível para identificar e localizar objetos em imagens ou vídeos, impulsionando avanços significativos em uma variedade de aplicações de visão computacional (HE et al., 2017).

2.3 Detecção de Pedestres

A detecção de pedestres é o primeiro passo para identificar sua presença em uma cena. Vários métodos de detecção de pedestres foram desenvolvidos ao longo dos anos. Os métodos de detecção de pedestres em áreas de interesse variam em complexidade e eficácia. A detecção de pessoas utilizando múltiplas câmeras envolve técnicas sofisticadas que visam a integração de informações provenientes de diferentes pontos de vista (SUN et al., 2020).

Podemos utilizar métodos que contribuem para a detecção, em muitos casos, a integração das aplicações poderá gerar resultados que indiquem quais as melhores técnicas poderão ser agrupadas, com o intuito de resolver problemas do mundo real. Segue abaixo os métodos para detecção, que poderão ser combinados e testados para os cenários desejados:

Métodos como Histograma de Gradientes Orientados (HOG): é uma técnica que extrai características das imagens baseadas em gradientes de intensidade. É frequentemente utilizado em métodos de detecção de pedestres (DALAL; TRIGGS, 2005).

Detecção de poses: Um sistema de detecção de pose em tempo real é uma solução tecnológica projetada para identificar e rastrear as posições específicas das articulações e partes do corpo humano em tempo real, a partir de dados de entrada como imagens de câmeras ou vídeos em tempo real. Esses sistemas são utilizados em diversas aplicações, como realidade aumentada, jogos, interações homem-máquina, monitoramento de movimentos, análise de postura e muitas outras áreas.

Tecnologias como *AlphaPose* (FANG et al., 2022) sistema avançado de estimativa e rastreamento de pose multi-pessoa em tempo real, projetado para capturar poses corporais detalhadas de várias pessoas em cenas complexas, conforme a Figura 3. Como o *OpenPose* (Cao et al., 2019) avançado sistema de detecção de pose em tempo real, que tem como objetivo identificar a posição de pontos-chave em todo o corpo humano a partir de imagens ou vídeos. Desenvolvido pela *Carnegie Mellon University* e pela *CMU Perceptual Computing Lab*, o *OpenPose* é uma implementação de código aberto que utiliza técnicas de visão computacional e aprendizado profundo para realizar essa tarefa de forma eficiente e precisa.

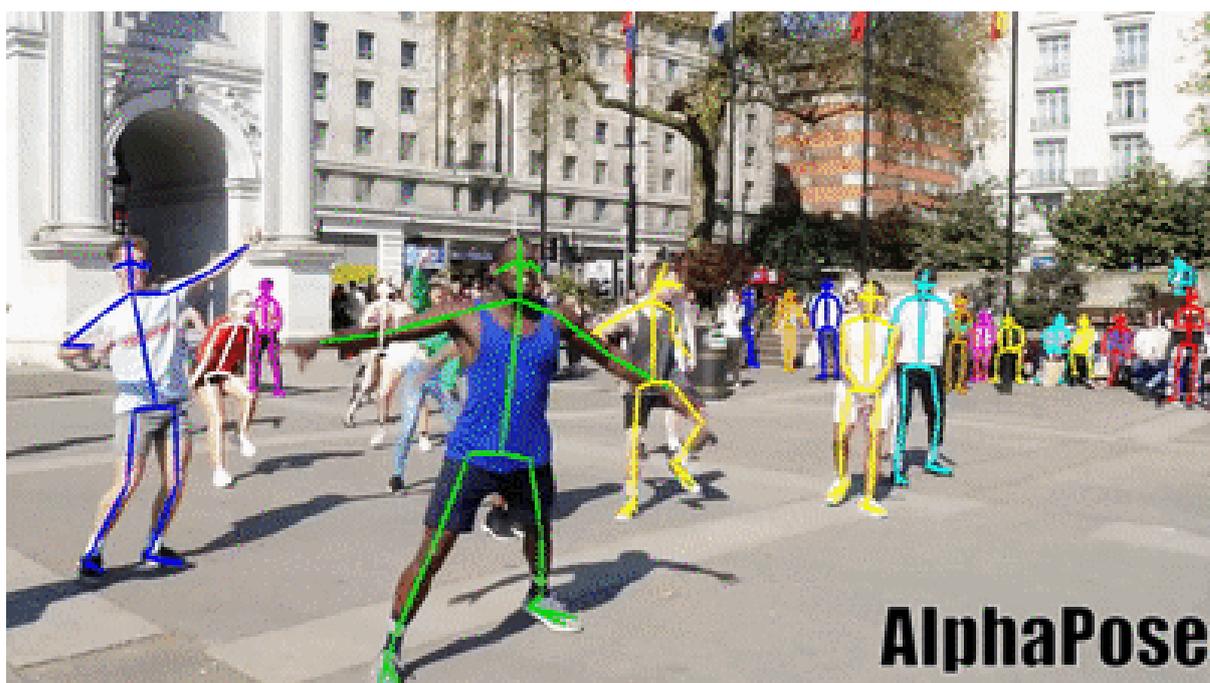


Figura 3 – Detecções de Poses - *AlphaPose* (FANG et al., 2022)

Redes Neurais Convolucionais (CNNs): As CNNs revolucionaram a detecção de objetos, incluindo pedestres. Arquiteturas como YOLO (*You Only Look Once*) e *Faster R-CNN* têm se mostrado eficazes para essa tarefa (REDMON et al., 2016).

Redes Neurais Recorrentes (RNNs) para Sequências Temporais: Em cenas de vídeo ou sequências temporais, as RNNs podem ser empregadas para rastrear pedestres ao longo do tempo. (LIANG; HU, 2015).

Aprendizado Profundo e Detecção de Padrões Complexos: Técnicas baseadas em aprendizado profundo são capazes de detectar padrões complexos e representações de alto nível, o que pode melhorar a detecção de pedestres (LIU et al., 2020).

Detecção Baseada em Densidade e Segmentação de Instâncias: Métodos que detectam pedestres com base em mapas de densidade ou segmentação de instâncias têm ganhado atenção significativa (DOLLAR et al., 2011).

Detecção de Pedestres com Métodos de Aprendizado Não Supervisionado: Abordagens não supervisionadas, como detecção de anomalias, podem ser aplicadas na identificação de pedestres em contextos inesperados ou não conformes (SCHLEGL et al., 2017).

Uso de Sensores Complementares: Além de dados de imagem, sensores como LiDAR (*Light Detection and Ranging*) têm sido integrados para melhorar a precisão da detecção de pedestres (LI; ZHANG; XIA, 2016).

Esses métodos e referências representam uma visão geral dos avanços e estratégias comuns na detecção de pedestres, mas a escolha adequada dependerá do contexto da aplicação, disponibilidade de dados e requisitos específicos do projeto de pesquisa.

2.4 Rastreamento de Pedestres

O rastreamento de pedestres é uma área fundamental em visão computacional e sistemas de vigilância. Existem diversas abordagens e técnicas desenvolvidas para lidar com os desafios específicos relacionados ao rastreamento de pessoas.

Após a detecção, o rastreamento de pedestres se torna essencial para acompanhar seus movimentos. O rastreamento preciso é fundamental para evitar a perda de alvos e manter um histórico contínuo de suas trajetórias. Rastrear pedestres é essencial para entender o movimento e o comportamento em tempo real. Métodos que combinam detecção com algoritmos de rastreamento, como o *Deep Tracking System (DeepSort)* (WOJKE; BEWLEY, 2018), melhoram a continuidade do acompanhamento em cenários complexos.

Seguem abaixo exemplos de técnicas aplicadas ao método de detecção que são aplicados aos métodos de rastreamento, são eles:

Re-identificação de Pedestres: Quando um pedestre é detectado em uma câmera,

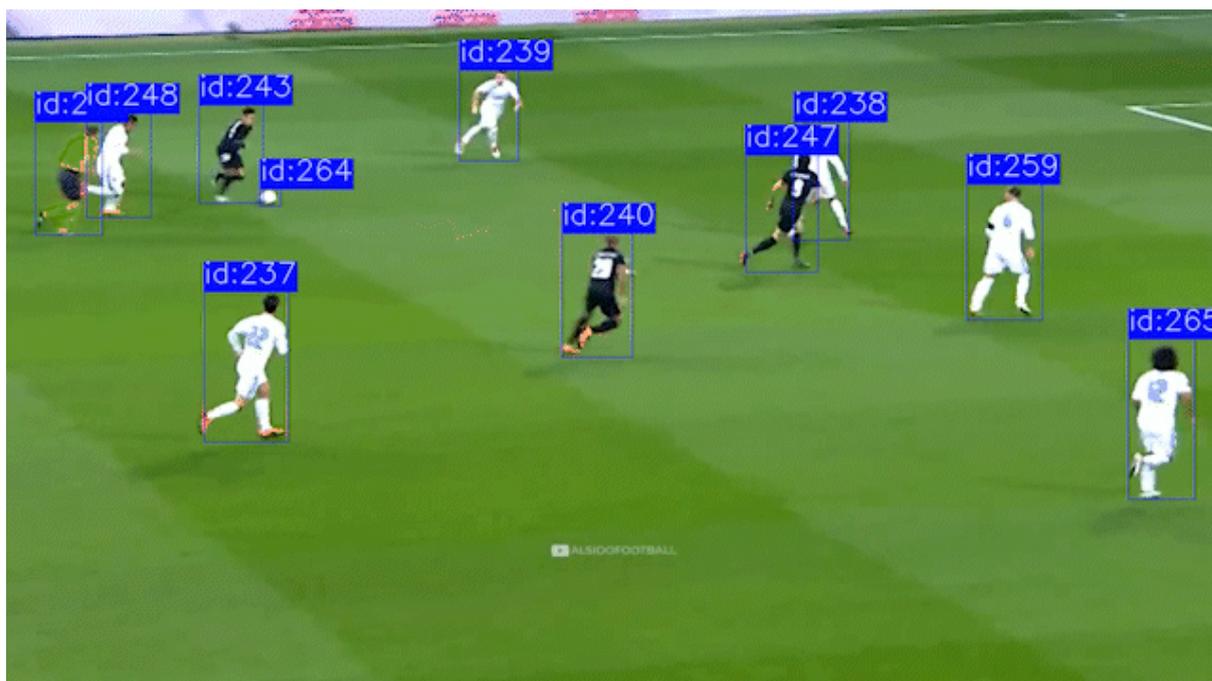


Figura 4 – Tecnologia *Deepsort* (WOJKE; BEWLEY, 2018)

a re-identificação é usada para confirmar sua identidade nas outras câmeras. Isso é especialmente útil para superar oclusões e mudanças de perspectiva. (ZHENG et al., 2015).

Métodos Baseados em Aprendizado Profundo: Redes neurais convolucionais (CNNs) têm sido amplamente empregadas para tarefas de detecção e rastreamento de pedestres. Modelos como *Faster R-CNN*, *YOLO (You Only Look Once)* (ADARSH; RATHI; KUMAR, 2020), e *SSD (Single Shot Multibox Detector)* são comuns nesse contexto (WOJKE; BEWLEY, 2018).

Rastreamento Baseado em Características de Aparência: Modela as características visuais únicas de um pedestre para identificação e rastreamento contínuo. Pode incluir métodos de aprendizado de máquina para aprendizado de características (COMANICIU; RAMESH; MEER, 2003).

Rastreamento Baseado em Contornos: Segmenta o contorno de pedestres para rastreamento. É eficaz em situações em que a forma do objeto é crucial para a identificação (WU; LIM; YANG, 2013).

Rastreamento Multi-Câmera: Utiliza informações de várias câmeras para rastrear um mesmo pedestre em diferentes pontos de vista. Essa abordagem é crucial para ambientes de vigilância complexos (BUTT; COLLINS, 2013).

Abordagens Baseadas em Grafos: Modela o rastreamento como um problema de grafos, onde os nós representam detecções e as arestas indicam associações entre detecções consecutivas (WOJKE; BEWLEY, 2018).

Rastreamento 3D de Pedestres: Estende o rastreamento para o domínio tridimen-

sional, utilizando informações de profundidade para estimar a localização precisa dos pedestres (CHEN et al., 2016).

Avaliação de Desempenho: Métricas como Precisão de Rastreamento MODP (*Multi-Object Detection Precision*), Taxa de Erro de Rastreamento, e MODA (*Multi-Object Detection Accuracy*) são comumente usadas para avaliar o desempenho de algoritmos de rastreamento (MUELLER; SMITH; GHANEM, 2016).

2.5 Múltiplas Câmeras

Múltiplas câmeras são frequentemente implantadas em áreas de interesse para cobrir uma ampla gama de perspectivas e garantir a detecção e rastreamento eficazes. Conseguimos observar a importância do uso de múltiplas câmeras para aprimorar a precisão do rastreamento, pois o uso de múltiplas câmeras oferece uma vantagem significativa nas áreas extensas. Sistemas que integram informações de várias fontes, como o uso de Redes de Convolução em Grafos (GCNs) e/ou Redes Neurais de Grafos (GNNs - *Graph Neural Networks*), proporcionam uma compreensão mais abrangente do ambiente.

Para utilização de múltiplas câmeras, existem algumas técnicas e métodos que são aplicados para apoiar na identificação e combinação de dados entre as imagens, abaixo segue uma lista das técnicas aplicadas a múltiplas câmeras.

Filtro de Kalman: Uma técnica amplamente utilizada para prever a localização futura de pedestres com base em seus estados anteriores. O Filtro de Kalman é eficaz, especialmente quando combinado com algoritmos de detecção (KALMAN, 1960).

Calibração de Câmeras: Para integrar informações de múltiplas câmeras, é crucial realizar a calibração, que envolve estimar parâmetros intrínsecos e extrínsecos de cada câmera (Figura 5). Essa etapa garante a correspondência correta entre os diferentes pontos de vista (ZHANG, 2000)

Fusão de Dados de Detecção: As detecções de pedestres feitas individualmente por cada câmera podem ser fundidas para obter uma visão abrangente. Métodos de fusão incluem média ponderada, votação majoritária e técnicas mais avançadas, como fusão baseada em probabilidade (HALL et al., 2017).

Modelagem de Trajetórias: Para rastrear pedestres em um ambiente coberto por várias câmeras, a modelagem de trajetórias se torna essencial. Métodos como filtragem de Kalman estendido ou modelos de espaço de estados são comumente empregados (RASMUSSEN; HAGER, 2001).

Sincronização Temporal: Garantir a sincronização temporal entre as câmeras é crucial para uma fusão de dados eficaz. Mecanismos de sincronização, como carimbos de data/hora precisos, são aplicados para alinhar os frames corretamente (LASASSMEH;

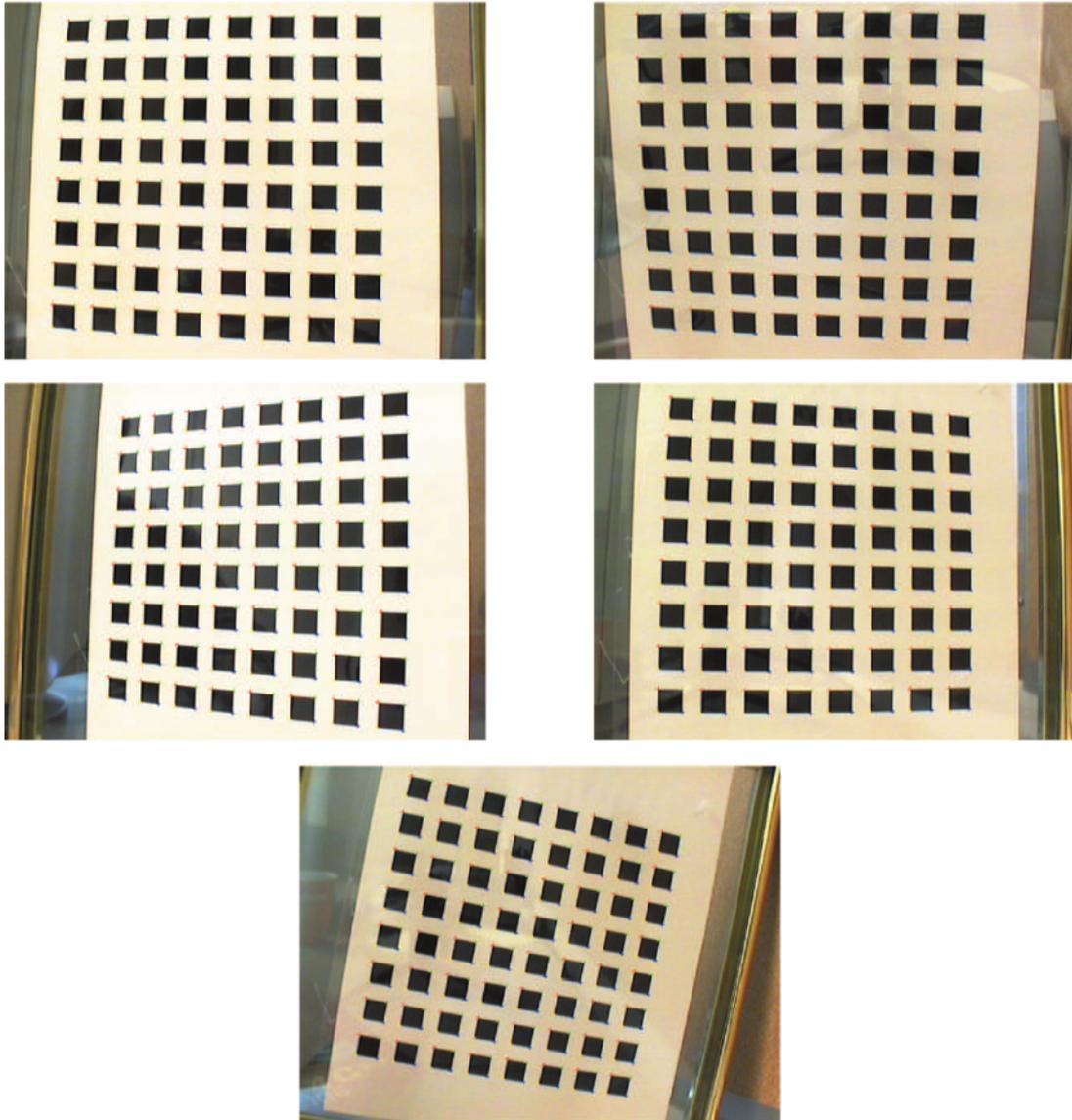


Figura 5 – Metodologia de calibração de câmeras (ZHANG, 2000)

CONRAD, 2010).

Redes de Sensores Sem Fio (RSSF): A integração de RSSFs pode ser benéfica para a troca eficiente de informações entre câmeras distribuídas. Essa abordagem contribui para uma cobertura mais ampla da área de interesse (AKYILDIZ et al., 2002).

Aprendizado Profundo para Fusão de Dados: Técnicas baseadas em redes neurais profundas são cada vez mais aplicadas para a fusão de dados de múltiplas câmeras. Redes siamesas, redes de atenção e abordagens de aprendizado profundo end-to-end podem ser exploradas (DEVLIN et al., 2018), (XIAO et al., 2016), (KOCH et al., 2015).

Métodos Baseados em Grafos: Modelos de grafos podem ser usados para representar relações entre diferentes pontos de vista das câmeras. Isso facilita a associação e rastreamento de pedestres em tempo real (YANG et al., 2022).

Considerações Éticas e Privacidade: Ao implementar sistemas de vigilância com

múltiplas câmeras, considerações éticas e de privacidade devem ser levadas em conta. Políticas claras de proteção de dados e anonimização podem ser aplicadas (MÜLLER, 2020).

2.6 Integração de Dados de Detecção para o Rastreamento

A incorporação de informações de profundidade, seja por meio de sensores dedicados ou técnicas de visão estéreo, auxilia na detecção de oclusões e contribui para uma análise mais precisa em três dimensões, melhorando a compreensão do ambiente.

Além da detecção, atributos e trajetórias dos pedestres são cruciais para entender seu comportamento. O uso de técnicas como a Re-identificação de Pessoa (ReID) e modelos de trajetória melhora a identificação e previsão de movimentos, como mostra a Figura 6 (ZHENG et al., 2015). O problema da re-identificação de pessoas envolve a correspondência de uma pessoa detectada em quadros diferentes. Isso é essencial para manter trilhas de pedestres em ambientes complexos.

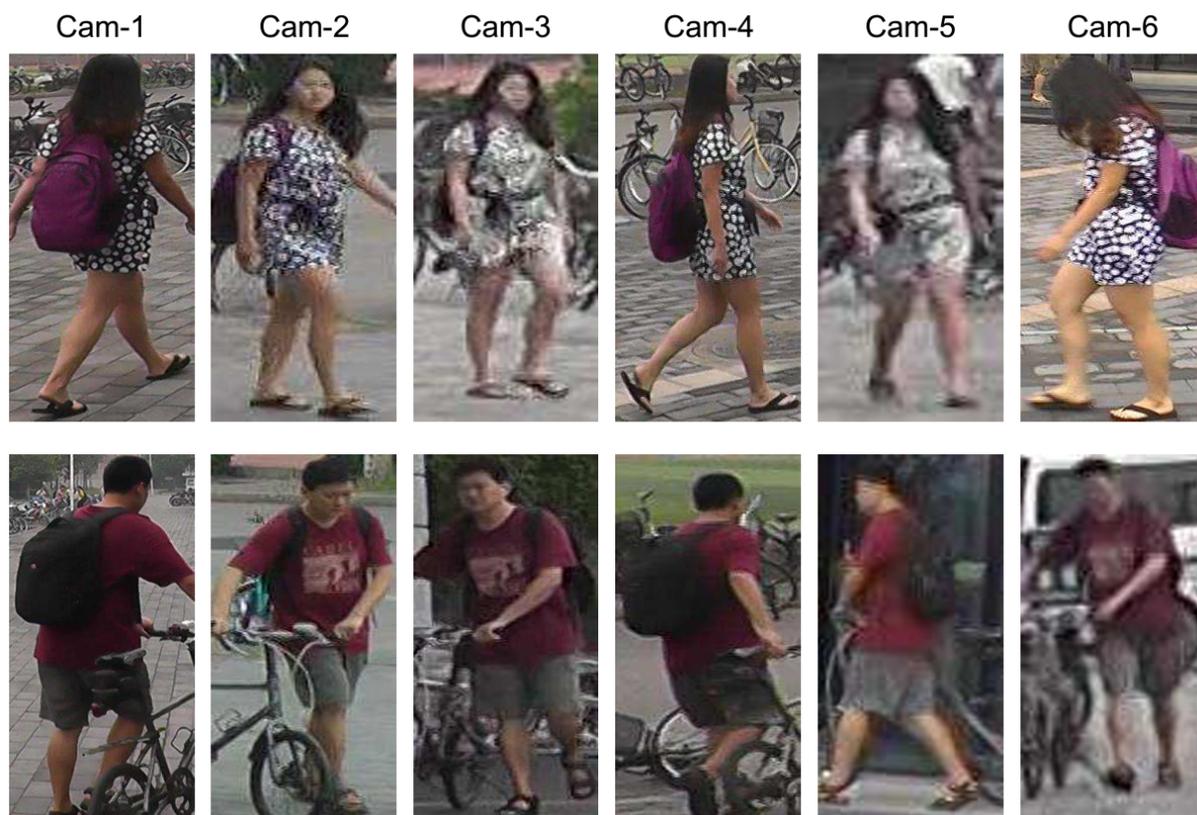


Figura 6 – Re-identificação de pessoas em múltiplas câmeras (ZHENG et al., 2015)

2.7 Conjuntos de Dados (Datasets)

A avaliação de algoritmos de detecção e rastreamento de pedestres é fundamental. Muitos trabalhos usam conjuntos de dados relevantes, como o *Wildtrack* (CHAVDAROVA

et al., 2018) para cenas não roteirizadas e o *MOTChallenge* (LEAL-TAIXÉ et al., 2015) para cenários desafiadores de múltiplos objetos. Normalmente oferecem desafios específicos e oportunidades para validar e aprimorar os algoritmos de detecção e rastreamento em ambientes diversificados, projetando desafios encontrados em ambientes urbanos. Ao compreender a singularidade dos conjuntos de dados, os pesquisadores e profissionais poderão aprimorar suas abordagens e modelos, resultando em soluções mais eficazes e robustas para a detecção de pedestres em áreas de interesse.

Podemos observar mais datasets que poderão ajudar a realizar a avaliação da detecção e rastreamento de pedestres. Abaixo listamos os principais datasets.

MOT (Multiple Object Tracking) Datasets: Conjunto de datasets projetados para avaliar algoritmos de rastreamento de múltiplos objetos, incluindo pedestres, em uma variedade de cenários e condições (MUELLER; SMITH; GHANEM, 2016).

KITTI Vision Benchmark Suite: Originalmente criado para avaliar algoritmos de visão veicular, o KITTI inclui dados para rastreamento de pedestres, oferecendo cenas urbanas desafiadoras (GEIGER et al., 2013).

DukeMTMC (Duke Multi-Target, Multi-Camera Tracking) Dataset: Dataset que visa rastreamento de múltiplos alvos em múltiplas câmeras, fornecendo desafios como oclusões e variações de iluminação (RISTANI et al., 2016).

ETH Pedestrian Detection Dataset: Contém sequências de vídeos capturadas por câmeras montadas em veículos, oferecendo dados para o rastreamento de pedestres em diferentes condições, exemplo na Figura 7 (GEIGER; LENZ; URTASUN, 2012).

PETS (Performance Evaluation of Tracking and Surveillance) Datasets: Séries de datasets focados na avaliação de algoritmos de rastreamento e vigilância. Inclui cenas internas e externas (FERRYMAN; SHAHROKNI, 2009).

Wildtrack: A Multi-Camera HD Dataset for Dense Unscripted Pedestrian Detection: Desenvolvido para detecção de pedestres densa em ambientes não roteirizados. Contém múltiplas séries de vídeos anotados (CHAVDAROVA et al., 2018).

CityPersons: Projetado para detecção de pedestres em ambientes urbanos, este dataset também fornece dados úteis para pesquisas em rastreamento de pedestres (ZHANG; BENENSON; SCHIELE, 2017).

MOTChallenge: Uma série de desafios anuais que fornecem datasets específicos para avaliar algoritmos de rastreamento de múltiplos objetos, incluindo pedestres (LEAL-TAIXÉ et al., 2015).

Esses datasets são fundamentais para o desenvolvimento e avaliação de algoritmos de rastreamento de pedestres em uma variedade de contextos, desde ambientes urbanos até situações mais desafiadoras. A escolha do dataset dependerá dos objetivos específicos

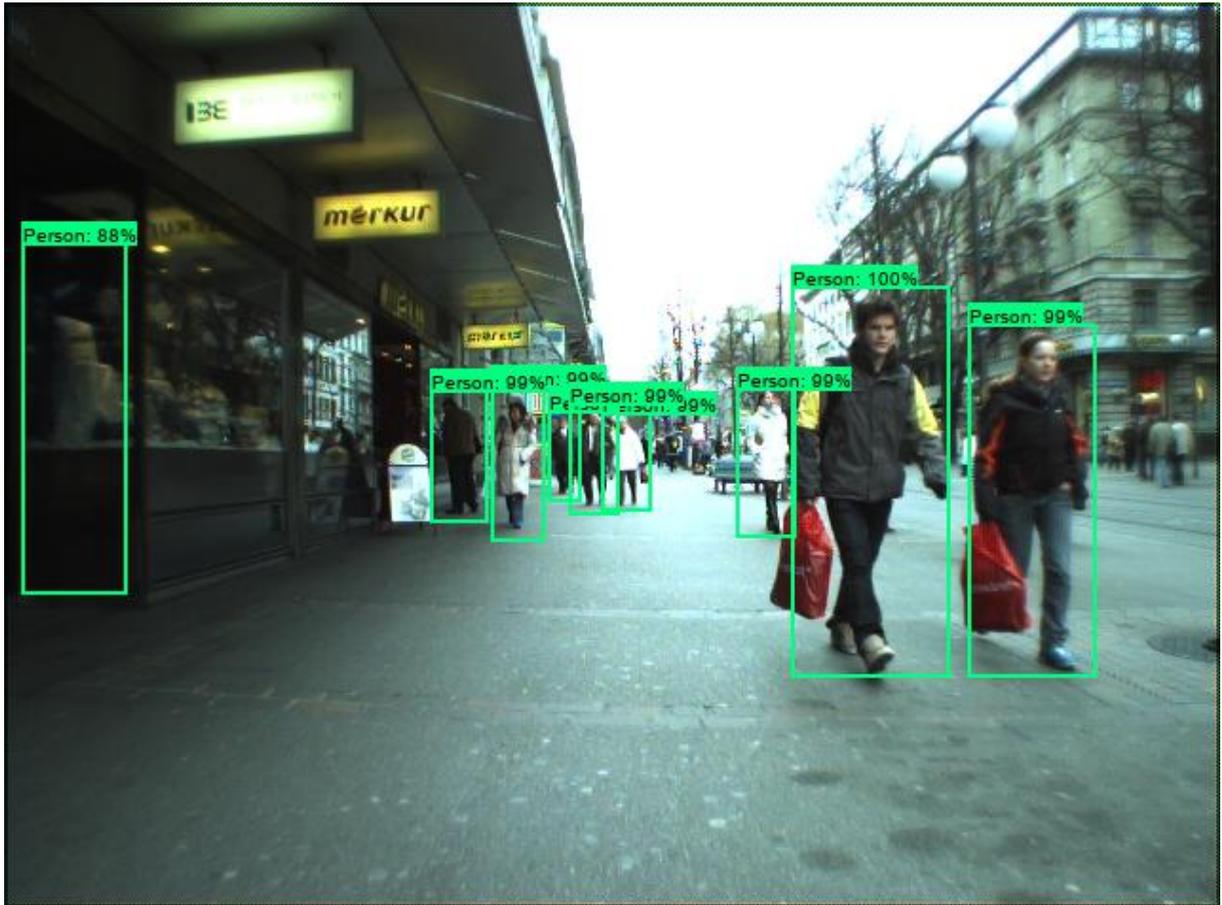


Figura 7 – ETH *Pedestrian Detection Dataset* (GEIGER; LENZ; URTASUN, 2012)

da pesquisa e das condições que se pretende abordar.

2.8 Paradigmas e Abordagens

Há vários paradigmas e abordagens utilizados na detecção e rastreamento de pedestres em visão computacional e sistemas de inteligência artificial. Alguns dos paradigmas mais comuns incluem:

Detecção baseada em Redes Neurais Convolucionais (CNNs): Utilização de arquiteturas de redes neurais convolucionais, como YOLO (*You Only Look Once*), SSD (*Single Shot MultiBox Detector*) e Faster R-CNN, para detectar pedestres em imagens e vídeos (REDMON et al., 2016).

Extração de Características (*Feature-based*): Métodos tradicionais que se concentram na extração de características específicas de pedestres, como Histogramas de Gradientes Orientados (HOG), Histogramas de Fluxo Óptico e Descritores Locais Invariantes a Escala (SIFT, SURF) (DALAL; TRIGGS, 2005).

Aprendizado Semi-Supervisionado e Não Supervisionado: Algoritmos que podem aprender com conjuntos de dados parcialmente rotulados ou sem supervisão direta, como al-

goritmos de agrupamento (ex: K-means) e aprendizado semi-supervisionado (ex: algoritmos de propagação de rótulo) (ZHU, 2005).

Aprendizado por Transferência (*Transfer Learning*): Utilização de modelos pré-treinados em conjuntos de dados massivos para tarefas de detecção e rastreamento de pedestres, adaptando e ajustando os modelos para novos cenários ou domínios específicos (HE et al., 2016).

Detecção em Tempo Real e Eficiência Computacional: Desenvolvimento de algoritmos e arquiteturas que priorizam a eficiência computacional para possibilitar a detecção e rastreamento de pedestres em tempo real, especialmente em sistemas embarcados ou com recursos computacionais limitados (SANDLER et al., 2018).

Modelos baseados em Geometria e Aparência: Combinação de informações geométricas (ex: formas e contornos) e características visuais (ex: cor, textura) para identificar pedestres em diferentes condições de iluminação, pose e escala (ZHANG; LI; OGUNBONA, 2017).

Aprendizado Robusto e Adversarial: Desenvolvimento de modelos capazes de resistir a perturbações deliberadas ou ruídos no ambiente, como adversidades de iluminação, oclusões parciais e variações climáticas (KURAKIN; GOODFELLOW; BENGIO, 2018).

Redes Neurais Recorrentes e Redes Neurais Convolucionais 3D (3D CNNs): Utilização de arquiteturas temporais para lidar com o rastreamento de pedestres em sequências de vídeo, capturando informações temporais e espaciais (DONAHUE et al., 2015).

Aprendizado baseado em Grafos e Estruturas Hierárquicas: Modelos que incorporam informações contextuais e de relacionamento entre pedestres e seu ambiente, utilizando grafos e estruturas hierárquicas para melhorar a precisão do rastreamento e detecção (BERG; KIPF; WELLING, 2017).

Combinação de Métodos: Alguns trabalhos propõem abordagens híbridas, combinando técnicas de detecção e regressão para aprimorar o rastreamento de pedestres em situações desafiadoras. A combinação pode incluir a integração de informações de múltiplos sensores, como câmeras RGB e sensores de profundidade, para aprimorar a detecção e o rastreamento em diferentes condições ambientais. (CHENG; HAN, 2016).

Esses paradigmas e abordagens são frequentemente combinados e adaptados para atender a diferentes requisitos e desafios específicos encontrados na detecção e rastreamento de pedestres em ambientes variados.

2.9 Desafios Atuais e Futuros

Ao abordar o rastreamento de pedestres é essencial considerar a natureza específica do ambiente, as condições de iluminação, o uso de múltiplas câmeras e outras caracte-

rísticas que podem impactar o desempenho dos métodos escolhidos. Apesar dos avanços significativos, existem desafios contínuos na detecção e rastreamento de pedestres em áreas de interesse, como o tratamento de oclusões, movimentos abruptos e mudanças de postura. Além disso, a escalabilidade para ambientes de grande escala e a privacidade são preocupações crescentes.

A implementação de sistemas de detecção em ambientes públicos levanta questões éticas e de privacidade. Estratégias que equilibram a eficácia do sistema com o respeito aos direitos individuais são fundamentais. A detecção e rastreamento de pedestres em áreas de interesse apresentam vários desafios, incluindo:

Oclusões: A oclusão parcial ou completa de pedestres por objetos, outros pedestres ou estruturas arquitetônicas é um desafio comum que requer técnicas avançadas de rastreamento.

Variações de Iluminação e Condições Meteorológicas: Mudanças nas condições de iluminação e variações climáticas podem afetar a precisão da detecção.

Multi-câmera: A integração de informações de múltiplas câmeras em um ambiente de vigilância aumenta a complexidade do rastreamento.

Re-identificação: A re-identificação de pedestres entre câmeras ou após oclusões é um desafio crucial para manter a continuidade das trilhas.

Eficiência Computacional: A capacidade de realizar o rastreamento em tempo real é essencial, especialmente em cenários de segurança.

3 Trabalhos Relacionados

Neste capítulo, são apresentados trabalhos relacionados ao rastreamento de pedestres em ambientes diversos. A análise abrange diferentes métodos e paradigmas utilizados na detecção e no acompanhamento de pedestres por meio de abordagens baseadas em detecção e regressão.

Apresentamos uma revisão abrangente dos trabalhos relacionados à detecção da presença de pedestres em áreas de interesse usando múltiplas câmeras. Os estudos discutidos abordam diversos aspectos da detecção e rastreamento de pedestres, métodos de aprendizado de máquina e o uso de múltiplas câmeras para melhorar a precisão do rastreamento.

A análise dos artigos inclui técnicas baseadas em detecção, regressão e métodos combinados, todos aplicados ao rastreamento de pedestres. Os métodos variam desde abordagens baseadas em CNNs para detecção até o uso de filtros de Kalman, redes de identificação de características (ReID) e modelos de correspondência de movimento. Os trabalhos relacionados serve como base para o desenvolvimento da pesquisa nesta dissertação.

3.1 Abordagens de Detecção e Rastreamento

(VO et al., 2020), explora o conceito de associação de pessoas em múltiplas visualizações (*Multi-View*) de maneira auto-supervisionada, isto é, sem a necessidade de anotações humanas, pode ser visualizada na Figura 8. Baseado na criação de um método autônomo para associar identidades de pessoas detectadas em diferentes vistas de câmeras, sem depender de dados anotados manualmente para treinamento. Isso significa que o sistema é projetado para aprender automaticamente como associar corretamente indivíduos vistos em diferentes câmeras sem a necessidade de supervisão humana. Esta abordagem inclui técnicas de auto-supervisão, como aprendizado por contraste ou métodos similares que exploram as relações entre diferentes instâncias de pessoas em múltiplas visualizações para aprender representações robustas. Essas representações aprendidas são então utilizadas para associar automaticamente as identidades das pessoas rastreadas nas diferentes câmeras.

A proposta se concentra na auto-supervisão para resolver o desafio de associar corretamente pessoas detectadas em diferentes vistas de câmeras, o que pode ser crucial em ambientes de vigilância multi-câmera ou sistemas de monitoramento para garantir a precisão na identificação e rastreamento de indivíduos, visando melhorar a precisão e

eficiência no rastreamento de indivíduos detectados em diferentes câmeras, sem depender de anotações humanas para treinamento do sistema.

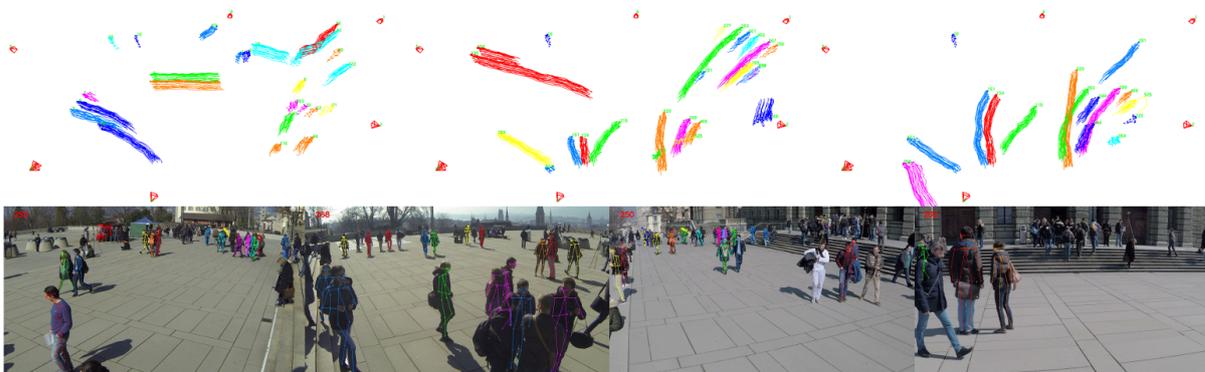


Figura 8 – Rastreamento 3D de 7 câmeras do conjunto de dados Wildtrack. (VO et al., 2020)

(YOU; JIANG, 2020), aborda o desenvolvimento de um sistema de rastreamento multi-câmera em tempo real utilizando técnicas de aprendizado profundo para a análise tridimensional. Focado na criação de um sistema capaz de rastrear objetos em um ambiente tridimensional usando informações provenientes de múltiplas câmeras em tempo real.

A abordagem proposta utiliza técnicas de aprendizado profundo para analisar as informações fornecidas por várias câmeras e rastrear objetos em 3D. Esse sistema provavelmente envolve a aplicação de redes neurais convolucionais (CNNs) ou arquiteturas similares para extrair e processar características dos objetos rastreados em diferentes câmeras. Obtendo a capacidade de rastrear objetos em 3D usando informações de várias câmeras pode ser fundamental para aplicações como vigilância, monitoramento de tráfego, realidade aumentada, entre outras, permitindo uma compreensão mais abrangente e detalhada do ambiente.

(BRASÓ; LEAL-TAIXÉ, 2020), trata do desenvolvimento de um resolvedor neural para o rastreamento de múltiplos objetos. Com foco na aplicação de técnicas de aprendizado de máquina, especificamente redes neurais, para resolver o desafio do rastreamento de múltiplos objetos em vídeos.

Esta abordagem foca na utilização de redes neurais para solucionar problemas associados ao rastreamento de múltiplos objetos, que tradicionalmente envolvem a associação correta de objetos detectados em diferentes quadros de vídeo ao longo do tempo. O *neural solver* (resolvedor neural) mencionado pode ser interpretado como um modelo de rede neural desenvolvido para aprender a associar e rastrear múltiplos objetos em um vídeo. Essa rede pode ser treinada com dados de entrada, como informações de detecção de objetos em quadros de vídeo sequenciais, e aprender a associar corretamente as identidades dos objetos ao longo do tempo.

(HE et al., 2020) aborda um método de rastreamento de múltiplos alvos em múltiplas câmeras por meio da atribuição de *tracklets* aos alvos. Realizando a implementação de um sistema capaz de rastrear simultaneamente múltiplos alvos em um ambiente com várias câmeras. A abordagem proposta utiliza *tracklets*, que são sequências de detecções temporais de um mesmo objeto, para realizar a associação entre esses *tracklets* e os alvos reais.

A atribuição de *tracklets* aos alvos pode ser realizada através de algoritmos de associação de dados ou métodos de aprendizado de máquina para identificar e agrupar adequadamente as sequências de detecções que pertencem a um mesmo alvo. Essa abordagem pode ser relevante em aplicações de vigilância e monitoramento, especialmente em situações onde é necessário rastrear vários alvos simultaneamente em um espaço coberto por múltiplas câmeras. O uso de *tracklets* pode oferecer uma maneira eficaz de gerenciar e associar os dados de detecção ao longo do tempo, resultando em um rastreamento mais preciso e robusto dos múltiplos alvos.



Figura 9 – Multi-Target Multi-Camera Tracking by Tracklet-to-Target Assignment. (HE et al., 2020)

(YOU; YAO; XU, 2020), aborda um método de rastreamento de múltiplos alvos utilizando associação de pose baseada em óptica em um ambiente multi-câmera. Um sistema capaz de rastrear simultaneamente múltiplos alvos, utilizando informações de pose (posição ou orientação) obtidas a partir de diferentes câmeras em um ambiente com várias

delas.

A proposta faz uso de informações visuais e geométricas de múltiplas câmeras para identificar e associar corretamente as poses dos alvos rastreados. O uso de dados de várias câmeras pode permitir uma visão mais abrangente da cena, facilitando a identificação e o rastreamento de alvos em diferentes perspectivas e condições de iluminação. A associação de pose baseada em óptica sugere o uso de características visuais das poses dos alvos para realizar a correspondência e manter a consistência no rastreamento, possivelmente utilizando algoritmos de correspondência e técnicas de visão computacional.

3.2 Avanços em Rastreamento de Pedestres

(STADLER; BEYERER, 2021a) aborda estratégias para aprimorar o rastreamento de múltiplos pedestres por meio do gerenciamento de trilhas e do tratamento de oclusões. O estudo visa melhorar a precisão e a continuidade do rastreamento de pedestres em ambientes desafiadores, nos quais há a presença de várias pessoas e podem ocorrer situações de oclusão, ou seja, quando os pedestres estão temporariamente obstruídos por outros objetos ou pessoas na cena.

Com o uso de técnicas avançadas de visão computacional e algoritmos de rastreamento para gerenciar as trilhas dos pedestres ao longo do tempo. O gerenciamento de trilhas refere-se à capacidade do sistema de rastreamento atribuir e manter consistentemente as identidades dos pedestres, mesmo quando eles saem e entram no campo de visão das câmeras ou estão temporariamente ocultos, concentrando-se no gerenciamento de trilhas e no tratamento de oclusões, visando melhorar a eficácia e a confiabilidade do rastreamento em ambientes desafiadores.

(SUNDARARAMAN et al., 2021), aborda a tarefa de rastreamento de cabeças de pedestres em multidões densas. A abordagem específica de rastreamento de cabeças de pedestres em ambientes onde há uma grande aglomeração de pessoas, ou seja, em multidões densas. Rastrear indivíduos em tais condições pode ser desafiador devido à proximidade física entre as pessoas, oclusões frequentes e variações significativas de movimento.

Utilizando técnicas de visão computacional e processamento de imagem para detectar e rastrear as cabeças dos pedestres. Isso pode envolver o uso de algoritmos de detecção de objetos e técnicas de rastreamento específicas para localizar e acompanhar as cabeças das pessoas na multidão. O objetivo desse tipo de pesquisa é oferecer uma solução para identificar e rastrear indivíduos específicos em multidões densas, o que pode ser valioso em cenários como eventos em estádios, manifestações, ou locais públicos movimentados.

(LI et al., 2021) trata de um modelo baseado em SSD (*Single Shot Multibox Detector*) para o rastreamento de pedestres em ambientes de emergência súbita ou violentos. É um

modelo baseado na arquitetura SSD, uma técnica de detecção de objetos em tempo real, para o rastreamento de pedestres em cenários considerados críticos, como ambientes de emergência ou violentos.

Incorporando redes neurais convolucionais profundas e técnicas de aprendizado profundo para detectar e rastrear pedestres em situações desafiadoras. O modelo SSD é conhecido por sua eficiência e capacidade de detectar objetos em tempo real. Com ênfase na adaptabilidade do modelo em condições extremas, onde há potencial para ocorrência de emergências ou situações violentas. Este tipo de sistema de rastreamento pode ser fundamental para monitorar e detectar pedestres em ambientes onde a segurança é uma preocupação crítica.



Figura 10 – Detecção com modelo SSD. (HE et al., 2020)

(STADLER; BEYERER, 2021b), aborda um método de rastreamento de múltiplos pedestres utilizando clusters. A implementação de um sistema de rastreamento capaz de acompanhar múltiplos pedestres em uma cena complexa. A abordagem proposta se baseia no conceito de agrupamento (clusters), que visa organizar os dados dos pedestres em grupos ou clusters semelhantes com base em características específicas, como localização, movimento, entre outras.

A utilização de clusters para o rastreamento de pedestres pode envolver técnicas de análise de agrupamento, onde os pedestres são associados a grupos com base em suas características comuns. Esse método pode oferecer uma forma de gerenciar e rastrear pedestres individualmente ou em grupos, auxiliando na identificação e na continuidade do rastreamento mesmo em cenários desafiadores, como oclusões temporárias ou movimentos complexos. A eficácia desse método depende da capacidade de identificar corretamente os clusters de pedestres e de manter a consistência na atribuição de indivíduos aos grupos ao longo do tempo.

(SHAHZAD; JALAL, 2021), descreve um sistema de vigilância inteligente para o rastreamento e contagem de pedestres utilizando correspondência de modelos (*template matching*). Esta abordagem implementa um sistema de vigilância inteligente que usa a técnica de correspondência de modelos para identificar e rastrear pedestres em um ambiente monitorado. A correspondência de modelos é uma técnica que compara um

modelo pré-definido (*template*, conforme ilustrado na Figura 11) com regiões de interesse nas imagens para detectar objetos semelhantes ao modelo.



Figura 11 – Identificação de limitações. (SHAHZAD; JALAL, 2021)

A pesquisa de (SHAHZAD; JALAL, 2021) utiliza uma abordagem baseada em visão computacional para realizar a detecção e o rastreamento de pedestres. O *template matching* é uma técnica simples, porém pode ser eficaz para tarefas de rastreamento em cenários controlados, onde as condições de iluminação e o fundo permanecem constantes ou variam de maneira previsível. A contagem de pedestres é outra funcionalidade chave deste sistema, permitindo registrar e quantificar o número de pedestres que entram ou saem de uma determinada área monitorada.

3.3 Detecção e Rastreamento de Pedestres

(YOON et al., 2021) apresenta um método para o rastreamento online de múltiplos pedestres utilizando uma abordagem baseada em correspondência de aparência temporal profunda. Desenvolvendo um algoritmo capaz de rastrear simultaneamente vários pedestres em tempo real, considerando as informações temporais e a correspondência de aparência entre diferentes quadros ou imagens sequenciais.

A metodologia proposta provavelmente utiliza técnicas de aprendizado profundo para extrair características visuais dos pedestres ao longo do tempo. Isso pode incluir o uso de CNNs ou RNNs para capturar e aprender as informações de aparência dos pedestres em diferentes quadros. A abordagem de correspondência de aparência temporal profunda busca identificar e associar corretamente os pedestres ao longo do tempo, levando em consideração as variações de aparência e movimento que podem ocorrer. Essa técnica

é especialmente útil para lidar com desafios como oclusões temporárias, mudanças de iluminação e variações de pose entre diferentes frames de vídeo.

O método proposto neste artigo tem relevância em aplicações de monitoramento de segurança, sistemas de vigilância urbana e análise de vídeos, onde o rastreamento preciso e contínuo de múltiplos pedestres é fundamental para garantir a segurança e compreender o comportamento em ambientes dinâmicos. Ao utilizar a correspondência de aparência temporal profunda, essa abordagem oferece uma solução promissora para o rastreamento eficiente e confiável de pedestres em tempo real.

(ZHANG; WANG; GU, 2021) aborda um método de rastreamento online de múltiplos objetos com foco em re-identificação de pedestres e processamento de oclusões. O estudo concentra-se na melhoria do rastreamento de objetos em tempo real, especialmente de pedestres, lidando com desafios como oclusões, variações de iluminação e mudanças de perspectiva, que frequentemente afetam a precisão do rastreamento.



Figura 12 – Rastreamento online de múltiplos objetos(ZHANG; WANG; GU, 2021)

A metodologia proposta pode envolver o uso de técnicas de re-identificação de pedestres, que são métodos para identificar corretamente o mesmo pedestre em diferentes quadros ou em várias câmeras, mesmo quando ocorrem oclusões ou mudanças de aparência. Além disso, o processamento de oclusões pode incluir estratégias para lidar com situações em que um objeto rastreado está temporariamente obstruído por outros elementos da cena. A abordagem provavelmente utiliza algoritmos e técnicas específicas para rastreamento de objetos, como filtros de Kalman, redes neurais, métodos de correspondência de características, entre outros, a fim de melhorar a precisão e a continuidade do rastreamento dos objetos ao longo do tempo.

Essa pesquisa tem relevância em aplicações de vigilância, segurança e monitoramento, onde o rastreamento preciso de objetos, especialmente de pedestres, é fundamental. O uso de técnicas de re-identificação e processamento de oclusões pode contribuir significativamente para aprimorar a confiabilidade e a precisão do rastreamento de objetos em

ambientes dinâmicos e desafiadores.

(HUO; OU; LI, 2021) aborda um algoritmo de rastreamento de múltiplos alvos baseado em aprendizado profundo (*Deep Learning*). Desenvolvimento de um método de rastreamento de múltiplos alvos que faz uso de técnicas e modelos de aprendizado profundo para melhorar a precisão e eficiência do rastreamento em ambientes desafiadores.

O método proposto provavelmente incorpora redes neurais profundas ou arquiteturas similares de aprendizado profundo, como redes convolucionais (CNNs) ou redes neurais recorrentes (RNNs), para realizar a detecção e rastreamento dos múltiplos alvos ao longo do tempo. A utilização de técnicas de aprendizado profundo pode permitir uma melhor compreensão das características dos alvos, facilitando a detecção e a associação correta das diferentes instâncias dos alvos em diferentes quadros sequenciais ou câmeras.

(XU; BO; WANG, 2021), apresenta uma abordagem inovadora para o rastreamento *multi-target* em múltiplas câmeras, fundamentada no conceito de agrupamento de características. Com rastreamento simultâneo de múltiplos alvos em ambientes com várias câmeras. A abordagem proposta baseia-se no agrupamento de características, um método que se destina a identificar e seguir múltiplos alvos por meio da extração e agrupamento de características específicas, como padrões visuais ou comportamentais, nos dados capturados pelas câmeras.

O método desenvolvido visa superar desafios como oclusões, variações de iluminação e mudanças de perspectiva entre diferentes câmeras. Isso é alcançado pela identificação e agrupamento de características relevantes para os alvos rastreados, permitindo o acompanhamento preciso, mesmo quando os alvos se movem entre diferentes campos de visão das câmeras. A abordagem de agrupamento de características contribui para uma melhor associação de alvos e uma redução nos erros de identificação, melhorando a precisão do rastreamento multi-target em ambientes com múltiplas câmeras.

3.4 Estudos Integrados sobre Detecção e Rastreamento de Pedestres

(XIAO et al., 2022), aborda o desenvolvimento de um sistema de rastreamento de pedestres em tempo real usando uma única câmera (monocular) para capturar informações visuais do ambiente. Método capaz de rastrear pedestres em 3D usando apenas uma câmera monocular, ou seja, uma câmera que captura imagens de uma única perspectiva. Esse tipo de sistema é desafiador devido à falta de informações de profundidade direta fornecidas por uma câmera monocular.

Utilizando técnicas de visão computacional e aprendizado profundo para estimar a posição tridimensional dos pedestres em tempo real. Isso pode envolver a utilização de redes neurais convolucionais (CNNs) ou outras arquiteturas de aprendizado profundo para

extrair informações de profundidade e características visuais dos pedestres a partir de uma única imagem.

Esse tipo de sistema pode ser útil em uma ampla gama de aplicações, incluindo segurança, sistemas de assistência ao motorista, monitoramento de tráfego, entre outros, oferecendo uma solução para rastreamento de pedestres que não depende de múltiplas câmeras ou sensores para fornecer informações de profundidade.

([YANG et al., 2022](#)), discute um método para o rastreamento online de múltiplos objetos (pedestres) com refinamento de previsão e classificação de trilhas. Com foco na implementação de um sistema capaz de rastrear online, em tempo real, múltiplos pedestres em um ambiente dinâmico, onde o movimento e a interação entre os pedestres podem gerar desafios para o rastreamento preciso.

A abordagem proposta provavelmente utiliza técnicas de previsão refinada, onde modelos são empregados para antecipar as trajetórias futuras dos pedestres com base em seus movimentos passados, aprimorando assim a capacidade do sistema de rastreamento em antecipar suas posições.

Além disso, a classificação de trilhas pode envolver o uso de algoritmos ou modelos para avaliar a qualidade das trilhas rastreadas, permitindo diferenciar corretamente entre os pedestres e potenciais falsas detecções, ajudando a melhorar a precisão e confiabilidade do sistema de rastreamento.

([WANG; YANG, 2022](#)), aborda um método de rastreamento de pedestres múltiplos utilizando as arquiteturas *YOLOv5* e *DeepSORT*. Sistema capaz de detectar e rastrear múltiplos pedestres em tempo real, utilizando a arquitetura *YOLOv5* para detecção de pedestres e a técnica *DeepSORT* para realizar o rastreamento.

O *YOLOv5* é uma arquitetura de detecção de objetos que se destaca pela sua rapidez e precisão na detecção em tempo real, enquanto o *DeepSORT* é um algoritmo de rastreamento que utiliza técnicas de aprendizado profundo para atribuir identidades consistentes aos objetos detectados ao longo de diferentes quadros de vídeo.

A combinação dessas duas arquiteturas permite a detecção precisa de pedestres em imagens ou vídeos em tempo real pelo *YOLOv5* e, em seguida, o rastreamento eficiente desses pedestres ao longo do tempo pelo *DeepSORT*, mantendo suas identidades consistentes, mesmo quando estão em movimento ou temporariamente ocultos do campo de visão da câmera. Enfatizando a sinergia entre o *YOLOv5* e o *DeepSORT* para oferecer um sistema de detecção e rastreamento de pedestres múltiplos em tempo real.

([LIMA et al., 2022](#)), descreve um método de detecção de pedestres em 3D projetado para ambientes com múltiplas câmeras. Essa abordagem é notável por sua capacidade de generalização e eficácia na detecção de pedestres em diversas condições de visualização.

O foco principal do estudo é a localização precisa de pedestres em ambientes complexos, onde várias câmeras fornecem diferentes perspectivas da cena. O método proposto visa localizar pedestres no espaço tridimensional, aproveitando as informações coletadas por essas múltiplas câmeras.

A contribuição chave do trabalho está na capacidade do método em ser generalizável, ou seja, em sua habilidade de se adaptar e funcionar consistentemente em variadas condições de visualização. Isso inclui cenários com mudanças de iluminação, oclusões, variações de perspectiva e outros desafios frequentes encontrados em ambientes dinâmicos.

Ao fazer uso de múltiplas câmeras, Figura 13, a abordagem proposta busca criar uma representação tridimensional abrangente do ambiente, identificando e localizando pedestres em diferentes posições e ângulos de visão. A eficácia do método reside na sua capacidade de oferecer resultados precisos e consistentes, independentemente das variações nas condições de visualização.



Figura 13 – Resultados de detecção de pedestres em 3D. (LIMA et al., 2022)

O método de (VORA et al., 2023) aborda o desafio da detecção de pedestres em ambientes urbanos inteligentes, vigilância, direção autônoma e robótica, com ênfase na estimativa da localização tridimensional (3D) dos pedestres. Destaca-se a importância da detecção precisa para o desenvolvimento de cidades inteligentes sustentáveis, sendo comum o uso de múltiplas câmeras para melhorar a precisão da detecção 3D. No entanto, a detecção em ambientes lotados continua sendo um desafio, e métodos baseados em IA têm progredido significativamente nesse campo.

A proposta é investigar como a IA generalizável pode melhorar a detecção 3D de

pedestres em ambientes multi-câmeras sem a necessidade de dados rotulados do cenário-alvo, exibido no exemplo da Figura 14. São discutidas abordagens como a pseudo-rotulagem e a obtenção automática de rótulos para dados de detecção. Contribui ao explorar métodos de IA generalizável para tornar a localização de pedestres robusta a mudanças de domínio, aumentando a aplicabilidade prática da detecção de pedestres em tempo real. Além disso, são apresentadas estratégias para treinamento de detectores de pedestres 3D multi-câmeras usando procedimentos de rotulagem automática, com avaliações quantitativas e qualitativas demonstrando melhorias no desempenho de detecção.

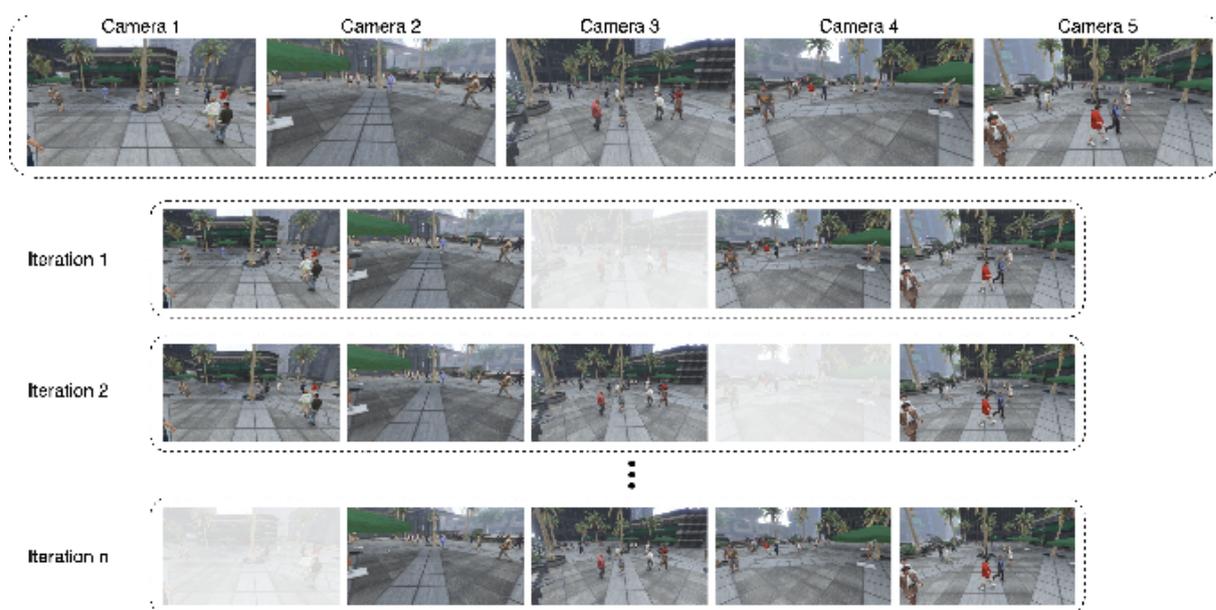


Figura 14 – Regularização *DropView*. (VORA et al., 2023)

(HOU; ZHENG; GOULD, 2020) descreve em sua investigação a detecção de pedestres em ambientes lotados, utilizando múltiplas visualizações de câmera. O estudo aborda a questão da oclusão em tarefas de visão computacional, especialmente em problemas de detecção, e propõe o método MVDet para lidar com essas ambiguidades. O foco principal é na detecção de pedestres a partir de várias câmeras RGB (multivisão), onde os quadros são sincronizados e as câmeras têm campo de visão sobreposto, Figura 15. O MVDet enfrenta desafios devido a oclusões e multidões, propondo duas abordagens principais: agregação de informações de multivisão e agregação espacial.

Para a agregação multivisão, o MVDet utiliza uma abordagem livre de âncoras, combinando mapas de características projetados, ao invés da abordagem baseada em âncoras adotada por métodos anteriores. Já para a agregação espacial, o MVDet emprega uma abordagem totalmente convolucional, minimizando a dependência de técnicas como inferência de campo médio. O MVDet é treinado de maneira end-to-end e demonstrou eficácia em dois conjuntos de dados em grande escala, *Wildtrack* e *MultiviewX*, obtendo resultados competitivos mesmo em níveis variados de oclusões. No *Wildtrack*, alcançou 88,2% de MODA, um aumento de 14,1% em relação ao estado-da-arte anterior. No

MultiviewX, também obteve resultados competitivos sob diferentes níveis de oclusão.

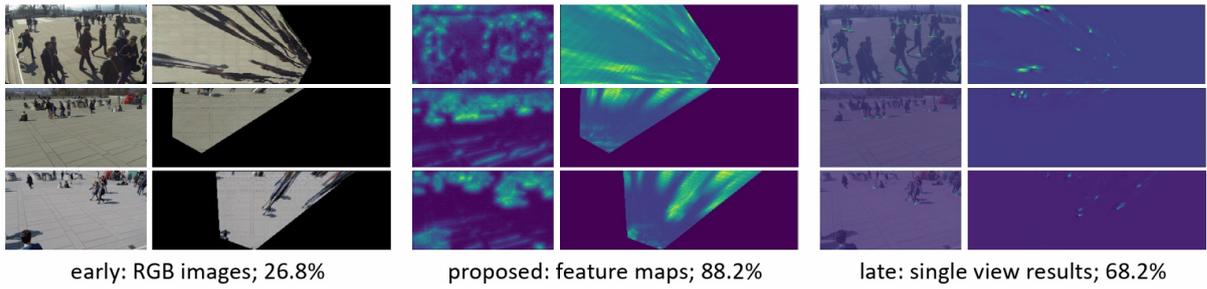


Figura 15 – Projeções de características e seu desempenho. (HOU; ZHENG; GOULD, 2020)

A pesquisa de (LIMA et al., 2023) aborda um método sobre o reconhecimento de pedestres em ambientes tridimensionais, que é essencial para diversas áreas como cidades inteligentes, vigilância e direção autônoma, enfrenta desafios persistentes, especialmente na detecção precisa em 3D. Atuando sobre a necessidade de dados rotulados do cenário-alvo que tem sido uma limitação. Nesse sentido, a IA generalizável surge como uma solução promissora, buscando modelos capazes de lidar com novos cenários e domínios. Investigando abordagens de detecção de pedestres 3D de multi-visão não rotulada, destacando a rotulagem automática como uma técnica que pode melhorar a precisão da detecção.

Dois abordagens de rotulagem automática são discutidas: a pseudo-rotulagem, que utiliza um detector supervisionado para gerar rótulos para os dados-alvo, e a rotulagem automática, que emprega um detector não treinado para criar os rótulos, a rotulagem é possível ser vista na Figura 16. Limitações são identificadas, como a incapacidade de detectar pedestres fora do plano terrestre e desafios na generalização dos dados de treinamento. Os resultados mostram que a rotulagem automática pode melhorar a detecção de pedestres, especialmente quando combinada com ajuste fino do modelo.

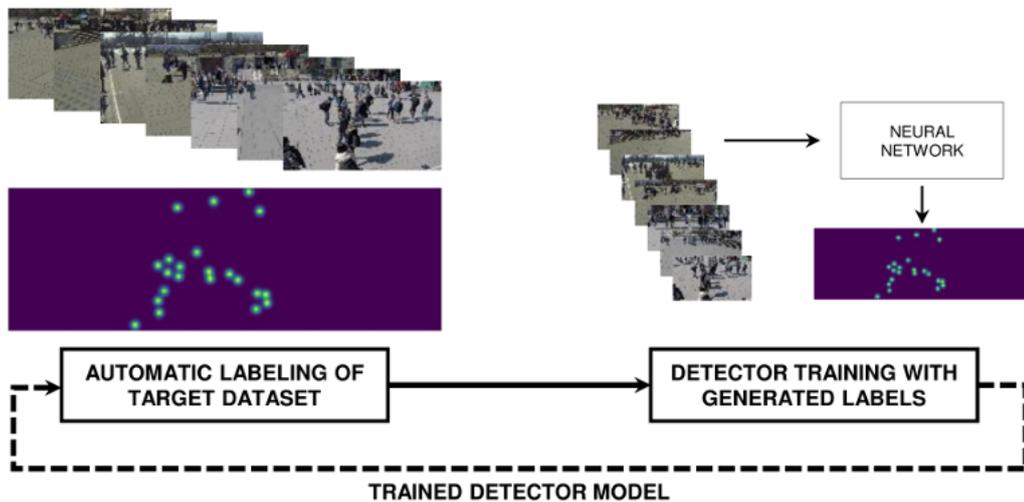


Figura 16 – Abordagem da técnica com multi visualização não rotulada. (LIMA et al., 2023)

(LYRA et al., 2022), foca na implementação de um sistema de rastreamento de pedestres em 3D em tempo real, Figura 17, utilizando informações provenientes de múltiplas câmeras. Visando desenvolver um método capaz de rastrear pedestres em 3D de forma online, ou seja, em tempo real, em um ambiente que envolve várias câmeras. A abordagem proposta visa garantir que o sistema seja capaz de rastrear pedestres independentemente das condições variáveis de iluminação, mudanças ambientais e diferentes perspectivas capturadas pelas múltiplas câmeras.

O método utiliza técnicas de fusão de dados provenientes das múltiplas câmeras para estimar as trajetórias tridimensionais dos pedestres em tempo real. Isso pode envolver o uso de algoritmos de correspondência e técnicas de visão computacional para unificar as informações capturadas por diferentes câmeras e construir uma representação 3D dos pedestres rastreados.

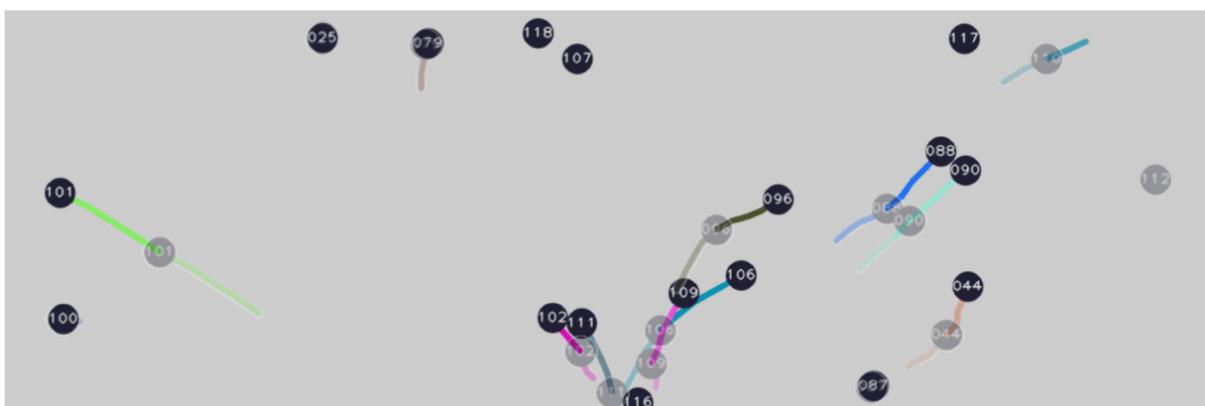


Figura 17 – Rastreamento de pedestres em 3D. (LYRA et al., 2022)

3.5 Detecção de Presença de Pedestres em Áreas de Interesse

Detectar a presença de pedestres em locais pré-determinados não é uma abordagem adotada na maioria das técnicas de detecção de pedestres. Porém, normalmente é possível encontrar informações relacionadas à detecção de pedestres em imagens completas na literatura, identificando sua localização. Além disso, a caixa delimitadora é frequentemente usada para determinar a posição ou localização do pedestre.

Técnicas de identificação de pedestres para proteção ou possível ameaça podem ser encontradas na literatura. Broggi et al. (BROGGI et al., 2009) utilizam um modelo baseado em cenários, no qual é necessário conhecer a localização do pedestre e o destino do veículo na cena para validação entre as informações. No entanto, esta abordagem identificará inicialmente o cenário completo, incluirá o veículo como objeto detectado e, em seguida, detectará o pedestre. Além disso, esse cenário exige muita informação e coleta de dados, o que torna a técnica bastante robusta e cara.

O modelo de detecção de pedestres utilizado por Pustokhina et al. (PUSTOKHINA et al., 2021) segue um procedimento de identificação diferente quanto à presença de pedestres. Ele identifica pedestres com muitas informações para treinamento e utiliza um pré-processo para treinamento de redes neurais, Figura 18.



Figura 18 – Detecção de anomalias por Pustokhina (PUSTOKHINA et al., 2021)

A detecção de pedestres em busca de sua proteção é um dos cenários de grande relevância para a sociedade quando falamos em segurança. Pudemos observar muitos estudos na obra de Gandhi & Trivedi (GANDHI; TRIVEDI, 2007) que apresenta sistemas de proteção de pedestres. O estudo mostra um cenário de proteção vital, mas em muitos métodos o ambiente em que o pedestre está localizado pode sofrer colisões com veículos. Entendemos como é importante proteger os pedestres neste ambiente. Mesmo assim, observamos como é difícil encontrar estudos que citem outros ambientes que possam ter locais de risco para pedestres. A determinação de localizações pode gerar uma nova visão dos dados, o que pode trazer melhorias significativas na prevenção de acidentes com pedestres.

A implementação descrita por Ansarnia et al. (ANSARNIA et al., 2022) mostra um cenário que pode ser comparado usando algoritmos de aprendizagem profunda para identificar pessoas e objetos em situações complementares. É mencionada a combinação de três algoritmos para validação de possíveis colisões entre veículos e pedestres. A segmentação adotada no experimento mostra detecção de objetos, detecção de pedestres e detecção da rodovia onde os objetos e pedestres podem estar, exibido na Figura 19. No entanto, apenas uma câmera fixa é usada e o algoritmo identifica muitas áreas onde as detecções são colocadas.

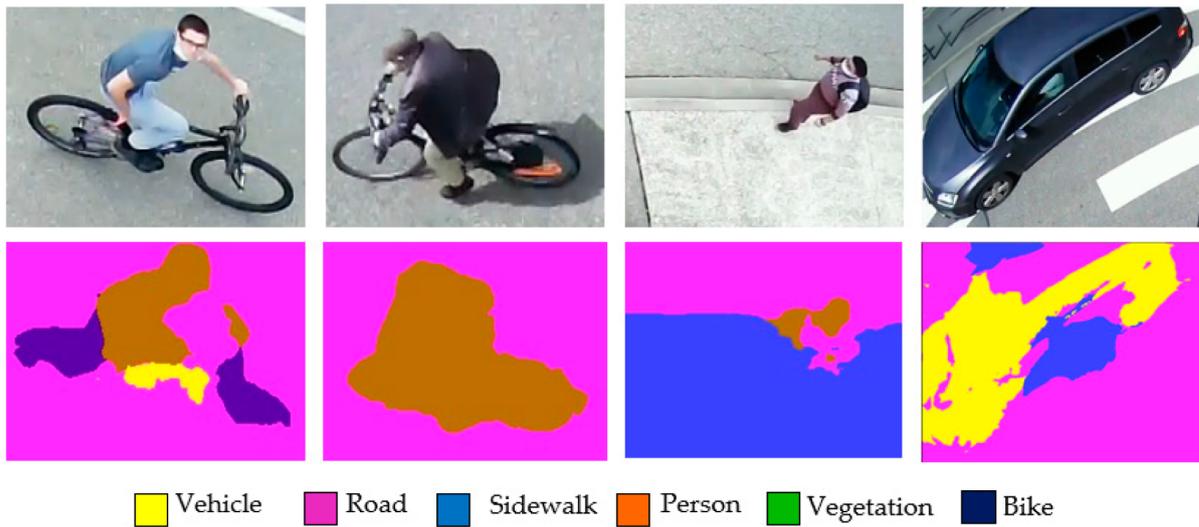


Figura 19 – Combinação implementada por Asarnia et al. (ANSARNIA et al., 2022)

3.6 Análise e Tendências

Os estudos revisados demonstram uma variedade de métodos, desde redes neurais para detecção até técnicas de filtragem e correspondência para rastreamento. A tendência predominante envolve a combinação de métodos de detecção de alta precisão com abordagens de rastreamento robustas para lidar com desafios como oclusão, variações de aparência e ambientes complexos.

Esses estudos recentes indicam um avanço significativo na área de rastreamento de pedestres e objetos em ambientes complexos, buscando soluções mais precisas, rápidas e robustas para lidar com desafios específicos de vigilância e monitoramento.

4 Materiais e Métodos

Neste capítulo serão apresentados os materiais utilizados na aquisição e processamento dos dados, bem como os métodos e métricas de avaliação utilizados nas experiências e estudos dessa dissertação. Serão detalhados os métodos empregados para a detecção e o rastreamento de pedestres em áreas de interesse. O estudo visa comparar e analisar técnicas de detecção e rastreamento de pedestres utilizando múltiplas câmeras, avaliando sua eficácia na identificação da presença de pedestres em locais predefinidos nas imagens.

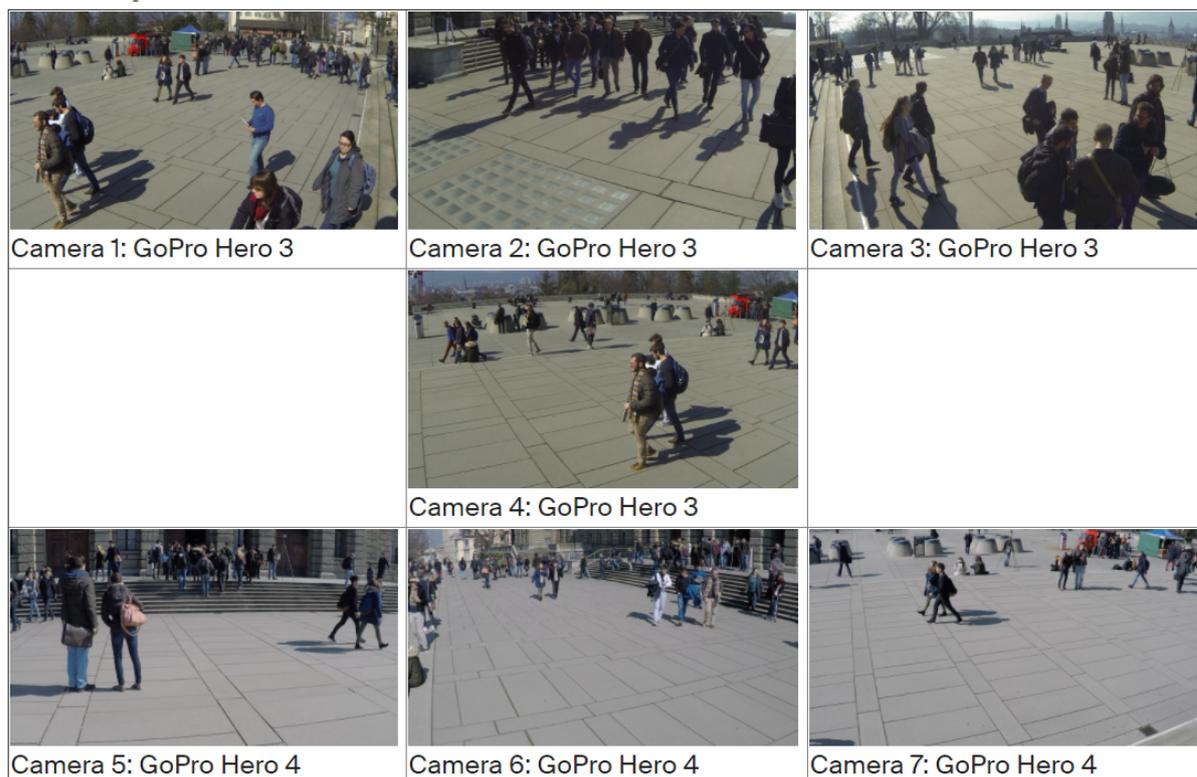
4.1 Datasets Utilizados

Antes de iniciar os métodos de detecção e rastreamento, normalmente ocorrem ações sobre os dados, é a realização do pré-processamento dos dados. Os dados são pré-processados para remoção de ruídos, análise sobre as informações para ter a certeza que os dados estão completos, e é realizada a calibração das câmeras para garantir a consistência entre os diferentes conjuntos de dados das câmeras.

Dataset *Wildtrack*: No cenário em constante evolução da visão computacional, o dataset *Wildtrack* se destaca como uma ferramenta essencial para a pesquisa e desenvolvimento de sistemas de detecção de pedestres.

O *Wildtrack* oferece uma variedade única de características e complexidades que refletem as condições do mundo real. O conjunto de dados consiste em dados provenientes de sete câmeras distintas, cada uma capturando imagens de um ambiente específico, como mostra a Figura 20. Além disso, o conjunto de dados fornece informações detalhadas sobre parâmetros intrínsecos e extrínsecos das câmeras, essenciais para a calibração e compreensão do ambiente de captura. Esse conjunto de dados oferece anotações reais de cada pedestre em 400 quadros, juntamente com parâmetros intrínsecos e extrínsecos das câmeras utilizadas.

Os parâmetros extrínsecos definem a localização e orientação de um ponto no sistema de coordenadas da câmera em relação ao mundo. Já os parâmetros intrínsecos são necessários para relacionar o ponto nas coordenadas do pixel com o sistema de coordenadas da câmera.

Figura 20 – Dataset *Wildtrack*

Dataset Si.U: O Si.U, *Spatial Intelligence for Urban Intelligence*, é um novo conjunto de dados desenvolvido pelo time do Voxar Labs, destinado a apoiar essa abordagem. Ele consiste em múltiplos vídeos capturados por diversas câmeras, permitindo a detecção de pedestres em ambientes urbanos de forma mais precisa e detalhada. Esse dataset foi criado com o objetivo de melhorar a compreensão da utilização dos espaços urbanos, reduzindo o tempo gasto na coleta manual de informações e fornecendo dados mais ricos sobre os padrões de movimento e permanência das pessoas nos locais de interesse.

Este dataset inclui vídeos capturados por quatro câmeras posicionadas em locais estratégicos, locais considerando a visualização em comum entre as câmeras, podendo assim abranger uma grande área dentro do cenário a ser monitorado, neste contexto foi atribuída uma área de interesse de 12m x 22m. Os vídeos foram gravados em diferentes condições de iluminação para simular cenários variados, refletindo as condições reais de utilização dos espaços urbanos, Figura 21. As anotações fornecidas nesses vídeos indicam a localização dos pedestres, seus movimentos e o tempo de permanência em áreas específicas.

Com o Si.U, é possível mapear o fluxo de pessoas e suas permanências em diferentes áreas, oferecendo informações importantes sobre as rotas preferenciais, áreas mais frequentadas e o comportamento dos pedestres. Essas informações são cruciais para identificar pontos de melhoria e oportunidades de intervenção que promovam espaços urbanos mais funcionais e inclusivos.

Momentos Si.U: A iluminação das cenas também foram variadas, uma vez que



Figura 21 – Capturas do Dataset Si.U.

as observações feitas durante a etapa de sensoriamento de uma área são realizadas em diferentes horários do dia, que podem ser à luz da manhã ou à luz fraca da noite.

Os horários de gravação foram:

- LA01 - Início da tarde (14h)
- LA02 - Final da tarde (16h)
- NI01 - Noite (18h)

Na Figura 22 exibimos a diferença entre os momentos utilizados com o dataset Si.U, podemos observar que a iluminação é um ponto crucial para realizar a detecção dos pedestres na cena, na sequência LA01, LA02 e NI01.



Figura 22 – Momentos Dataset Si.U.

4.2 Metodologia de Detecção em Áreas de Interesse

O estudo emprega duas abordagens principais: a detecção de pedestres em áreas de interesse e o rastreamento dos mesmos. Ambos os métodos visam identificar a presença

de pedestres em locais específicos nas imagens capturadas pelas múltiplas câmeras. Os algoritmos descritos a seguir, serão os métodos utilizados integrados ao objetivo da pesquisa que é detectar pedestres em áreas de interesses.

Método de Detecção de Pedestres 3D Multi Câmera: O método de detecção de pedestres de (LIMA et al., 2022) utiliza técnicas avançadas de visão computacional e aprendizado de máquina para identificar regiões de interesse nas imagens capturadas. O algoritmo aplica a identificação de poses, integradas a captura de keypoints e identificação de pessoas. Além de incluir uma técnica de re-identificação de pedestres que são identificados em mais de uma câmera, as detecções são exibidas na Figura 23. A pesquisa aplica o algoritmo de detecção que foi implementado para localizar pedestres nas áreas pré-definidas, calculando coordenadas 3D no plano de solo para cada detecção.



Figura 23 – Detecção de pedestres ao mesmo tempo instantânea e visualização do plano de solo de cada pedestre. (LIMA et al., 2022)

Detecção Multi Visualização Generalizado - GMVD: O método inovador de Detecção Multi-Visão (MVD) que visa melhorar a capacidade de generalização dos modelos. Esse método proposto modifica o objetivo de treinamento e utiliza uma estratégia de *pooling* médio nos mapas de características para permitir que o modelo aceite um número variável de visualizações de câmera como entrada (VORA et al., 2023). Além disso, introduz uma regularização chamada DropView e utiliza uma combinação de métricas de Divergência de *Kullback-Leibler* (KLDiv) e Correlação Cruzada de *Pearson* (CC) para comparar os resultados previstos com os dados reais.

A pesquisa identifica importantes limitações na configuração atual da Detecção Multi-Visão (MVD), que tendem a levar os modelos a se adaptarem excessivamente a configurações de treinamento específicas. Aplicando o projeto em configurações experimentais inovadoras para avaliação dos modelos MVD em cenários mais realistas.

Sistema de Detecção Multi visualização - MVDet: O método de detecção

de pedestres de (HOU; ZHENG; GOULD, 2020) utiliza técnicas avançadas de visão computacional e aprendizado de máquina para identificar regiões de interesse nas imagens capturadas. O algoritmo aplica a identificação de poses, integradas a captura de keypoints e identificação de pessoas. Além de incluir uma técnica de re-identificação de pedestres que são identificados em mais de uma câmera, as detecções são exibidas na Figura 24. A pesquisa aplica o algoritmo de detecção que foi implementado para localizar pedestres nas áreas pré-definidas, calculando coordenadas 3D no plano de solo para cada detecção.



Figura 24 – Detecção de pedestres ao mesmo tempo instantânea e visualização do plano de solo de cada pedestre. (HOU; ZHENG; GOULD, 2020)

Detecção de Pedestres 3D Multi visualização não Rotulada: O método citado possui resultados aplicados ao dataset *Wildtrack* que puderam ser aplicados e comparados as demais técnicas. Foram exploradas estratégias para identificar pedestres em três dimensões em várias perspectivas sem depender de rótulos pré-existentes, utilizando algoritmos de inteligência artificial capazes de generalizar (LIMA et al., 2023).

A técnica de pseudo-rotulagem mostrou-se eficaz ao aprimorar os resultados de um detector supervisionado, desde que este conseguisse se adaptar satisfatoriamente aos dados sem rótulos. Além disso, a abordagem de rotulagem automática, baseada em um detector sem treinamento prévio e ajuste fino, apresentou resultados superiores em todas as configurações de teste de origem para destino em comparação com o uso direto de um detector sem treinamento ou treinado com dados previamente rotulados.

Algoritmo de Rastreamento de Pedestres: O método de rastreamento visa complementar as informações obtidas pela detecção, validando a continuidade do movimento dos pedestres ao longo do tempo. O algoritmo do (LYRA et al., 2022) utiliza informações previamente detectadas para rastrear o movimento dos pedestres em várias imagens sequenciais. Executamos o algoritmo para realizar o rastreamento que também identifica a passagem dos pedestres nas áreas de interesse, validando a continuidade do movimento dos mesmos através de um limiar.

Visualização de Caminhos de Pedestres: A visualização dos caminhos pedestres é um aspecto crucial para compreender os padrões de movimento das pessoas em espaços

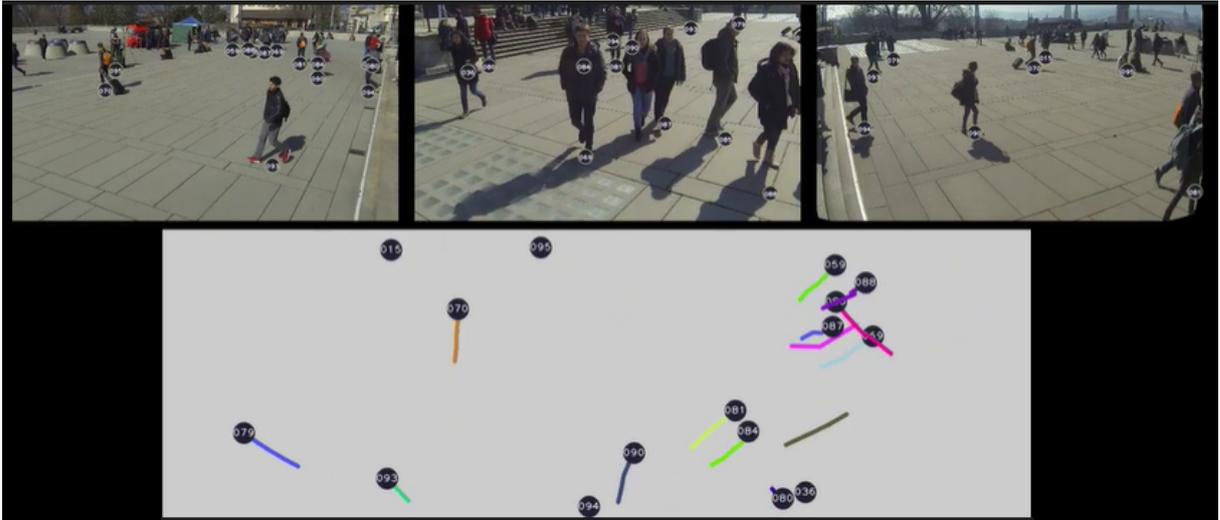


Figura 25 – Resultados de rastreamento de pedestres vistos em múltiplas câmeras ao mesmo tempo instantaneamente t e sua trilha no plano terrestre. (LYRA et al., 2022)

urbanos. Métodos como a visualização baseada em pontos, heatmaps, e visualizações baseadas em linhas são utilizados para representar trajetórias e a ocupação de espaços por pedestres.

Comparação e Fusão de Dados de Detecção e Rastreamento: Os resultados obtidos pela detecção e pelo rastreamento são comparados e fundidos para avaliar a consistência e a precisão das detecções em áreas de interesse. Essa etapa visa identificar discrepâncias entre os métodos e melhorar a qualidade geral do rastreamento.

4.3 Definição e Verificação das Áreas de Interesse

A definição das áreas de interesse é feita através da especificação de coordenadas que delimitam estas áreas nas imagens. A determinação das áreas de interesse é realizada fornecendo dois vértices que definem os pontos diagonais dessas áreas, como pode ser visto na Figura 26.

Esses dados são alinhados para formar um retângulo entre os vértices, fazendo a validação entre a menor distância entre os pontos do vértice, a ação é determinada pela equação 4.1, isso significa que a região de interesse é o conjunto de todos os pontos (x, y) que têm coordenadas x dentro do intervalo $[x1, x2]$ e coordenadas y dentro do intervalo $[y1, y2]$. O algoritmo de verificação é então aplicado para verificar a presença de pedestres dentro dessas áreas, gerando alertas sempre que um pedestre é detectado.

$$R = (x, y) | x1 \leq x \leq x2, y1 \leq y \leq y2 \quad (4.1)$$

Executamos o algoritmo de detecção ou a leitura dos dados salvos contendo as

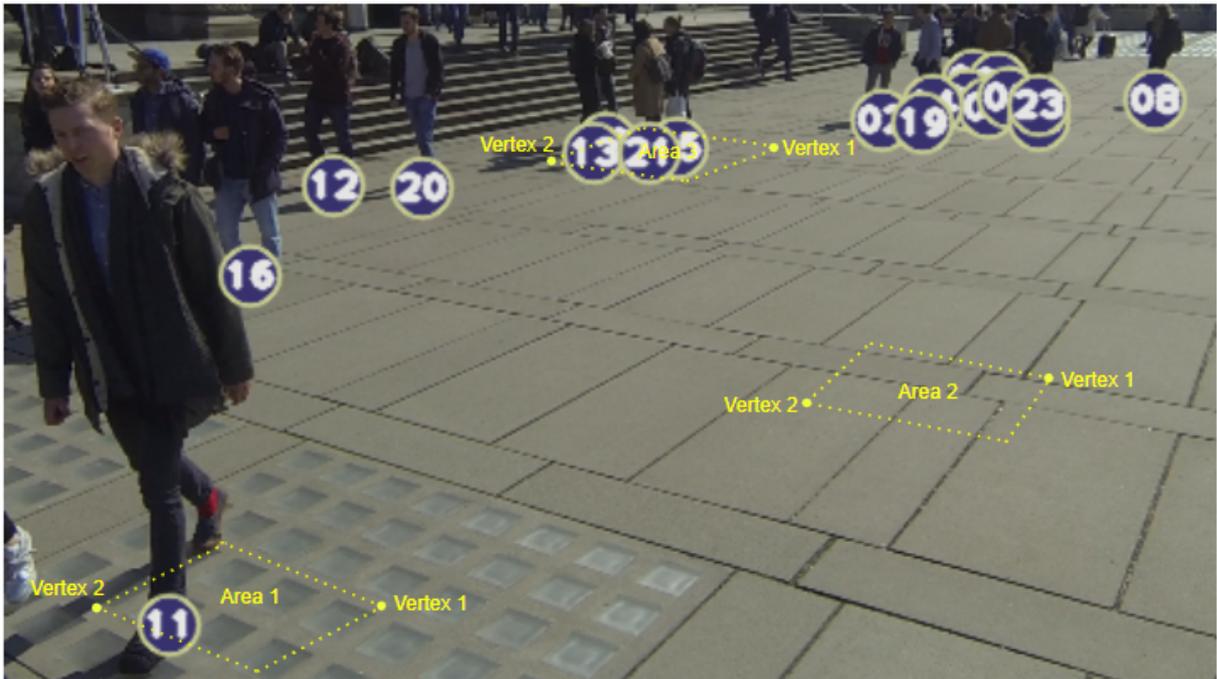


Figura 26 – Áreas de interesse (SILVA; LIMA; TEICHRIB, 2023)

detecções de pedestres para o método de detecção. Os resultados obtidos possuem a nomenclatura (t, x, y) , que são o instante de tempo t , a posição x e a posição y .

Para o rastreamento, os dados de um objeto são obtidos com as seguintes informações: o instante de tempo t , a identificação do pedestre e sua respectiva posição no plano terra (x, y) , com posição $z = 0$. Porém, para o rastreamento é criada uma estrutura de dados que inclui todos os dados de posicionamento de uma mesma pessoa, pois o rastreamento possui o ID para identificar o mesmo pedestre em vários cenários.

Para ambos os métodos, lemos as informações básicas para realizar a avaliação ao final do processo, identificando assim os sucessos de cada método.

Algoritmo de Detecção de Presença de Pedestres em Áreas de Interesse: Criamos um algoritmo 1 para verificar a passagem na área de interesse. A função verifica os valores (x, y) do pedestre, comparando-os com os vértices da área. Caso o pedestre esteja na área, um alerta sobre o quadro atual é adicionado no tempo t .

Para que o algoritmo de verificação valide os resultados entre detecção e rastreamento, modificamos a linha 4 do algoritmo, que terá o mesmo fluxo para as demais validações, então alteramos *deteccao* em *detecoes* para *rastreado* em *rastreados*, onde *rastreados* é o rastreamento de pedestres.

4.4 Avaliação e Métricas de Desempenho

A avaliação consistiu em comparar as previsões dos métodos com a informação real de cada exemplo. Em seguida, para determinar a presença de pedestres nas áreas de

Algoritmo 1 Validação de Pedestres em Áreas de Interesse

```

1: Após executar o método de Detecção
2: Para area in areas Faça
3:   Para deteccao in deteccoes Faça
4:     Se deteccao está na área Então
5:       Salva informação do pedestre
6:       Salva informação do frame (momento t)
7:     Fim Se
8:   Fim Para
9:   Validação informação real do dataset
10:  Para anotacao in anotacoes Faça
11:    Se anotacao está na área Então
12:      Salva informação do pedestre
13:      Salva informação do frame (momento t)
14:    Fim Se
15:  Fim Para
16: Fim Para
17: Validação da informação real comparada com a informação de detecções dos métodos
18: Criação da Matriz de Confusão
19: Avaliação das Métricas

```

interesse, observe que caso seja detectado apenas um pedestre no local pré-determinado, a informação da posição do pedestre é salva e um alerta é emitido. O alerta é a identificação para aplicação que está executando o algoritmo. Avaliamos o método de rastreamento da mesma forma que fizemos para o método de detecção usando as anotações precisas do conjunto de dados.

Métricas Utilizadas na Avaliação: As métricas de desempenho incluídas são acurácia, precisão, recall e F-Score, para avaliar a consistência e a coesão entre os métodos de detecção e rastreamento.

- Acurácia

$$Acc = \frac{VP + VN}{VP + VN + FP + FN} \quad (4.2)$$

- Precisão

$$Prec = \frac{VP}{VP + FP} \quad (4.3)$$

- Recall

$$Recall = \frac{VP}{VP + FN} \quad (4.4)$$

- F-Score

$$F_{Score} = 2 \cdot \frac{Acuracia \cdot Recall}{Acuracia + Recall} \quad (4.5)$$

Onde VP é verdadeiro positivo, VN é verdadeiro negativo, FP é falso positivo e FN é falso negativo.

5 Resultados

Neste capítulo serão apresentados os resultados obtidos utilizando os métodos de Detecção e Rastreamento aplicados aos datasets *Wildtrack* e Si.U apresentados no Capítulo 4.

Os resultados obtidos são analisados detalhadamente, destacando o desempenho de cada método em diferentes cenários e identificando os pontos fortes e fracos de cada abordagem. Utilizando dois datasets *Wildtrack* e Si.U, a análise inclui a comparação entre os métodos para identificar como melhorar a precisão e a eficácia da detecção e rastreamento de pedestres, além de confirmar o objetivo principal que é realizar a detecção em áreas determinadas pelo usuá.

Avaliamos os métodos mencionados acima em um cenário 3D multi-câmera de detecção e rastreamento de pedestres para observar o comportamento de cada técnica na detecção de pedestres em locais pré-definidos para utilizar aquela que apresenta os melhores comportamentos, além de ser possível identificar erros nos métodos de detecção e rastreamento. Apresentamos os resultados obtidos nas subseções a seguir e os detalhes dos experimentos.

5.1 Dataset *Wildtrack*

Os pontos possuem coordenadas no plano terra, que na vista 3D possuem o valor $z = 0$.

Tabela 1 – Grupos de áreas de interesse 1 - Wildtrack

Testes	Teste 1		Teste 2	
	Vértice 1	Vértice 2	Vértice 1	Vértice 2
Área 1	(7.0, 5.0)	(4.0, 1.0)	(-5.0, -3.0)	(0.0, -5.0)
Área 2	(2.0, 10.0)	(7.0, 7.0)	(-3.0, 18.0)	(0.0, 15.0)
Área 3	(1.0, 16.0)	(4.0, 12.0)	(10.0, 16.0)	(6.0, 12.0)

Tabela 2 – Grupos de áreas de interesse 2 - Wildtrack

Testes	Teste 3		Teste 4	
	Vértice 1	Vértice 2	Vértice 1	Vértice 2
Área 1	(-3.0, -3.0)	(-5.0, -5.0)	(-3.0, -4.0)	(0.0, -5.0)
Área 2	(-1.0, 21.0)	(0.0, 18.0)	(3.0, -5.0)	(0.0, -8.0)
Área 3	(5.0, 16.0)	(8.0, 12.0)	(2.0, 10.0)	(6.0, 6.0)

Avaliação das Métricas: De acordo com os resultados obtidos nas detecções, avaliamos as métricas citadas acima. Os resultados obtidos para cada método estão descritos na Tabela 3. Na tabela também apresentamos outros resultados além das métricas de avaliação. Os resultados também são compostos pelo número de pedestres dentro das áreas de interesse detectados por cada método, o número real de pedestres dentro das áreas de interesse, ou seja, com as informações reais de cada pedestre, e a diferença entre eles, para identificar o número de erros.

Após a conclusão do processo, comparamos os resultados do método de detecção e do método de rastreamento. Na comparação direta, identificamos que os valores obtidos por detecção são melhores. Também realizamos uma comparação entre os frames que foram selecionados em um método e não em outro.

Tabela 3 – Métricas de avaliação dos métodos de Detecção e Rastreamento de pedestres dentro das áreas de interesse para dataset *Wildtrack*

Teste	Método	Acurácia	Precisão	Recall	F-Score	Detecção	Info. Real	Dif. Total
1	Detecção	0.973	0.986	0.984	0.985	1144	1150	6
	Rastreamento	0.968	0.992	0.973	0.982	1104		46
2	Detecção	0.905	0.942	0.936	0.939	606	769	163
	Rastreamento	0.892	0.956	0.904	0.930	526		243
3	Detecção	0.953	0.987	0.954	0.970	720	893	173
	Rastreamento	0.955	0.994	0.951	0.972	632		261
4	Detecção	0.978	0.994	0.979	0.986	811	830	19
	Rastreamento	0.975	0.994	0.976	0.985	799		31
Média	Detecção	0.952	0.977	0.963	0.970	820	910	90
	Rastreamento	0.948	0.984	0.951	0.967	765		145

Comparamos os resultados para identificar porque a Detecção apresentou melhores resultados que o Rastreamento. Por fim, finalizadas as execuções, utilizamos os resultados para comparar quais instantes de tempo t (frames) eram incompatíveis entre os métodos.

Como a técnica de Rastreamento utiliza a Detecção como base para detecção de pedestres, obtivemos um resultado de 100% de compatibilidade para três dos quatro grupos de interesse entre Rastreamento e Detecção. Em todos os frames, o método Rastreamento gerou um alarme, e o método Detecção também gerou um alarme.

Esses casos ocorreram nas execuções 1, 2 e 4. Na execução 3, tivemos dois frames em que o Rastreamento gerou alarme, mas a Detecção não, e cinco frames em que o Rastreamento não gerou alarme, mas a Detecção sim. No entanto, dois dos cinco quadros identificados pela Detecção eram falsos positivos, já que a saída do Rastreamento estava correta para esses cinco quadros.

Apresentamos esta comparação na tabelas 4 e 5, que descreve o número total de frames que tiveram pedestres detectados nas áreas de interesse (coluna “Frames

Detectados”), a diferença entre Detecção e Rastreamento, que são os diferentes frames entre cada uma das técnicas (coluna “Dif. Métodos”). Depois, na coluna “Frames Reais”, temos o número real de quadros que possuíam travessias de pedestres nas áreas de interesse.

Por fim, na última coluna, temos uma comparação dos casos onde as saídas de ambos os métodos foram diferentes (coluna “Dif. Erro”) para verificar quais estão corretas. Isso visa verificar se a Detecção foi de fato melhor que o Rastreamento, pois poderíamos ter mais detecções em um método mas todos Falsos Positivos, assim analisamos a quantidade de frames extras detectados pela Detecção e quantos deles estavam corretos. As informações são baseadas no total de frames disponíveis, que são 400 frames para o conjunto de dados *Wildtrack*.

Tabela 4 – Métricas de avaliação de frames com detecção de pedestres entre métodos de Detecção e Rastreamento

Teste	Método	Acurácia	Precisão	Recall	F-Score
1	Detecção	0.993	0.994	0.997	0.995
	Rastreamento	0.983	1.000	0.981	0.982
2	Detecção	0.995	1.000	0.994	0.994
	Rastreamento	0.960	1.000	0.949	0.954
3	Detecção	0.968	0.994	0.966	0.967
	Rastreamento	0.965	1.000	0.957	0.961
4	Detecção	0.985	1.000	0.982	0.983
	Rastreamento	0.985	1.000	0.982	0.983
Média	Detecção	0.985	0.997	0.985	0.985
	Rastreamento	0.973	1.000	0.967	0.970

Tabela 5 – Comparação de frames com detecção de pedestres entre métodos de Detecção e Rastreamento

Teste	Método	Frames Detectados	Dif. Método	Frames Reais	Dif. Erro
1	Detecção	363	6	364	2
	Rastreamento	357	0		0
2	Detecção	311	14	313	0
	Rastreamento	297	0		0
3	Detecção	315	5	326	2
	Rastreamento	312	2		0
4	Detecção	324	0	330	0
	Rastreamento	324	0		0
Média	Detecção	328	6	333	1
	Rastreamento	323	1		0

O método Rastreamento, obteve resultados melhores que o método de Detecção na métrica de precisão, a métrica foi melhor em todos os cenários de execução. Porém, nas demais métricas, teve resultados piores. Faz sentido que o Rastreamento tenha sido pior que

a Detecção porque o Rastreamento às vezes descarta as detecções corretas, causando um aumento nos falsos negativos, por outro lado, o Rastreamento também descarta detecções erradas, causando uma diminuição de falsos positivos.

Avaliação Comparativa dos Métodos de Detecção: Realizamos a comparação dos métodos de detecção, conforme descritos anteriormente, os métodos são (LIMA et al., 2022), (LIMA et al., 2023), (VORA et al., 2023), (HOU; ZHENG; GOULD, 2020), utilizamos os últimos 40 frames do dataset *Wildtrack* para realizar este comparativo, tendo em vista estudos aplicados para a quantidade de informações em cada frame, além da utilização de resultados comparativos entre outros métodos para o mesmo dataset. Adicionamos o nosso algoritmo de validação de passagem de pedestres em áreas de interesses com os resultados de detecção obtido por cada método. Os resultados obtidos pelos métodos de detecção são exibidos na tabela 6.

Tabela 6 – Métricas de avaliação dos métodos de Detecção de pedestres dentro das áreas de interesse para os últimos 40 frames do dataset *Wildtrack*

Teste	Acurácia	Precisão	Recall	F-Score	Frames Detectados	Frames Reais	Diferença Erro
1	1.000	1.000	1.000	1.000	40	40	0
2	1.000	1.000	1.000	1.000	40	40	0
3	1.000	1.000	1.000	1.000	40	40	0
4	1.000	1.000	1.000	1.000	40	40	0
Média	1.000	1.000	1.000	1.000	40	40	0

Os resultados obtidos para a inclusão do algoritmo de passagem de pedestres foram bons. Tendo em vista que o objetivo da técnica é verificar o frame que tiveram passagem de pedestres em áreas de interesse, nossa identificação do frame foi realizada com sucesso para todos métodos de detecção. Conseguimos confirmar assim que a nossa técnica poderá ser integrada a qualquer método de detecção de pedestres. Os métodos testados possuem semelhanças sobre as suas avaliações das métricas MODA, MODP, Precisão e Recall. Podemos garantir assim que técnicas com bons resultados poderão ser utilizados como detector junto a técnica proposta.

5.2 Dataset Si.U

Definição das Áreas de Interesse: Do mesmo modo que realizamos com o dataset *Wildtrack*, configuramos quatro grupos de áreas de interesse contendo três áreas em cada grupo, como pode ser visto nas tabelas 7 e 8 que contêm os vértices (x,y) de cada área em metros.

Tabela 7 – Grupos de áreas de interesse 1 - Si.U

Testes Áreas	Teste 1		Teste 2	
	Vértice 1	Vértice 2	Vértice 1	Vértice 2
Área 1	(5.0, 1.0)	(8.0, 9.0)	(4.0, 2.0)	(7.0, 8.0)
Área 2	(3.0, 2.0)	(1.0, 7.0)	(1.0, 3.0)	(2.0, 6.0)
Área 3	(9.0, 3.0)	(8.0, 3.0)	(10.0, 4.0)	(5.0, 1.0)

Tabela 8 – Grupos de áreas de interesse 2 - Si.U

Testes Áreas	Teste 3		Teste 4	
	Vértice 1	Vértice 2	Vértice 1	Vértice 2
Área 1	(6.0, 1.0)	(9.0, 5.0)	(2.0, 1.0)	(4.0, 3.0)
Área 2	(5.0, 1.0)	(1.0, 9.0)	(5.0, 1.0)	(6.0, 6.0)
Área 3	(12.0, 3.0)	(1.0, 6.0)	(3.0, 2.0)	(7.0, 4.0)

Avaliação das Métricas: Avaliamos as métricas citadas anteriormente. Os resultados obtidos para cada momento estão descritos nas seguintes tabelas, Tabela 9, Tabela 10 e Tabela 11. Na tabela também apresentamos outros resultados além das métricas de avaliação, da mesma forma que foi realizado com o dataset *Wildtrack*. Demonstramos as últimas três colunas com os dados relacionados aos frames que tiveram alertas, e comparamos os dados obtidos com as informações reais e a identificação de dados errados.

Como citado na avaliação realizada no dataset *Wildtrack*, os valores obtidos pelo método Rastreamento só foram melhores na métrica de Precisão. Diante desta análise, avaliamos o dataset Si.U apenas com o método de Detecção.

Tabela 9 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - LA01

Teste	Acurácia	Precisão	Recall	F-Score	Frames Detectados	Frames Reais	Diferença Erro
1	0.585	0.750	0.103	0.182	4	29	25
2	0.785	0.667	0.133	0.222	3	15	12
3	0.369	1.000	0.317	0.481	19	60	41
4	0.969	0.909	0.909	0.909	11	11	0
Média	0.677	0.832	0.366	0.449	9	29	20

Os resultados da avaliação da detecção de pedestres realizada pela tecnologia para o momento LA01 foram:

- MODA: 45.3 / MODP: 75.5 / PREC: 98.5 / RCLL: 46.0

Tabela 10 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - LA02

Teste	Acurácia	Precisão	Recall	F-Score	Detecção	Informação Real	Diferença Total
1	0.714	1.000	0.714	0.833	30	42	12
2	0.524	0.636	0.304	0.412	11	23	12
3	0.881	1.000	0.881	0.937	37	42	5
4	0.762	0.793	0.852	0.821	29	27	-2
Média	0.720	0.857	0.688	0.751	27	34	7

Os resultados da avaliação da detecção de pedestres realizada pela tecnologia para o momento LA02 foram:

- MODA: 40.1 / MODP: 74.5 / PREC: 77.5 / RCLL: 56.4

Tabela 11 – Métricas de avaliação do método de Detecção de pedestres dentro das áreas de interesse para dataset Si.U - NI01

Teste	Acurácia	Precisão	Recall	F-Score	Detecção	Informação Real	Diferença Total
1	0.367	1.000	0.136	0.240	6	44	38
2	0.367	1.000	0.050	0.095	2	40	38
3	0.467	1.000	0.289	0.448	13	45	32
4	0.850	0.500	0.333	0.400	6	9	3
Média	0.513	0.875	0.202	0.296	7	35	28

Os resultados da avaliação da detecção de pedestres realizada pela tecnologia para o momento NI01 foram:

- MODA: 55.8 / MODP: 71.5 / PREC: 97.7 / RCLL: 57.1

Os testes realizados para cada momento conseguem ser satisfatórios quando tratamos da inclusão de outro dataset diferente do *Wildtrack*, conseguimos fazer a inclusão com resultados que conseguem ser visto e comparados entre os datasets com dados semelhantes. Conseguimos observar que os resultados estão parecidos quando trata-se de precisão, pois validamos as passagens dos pedestres de forma correta.

Nossos resultados são satisfatórios quando validamos as informações com relação ao frame (instante t), quando o pedestre foi detectado na área de interesse, atingindo assim o objetivo de realizar a detecção de pedestres em áreas determinadas pelo usuário. Observando os dados dos frames com alertas, tivemos casos em que as informações estavam todas corretas, e ficamos com casos faltantes, podendo ser aplicadas a validação da detecção, já que não temos uma precisão totalmente correta.

6 Conclusões

Apresentamos nesta dissertação um estudo sobre detecção de presença de pedestres em locais pré-determinados na imagem utilizando técnicas de Detecção e Rastreamento multi câmeras, realizando uma comparação entre as técnicas.

Conforme observado, tivemos um resultado melhor na métrica de precisão através do método de Rastreamento. Contudo, nas demais métricas os resultados não foram bons, mostrando assim que o método Detecção obtém uma operação quadro a quadro melhor que o Rastreamento, portanto para o problema de travessia de pedestres, ou ter pelo menos um pedestre em uma área de interesse, utilizar Detecção é mais adequado e assertivo.

Em cada teste é possível identificar uma grande quantidade de pedestres detectados, contudo a informação dos frames reais comparados ao frames detectados mostram para nós que os instantes que foram identificados pedestres poderão ajudar a manter o ambiente mais seguro para o pedestre.

Realizamos a comparação entre detectores de pedestres, realizar a detecção do pedestre integrada a validação de passagem do pedestre nas áreas de interesse, necessita de uma eficiência computacional, sendo necessário utilizar uma máquina com uma configuração mais robusta e específica, de acordo com a especificação do artigo ([LIMA et al., 2022](#)).

Os métodos de detecção destacados foram comparados entre si, podendo obter uma validação sobre o algoritmo criado. Desta forma os métodos de detecções existentes no estado da arte, poderão ser aplicados de forma integrada a validação de passagem de pedestres nas áreas de interesses do usuário, conseguimos abordar uma aplicação que poderá resultar ações para a segurança dos pedestres.

Do mesmo modo, a utilização da tecnologia de Rastreamento do artigo ([LYRA et al., 2022](#)) utiliza dados da detecção, sendo necessário realizar modificações e adaptações para utilização de um novo dataset.

A inclusão do dataset Si.U, foi um grande desafio, devido o conhecimento sobre o dataset, a quantidade de informações detalhadas de cada pedestres. O dataset do Si.U contempla a presença de objetos também na área que está sendo aplicada a área de detecção de pedestres e passar pelos objetos. Para os testes, utilizamos apenas uma amostra do dataset completo, pois a avaliação do time Voxar Labs sobre a completude do dataset é válido para amostra em relação ao dataset.

6.1 Contribuições

Avaliamos os métodos citados acima em uma multi câmera 3D para utilizar aquele que apresenta melhores comportamentos, além de ser possível identificar erros nos métodos de detecção e rastreamento.

Nossas contribuições são:

- Criação de algoritmo de validação de passagem de pedestres em áreas de interesses determinados pelo usuário.
- Demonstração de um algoritmo para validação de travessias de pedestres em locais pré-determinados, sendo possível validar a passagem do pedestre em tempo de detecção de pedestres ou com os dados de detecções já realizadas.
- Inclusão em métodos de detecção e rastreamento de pedestres com múltiplas câmeras. Esta análise pode ser complementar para identificar erros no rastreamento de um pedestre e também melhorar o rastreamento geral de pedestres.
- Comparação de resultados e avaliações quantitativas e qualitativas de ambas as técnicas em algumas áreas de interesse.
- Realizamos a inclusão de um novo dataset, podendo assim avaliar a possibilidade de integração da tecnologia com outros datasets. Identificamos lacunas que precisam ser avaliadas com relação ao novo dataset, devido às métricas com valor baixo.
- Publicação do artigo *Pedestrian Presence Detection in Areas of Interest Using Multiple Cameras* na *HCI INTERNATIONAL 2023. 25TH International Conference on Human-Computer Interaction*.

6.2 Trabalhos futuros

Para trabalhos futuros, podemos listar algumas implementações que poderão contribuir com a evolução da tecnologia atual.

- Realizar a exibição das áreas de riscos e alertas em tempo de detecção de pedestres.
- Avaliar os métodos em ambientes reais identificando pessoas em locais de risco, com o objetivo de melhorar a segurança ambiental.
- Integrar a tecnologia atual de detecção de pedestres com tecnologias de detecção de objetos.
- Aplicar tecnologia em parceiros da indústria, podendo ser analisados problemas do mundo real e visibilidade dos ambientes.

Referências

- ADARSH, P.; RATHI, P.; KUMAR, M. Yolo v3-tiny: Object detection and recognition using one stage improved model. In: IEEE. *2020 6th international conference on advanced computing and communication systems (ICACCS)*. [S.l.], 2020. p. 687–694. Citado 2 vezes nas páginas 5 e 8.
- AKYILDIZ, I. F. et al. Wireless sensor networks: a survey. *Computer networks*, Elsevier, v. 38, n. 4, p. 393–422, 2002. Citado na página 10.
- ANSARNIA, M. S. et al. Contextual detection of pedestrians and vehicles in orthophotography by fusion of deep learning algorithms. *Sensors*, MDPI, v. 22, n. 4, p. 1381, 2022. Citado 3 vezes nas páginas 8, 29 e 30.
- BERG, R. v. d.; KIPF, T. N.; WELLING, M. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263*, 2017. Citado na página 14.
- BRASÓ, G.; LEAL-TAIXÉ, L. Learning a neural solver for multiple object tracking. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 6247–6257. Citado na página 17.
- BROGGI, A. et al. A new approach to urban pedestrian detection for automatic braking. *IEEE Transactions on intelligent transportation systems*, IEEE, v. 10, n. 4, p. 594–605, 2009. Citado na página 28.
- BUTT, A. A.; COLLINS, R. T. Multi-target tracking by lagrangian relaxation to min-cost network flow. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2013. p. 1846–1853. Citado na página 8.
- Cao, Z. et al. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019. Citado na página 6.
- CHAVDAROVA, T. et al. Wildtrack: A multi-camera hd dataset for dense unscripted pedestrian detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 5030–5039. Citado na página 12.
- CHEN, X. et al. Monocular 3d object detection for autonomous driving. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 2147–2156. Citado na página 9.
- CHENG, G.; HAN, J. A survey on object detection in optical remote sensing images. *ISPRS journal of photogrammetry and remote sensing*, Elsevier, v. 117, p. 11–28, 2016. Citado na página 14.
- COMANICIU, D.; RAMESH, V.; MEER, P. Kernel-based object tracking. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 25, n. 5, p. 564–577, 2003. Citado na página 8.

- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: IEEE. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. [S.l.], 2005. v. 1, p. 886–893. Citado 2 vezes nas páginas 6 e 13.
- DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. Citado na página 10.
- DOLLAR, P. et al. Pedestrian detection: An evaluation of the state of the art. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 34, n. 4, p. 743–761, 2011. Citado na página 7.
- DONAHUE, J. et al. Long-term recurrent convolutional networks for visual recognition and description. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 2625–2634. Citado na página 14.
- FANG, H.-S. et al. Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, 2022. Citado 2 vezes nas páginas 8 e 6.
- FERRYMAN, J. M.; SHAHROKNI, A. Pets2009: Dataset and challenge. *2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, p. 1–6, 2009. Disponível em: <<https://api.semanticscholar.org/CorpusID:34646357>>. Citado na página 12.
- GANDHI, T.; TRIVEDI, M. M. Pedestrian protection systems: Issues, survey, and challenges. *IEEE Transactions on intelligent Transportation systems*, IEEE, v. 8, n. 3, p. 413–430, 2007. Citado na página 29.
- GEIGER, A. et al. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, Sage Publications Sage UK: London, England, v. 32, n. 11, p. 1231–1237, 2013. Citado na página 12.
- GEIGER, A.; LENZ, P.; URTASUN, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In: IEEE. *2012 IEEE conference on computer vision and pattern recognition*. [S.l.], 2012. p. 3354–3361. Citado 3 vezes nas páginas 8, 12 e 13.
- GIRSHICK, R. Fast r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2015. p. 1440–1448. Citado na página 4.
- HALL, D. et al. *Distributed data fusion for network-centric operations*. [S.l.]: Crc Press, 2017. Citado na página 9.
- HE, K. et al. Mask r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2017. p. 2961–2969. Citado na página 5.
- HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778. Citado na página 14.
- HE, Y. et al. Multi-target multi-camera tracking by tracklet-to-target assignment. *IEEE Transactions on Image Processing*, IEEE, v. 29, p. 5191–5205, 2020. Citado 3 vezes nas páginas 8, 18 e 20.

- HOU, Y.; ZHENG, L.; GOULD, S. Multiview detection with feature perspective transformation. In: SPRINGER. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16*. [S.l.], 2020. p. 1–18. Citado 6 vezes nas páginas 8, 1, 26, 27, 35 e 42.
- HUO, W.; OU, J.; LI, T. Multi-target tracking algorithm based on deep learning. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2021. v. 1948, n. 1, p. 012011. Citado na página 23.
- JABBAR, R. et al. Urban traffic monitoring and modeling system: An iot solution for enhancing road safety. In: IEEE. *2019 international conference on internet of things, embedded systems and communications (iintec)*. [S.l.], 2019. p. 13–18. Citado na página 1.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. 1960. Citado na página 9.
- KOCH, G. et al. Siamese neural networks for one-shot image recognition. In: LILLE. *ICML deep learning workshop*. [S.l.], 2015. v. 2, n. 1. Citado na página 10.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, AcM New York, NY, USA, v. 60, n. 6, p. 84–90, 2017. Citado na página 4.
- KURAKIN, A.; GOODFELLOW, I. J.; BENGIO, S. Adversarial examples in the physical world. In: *Artificial intelligence safety and security*. [S.l.]: Chapman and Hall/CRC, 2018. p. 99–112. Citado na página 14.
- LASASSMEH, S. M.; CONRAD, J. M. Time synchronization in wireless sensor networks: A survey. In: IEEE. *Proceedings of the IEEE SoutheastCon 2010 (SoutheastCon)*. [S.l.], 2010. p. 242–245. Citado na página 10.
- LEAL-TAIXÉ, L. et al. Motchallenge 2015: Towards a benchmark for multi-target tracking. *arXiv preprint arXiv:1504.01942*, 2015. Citado na página 12.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado 2 vezes nas páginas 4 e 5.
- LI, B.; ZHANG, T.; XIA, T. Vehicle detection from 3d lidar using fully convolutional network. *arXiv preprint arXiv:1608.07916*, 2016. Citado na página 7.
- LI, Z. et al. A deep pedestrian tracking ssd-based model in the sudden emergency or violent environment. *Journal of advanced transportation*, Hindawi Limited, v. 2021, p. 1–13, 2021. Citado na página 19.
- LIANG, M.; HU, X. Recurrent convolutional neural network for object recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 3367–3375. Citado na página 7.
- LIMA, J. P. et al. 3d pedestrian localization using multiple cameras: A generalizable approach. *Machine Vision and Applications*, Springer, v. 33, n. 4, p. 61, 2022. Citado 7 vezes nas páginas 8, 1, 24, 25, 34, 42 e 45.

- LIMA, J. P. et al. Toward unlabeled multi-view 3d pedestrian detection by generalizable ai: techniques and performance analysis. In: IEEE. *2023 36th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. [S.l.], 2023. p. 1–6. Citado 5 vezes nas páginas 8, 1, 27, 35 e 42.
- LIU, L. et al. Deep learning for generic object detection: A survey. *International journal of computer vision*, Springer, v. 128, p. 261–318, 2020. Citado na página 7.
- LIU, W. et al. Ssd: Single shot multibox detector. In: SPRINGER. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. [S.l.], 2016. p. 21–37. Citado na página 4.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 3431–3440. Citado na página 5.
- LYRA, V. G. d. M. et al. Generalizable online 3d pedestrian tracking with multiple cameras. In: *VISIGRAPP (5: VISAPP)*. [S.l.: s.n.], 2022. p. 820–827. Citado 6 vezes nas páginas 8, 1, 28, 35, 36 e 45.
- MECALUX. *Sinalização no chão do armazém*. 2024. Acesso em: 07 fev. 2024. Disponível em: <<https://www.mecalux.com.br/blog/sinalizacao-chao-armazem>>. Citado 2 vezes nas páginas 8 e 2.
- MUELLER, M.; SMITH, N.; GHANEM, B. A benchmark and simulator for uav tracking. In: SPRINGER. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. [S.l.], 2016. p. 445–461. Citado 2 vezes nas páginas 9 e 12.
- MÜLLER, V. C. Ethics of artificial intelligence and robotics. 2020. Citado na página 11.
- ONFLEX. *Como Evitar Acidentes de Trabalho na Indústria*. 2024. <<https://www.onflex.com.br/como-evitar-acidentes-de-trabalho-na-industria/>>. Artigo disponível online. Disponível em: <<https://www.onflex.com.br/como-evitar-acidentes-de-trabalho-na-industria/>>. Citado 2 vezes nas páginas 8 e 3.
- PUSTOKHINA, I. V. et al. An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety. *Safety science*, Elsevier, v. 142, p. 105356, 2021. Citado 2 vezes nas páginas 8 e 29.
- RASMUSSEN, C.; HAGER, G. D. Probabilistic data association methods for tracking complex visual objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 23, n. 6, p. 560–576, 2001. Citado na página 9.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 779–788. Citado 2 vezes nas páginas 7 e 13.
- REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 39, n. 6, p. 1137–1149, 2016. Citado 2 vezes nas páginas 4 e 5.

RISTANI, E. et al. Performance measures and a data set for multi-target, multi-camera tracking. In: SPRINGER. *European conference on computer vision*. [S.l.], 2016. p. 17–35. Citado na página 12.

SANDLER, M. et al. Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 4510–4520. Citado na página 14.

SCHLEGL, T. et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: SPRINGER. *International conference on information processing in medical imaging*. [S.l.], 2017. p. 146–157. Citado na página 7.

SHAHZAD, A. R.; JALAL, A. A smart surveillance system for pedestrian tracking and counting using template matching. In: IEEE. *2021 International Conference on Robotics and Automation in Industry (ICRAI)*. [S.l.], 2021. p. 1–6. Citado 3 vezes nas páginas 8, 20 e 21.

SILVA, K. F. dos Santos da; LIMA, J. P. Silva do M.; TEICHRIEB, V. Pedestrian presence detection in areas of interest using multiple cameras. In: SPRINGER. *International Conference on Human-Computer Interaction*. [S.l.], 2023. p. 93–105. Citado 2 vezes nas páginas 8 e 37.

STADLER, D.; BEYERER, J. Improving multiple pedestrian tracking by track management and occlusion handling. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2021. p. 10958–10967. Citado na página 19.

STADLER, D.; BEYERER, J. Multi-pedestrian tracking with clusters. In: IEEE. *2021 17th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. [S.l.], 2021. p. 1–10. Citado na página 20.

SUN, Z. et al. A survey of multiple pedestrian tracking based on tracking-by-detection framework. *IEEE Transactions on Circuits and Systems for Video Technology*, IEEE, v. 31, n. 5, p. 1819–1833, 2020. Citado na página 5.

SUNDARARAMAN, R. et al. Tracking pedestrian heads in dense crowd. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2021. p. 3865–3875. Citado na página 19.

VO, M. et al. Self-supervised multi-view person association and its applications. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 43, n. 8, p. 2794–2808, 2020. Citado 3 vezes nas páginas 8, 16 e 17.

VORA, J. et al. Bringing generalization to deep multi-view pedestrian detection. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. [S.l.: s.n.], 2023. p. 110–119. Citado 6 vezes nas páginas 8, 1, 25, 26, 34 e 42.

WANG, Y.; YANG, H. Multi-target pedestrian tracking based on yolov5 and deepsort. In: IEEE. *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*. [S.l.], 2022. p. 508–514. Citado na página 24.

WOJKE, N.; BEWLEY, A. Deep cosine metric learning for person re-identification. In: IEEE. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. [S.l.], 2018. p. 748–756. Citado 2 vezes nas páginas 8 e 7.

- WU, Y.; LIM, J.; YANG, M.-H. Online object tracking: A benchmark. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2013. p. 2411–2418. Citado na página 8.
- XIAO, P. et al. Real-time 3d pedestrian tracking with monocular camera. *Wireless Communications and Mobile Computing*, Hindawi, v. 2022, 2022. Citado na página 23.
- XIAO, T. et al. End-to-end deep learning for person search. *arXiv preprint arXiv:1604.01850*, v. 2, n. 2, p. 4, 2016. Citado na página 10.
- XU, J.; BO, C.; WANG, D. A novel multi-target multi-camera tracking approach based on feature grouping. *Computers & Electrical Engineering*, Elsevier, v. 92, p. 107153, 2021. Citado na página 23.
- YANG, J. et al. Online pedestrian multiple-object tracking with prediction refinement and track classification. *Neural Processing Letters*, Springer, v. 54, n. 6, p. 4893–4919, 2022. Citado 2 vezes nas páginas 10 e 24.
- YOON, Y.-C. et al. Online multiple pedestrians tracking using deep temporal appearance matching association. *Information Sciences*, Elsevier, v. 561, p. 326–351, 2021. Citado na página 21.
- YOU, Q.; JIANG, H. Real-time 3d deep multi-camera tracking. *arXiv preprint arXiv:2003.11753*, 2020. Citado na página 17.
- YOU, S.; YAO, H.; XU, C. Multi-target multi-camera tracking with optical-based pose association. *IEEE Transactions on Circuits and Systems for Video Technology*, IEEE, v. 31, n. 8, p. 3105–3117, 2020. Citado na página 18.
- ZHANG, J.; LI, W.; OGUNBONA, P. Joint geometrical and statistical alignment for visual domain adaptation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 1859–1867. Citado na página 14.
- ZHANG, S.; BENENSON, R.; SCHIELE, B. Citypersons: A diverse dataset for pedestrian detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 3213–3221. Citado na página 12.
- ZHANG, X.; WANG, X.; GU, C. Online multi-object tracking with pedestrian re-identification and occlusion processing. *The Visual Computer*, Springer, v. 37, p. 1089–1099, 2021. Citado 2 vezes nas páginas 8 e 22.
- ZHANG, Z. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 22, n. 11, p. 1330–1334, 2000. Citado 3 vezes nas páginas 8, 9 e 10.
- ZHENG, L. et al. Person re-identification meets image search. *arXiv preprint arXiv:1502.02171*, 2015. Citado 2 vezes nas páginas 8 e 11.
- ZHU, X. J. Semi-supervised learning literature survey. University of Wisconsin-Madison Department of Computer Sciences, 2005. Citado na página 14.