



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO  
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA  
CURSO DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

Ali Castilho Santos

**Avaliação de viabilidade da utilização de ferramentas de inteligência artificial e  
visão computacional para a identificação de obstáculos e avaliação das  
condições de acessibilidade em calçadas para pessoas com mobilidade  
reduzida.**

Florianópolis

2024

Universidade Federal de Santa Catarina  
Departamento de Informática e Estatística

**Avaliação de viabilidade da utilização de ferramentas de inteligência artificial e visão computacional para a identificação de obstáculos e avaliação das condições de acessibilidade em calçadas para pessoas com mobilidade reduzida.**

Trabalho de Conclusão de Curso submetido ao curso de graduação em Sistemas de Informação do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Sistemas de Informação.

Autor: Ali Castilho Santos

Orientador: Prof .Elder Rizzon Santos, Dr.

Florianópolis

2024

## **RESUMO**

A proposta deste trabalho é analisar a viabilidade da utilização de tecnologias de visão computacional e geoprocessamento para melhorar a acessibilidade de pessoas com mobilidade reduzida. Foram revisados trabalhos anteriores para selecionar tecnologias relevantes e aplicá-las neste estudo. A coleta de dados foi realizada através de vídeos gravados em diferentes condições urbanas. Utilizou-se a plataforma Roboflow para rotulação e treinamento dos modelos, seguida da avaliação de desempenho em diferentes cenários. Os resultados demonstraram a eficácia do uso desta metodologia, indicando que há possibilidade de atingir resultados ainda mais assertivos ao expandir o dataset, refinar os modelos ou integrar com sistemas em tempo real, promovendo cidades mais acessíveis e inclusivas.

Palavras-chave: Mobilidade reduzida, visão computacional, geoprocessamento, acessibilidade.

## **ABSTRACT**

This study aims to analyze the feasibility of using computer vision and geoprocessing technologies to improve accessibility for people with reduced mobility. Previous works were reviewed to select relevant technologies and apply them in this study. Data collection was conducted through videos recorded under various urban conditions. The Roboflow platform was used for labeling and model training, followed by performance evaluation in different scenarios. Results demonstrated the effectiveness of this methodology, suggesting that more accurate results could be achieved by expanding the dataset, refining the models, or integrating with real-time systems to promote more accessible and inclusive cities.

Keywords: Reduced mobility, computer vision, geoprocessing, accessibility, inclusive cities.

## LISTA DE FIGURAS

Figura 1 - Modelo proposto baseado em grafo. Fonte: Adaptado de (LI; TORRES, 2014)	24
Figura 2 – Camadas da arquitetura de um GIS Fonte: Adaptado de (LI; TORRES, 2014)	25
Figura 3 – Camadas do sistema proposto. Fonte: Adaptado de (LI; TORRES, 2014)	28
Figura 4 – Distribuição de classificações.	41
Figura 5 – Gráficos de treinamento do modelo 1.	43
Figura 6 – Imagem exemplificando Augmentations, apresentando rotação.	44
Figura 7 – Gráficos de treinamento do modelo 2.	45
Figura 8 – Imagem exemplificando Augmentations, apresentando rotação, shear e brightness.	46
Figura 9 – Gráficos de treinamento do modelo 3.	47

## LISTA DE TABELAS

Tabela 1 – Exemplos de rotulações.	40
Tabela 2 – Comparação de detecções entre modelos.	50
Tabela 3 – Especificidades de modelos.	52

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>16</b>
1.1. OBJETIVO	18
1.1.1. Objetivo Geral	18
1.1.2. Objetivos Específicos	18
<b>2. FUNDAMENTAÇÃO TEÓRICA</b>	<b>20</b>
2.1 Visão computacional	20
2.2 Geoprocessamento	21
2.3 Cidades inteligentes	22
<b>3. TRABALHOS RELACIONADOS</b>	<b>24</b>
3.1 Sistema colaborativo para planejamento de rotas para cadeirantes	24
3.2 Measuring visual enclosure for street walkability: using machine learning algorithms and google street view imagery	31
3.3 A deep learning approach for street pothole detection	33
3.4 Discussão	35
<b>4. DESENVOLVIMENTO</b>	<b>37</b>
4.1 Seleção de Ferramentas	37
4.1.1 Roboflow	38
4.2 Elaboração do conjunto de dados	38
4.2.1 Aquisição de dados	38
4.2.2 Rotulação e Ferramentas	39
4.2.3 Conjunto de Dados	43
4.2.4 Treinamento e Resultados	44
4.2.5 Avaliação de Desempenho	49
<b>5. CONCLUSÃO</b>	<b>55</b>
<b>REFERÊNCIAS</b>	<b>58</b>
<b>APÊNDICE A - ARTIGO NO FORMATO SBC</b>	<b>61</b>

## 1. INTRODUÇÃO

A Universidade Federal de Santa Catarina abriga a Coordenadoria de Acessibilidade Educacional (CAE), vinculada à Secretaria de Ações Afirmativas e Diversidades (SAAD). No entanto, o acesso à própria CAE, localizada no prédio da Reitoria 1, no campus Florianópolis, pode representar um desafio para pessoas com mobilidade reduzida devido às condições precárias das vias de pedestres nos arredores da universidade, representando riscos evidentes para aqueles que as utilizam.

Diante dessa realidade, é fundamental buscar soluções inovadoras e eficientes que possam melhorar a acessibilidade e a segurança não apenas na universidade, mas em toda a cidade.

Existem diversas ferramentas e abordagens para enfrentar esse desafio. Melhorias físicas envolvem a construção e manutenção de rampas, calçadas niveladas e a implementação de sinalização adequada. Campanhas de conscientização são essenciais para educar a população sobre a importância da acessibilidade e os desafios enfrentados por pessoas com mobilidade reduzida. Políticas públicas, como regulamentações e incentivos para a construção de espaços acessíveis e a adaptação dos já existentes, também são fundamentais.

Nesse contexto, a utilização de ferramentas de inteligência artificial e visão computacional desponta como uma possibilidade promissora. Essas tecnologias têm a capacidade de extrair informações confiáveis de diversas fontes, como vídeos gravados por celulares, e apresentá-las aos órgãos responsáveis.

A inteligência artificial pode analisar grandes volumes de dados de forma rápida e precisa, identificando padrões e áreas problemáticas que necessitam de intervenção. A visão computacional, em particular, pode ser utilizada para identificar e mapear obstáculos de forma detalhada, como buracos, desníveis e obstruções nas calçadas.

Essa tecnologia pode gerar relatórios detalhados e visuais, que são facilmente compreensíveis pelos responsáveis pela tomada de decisões, promovendo a conscientização e contribuindo para a formulação de políticas públicas mais efetivas. Além disso, a implementação dessas tecnologias pode ser escalada para cobrir grandes áreas urbanas, tornando o processo de identificação de problemas mais eficiente e abrangente.

Este trabalho de conclusão de curso tem como objetivo avaliar a viabilidade da aplicação de tecnologias de visão computacional e inteligência artificial para extrair dados sobre condições de acessibilidade urbana, focando na identificação de obstáculos, mapeamento de condições precárias das calçadas e análise da infraestrutura existente.

Serão exploradas técnicas de inteligência artificial em conjunto com ferramentas de visão computacional, como processamento de imagens e detecção de objetos. Além da revisão de literatura, o estudo inclui a coleta de dados, desenvolvimento e avaliação de modelos práticos. Este trabalho proporcionará uma análise da eficácia dessas tecnologias, demonstrando seu potencial para fomentar o diálogo entre poder público, comunidade acadêmica e sociedade, visando cidades mais acessíveis e inclusivas.

Por meio da coleta e análise de dados obtidos a partir de vídeos gravados pessoalmente em diversas condições urbanas, busca-se identificar e mapear obstáculos, como buracos e condições inadequadas das calçadas, que impactam negativamente a mobilidade de pessoas com deficiência ou mobilidade reduzida. A utilização dessas tecnologias permitirá a geração de dados precisos e detalhados, contribuindo para a melhoria da infraestrutura urbana e promovendo um ambiente mais seguro e acessível para todos.

Os resultados obtidos serão apresentados em uma comparação de análises de aplicações de modelos de identificação de objetos, fornecendo indícios promissores do uso das tecnologias e ferramentas utilizadas. Além disso, espera-se que os resultados possam ser compartilhados com os entes mencionados anteriormente, como a Coordenadoria de Acessibilidade Educacional e a Secretaria de Ações Afirmativas e Diversidades, a fim de promover o diálogo e a cooperação entre as partes interessadas na busca por soluções efetivas e inclusivas.

Ao avaliar a viabilidade e eficácia dessas tecnologias, espera-se contribuir para a melhoria da acessibilidade e mobilidade de pessoas com deficiência ou mobilidade reduzida, não apenas na universidade, mas também em outros contextos urbanos. Além disso, os resultados obtidos poderão servir como referência para futuros estudos e incentivar a adoção de soluções baseadas em inteligência artificial e visão computacional em outras cidades que enfrentam desafios semelhantes.

Dessa forma, a proposta deste trabalho é ampliar o escopo de atuação, indo além do ambiente universitário, e buscar soluções tecnológicas que possam contribuir para a criação de cidades mais acessíveis, inclusivas e seguras para todos os cidadãos, especialmente para aqueles com mobilidade reduzida. A implementação de tais soluções pode fomentar um ambiente colaborativo entre academia, poder público e sociedade civil, promovendo a inovação contínua e a aplicação prática das tecnologias desenvolvidas. Isso não apenas melhora a infraestrutura urbana, mas também promove uma maior conscientização sobre a importância da acessibilidade, incentivando outras cidades a adotarem práticas semelhantes.

## **1.1. OBJETIVO**

### **1.1.1. Objetivo Geral**

O objetivo geral deste estudo é explorar o potencial de utilização de ferramentas de inteligência artificial e visão computacional no mapeamento de condições de calçamentos e vias para pedestres, buscando fazer uso de uma ferramenta completa e de fácil uso, para que o enfoque seja nos resultados e não nas ferramentas em si. Serão analisadas as possibilidades oferecidas por essas tecnologias, aprofundando-se em tópicos que contribuam para a produção de informações relevantes nessa área específica.

O intuito é investigar como a inteligência artificial e a visão computacional podem auxiliar na identificação de problemas e na avaliação da qualidade das estruturas de calçamento, fornecendo dados importantes para o planejamento e manutenção adequada das vias destinadas aos pedestres. Ao final deste estudo, espera-se oferecer conhecimentos valiosos para aprimorar a infraestrutura urbana, promovendo a segurança e a mobilidade das pessoas.

### **1.1.2. Objetivos Específicos**

- Realizar uma análise de trabalhos relevantes que abordam o uso de inteligência artificial e visão computacional para a identificação de objetos e o mapeamento de vias para pedestres, utilizando imagens e/ou dados brutos ou previamente processados.
- Analisar a viabilidade e identificar possíveis formas de empregar e/ou aprimorar

tecnologias para a análise de imagens e dados relacionados aos calçamentos e vias para pedestres.

- Destacar as oportunidades, avanços e áreas de melhoria para futuros trabalhos e interesses de terceiros.
- Contribuir para o avanço do conhecimento na área de acessibilidade em mobilidade urbana e inclusão, por meio da geração de dados precisos e detalhados que informem a formulação de políticas públicas e melhorias na infraestrutura, além de fomentar a conscientização sobre a importância da acessibilidade urbana.

## 2. FUNDAMENTAÇÃO TEÓRICA

### 2.1 Visão computacional

A visão computacional é uma área de pesquisa ativa, focada em desenvolver técnicas e algoritmos avançados para a análise e interpretação de imagens. Um marco significativo foi o trabalho de Krizhevsky, Sutskever e Hinton (2012), que introduziu a arquitetura de redes neurais convolucionais profundas, conhecida como AlexNet. Esse modelo revolucionou o campo, estabelecendo novos padrões para o reconhecimento de imagens.

AlexNet é composta por oito camadas treináveis: cinco camadas convolucionais e três camadas totalmente conectadas. As funções de ativação ReLU (Rectified Linear Unit) são usadas para introduzir não-linearidade, acelerando o treinamento e melhorando o desempenho. Além disso, AlexNet emprega técnicas de normalização local e dropout para reduzir o overfitting, tornando o modelo mais robusto e eficiente.

Após o trabalho de Girshick et al. (2014) com a técnica R-CNN, a área de detecção de objetos avançou significativamente. Novas abordagens como Faster R-CNN (Ren et al., 2015), YOLO (You Only Look Once) (Redmon et al., 2016) e SSD (Single Shot MultiBox Detector) (Liu et al., 2016) foram desenvolvidas, melhorando a precisão e a velocidade de detecção de objetos em imagens. Essas técnicas são amplamente aplicadas em diversos campos, incluindo segurança, veículos autônomos e reconhecimento facial.

No contexto de aplicação em imagens de ruas, o trabalho de Zhang et al. (2017) é relevante. Eles exploraram a detecção de semáforos em imagens capturadas por veículos autônomos, utilizando uma combinação de redes neurais convolucionais e técnicas de processamento de imagens para detectar e reconhecer semáforos em tempo real. Esse estudo contribuiu para o desenvolvimento de sistemas de condução autônoma mais seguros e eficientes.

Esses são apenas alguns exemplos das muitas contribuições científicas que têm impulsionado o campo da visão computacional. Através desses trabalhos e de outros estudos inovadores, a inteligência artificial e a visão computacional têm se fortalecido, possibilitando avanços notáveis na análise e interpretação de imagens. Essas tecnologias são cruciais no contexto de imagens de ruas, como as disponibilizadas no Google Street View, oferecendo novas possibilidades para a

análise urbana e a melhoria da acessibilidade.

## **2.2 Geoprocessamento**

O Geoprocessamento é uma área multidisciplinar que envolve a coleta, análise, interpretação e visualização de dados geográficos. Com o avanço da tecnologia e o uso crescente de inteligência artificial (IA) e visão computacional, o campo do Geoprocessamento tem se beneficiado dessas técnicas para extrair informações valiosas dos dados espaciais.

A integração de IA e visão computacional no geoprocessamento permite a automação e a melhoria da precisão em diversas tarefas. Um exemplo significativo é a detecção automática de objetos em imagens de sensoriamento remoto. Zhang et al. (2017) propuseram o uso de redes neurais convolucionais para segmentação e classificação de objetos em imagens de satélite, possibilitando a identificação automatizada de características geográficas, como edifícios, estradas e corpos d'água. Este trabalho exemplifica como o Geoprocessamento pode se beneficiar da IA para melhorar a análise e a precisão dos dados geográficos.

Além disso, a IA tem sido aplicada no Geoprocessamento para a análise e previsão de padrões espaciais. Oliveira et al. (2020) utilizaram técnicas de aprendizado de máquina para analisar a relação entre dados geográficos e a ocorrência de deslizamentos de terra. O estudo demonstrou que os modelos de IA foram capazes de identificar padrões ocultos nos dados, contribuindo para a prevenção e mitigação de desastres naturais. Isso mostra como o Geoprocessamento, combinado com a IA, pode ser uma ferramenta poderosa para a gestão de riscos e a previsão de eventos geográficos.

Outra área de pesquisa relevante é a modelagem preditiva de demanda por serviços públicos. Silva et al. (2019) desenvolveram um modelo baseado em IA para prever a demanda por serviços de transporte público, utilizando dados geográficos e socioeconômicos. Essa abordagem permitiu um planejamento mais eficiente dos recursos e uma melhor alocação de serviços nas áreas urbanas. Aqui, o Geoprocessamento é essencial para entender a distribuição espacial da demanda e otimizar a prestação de serviços públicos.

A combinação de Geoprocessamento, IA e visão computacional tem sido amplamente aplicada na agricultura de precisão. Pesquisadores como Kussul et al.

(2017) exploraram o uso de imagens de satélite e técnicas de aprendizado de máquina para monitorar a saúde das plantações, identificar áreas de estresse vegetal e otimizar o uso de insumos agrícolas. Essas tecnologias permitem um monitoramento mais detalhado e preciso das condições agrícolas, melhorando a produtividade e sustentabilidade das práticas agrícolas.

Esses exemplos ilustram como o Geoprocessamento tem se beneficiado das técnicas de IA e visão computacional para extrair informações valiosas dos dados geográficos. Com o contínuo avanço dessas tecnologias, espera-se que o Geoprocessamento se torne ainda mais preciso, eficiente e capaz de fornecer insights relevantes para uma variedade de aplicações, desde a gestão de desastres naturais até o planejamento urbano e a agricultura de precisão.

### **2.3 Cidades inteligentes**

As cidades inteligentes representam uma evolução no planejamento urbano, integrando tecnologias avançadas para melhorar a qualidade de vida dos cidadãos, otimizar a gestão urbana e promover a sustentabilidade. Este conceito abrange diversas áreas, incluindo mobilidade urbana, gestão de energia, segurança pública e saneamento, entre outras.

Cidades inteligentes são definidas como áreas urbanas que utilizam tecnologia da informação e comunicação (TIC) para aumentar a eficiência dos serviços urbanos, melhorar a qualidade de vida dos cidadãos e promover a sustentabilidade. A ideia central é criar um ambiente urbano mais conectado, onde a infraestrutura e os serviços são otimizados por meio do uso de dados e tecnologia.

A mobilidade urbana inteligente envolve a gestão do tráfego, onde sensores e câmeras monitoram o fluxo em tempo real, permitindo ajustes dinâmicos nos sinais de trânsito e sugerindo rotas alternativas para reduzir congestionamentos. No transporte público, aplicativos inteligentes fornecem informações sobre horários, rotas e disponibilidade, enquanto sistemas preditivos ajudam a otimizar o planejamento de rotas.

Na gestão de energia, as redes elétricas inteligentes, ou smart grids, monitoram e gerenciam a distribuição de energia de forma eficiente, integrando fontes renováveis e reduzindo perdas. Edifícios inteligentes utilizam sensores para monitorar e controlar o consumo de energia em tempo real, ajustando

automaticamente iluminação e climatização para economizar energia.

A segurança pública também se beneficia das cidades inteligentes. A vigilância inteligente utiliza câmeras de segurança com algoritmos de IA para detectar atividades suspeitas, objetos abandonados e padrões de comportamento anormais. Sistemas de comunicação integrados coordenam rapidamente a resposta a emergências, utilizando dados em tempo real para alocar recursos eficientemente.

As cidades inteligentes melhoram a qualidade de vida por meio da prestação eficiente e responsiva dos serviços públicos, facilitando o acesso dos cidadãos a esses serviços e contribuindo para um ambiente sustentável. A gestão urbana é otimizada através da coleta e análise de dados em tempo real, permitindo que os gestores urbanos tomem decisões mais informadas e eficazes e facilitem o planejamento e a implementação de melhorias na infraestrutura urbana. Além disso, a segurança e resiliência são aprimoradas com tecnologias de vigilância e comunicação, melhorando a capacidade de resposta a emergências e a análise preditiva que ajuda a identificar e mitigar riscos.

No entanto, as cidades inteligentes enfrentam desafios significativos, como a coleta e o gerenciamento de grandes volumes de dados, que levantam questões sobre privacidade e segurança. A implementação dessas tecnologias requer investimentos substanciais e uma infraestrutura robusta. No futuro, espera-se que as cidades inteligentes se tornem ainda mais integradas e eficientes, com a adoção crescente de tecnologias emergentes como a Internet das Coisas (IoT), 5G e blockchain. A colaboração entre governos, empresas e a academia será crucial para superar os desafios e maximizar os benefícios das cidades inteligentes.

### **3. TRABALHOS RELACIONADOS**

Durante o processo de pesquisa, foram realizadas investigações detalhadas sobre a viabilidade da utilização de ferramentas de inteligência artificial e visão computacional para a identificação de obstáculos e avaliação das condições de acessibilidade em calçadas destinadas a pessoas com mobilidade reduzida.

A revisão da literatura abrangeu conceitos, técnicas, abordagens e possíveis tecnologias aplicáveis ao processamento de imagens e análise de dados. Visando evidenciar a possibilidade de aprofundamento de estudos e uso de ferramentas que até poucos anos atrás eram inconcebíveis, mas que estão se tornando cada vez mais acessíveis e usuais, como o geoprocessamento e as aplicações de visão computacional com bases de dados existentes. Esses estudos convergem em dados que abrem novas portas e removem barreiras na área objeto deste estudo.

A pesquisa, feita em setembro de 2022, foi conduzida exclusivamente no Google Scholar devido à sua ampla cobertura de artigos acadêmicos e materiais de diversas fontes. O Google Scholar foi especialmente útil por indexar uma ampla gama de fontes. As palavras-chave utilizadas incluíram: acessibilidade, pessoas com deficiência, YOLO, processamento de imagens, detecção de objetos, aprendizado de máquina e detecção de buracos.

Foram considerados apenas artigos que tratam da identificação de obstáculos e análise de condições de calçadas utilizando inteligência artificial e visão computacional. Trabalhos que não se enquadram nesses critérios ou que não fornecem informações detalhadas foram excluídos. Foram incluídos apenas estudos publicados nos últimos dez anos em português e inglês, abrangendo imagens e vídeos de diferentes tipos.

A avaliação dessas tecnologias visa fornecer dados para possíveis esforços de melhoria da acessibilidade e mobilidade de pessoas com deficiência ou mobilidade reduzida em contextos urbanos. Os resultados podem servir como referência para futuros estudos e incentivar a adoção dessas soluções em outras cidades que enfrentam desafios semelhantes.

#### **3.1 Sistema colaborativo para planejamento de rotas para cadeirantes**

BARCYZSZYN, Guilherme Lincon. Sistema colaborativo para planejamento de rotas para cadeirantes. 2019. 71 f. Dissertação (Mestrado em Computação

Aplicada) - Universidade Tecnológica Federal do Paraná, Curitiba, 2019. Em se tratando do desafio do planejamento de rotas na computação urbana, que requer o processamento de grandes volumes de dados e o feedback dos usuários. Tradicionalmente, os serviços de planejamento de rotas são baseados em ruas, o que não é adequado para usuários com deficiências. Para solucionar esse problema é apresentado um modelo baseado em calçadas, que atende às necessidades de cadeirantes por meio de um serviço de planejamento de rotas. O modelo é definido matematicamente como um grafo, com vértices representando esquinas e arestas representando calçadas ou cruzamentos. Os custos de cada aresta são determinados por diversos fatores, como distância, inclinação e condição/manutenção de rampas, cruzamentos e calçadas.

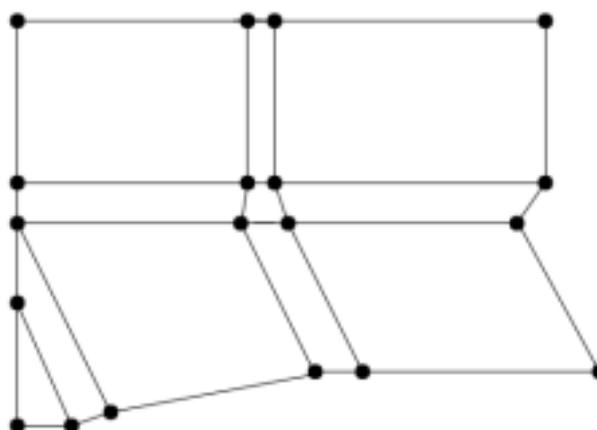


Figura 1 - Modelo proposto baseado em grafo. Fonte: Adaptado de (LI; TORRES, 2014)

Os custos das arestas foram determinados considerando fatores de acessibilidade, como distâncias, inclinação e condições das rampas e calçadas. O modelo desenvolvido leva em consideração o feedback do usuário, atualizando suas informações com dados fornecidos por cadeirantes sobre problemas de acessibilidade e soluções do departamento de planejamento urbano. Além disso, é apresentado um algoritmo de planejamento de rotas que gera um conjunto de rotas alternativas com base em critérios de acessibilidade, juntamente com um algoritmo de recomendação de atalhos para auxiliar o departamento de planejamento urbano nas tomadas de decisão. O autor desenvolveu e disponibilizou também uma API para fornecer os serviços de caminho mínimo em calçadas.

Barczyszyn, para alcançar seu objetivo trabalhou em etapas: revisão bibliográfica abrangendo Sistemas de Informação Geográfica, algoritmos de caminhos mínimos e cidades inteligentes; exploração de trabalhos relacionados para

identificar modelos e sistemas existentes de planejamento de rotas para pessoas com necessidades especiais; elaboração de modelo teórico baseado em calçadas, seguindo arquitetura proposta por Zheng et al. (2014); implementação de modelo utilizando dados geoespaciais abertos da cidade de Curitiba; fornecimento de serviço de planejamento de rotas para cadeirantes; e validação da implementação do modelo por meio de testes focados nos objetivos definidos para garantir deslocamentos acessíveis.

Dentre os temas estudados em trabalhos relacionados, o autor também destaca os Sistemas de Informação Geográfica (SIG), que desempenham um papel fundamental na manipulação de dados geográficos e seguem uma arquitetura específica. Essa arquitetura é composta por três camadas distintas: a camada de apresentação, que lida com a visualização e interação com os dados geográficos; a camada de processamento, que realiza análises espaciais e aplica algoritmos para extrair informações relevantes; e a camada de dados, que armazena e gerencia os dados geográficos em formatos apropriados.



Figura 2 – Camadas da arquitetura de um GIS Fonte: Adaptado de (LI; TORRES, 2014)

Outro importante tema exposto pelo autor é o geoprocessamento, que trata de uma área de conhecimento recente que desempenha um papel importante na resolução de problemas urbanos. Surgiu no final do século XX como uma ferramenta avançada para interpretações relacionadas ao espaço. Envolve técnicas que permitem análises espaciais, manipulação e gerenciamento de informações georreferenciadas com agilidade e precisão. Os dados geográficos estão sempre associados a atributos que os relacionam a uma localização geográfica existente ou

a outras entidades geográficas com localização conhecida. O geoprocessamento fornece ferramentas essenciais para manipular esses dados, auxiliando no desenvolvimento do trabalho proposto com base nos dados fornecidos pelo departamento urbano da cidade.

No contexto da informação geográfica, as relações espaciais desempenham um papel fundamental. Elas podem ser expressas de várias maneiras, como relações topográficas, métricas e direcionais. As relações topográficas indicam conexões entre objetos, como adjacência, inclusão, sobreposição e cruzamento. As relações métricas envolvem medidas quantitativas, como área, distância, comprimento e perímetro. As relações direcionais expressam orientações baseadas em pontos cardeais ou posições relativas. O conhecimento dessas relações espaciais é indispensável para o entendimento e desenvolvimento do trabalho, pois permite a transformação de informações contendo objetos geográficos em dados úteis para o modelo proposto. Por exemplo, ao buscar o ponto central de um cruzamento, utiliza-se a operação "cruza" entre duas ruas. Ao procurar todas as esquinas dentro de um raio em torno desse ponto central, a operação "contém" é utilizada para agrupá-las.

Além disso, o entendimento das relações espaciais é essencial para o uso efetivo dos algoritmos de menor caminho, como o Algoritmo de Dijkstra e o A\*. As relações topológicas e métricas são fundamentais na representação e análise dos objetos geográficos, permitindo a identificação de conexões, sobreposições, inclusões e outras interações entre entidades. Ao aplicar o Algoritmo de Dijkstra ou o A\* para encontrar a rota ideal para um cadeirante considerando diferentes custos, é necessário utilizar as relações espaciais para definir as restrições e considerar os elementos relevantes para o planejamento da rota. Dessa forma, o conhecimento das relações espaciais e a familiaridade com os algoritmos de menor caminho são cruciais para o desenvolvimento de soluções eficientes no contexto de cidades inteligentes, onde a otimização do trajeto e a acessibilidade são aspectos-chave para melhorar a qualidade de vida dos cidadãos.

Os dados utilizados pelo autor foram adquiridos principalmente de três fontes: sensoriamento urbano, feedback de cadeirantes e outros cidadãos, e feedback de autoridades de planejamento urbano. A agência de planejamento urbano fornece informações geoespaciais sobre a rede de transportes, pontos de interesse, topografia, tráfego, ruído, acidentes e meteorologia.

Especificamente, no trabalho, foi utilizado um conjunto de camadas de mapas base fornecido pelo Instituto de Pesquisa e Planejamento Urbano (IPPUC), que inclui as geometrias das ruas e quadras, bem como objetos de baixo nível com suas coordenadas geográficas e dados de elevação.

O sistema de banco de dados é construído utilizando o PostGIS, um plugin para o PostgreSQL que permite a manipulação de dados geográficos. O modelo baseado em calçadas e cruzamentos é utilizado para o planejamento de rotas para cadeirantes, e as informações de acessibilidade são atualizadas por meio do feedback dos usuários e das ações do departamento de planejamento urbano para melhorar as condições de acessibilidade da cidade.

A arquitetura do sistema proposto segue o framework para computação urbana proposto por Zheng et al. (2014). A arquitetura é composta por várias camadas, incluindo sensoriamento urbano e aquisição de dados, gerenciamento de dados urbanos, análise de dados e fornecimento de serviços. O sensoriamento urbano coleta dados por meio de sensores, dispositivos móveis, redes sociais e feedback dos usuários. O gerenciamento de dados urbanos estrutura os dados de forma eficiente para análise. A análise de dados envolve algoritmos de otimização, mineração de dados, aprendizado de máquina e visualização, e, neste trabalho, também inclui o planejamento de rotas.

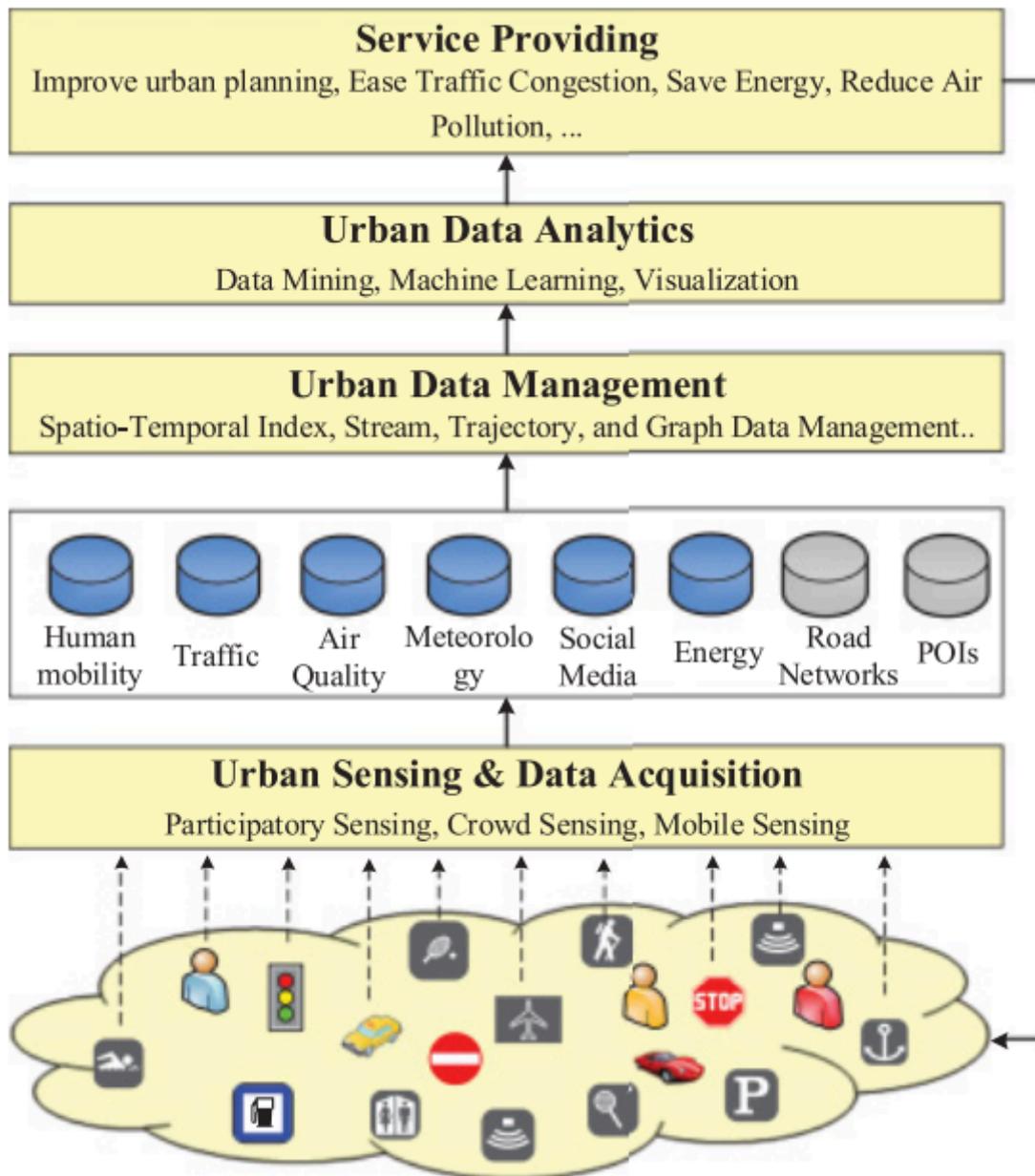


Figura 3 – Camadas do sistema proposto. Fonte: Adaptado de (LI; TORRES, 2014)

A camada de fornecimento de serviços fornece serviços específicos aos usuários, como a visualização da rota mais curta e a possibilidade de reportar problemas. O trabalho realizado integra as diferentes camadas de dados fornecidos pelo IPPUC e forma um banco de dados geográfico que é utilizado para a criação do banco de dados específico para o planejamento de rotas. As figuras mostram visualizações dos dados do domínio de quadras e altimetria combinados com o OpenStreetMaps.

O autor explicita que foram utilizadas as seguintes tecnologias: PostgreSQL

9.5 como Sistema Gerenciador de Banco de Dados (SGBD), a extensão PostGIS 2.2 para manipulação de dados geoespaciais, QuantumGIS 2.16.3 para visualização dos dados em mapas e o plugin PgRouting 2.2.2 para cálculo de rotas. O PgRouting é uma extensão do PostGIS que fornece várias funções para cálculo de caminhos mínimos usando dados geográficos, como algoritmos de Dijkstra e A\*.

A implementação do grafo neste trabalho seguiu o modelo teórico apresentado. Inicialmente, foram obtidos os contornos das calçadas a partir das geometrias dos blocos das quadras da cidade. Em seguida, os cruzamentos foram identificados e linhas foram criadas para representá-los. O banco de dados fornecido pelo IPPUC continha geometrias de polígonos das quadras da cidade, que foram utilizadas para criar o grafo. A função `STExteriorRing` do PostGIS foi usada para obter o contorno dessas geometrias, que representa as calçadas.

Para criar os cruzamentos, foram identificados os pontos iniciais e finais das travessias e linhas foram geradas a partir dessas travessias. Primeiro, foram identificados os pontos de cruzamento utilizando geometrias de linhas que representam as ruas da cidade. A função `STIntersects` foi usada para gerar pontos onde havia cruzamentos. Em seguida, os pontos de esquina nas linhas das calçadas foram identificados usando a função `STDump`, e as linhas das calçadas foram simplificadas e pequenas quebras foram ignoradas usando a função `STSimplify`.

Depois de identificar os pontos de esquina, foi necessário agrupá-los por cruzamento. Isso foi feito gerando um raio em torno de cada ponto central de cruzamento e usando a função `STWithin` para identificar quais pontos de esquina estavam dentro do raio. Os pontos de esquina foram agrupados por cruzamento. Por fim, as linhas das travessias foram criadas ligando os pontos de esquina usando a função `STConvexhull` para gerar polígonos entre os pontos e, em seguida, retirando o contorno desses polígonos para obter as linhas das travessias.

Conclui-se que um planejamento de rotas é um problema desafiador que envolve o processamento de uma grande quantidade de dados - como mapas, imagens e topografia da cidade - além de uma grande quantidade de informações complexas oriundas de feedbacks de usuários e de agências municipais, a fim de definir corretamente caminhos confiáveis.

Segundo Barczynszyn, o objetivo desse trabalho é utilizar a computação para proporcionar uma maior comodidade e qualidade de vida aos cadeirantes, tornando

o seu deslocamento menos custoso. Para tanto, este trabalho apresentou um modelo teórico que contempla os principais pontos que interferem no deslocamento dos cadeirantes em um cenário urbano como rampas, inclinação, acessibilidade, fatores esses que trabalhos anteriores não apresentam.

Os objetivos específicos alcançados por este trabalho foram, portanto: a realização de uma revisão bibliográfica nos seguintes conceitos: SIG, algoritmos e cidades inteligentes; uma pesquisa explorando os principais trabalhos que elaboraram modelos semelhantes ao apresentado aqui; a elaboração de um novo modelo teórico utilizando a arquitetura proposta por Zheng et al. (2014); a implementação do modelo através de dados geográficos e um banco de dados geoespacial; a disponibilização do modelo através de uma API; e a validação do modelo através de testes.

As principais dificuldades retratadas foram a manipulação das funções corretas para fazer a transformação de dados, principalmente para a geração correta de arestas de cruzamentos, e a obtenção dos pontos de cruzamento entre aresta de calçada e curvas de altimetria. A posterior aplicação de todas as transformações para toda a cidade de Curitiba também foi uma das grandes dificuldades por envolver uma enorme quantidade de registros.

### **3.2 Measuring visual enclosure for street walkability: using machine learning algorithms and google street view imagery**

O estudo do design urbano e da mobilidade urbana enfrenta desafios em relação à consistência e objetividade dos dados, especialmente quando se analisam diversos tipos de vias, calçamentos, construções e componentes. No entanto, os avanços nos instrumentos e tecnologias de sensoriamento têm proporcionado melhorias significativas na obtenção de dados confiáveis, precisos, ágeis e em grande escala. Com a melhoria exponencial dos sensores e técnicas de digitalização de dados, é possível aplicar cada vez mais métodos de processamento de dados em diversas fontes e bases de dados, automatizando a obtenção de resultados e refinando os dados com alta precisão e economia de tempo e recursos.

Um exemplo de aplicação desses avanços é a análise de dados extraídos das imagens do Google Street View por meio de técnicas de aprendizado de máquina, bem exemplificado no artigo YIN, Li; WANG, Zhenxin. Measuring visual enclosure for

street walkability: Using machine learning algorithms and Google Street View imagery. 2016. Esses dados coletados por instrumentos de alta precisão podem ser objeto de estudo em diversas áreas.

Este artigo explora o potencial dessas técnicas para identificar, classificar e categorizar várias vias urbanas, bem como o impacto dos objetos e construções nelas presentes. No entanto, o estudo reconhece duas limitações significativas ao trabalhar com essas métricas. A primeira preocupação está relacionada à viabilidade do processamento em larga escala, uma vez que o processamento de imagens de alta resolução tende a ser custoso. Além disso, há frequentes inconsistências na confiabilidade dos resultados das análises, devido à variação apresentada e à falta de padronização da margem de erro.

O estudo foi realizado na cidade de Buffalo, nos Estados Unidos, ressaltando a importância de tornar a mobilidade não apenas possível, mas também atrativa, confiável, viável e efetiva. Quando a locomoção urbana é fluida, a qualidade de vida aumenta, e as pessoas têm a oportunidade de praticar exercícios físicos, seja no dia a dia ou como atividade esportiva, quando o ambiente urbano proporciona segurança e acomoda essas práticas.

Ao citar outros estudos, este artigo faz uso das variáveis D (densidade, design, diversidade, acessibilidade de destino e distância de trajeto) para abordar o assunto. Embora essas variáveis sejam abordadas no estudo, acredita-se que sejam insuficientes para refletir com alta precisão a qualidade da mobilidade urbana na perspectiva do pedestre. No entanto, essas variáveis fornecem dados importantes para pesquisas adicionais.

Utilizando 3592 imagens panorâmicas de 311 blocos (equivalentes a quarteirões), foram aplicadas as técnicas de rede neural artificial (ANN) e Máquina de Vetores de Suporte (SVM), seguindo métodos convencionais de visão computacional e reconhecimento de padrões. A análise abrangeu textura e cor usando ANN, além de detecção e segmentação combinadas com classificação do SVM.

No contexto deste estudo, foi explorado o potencial de utilização de ferramentas de inteligência artificial e visão computacional para o mapeamento e análise das condições das vias destinadas aos pedestres. Ao analisar o estado da arte nessa área e considerar abordagens já existentes, constatou-se que essas

tecnologias podem desempenhar um papel significativo na identificação de problemas e na avaliação da qualidade das vias, especialmente para grupos sensíveis como as Pessoas com Mobilidade Reduzida (PMR).

Através do processamento de imagens e da aplicação de algoritmos de inteligência artificial, foi possível identificar potenciais riscos e obter informações relevantes sobre as condições dos calçamentos e das vias para pedestres. Esses resultados podem ser utilizados para embasar a formulação de políticas públicas e para auxiliar na tomada de decisão em relação a melhorias e intervenções necessárias.

A disponibilização dos dados coletados e processados de maneira acessível e compreensível é um aspecto essencial para a efetividade dessas informações. Através da criação de relatórios detalhados e da apresentação visual dos resultados, busca-se oferecer uma forma clara e intuitiva de interação com os dados, tanto para o público em geral quanto para especialistas e responsáveis pela tomada de decisão.

É importante ressaltar que a aplicação dessas tecnologias deve ser feita de forma responsável e ética, considerando questões de privacidade e segurança. Além disso, a colaboração entre diferentes atores, como instituições acadêmicas, órgãos governamentais e comunidades locais, é fundamental para o sucesso da implementação de melhorias na acessibilidade das vias para pedestres.

Em suma, este estudo demonstrou o potencial das técnicas de processamento de dados, utilizando imagens do Google Street View, para analisar e compreender as características das vias urbanas. Embora as variáveis D tenham fornecido informações relevantes, é necessário continuar pesquisando e aprimorando as metodologias para uma avaliação mais precisa da qualidade da mobilidade urbana na perspectiva do pedestre. O estudo destaca a importância de considerar não apenas os aspectos físicos das vias, mas também as percepções e experiências dos usuários, a fim de promover um ambiente urbano inclusivo, seguro e sustentável.

### **3.3 A deep learning approach for street pothole detection**

Buracos em vias e rodovias, além de danificar os veículos, podem causar sérios acidentes de trânsito, tornando necessário encontrar uma forma de

mapeá-los para uma ação imediata. Este artigo aborda essa questão ao utilizar o Aprendizado Profundo (Deep Learning) para fornecer um método automatizado de mapeamento de vias. Quatro modelos foram treinados utilizando um conjunto de dados pré-processado, incluindo YOLO V3 (You Only Look Once), SSD (Single Shot Detector), HOG (Histogram of Oriented Gradients) com SVM (Support Vector Machine) e Faster R-CNN.

Os autores destacam o interesse significativo, especialmente em países desenvolvidos, na detecção rápida e eficiente de buracos nas vias. O artigo explica como algumas metodologias de detecção foram aplicadas, começando pelo SVM para detecção de buracos. Essa técnica envolve extrair a região da imagem com base no histograma da região e no kernel SVM.

Utilizando o aprendizado profundo baseado em Redes Neurais Convolucionais (CNN), o ruído causado pela iluminação e sombras é ignorado, e as detecções podem ser feitas a partir de fotos tiradas por celulares. O artigo também menciona um estudo de classificação binária baseado em redes neurais profundas e um modelo chamado Crack-net para detectar fraturas e rachaduras nas vias.

O esquema de detecção é dividido em duas partes: preparação dos dados e detecção. Na primeira parte, são coletadas imagens de vias que são divididas em duas categorias: com buracos e sem buracos, com dois níveis de dificuldade: fácil e difícil. Na segunda parte, os modelos de aprendizado profundo são alimentados com essas imagens devidamente preparadas e o processo de treinamento é iniciado. Convertendo a imagem para escala de cinza, também é possível extrair dados sobre o tamanho dos buracos. Embora a sombra em algumas imagens tenha causado interferências, os resultados ainda foram satisfatórios. As análises foram realizadas usando conjuntos de teste com 200, 650, 850, 1000, 1100 e 1500 imagens, com tempo máximo de espera de 8 horas e 30 minutos para o último grupo.

Comparando os resultados dos quatro modelos testados, os autores concluíram que, ao analisar uma quantidade pequena de dados, o Faster R-CNN apresentou o melhor desempenho. No entanto, à medida que o tamanho do banco de dados aumenta, o YOLO V3 se destaca, oferecendo resultados mais confiáveis e exigindo menos tempo de processamento. Embora todos os algoritmos tenham apresentado resultados semelhantes, exceto o HOG, que teve desempenho significativamente inferior e foi desconsiderado nas análises posteriores. Com 200

imagens, a técnica HOG levou 2 horas e obteve uma taxa de acerto de 24%, o SSD levou 4 horas e alcançou 47% de acerto, o Faster R-CNN levou 2 horas e atingiu 72% de acerto, enquanto o YOLO V3 precisou de 3 horas e alcançou apenas 53% de acerto.

Na amostra de 650 imagens, o HOG levou 3 horas e 30 minutos, com apenas 25% de acerto, o SSD levou 5 horas e obteve 59% de acerto, o Faster R-CNN levou apenas 2 horas e 30 minutos, demonstrando maior velocidade de processamento e liderando em acurácia, com 71%, enquanto o YOLO V3 levou 4 horas e alcançou 67% de acerto. Com 850 imagens, o HOG alcançou apenas 27% de acerto e não foi mais utilizado nas etapas seguintes. No entanto, o salto significativo em desempenho e acurácia foi observado na etapa com 1500 imagens de amostra, onde o YOLO V3 levou 5 horas para fornecer resultados com 82% de acerto, em comparação com 6 e 8 horas e 30 minutos para o Faster R-CNN e SSD, respectivamente, com taxas de acerto de 74% e 80%.

### **3.4 Discussão**

A revisão da literatura demonstra que existem pesquisas recentes sobre a aplicação de inteligência artificial e visão computacional para a identificação de obstáculos em calçadas. No entanto, diferem na coleta de dados, contando com imagens de terceiros, feedback de cidadãos ou capturas por carros. Este trabalho difere ao utilizar vídeos gravados pessoalmente com smartphones, focando em imagens frontais das calçadas. Não foram encontrados trabalhos anteriores que utilizassem essa abordagem específica até o momento.

Em relação aos modelos utilizados nas pesquisas, há uma predominância do uso de redes neurais convolucionais, especialmente o YOLO (You Only Look Once), para a detecção de objetos. Estudos indicam que modelos baseados em YOLO apresentam bons desempenhos em seus contextos específicos.

Na data de pesquisa dos trabalhos relacionados não foram identificadas pesquisas brasileiras que utilizem detecção de objetos para avaliar a acessibilidade urbana, o que evidencia uma lacuna de conhecimento. Este trabalho busca preencher essa lacuna, contribuindo para a pesquisa nacional na área de acessibilidade.

A seleção de trabalhos relacionados foi resultado de revisão criteriosa e

discussões entre autor e orientador, para assegurar que apenas artigos pertinentes fossem incluídos. Essas medidas garantem que a revisão seja abrangente e confiável, fornecendo uma base sólida para o desenvolvimento do trabalho e contribuindo significativamente para o avanço do conhecimento na área de acessibilidade e mobilidade urbana.

## **4. DESENVOLVIMENTO**

A utilização de inteligência artificial e visão computacional requer uma abordagem meticulosa e bem estruturada. A complexidade dessas tecnologias e sua aplicação prática demandam um planejamento detalhado, garantindo que cada etapa do projeto seja realizada com precisão e eficiência. Este capítulo detalha o processo de seleção das ferramentas utilizadas, a metodologia adotada para a criação e treinamento dos modelos, e as escolhas tecnológicas que fundamentaram o desenvolvimento do projeto. O objetivo é garantir que cada etapa seja transparente e baseada em decisões tecnológicas adequadas para alcançar resultados eficazes e replicáveis.

O uso de ferramentas que cobrem todas as etapas do desenvolvimento de modelos de visão computacional foi tratado com prioridade nesta abordagem. Essas etapas incluem a captura e organização de dados, a categorização dos dados, o treinamento dos modelos, a obtenção de relatórios detalhados e a aplicação dos modelos treinados em novos cenários. A integração completa dessas etapas em uma única plataforma simplifica o fluxo de trabalho e aumenta a eficiência do desenvolvimento.

A seguir, descrevemos a seleção das ferramentas e o motivo pelo qual o Roboflow foi a plataforma escolhida para este estudo. A escolha das ferramentas foi orientada pelo consenso entre o autor e o orientador, focando em plataformas que permitissem um processo de desenvolvimento completo, mesmo que em regime de avaliação gratuita.

### **4.1 Seleção de Ferramentas**

A seleção de ferramentas para este projeto foi um processo colaborativo entre o autor e o orientador, com o objetivo de utilizar plataformas que permitissem a criação de modelos e obtenção de resultados de maneira eficiente, mesmo em um regime de avaliação gratuita. O consenso foi de que as ferramentas escolhidas deveriam cobrir todas as etapas necessárias: montagem do dataset (separação de frames de vídeos), categorização, treinamento de modelos, obtenção de relatórios e aplicação dos modelos treinados em novos envios.

Seguindo essa ordem, a busca inicial no Google por "label YOLO" resultou no Roboflow como o primeiro resultado. Após testar a plataforma, concluiu-se que ela

atendia a todos os pré-requisitos acordados, embora oferecesse apenas 3 créditos gratuitos para o treinamento de modelos. A versatilidade do Roboflow e sua capacidade de lidar com todas as etapas do processo de desenvolvimento foram determinantes para sua escolha.

#### **4.1.1 Roboflow**

Roboflow é uma empresa dedicada a facilitar o desenvolvimento e implantação de modelos de visão computacional. Fundada em 2019 por Joseph Nelson e Brad Dwyer, a empresa tem como missão democratizar a criação de aplicações de visão computacional, tornando-a acessível para desenvolvedores de todos os níveis de habilidade.

A plataforma do Roboflow permite aos usuários importar, organizar, classificar e treinar modelos de visão computacional de maneira intuitiva. Além disso, oferece ferramentas para a implantação e monitoramento de modelos, proporcionando um fluxo de trabalho completo desde a criação do dataset até a aplicação prática dos modelos treinados.

Roboflow é uma ferramenta comercial que ganhou destaque rapidamente no campo da visão computacional, sendo frequentemente utilizada tanto em contextos acadêmicos quanto comerciais. A plataforma é aplicada em diversas áreas, incluindo saúde, agricultura, manufatura, segurança e muito mais. Sua utilização em diversos setores demonstra a flexibilidade e a robustez das soluções que oferece.

A empresa não possui referências bibliográficas tradicionais, mas é amplamente reconhecida na comunidade de desenvolvedores e pesquisadores por sua eficiência e facilidade de uso. Com uma forte comunidade de usuários e um compromisso contínuo com a inovação, o Roboflow se posiciona como uma ferramenta essencial para desenvolvedores que buscam implementar modelos de visão computacional de forma eficiente e eficaz.

## **4.2 Elaboração do conjunto de dados**

### **4.2.1 Aquisição de dados**

Para a criação do conjunto de dados, o próprio autor filmou 11 pequenos vídeos com durações variando entre 10 e 400 segundos. Esses vídeos foram capturados em diferentes pontos da cidade de Forquilha, no estado de Santa Catarina, abrangendo uma ampla variedade de bairros, tipos de terrenos e

condições urbanas. As filmagens ocorreram em diferentes horários do dia e sob diversas condições climáticas, formando uma base de dados diversificada. Essas medidas garantem a representatividade e a abrangência do conjunto de dados utilizado para a pesquisa.

A estratégia de capturar 1 frame por segundo de cada vídeo resultou em uma coleção significativa de imagens estáticas, permitindo a análise detalhada de cada trecho de calçada. Esta técnica de extração de frames foi escolhida para balancear a quantidade de dados e a variabilidade das cenas, diminuindo redundâncias enquanto se mantém a representatividade das condições reais das calçadas. A diversificação dos cenários de captura é crucial para treinar um modelo de visão computacional que seja generalizável e eficiente em diferentes contextos urbanos.

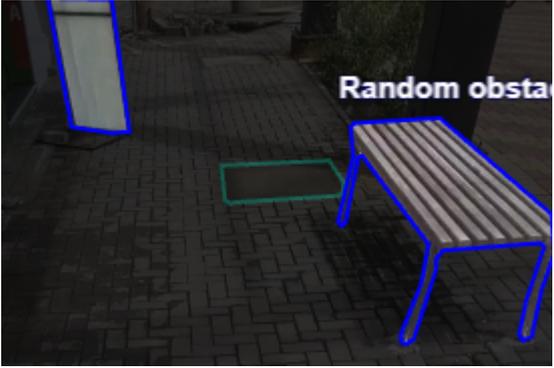
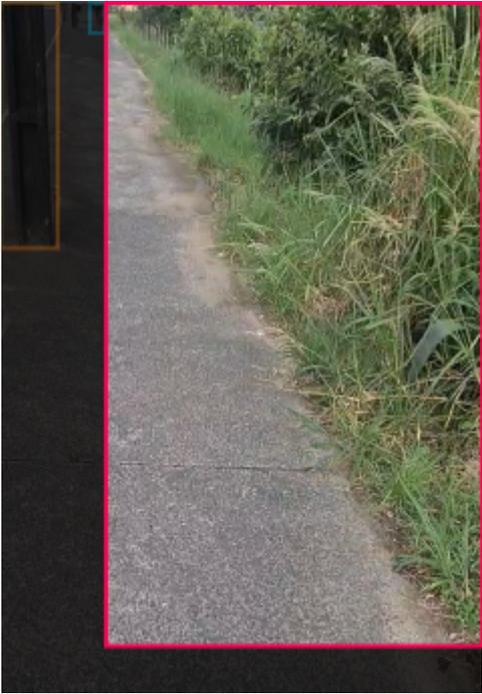
Além disso, a seleção dos pontos de filmagem foi planejada para incluir áreas com diferentes níveis de tráfego pedestre e variações de infraestrutura, como ruas residenciais, comerciais e avenidas principais e pontos em obras. Este planejamento visou garantir que o modelo resultante pudesse identificar e avaliar a qualidade das calçadas de forma precisa, independentemente do ambiente específico. Esta metodologia confere maior confiança nos dados coletados, assegurando que o treinamento do modelo se baseia em um conjunto de dados abrangente e representativo das condições reais encontradas em cidades.

#### **4.2.2 Rotulação e Ferramentas**

A plataforma Roboflow foi utilizada para rotular as imagens extraídas dos vídeos, com os seguintes rótulos discriminados na tabela abaixo. Car, Flat Obstacle, Grass, Gravel, Light Pole, Potential Hole, Random Obstacle, Person, e Tree. Esses rótulos foram escolhidos com base na relevância para a análise da qualidade das calçadas e a segurança dos pedestres.

<b>Classe</b>	<b>Exemplo</b>
---------------	----------------

<p>Potential Hole</p>	
<p>Gravel</p>	
<p>Light Pole</p>	

<p>Tree</p>	 Two vertical images showing tree trunks. The left image shows a tree trunk with a red bounding box. The right image shows a tree trunk with a red bounding box.
<p>Random Obstacle</p>	 A scene with a bench and a sign. The bench is outlined in blue. A sign is outlined in green. The text "Random obstacle" is visible in the top right corner.
<p>Grass</p>	 A path with grass on the right side. A red bounding box is drawn around the grass area.

<p>Car</p>	
<p>Person</p>	

Tabela 1 – Exemplos de rotulações.

O uso do Roboflow para a anotação e rotulação de imagens no projeto demonstrou várias facilidades e algumas limitações. A taxa de anotação manual foi de cerca de 1 classe por minuto em cada frame, a plataforma Roboflow mostrou-se eficiente na identificação automática de certas classes, como "Grass", "Person" e "Car", necessitando pouca intervenção manual. Após a catalogação manual de aproximadamente 15 frames, a ferramenta conseguiu identificar essas classes corretamente, o que reduziu significativamente o tempo e o esforço necessários para o processo de anotação dessas classes.

Para outras classes, no entanto, a necessidade de intervenção manual foi consideravelmente maior, classes como "Flat Obstacle" e "Potential Hole"

demandaram cerca de 100 intervenções manuais cada. Isso pois ocorreram falhas na detecção ou à atribuição incorreta de rótulos, evidenciando uma limitação do modelo na identificação precisa dessas classes.

A necessidade de extensivas intervenções manuais para as classes mais críticas evidencia um ponto de melhoria para a plataforma, sugerindo que modelos pré-treinados ou específicos para certos tipos de obstáculos urbanos poderiam ser desenvolvidos ou aprimorados. Em resumo, enquanto o Roboflow facilitou a anotação de classes mais comuns com alta precisão e baixo esforço manual, ele ainda apresenta desafios na detecção de classes mais complexas e menos representadas, exigindo maior intervenção manual para alcançar resultados satisfatórios.

### 4.2.3 Conjunto de Dados

O dataset contém um total de 995 imagens. A divisão do conjunto de dados seguiu as disposições padrão do Roboflow, Train Set: 71% (709 imagens), Valid Set: 19% (188 imagens) Test Set: 10% (98 imagens). As imagens foram anotadas com um total de 2.927 anotações, resultando em uma média de 2,9 anotações por imagem, distribuídas em nove classes e sua distribuição é explicitada na figura abaixo.

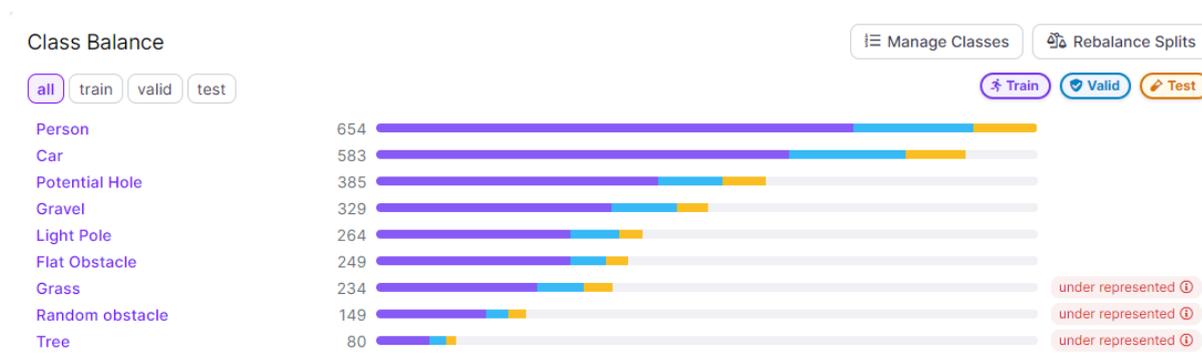


Figura 4 – Distribuição de classificações.

Em ordem de importância para o estudo, a categoria "Potential Hole" é a mais crítica. Esta categoria abrange buracos, fraturas no calçamento, tampões de companhias de água e energia, além de detritos que podem atrapalhar a locomoção de cadeirantes e pessoas com mobilidade reduzida. A identificação desses elementos é fundamental, pois representam riscos imediatos e significativos à segurança dos pedestres, podendo causar quedas ou acidentes graves.

As categorias "Gravel" e "Grass" também são de alta relevância para o estudo. Superfícies de cascalho e gramado representam uma dificuldade extrema ou

até impossibilidade de locomoção para cadeirantes e usuários de andadores. Essas superfícies podem causar instabilidade, dificultar a locomoção e aumentar significativamente o risco de acidentes, afetando diretamente a segurança e a acessibilidade dos grupos de interesse.

"Flat Obstacle" inclui elevações, desníveis ou pisos lisos, como azulejos, que também podem apresentar riscos à mobilidade, especialmente em condições de pouca visibilidade ou para pessoas com dificuldades de locomoção. A identificação desses obstáculos é essencial para prevenir acidentes e garantir a acessibilidade.

A categoria "Random Obstacle" engloba itens como placas móveis, bancos e outros obstáculos irregulares. Embora possam não representar riscos imediatos, sua presença pode obstruir a passagem e dificultar a mobilidade, especialmente em calçadas estreitas. Esses objetos requerem uma análise mais detalhada para avaliar seu impacto na segurança dos pedestres.

"Tree" e "Light Pole" geralmente não representam riscos isoladamente, mas podem obstruir a passagem em calçadas estreitas ou quando estão próximos a outros obstáculos. Sua identificação é importante para garantir que a passagem esteja livre e acessível.

Por fim, os rótulos "Car" e "Person" foram incluídos devido à sua frequente presença nas imagens durante o treinamento. A inclusão desses rótulos foi essencial para melhorar a precisão do modelo, evitando inconsistências nos resultados e garantindo uma análise mais completa do ambiente urbano com enfoque em acessibilidade.

#### **4.2.4 Treinamento e Resultados**

Para avaliar a eficácia do modelo de detecção de qualidade de calçadas, foram treinados três modelos distintos com diferentes abordagens. Para todos os modelos, foram utilizados os hiperparâmetros padrão do Roboflow, com uma taxa de aprendizado inicial ( $lr_0$ ) de 0,01 e uma taxa de aprendizado final ( $lr_f$ ) de 0,01. O primeiro modelo foi treinado sem a aplicação de técnicas de augmentation, utilizando o tipo de modelo YOLO-NAS Object Detection (Accurate) com o checkpoint coco/14.

Este modelo rodou por 294 épocas e apresentou um mAP de 73,2%, precisão de 67,9% e recall de 75,9%. A "class loss" começou em 3.376 na época 0 e

caiu abruptamente para 1.112 já na 10<sup>a</sup> época, mantendo-se abaixo de 0,95 até a última época. Esta rápida queda na "class loss" indica uma aprendizagem inicial eficaz, estabilizando-se posteriormente para garantir uma boa generalização.

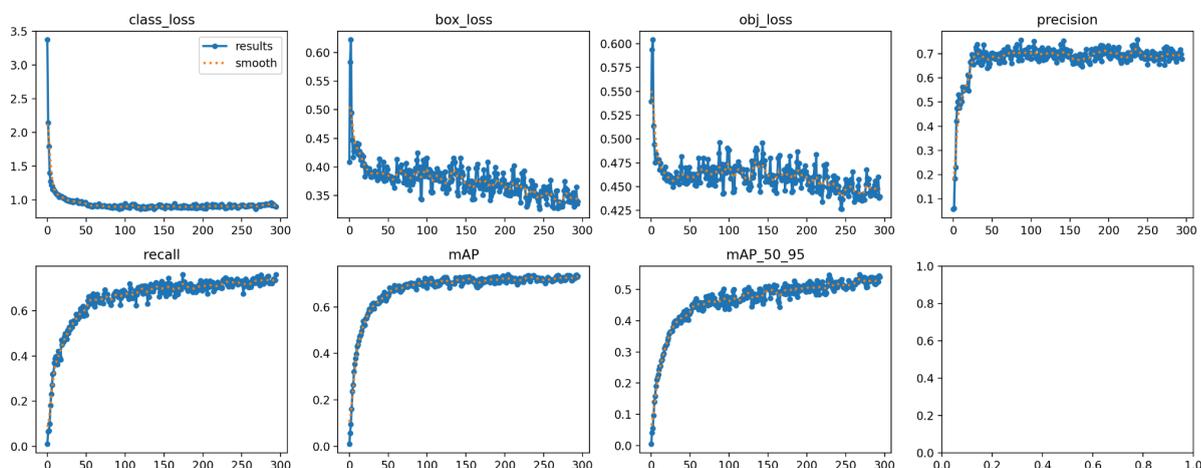


Figura 5 – Gráficos de treinamento do modelo 1.

A análise dos gráficos de treinamento deste modelo fornece uma visão detalhada de seu desempenho. A métrica "class loss" iniciou-se em 3,376 na época 0, apresentando uma queda abrupta para 1,112 já na 10<sup>a</sup> época, e estabilizando-se abaixo de 0,95 até a última época. Isso indica uma aprendizagem inicial rápida e eficaz, seguida por uma estabilização que sugere uma boa capacidade de generalização. O "box loss" também mostrou uma diminuição significativa ao longo das épocas, começando em cerca de 0,60 e estabilizando-se abaixo de 0,40 após 294 épocas. A "obj loss" seguiu um padrão semelhante, caindo de 0,60 para aproximadamente 0,45. Esses padrões indicam uma melhoria contínua na capacidade do modelo de localizar e classificar objetos corretamente.

A precisão do modelo aumentou rapidamente nas primeiras 50 épocas, atingindo cerca de 0,70 e mantendo-se estável a partir de então. O recall também mostrou uma tendência de crescimento constante, alcançando aproximadamente 0,75 ao final do treinamento. As métricas mAP e mAP\_50\_95 indicaram um desempenho robusto, com o mAP estabilizando em torno de 0,73 e o mAP\_50\_95 mostrando um padrão similar de crescimento e estabilização.

Os gráficos de loss de validação (val/box\_loss, val/cls\_loss, val/dfl\_loss) apresentaram uma queda consistente, indicando que o modelo não estava apenas aprendendo a partir dos dados de treinamento, mas também generalizando bem.

para os dados de validação. A precisão e recall nos dados de validação seguiram tendências positivas, corroborando a eficácia do modelo na tarefa proposta.

No entanto, a ausência de técnicas de augmentation pode ter limitado a capacidade do modelo de generalizar para novas condições e variações no ambiente, sugerindo que futuras iterações devem explorar técnicas de augmentation para potencializar ainda mais a robustez e adaptabilidade do modelo.

O segundo modelo utilizou o tipo de modelo Roboflow 3.0 Object Detection (Fast) com o checkpoint COCO. Neste modelo, os dados foram triplicados, resultando em um total de 2413 imagens, com o train set contendo 2127 imagens, o valid set com 188 imagens e o test set com 98 imagens. Diferente do primeiro modelo, que utilizou YOLO-NAS Object Detection (Accurate), este modelo se beneficia de uma arquitetura otimizada para detecção rápida, mas precisa.

Este modelo foi treinado por 260 épocas e apresentou um mAP de 77,0%, precisão de 74,9% e recall de 74,9%. As técnicas de augmentation (rotações entre  $-15^\circ$  e  $+15^\circ$ , *shear*  $\pm 10^\circ$  horizontal e vertical) introduziram variações nas imagens de treinamento, ajudando o modelo a lidar melhor com diferentes ângulos e distorções.



Figura 6 – Imagem exemplificando *Augmentations*, apresentando rotação.

A perda de localização das caixas delimitadoras (train/box\_loss) começou em 1,2 e diminuiu consistentemente ao longo das épocas, atingindo valores abaixo de 0,4, sugerindo que o modelo aprimorou progressivamente sua capacidade de localizar objetos com precisão. A perda de classificação (train/cls\_loss), que iniciou em 3,4, teve uma queda acentuada nas primeiras épocas e estabilizou-se em torno

de 0,5, indicando melhorias significativas na capacidade do modelo em classificar corretamente as classes.

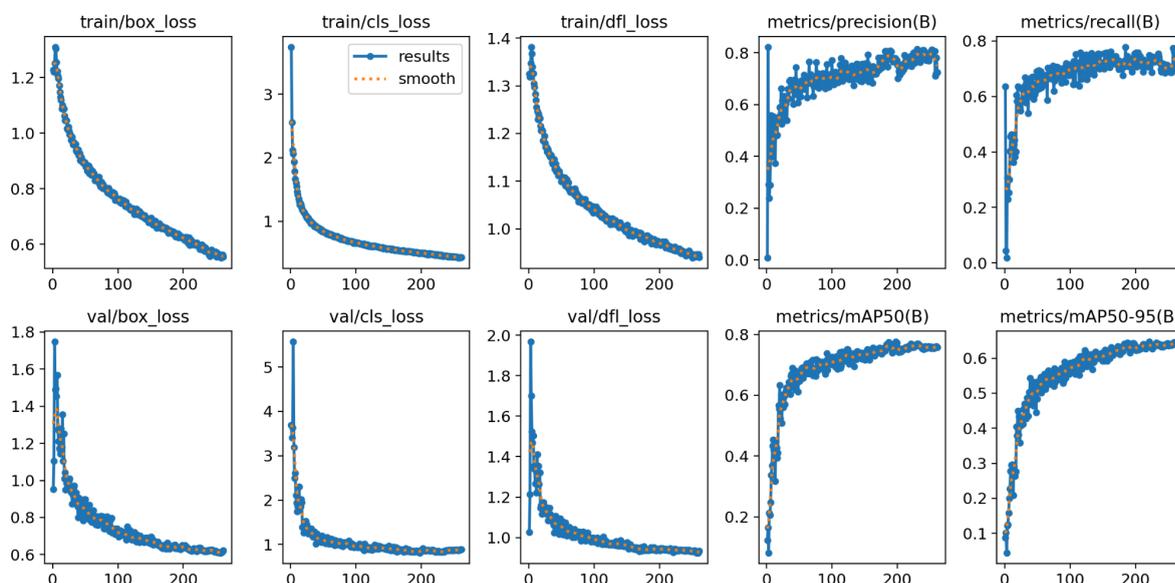


Figura 7 – Gráficos de treinamento do modelo 2.

A precisão do modelo aumentou rapidamente nas primeiras 50 épocas, estabilizando-se em torno de 0,75, refletindo uma alta proporção de verdadeiros positivos em relação aos resultados retornados pelo modelo. O recall também mostrou um aumento gradual, estabilizando-se próximo de 0,75, indicando uma boa cobertura das instâncias positivas.

As métricas de mAP (`metrics/mAP50(B)` e `metrics/mAP50-95(B)`) seguiram uma trajetória ascendente, com o mAP50 atingindo aproximadamente 0,8 e o mAP50-95 estabilizando-se próximo de 0,6, demonstrando a eficácia do modelo em diferentes limiares de IOU.

As perdas de validação (`val/box_loss`, `val/cls_loss`, `val/dfl_loss`) seguiram tendências semelhantes às perdas de treinamento, confirmando que o modelo não estava sofrendo de overfitting. A aplicação de técnicas de augmentation foi benéfica para a robustez e generalização do modelo, mostrando que o uso dessas técnicas melhorou a capacidade do modelo de se adaptar a novas condições e variações no ambiente, reforçando ainda mais a viabilidade da ferramenta para detecção de qualidade de calçadas.

O terceiro modelo foi treinado com o tipo de modelo Roboflow 3.0 Object Detection (Fast), utilizando o checkpoint `sidewalk-quality/2`. Este modelo também aplicou augmentation extra de *Brightness*: entre -25% e +25% porém, assim como

no modelo anterior, o dataset contém um total de 2413 imagens, o train set contendo 2127 imagens, o valid set com 188 imagens e o test set com 98 imagens.



Figura 8 – Imagem exemplificando *Augmentations*, apresentando rotação, *shear* e *brightness*.

Este modelo foi treinado por 41 épocas e apresentou um mAP de 76,8%, precisão de 75,8% e recall de 72,5%. Embora o tempo de treinamento tenha sido menor, este modelo apresentou desempenho satisfatório, destacando a eficiência do treinamento com menos épocas. O menor número de épocas de treinamento sugere que o modelo conseguiu aprender rapidamente as características dos dados, possivelmente devido à utilização de um checkpoint especializado que já tinha sido ajustado em um modelo anterior. Esta eficiência pode ser benéfica em situações onde o tempo de treinamento é limitado ou onde recursos computacionais são restritos.

A análise dos gráficos de treinamento mostra que a perda de localização das caixas delimitadoras (*train/box\_loss*) começou em 0,84 e diminuiu consistentemente ao longo das épocas, atingindo valores abaixo de 0,74. Isso sugere que o modelo

aprimorou sua capacidade de localizar objetos de forma eficiente. A perda de classificação (train/cls\_loss) começou em 2,0 e caiu rapidamente, estabilizando-se abaixo de 1,0, indicando melhorias significativas na capacidade do modelo em classificar corretamente as classes.

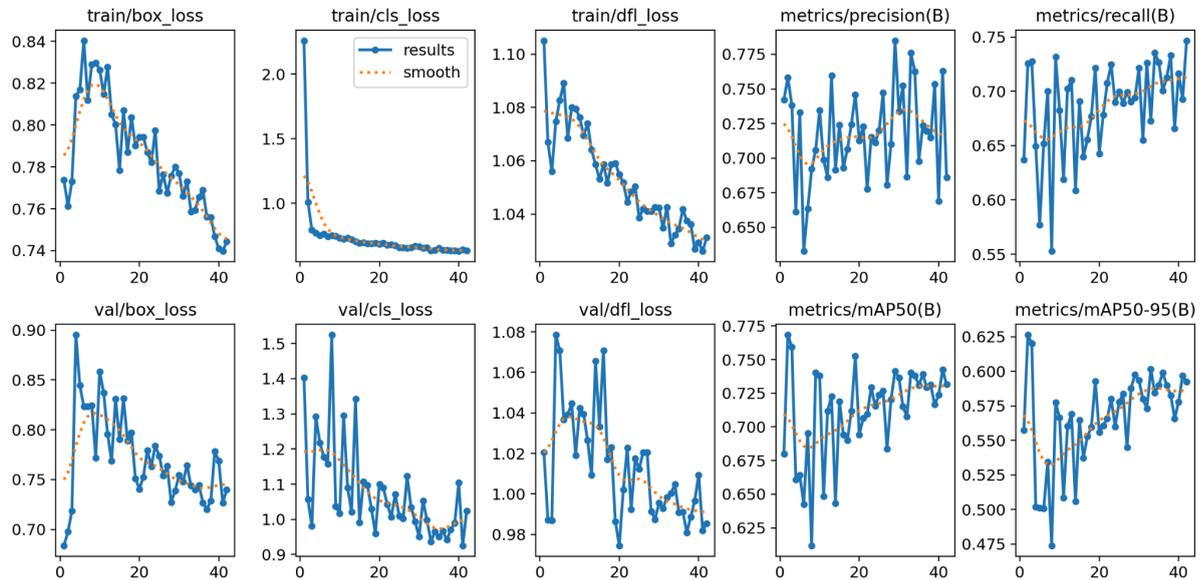


Figura 9 – Gráficos de treinamento do modelo 3.

A precisão do modelo mostrou flutuações nas primeiras épocas, mas estabilizou-se em torno de 0,70, refletindo uma proporção consistente de verdadeiros positivos. O recall também apresentou uma trajetória de crescimento, estabilizando-se próximo de 0,72, indicando uma boa cobertura das instâncias positivas.

As métricas de mAP (metrics/mAP50(B) e metrics/mAP50-95(B)) seguiram uma trajetória ascendente, com o mAP50 atingindo aproximadamente 0,77 e o mAP50-95 estabilizando-se próximo de 0,62, demonstrando a eficácia do modelo em diferentes limiares de IOU .

As perdas de validação (val/box\_loss, val/cls\_loss, val/dfl\_loss) apresentaram uma queda consistente, similar às perdas de treinamento, confirmando que o modelo não estava sofrendo de overfitting. Isso reforça a ideia de que o uso de um checkpoint especializado permitiu um aprendizado mais rápido e eficiente.

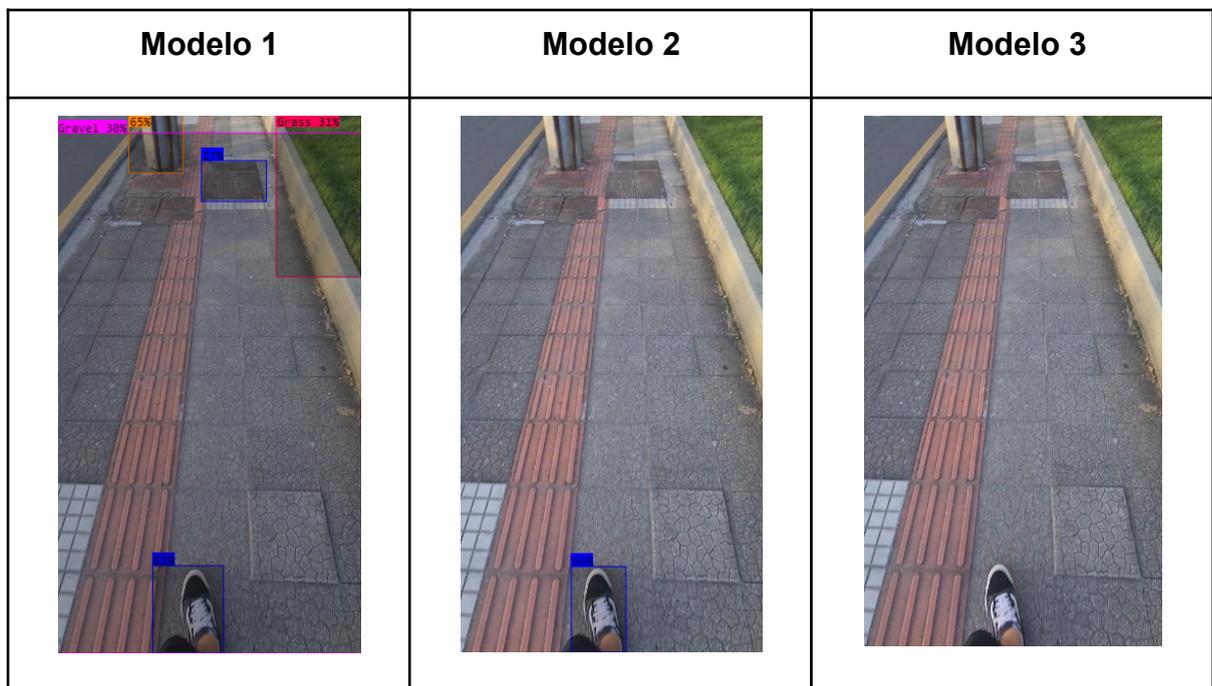
#### 4.2.5 Avaliação de Desempenho

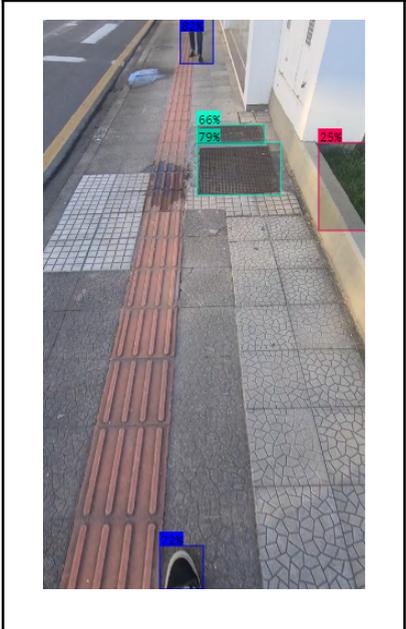
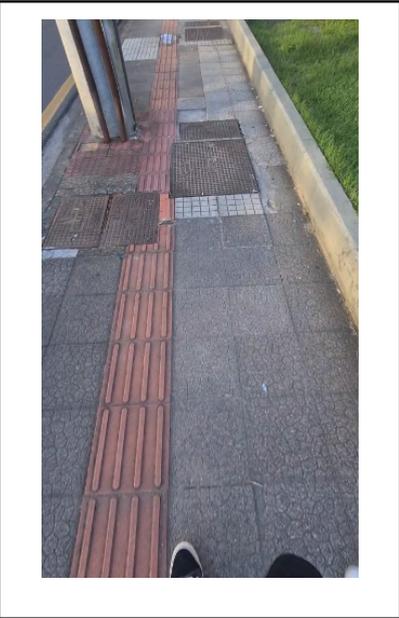
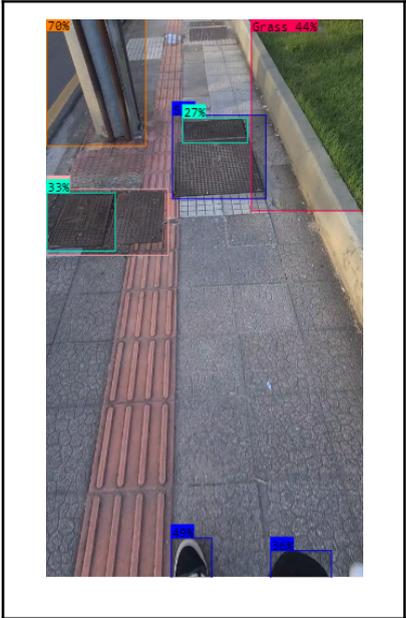
A análise dos resultados foi realizada utilizando a própria plataforma

Roboflow. Para a avaliação, foram utilizados os parâmetros "Confidence Threshold: 10%" e "Overlap Threshold: 50%". A análise foi feita frame a frame, com o "Confidence Threshold" de 10% sendo utilizado para avaliar as detecções menos precisas.

Foram gravados vídeos curtos em diferentes pontos da cidade de Criciúma, no estado de Santa Catarina, extraindo ao menos um frame de cada ponto com condições inexistentes no dataset inicial. Desse novo conjunto de vídeos, foram selecionados 20 frames com condições novas para análise. Nesse conjunto de 20 frames, houve a detecção de todas as classes do estudo, embora de forma indevida em diversos casos.

O conjunto de dados utilizado para treinamento incluiu um total de 995 imagens no modelo inicial, posteriormente triplicadas para 2413 imagens nos modelos subsequentes, com distribuições de treino, validação e teste de 2127, 188 e 98 imagens, respectivamente. As classes detectadas incluíram "Car", "Flat Obstacle", "Grass", "Gravel", "Light Pole", "Potential Hole", "Random Obstacle", "Person" e "Tree".





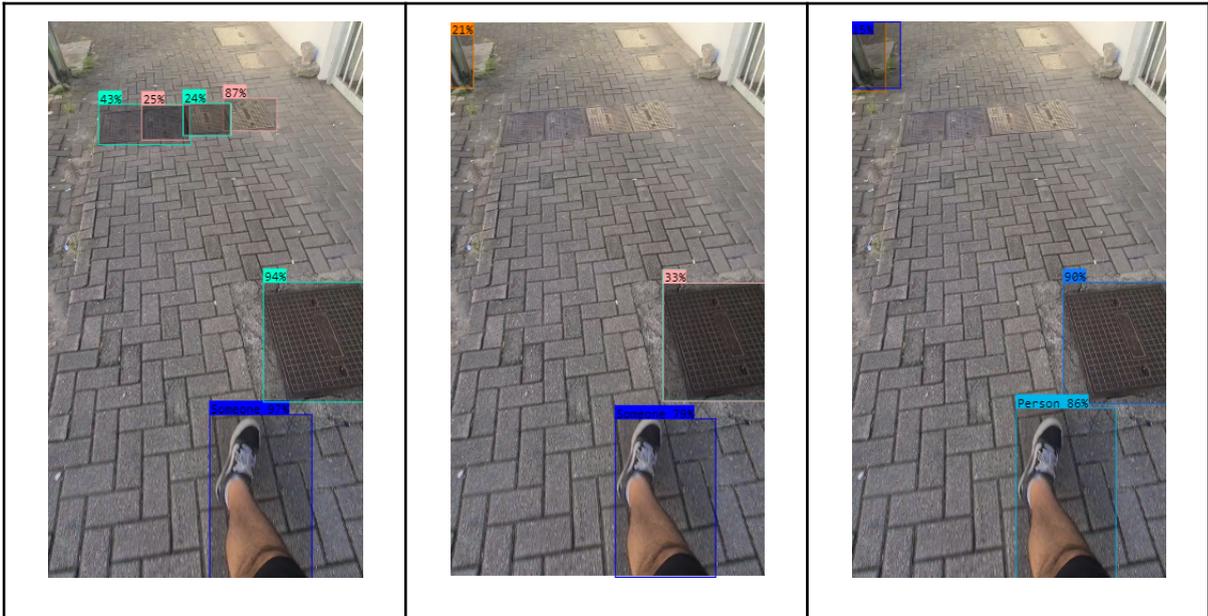


Tabela 2 – Comparação de detecções entre modelos.

No Modelo 1, em 20 frames foram conferidos 92 verdadeiros positivos, 4 falsos positivos e 24 falsos negativos. O cálculo da acurácia para este modelo é  $92 / (92+4+24) = 0,767$  ou 76,7%. Este modelo detectou com muito mais eficiência e precisão todas as classes. Isso pode ser atribuído à arquitetura YOLO-NAS Object Detection (Accurate), que, embora não tenha utilizado técnicas de augmentation, demonstrou uma capacidade de generalização superior. Este modelo foi treinado por mais épocas e conseguiu capturar melhor as variações presentes nos dados de teste, resultando em uma maior taxa de verdadeiros positivos e menor taxa de falsos positivos e negativos em comparação com os outros modelos.

O Modelo 1 se beneficiou de uma arquitetura robusta que, apesar de não usar técnicas de augmentation, mostrou-se eficiente na captura de variações de dados, resultando em uma alta taxa de detecção correta. A ausência de augmentation pode ter simplificado o aprendizado do modelo, focando-o em padrões mais consistentes nos dados, o que pode explicar sua superioridade em termos de precisão e recall.

No Modelo 2, foram detectados 34 verdadeiros positivos, 3 falsos positivos e 83 falsos negativos. O cálculo da acurácia para este modelo é  $34 / (34+3+83) = 0,288$  ou 28,8%. Este modelo utilizou a arquitetura Roboflow 3.0 Object Detection (Fast) com o checkpoint COCO e aplicou técnicas de augmentation como rotações e shear, triplicando o volume de dados. Apesar disso, o modelo não conseguiu alcançar a mesma precisão que o Modelo 1, sugerindo que a complexidade

introduzida pelas técnicas de augmentation pode ter dificultado o treinamento, resultando em uma maior taxa de erros.

O Modelo 2, apesar de utilizar augmentation para aumentar a diversidade dos dados de treinamento, não obteve um desempenho superior. Isso pode indicar que a arquitetura do modelo não conseguiu lidar de forma eficiente com a complexidade adicional dos dados aumentados, ou que os parâmetros de treinamento não foram otimizados para esta configuração, levando a uma menor precisão e maior taxa de falsos negativos.

No Modelo 3, foram detectados 41 verdadeiros positivos, 2 falsos positivos e 77 falsos negativos. O cálculo da acurácia para este modelo é  $41 / (41+2+77) = 0,345$  ou 34,5%. Este modelo também utilizou a arquitetura Roboflow 3.0 Object Detection (Fast) com o checkpoint sidewalk-quality/2 e aplicou técnicas de augmentation adicionais, incluindo ajustes de brilho. Este modelo mostrou uma ligeira melhoria em relação ao Modelo 2, mas ainda ficou atrás do desempenho do Modelo 1.

O Modelo 3, com augmentation extra de brilho, apresentou melhorias modestas em relação ao Modelo 2. A aplicação de ajustes de brilho poderia ter ajudado o modelo a generalizar melhor em condições de iluminação variáveis, mas a arquitetura e os parâmetros de treinamento ainda não foram suficientes para superar o desempenho do Modelo 1. Isso destaca a necessidade de um balanceamento cuidadoso entre a complexidade do modelo, a diversidade dos dados de treinamento e os métodos de augmentation aplicados.

O Modelo 1, com apenas 4 falsos positivos e 24 falsos negativos, demonstra um menor erro mínimo em comparação aos Modelos 2 e 3, que apresentaram mais falsos positivos e falsos negativos cada. Essa análise reforça a importância da escolha correta da arquitetura do modelo e da adequação dos hiperparâmetros e técnicas de treinamento para maximizar a eficácia na detecção de qualidade de calçadas.

	<b>Tipo de Modelo</b>	<b>Checkpoint</b>	<b>Augmentation</b>	<b>mAP</b>	<b>Precision</b>	<b>Recall</b>
<b>Modelo 1</b>	YOLO-NA S Object Detection	coco/14	-	73,2%	67,9%	75,9%

	(Accurate)					
<b>Modelo 2</b>	Roboflow 3.0 Object Detection (Fast)	COCOOn	<b>Rotation:</b> Between $-15^\circ$ and $+15^\circ$ <b>Shear:</b> $\pm 10^\circ$ Horizontal, $\pm 10^\circ$ Vertical	77.0%	74.9%	74.9%
<b>Modelo 3</b>	Roboflow 3.0 Object Detection (Fast)	sidewalk-quality/2 (Modelo 2)	<b>Rotation:</b> Between $-15^\circ$ and $+15^\circ$ <b>Shear:</b> $\pm 10^\circ$ Horizontal, $\pm 10^\circ$ Vertical <b>Brightness:</b> Between $-25\%$ and $+25\%$	76,8%	75.8%	72,5%

Tabela 3 – Especificidades de modelos.

## 5. CONCLUSÃO

O desenvolvimento do modelo de detecção de qualidade de calçadas apresentou resultados promissores, demonstrando a viabilidade do uso de técnicas de visão computacional e geoprocessamento para identificar obstáculos e potenciais riscos para pessoas com mobilidade reduzida em calçamentos. O uso da plataforma Roboflow foi crucial para a organização e rotulação do dataset, bem como para a implementação de técnicas de augmentation que podem vir a contribuir com a precisão e a generalização do modelo, embora não tenham apresentado tal resultado neste caso.

Os três modelos treinados, com e sem augmentation, mostraram desempenhos distintos. O primeiro modelo, que utilizou a arquitetura YOLO-NAS Object Detection (Accurate), apresentou a melhor performance com uma acurácia de 76,7%. Este modelo, embora não tenha utilizado técnicas de augmentation, demonstrou uma capacidade de generalização superior. Foi treinado por mais épocas e conseguiu capturar melhor as variações presentes nos dados de teste, resultando em uma maior taxa de verdadeiros positivos e menor taxa de falsos positivos e negativos.

O segundo modelo, utilizando a arquitetura Roboflow 3.0 Object Detection (Fast) com augmentation e o checkpoint COCO, apresentou uma acurácia de 28,8%. Apesar de aplicar técnicas de augmentation e triplicar o volume de dados, o desempenho foi inferior ao do primeiro modelo. Isso sugere que a complexidade introduzida pelas técnicas de augmentation pode ter dificultado o treinamento, resultando em uma maior taxa de erros.

O terceiro modelo, que também utilizou a arquitetura Roboflow 3.0 Object Detection (Fast) com augmentation adicional de brilho e o checkpoint sidewalk-quality/2, apresentou uma acurácia de 34,5%. Embora tenha mostrado uma ligeira melhoria em relação ao segundo modelo, ainda ficou atrás do desempenho do primeiro modelo. A aplicação de ajustes de brilho ajudou o modelo a generalizar melhor em condições de iluminação variáveis, mas a arquitetura e os parâmetros de treinamento ainda não foram suficientes para superar o desempenho do primeiro modelo.

A avaliação de desempenho em cenários diferentes do treinamento original indicou que, embora o primeiro modelo tenha apresentado um desempenho satisfatório na maioria das detecções, ainda há espaço significativo para melhorias. A não detecção de categorias críticas como "Potential Hole" e "Gravel" reforça a necessidade de um dataset mais abrangente e diversificado.

Para aprimorar o modelo e expandir sua aplicabilidade, são recomendadas várias frentes de trabalho futuro. Primeiramente, a expansão do dataset é crucial. Coletar mais dados de diferentes regiões, incluindo variados tipos de calçamento, condições climáticas e horários do dia, e aumentar a quantidade de imagens para classes sub-representadas como "Grass", "Random Obstacle" e "Tree" pode melhorar a precisão nessas categorias.

Além disso, o refinamento do modelo é necessário. Implementar e testar outras arquiteturas de modelos de detecção de objetos, além de YOLO e Roboflow 3.0, e experimentar diferentes técnicas de augmentation e ajuste de hiperparâmetros podem trazer ganhos adicionais em desempenho.

A integração do modelo em sistemas de análise em tempo real também é uma área promissora. Utilizar dispositivos móveis ou câmeras instaladas em veículos para avaliar a eficácia do modelo em situações dinâmicas e adaptar o treinamento conforme necessário pode aumentar a aplicabilidade prática do modelo.

Outra recomendação é desenvolver um sistema de feedback onde usuários possam reportar falhas ou detecções incorretas. Isso pode ajudar a refinar o modelo continuamente e adaptá-lo a novos padrões de dados. Além disso, explorar a integração do modelo com sistemas de navegação assistida para pessoas com deficiência visual ou mobilidade reduzida pode fornecer alertas em tempo real sobre potenciais obstáculos e condições adversas do calçamento.

Por fim, realizar estudos comparativos com outros modelos e métodos de detecção de qualidade de calçadas pode identificar pontos fortes e fracos relativos e promover uma melhoria contínua.

Os trabalhos relacionados revisados também destacaram a relevância do uso de tecnologias de IA e visão computacional em diversas áreas. Estudos sobre detecção de objetos e análise de condições urbanas corroboram os achados deste trabalho, mostrando que essas tecnologias são eficazes para aplicações em contextos urbanos e de mobilidade. A revisão da literatura ajudou a identificar

técnicas e abordagens que foram aplicadas com sucesso em outros contextos, reforçando a importância de um dataset diversificado e de técnicas de treinamento avançadas para melhorar a acurácia e a robustez dos modelos.

Além disso, a incorporação de técnicas de geoprocessamento foi essencial para mapear e analisar dados geográficos, contribuindo para uma melhor compreensão das condições urbanas e permitindo uma aplicação mais precisa das técnicas de visão computacional. O uso de geoprocessamento em cidades inteligentes destaca a importância de dados geoespaciais para a gestão urbana eficiente, melhorando a qualidade de vida dos cidadãos.

O trabalho realizado fornece uma base sólida para o desenvolvimento de soluções práticas e inovadoras para a melhoria da segurança e acessibilidade em ambientes urbanos. A continuidade e expansão deste trabalho poderão contribuir significativamente para a qualidade de vida das pessoas, especialmente aquelas com mobilidade reduzida ou outras necessidades específicas de acessibilidade.

## REFERÊNCIAS

- BARCZYSZYN, Guilherme Lincon. Sistema colaborativo para planejamento de rotas para cadeirantes. 2019. 71 f. Dissertação (Mestrado em Computação Aplicada) - Universidade Tecnológica Federal do Paraná, Curitiba, 2019.
- BENEVOLO, C.; DAI PRIMA, E.; VELASCO, M. A. Smart mobility in smart city: Action taxonomy, ICT intensity and public benefits. In: *TERRITORIAL SCIENCE*. Springer, Cham, 2016. p. 13-43.
- CARAGLIU, A.; DEL BO, C.; NIJKAMP, P. Smart cities in Europe. *Journal of Urban Technology*, v. 18, n. 2, p. 65-82, 2011.
- CHEN, J.; LIU, X.; WU, Q. Improving urban safety through AI and image analysis in smart cities. *International Journal of Information Management*, v. 45, p. 43-55, 2019.
- DALAL, Navneet; TRIGGS, Bill. Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE, 2005. p. 886-893.
- GIRSHICK, R.; DONAHUE, J.; DARRELL, T.; MALIK, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, 2014, p. 580-587.
- GOYAL, S.; MURTHY, S.; RADHAKRISHNAN, S. Smart grids: Current and future perspectives. *Journal of Energy Management*, v. 1, n. 1, p. 15-24, 2017.
- HOLLANDS, R. G. Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City*, v. 12, n. 3, p. 303-320, 2008.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. In: *PROCEEDINGS OF THE 25th INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS*, 2012, p. 1097-1105.

KUSSUL, Natalia; LAVRECHENKO, Olga; SHEVCHENKO, Ivan. Uso de imagens de satélite e técnicas de aprendizado de máquina para monitoramento de cultivos agrícolas. *Revista de Agricultura de Precisão*, v. 22, n. 3, p. 210-223, 2017.

LI, Qingquan; TORRES, Myriam. Real-time traffic monitoring and prediction using machine learning. In: *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2014. p. 2520-2525.

LIU, Wei; ANGUELOV, Dragomir; ERHAN, Dumitru; SZEGEDY, Christian; REED, Scott; FU, Cheng-Yang; BERG, Alexander C. SSD: Single shot multibox detector. In: *PROCEEDINGS OF THE EUROPEAN CONFERENCE ON COMPUTER VISION*, 2016, p. 21-37.

OLIVEIRA, Débora; CUNHA, Ulysses; NASCIMENTO, Ricardo. Aplicação de técnicas de aprendizado de máquina na detecção de padrões urbanos. In: *XXII Congresso Brasileiro de Automática (CBA)*. 2020. p. 139-144.

OLIVEIRA, S.; SANTOS, M.; ALMEIDA, F. Machine learning for landslide susceptibility mapping: Evaluating models for improved spatial prediction. *Geoscience Frontiers*, 2020.

REDMON, Joseph; DIVVALA, Santosh; GIRSHICK, Ross; FARHADI, Ali. You Only Look Once: Unified, Real-Time Object Detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. p. 779-788.

REN, Shaoqing; HE, Kaiming; GIRSHICK, Ross; SUN, Jian. Faster R-CNN: Towards real-time object detection with region proposal networks. In: *Advances in neural information processing systems*. 2015. p. 91-99.

SILVA, L.; PEREIRA, J.; COSTA, R. Predictive modeling of public transportation demand using geographic and socioeconomic data. *Transportation Research Part C: Emerging Technologies*, 2019.

SILVA, Thiago; CAMPOS, João; FERREIRA, José. Modelagem preditiva da demanda por serviços de transporte público usando dados geográficos e socioeconômicos. In: *XV Congresso de Pesquisa e Ensino em Transportes (COPET)*. 2019. p. 112-119.

YIN, Li; WANG, Zhenxin. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Environment and Planning B: Urban Analytics and City Science*, v. 43, n. 6, p. 921-939, 2016.

ZHANG, Congcong; ZHOU, Tao; SHI, Yusheng; LI, Zhiyi; ZHU, Hongzi. Detecting urban anomalies using multiple spatio-temporal data sources. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 2017. p. 1963-1966.

ZHANG, Junying; MA, Jin; LIU, Ping; ZHANG, Xusheng. CrackNet: A deep learning-based pavement crack detection method using convolutional neural network. In: *IEEE Transactions on Intelligent Transportation Systems*. 2016. v. 18, n. 4, p. 950-964.

ZHANG, Y.; XU, Y.; ZHANG, W.; YANG, J.; SUN, Y.; ZHANG, S. Object detection and classification in remote sensing imagery using convolutional neural networks. *Remote Sensing*, 2017.

ZHENG, Yu; FU, Haishan; XIAO, Guannan; FENG, Rui; XU, Zhengyu. Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, v. 5, n. 3, p. 38, 2014.

ZANDEGHEH, R.; KHAN, S. Privacy and security in smart cities. *Journal of Information Security and Applications*, v. 40, p. 1-7, 2018.

## APÊNDICE A - ARTIGO NO FORMATO SBC

# **Avaliação de viabilidade da utilização de ferramentas de inteligência artificial e visão computacional para a identificação de obstáculos e avaliação das condições de acessibilidade em calçadas para pessoas com mobilidade reduzida.**

**Ali Castilho Santos**

Departamento de Informática e Estatística, Universidade Federal de Santa Catarina,  
Florianópolis / SC, Brasil

[ali.cast.santos@gmail.com](mailto:ali.cast.santos@gmail.com)

***Abstract.** This study aims to analyze the feasibility of using computer vision and geoprocessing technologies to improve accessibility for people with reduced mobility. Previous works were reviewed to select relevant technologies and apply them in this study. Data collection was conducted through videos recorded under various urban conditions. The Roboflow platform was used for labeling and model training, followed by performance evaluation in different scenarios. Results demonstrated the effectiveness of this methodology, suggesting that more accurate results could be achieved by expanding the dataset, refining the models, or integrating with real-time systems to promote more accessible and inclusive cities.*

***Resumo.** A proposta deste trabalho é analisar a viabilidade da utilização de tecnologias de visão computacional e geoprocessamento para melhorar a acessibilidade de pessoas com mobilidade reduzida. Foram revisados trabalhos anteriores para selecionar tecnologias relevantes e aplicá-las neste estudo. A coleta de dados foi realizada através de vídeos gravados em diferentes condições urbanas. Utilizou-se a plataforma Roboflow para rotulação e treinamento dos modelos, seguida da avaliação de desempenho em diferentes cenários. Os resultados demonstraram a eficácia do uso desta metodologia, indicando que há possibilidade de atingir resultados ainda mais assertivos ao expandir o dataset, refinar os modelos ou integrar com sistemas em tempo real, promovendo cidades mais acessíveis e inclusivas.*

## **1. Introdução**

A Universidade Federal de Santa Catarina abriga a Coordenadoria de Acessibilidade Educacional (CAE), vinculada à Secretaria de Ações Afirmativas e Diversidades (SAAD). Contudo, o acesso à CAE no campus Florianópolis é desafiador para pessoas com mobilidade reduzida devido às condições precárias das vias de pedestres, representando riscos evidentes para esses usuários.

Melhorar a acessibilidade e segurança na universidade e na cidade requer soluções inovadoras e eficientes. As abordagens incluem melhorias físicas (rampas, calçadas niveladas, sinalização adequada), campanhas de conscientização e políticas públicas voltadas para a construção e adaptação de espaços acessíveis.

Ferramentas de inteligência artificial e visão computacional são promissoras nesse contexto. Essas tecnologias podem extrair dados de fontes como vídeos gravados por celulares e apresentá-los aos responsáveis, analisando grandes volumes de dados para identificar padrões e áreas problemáticas. A visão computacional pode mapear obstáculos detalhadamente, gerando relatórios compreensíveis para os tomadores de decisão e promovendo políticas públicas efetivas.

Este trabalho de conclusão de curso visa avaliar a viabilidade da aplicação de tecnologias de visão computacional e inteligência artificial para extrair dados sobre condições de acessibilidade urbana. Serão exploradas técnicas de processamento de imagens e detecção de objetos, com coleta de dados, desenvolvimento e avaliação de modelos práticos. A análise demonstrará o potencial dessas tecnologias para fomentar o diálogo entre poder público, comunidade acadêmica e sociedade, visando cidades mais acessíveis e inclusivas.

Por meio de vídeos gravados em diversas condições urbanas, busca-se mapear obstáculos que impactam a mobilidade de pessoas com deficiência ou mobilidade reduzida. A tecnologia permitirá a geração de dados precisos, contribuindo para a melhoria da infraestrutura urbana e promovendo um ambiente mais seguro e acessível para todos.

Os resultados serão apresentados em uma comparação de análises de modelos de identificação de objetos, fornecendo indícios promissores do uso das tecnologias. Espera-se que os resultados sejam compartilhados com a CAE e a SAAD, promovendo diálogo e cooperação para soluções inclusivas e efetivas.

Ao avaliar a viabilidade e eficácia dessas tecnologias, espera-se melhorar a acessibilidade e mobilidade, tanto na universidade quanto em outros contextos urbanos. Os resultados poderão servir como referência para futuros estudos e incentivar a adoção de soluções baseadas em inteligência artificial e visão computacional em outras cidades com desafios semelhantes.

A proposta deste trabalho é ampliar o escopo de atuação para além do ambiente universitário, buscando soluções tecnológicas para criar cidades mais acessíveis, inclusivas e seguras. A implementação dessas soluções pode fomentar a colaboração entre academia, poder público e sociedade civil, promovendo inovação contínua e aplicação prática das tecnologias, melhorando a infraestrutura urbana e conscientizando sobre a importância da acessibilidade.

## **2. Fundamentação teórica**

A análise de diversos trabalhos recentes demonstra avanços significativos em visão computacional, especialmente após o trabalho de Krizhevsky, Sutskever e Hinton (2012) com a AlexNet, que revolucionou o reconhecimento de imagens com redes neurais convolucionais profundas. Estudos como os de Girshick et al. (2014) com R-CNN e Ren et al. (2015) com Faster R-CNN, bem como as abordagens YOLO (Redmon et al., 2016) e SSD (Liu et al., 2016), melhoraram a precisão e velocidade na detecção de objetos. No contexto de imagens

urbanas, Zhang et al. (2017) utilizaram essas técnicas para detectar semáforos em tempo real, contribuindo para sistemas de condução autônoma mais seguros e eficientes.

No campo do geoprocessamento, a integração de inteligência artificial e visão computacional tem permitido automação e melhorias na precisão de diversas tarefas. Zhang et al. (2017) demonstraram a eficácia de redes neurais convolucionais para segmentação e classificação de objetos em imagens de satélite, identificando automaticamente características geográficas como edifícios e estradas. Oliveira et al. (2020) utilizaram aprendizado de máquina para analisar a relação entre dados geográficos e a ocorrência de deslizamentos de terra, mostrando que a IA pode identificar padrões ocultos nos dados e contribuir para a prevenção de desastres naturais. Silva et al. (2019) desenvolveram modelos preditivos para demanda de transporte público, utilizando dados geográficos e socioeconômicos, melhorando o planejamento urbano.

No âmbito das cidades inteligentes, o uso de tecnologias avançadas para otimizar serviços urbanos tem se destacado. Sensores e câmeras monitoram o fluxo de tráfego em tempo real, permitindo ajustes dinâmicos para reduzir congestionamentos e melhorar a mobilidade urbana. Redes elétricas inteligentes gerenciam a distribuição de energia eficientemente, integrando fontes renováveis e reduzindo perdas. Na segurança pública, câmeras com algoritmos de IA detectam atividades suspeitas e coordenam respostas a emergências de forma rápida e eficiente. Contudo, desafios como a gestão de grandes volumes de dados e questões de privacidade e segurança ainda persistem. A colaboração entre governos, empresas e a academia é crucial para superar esses desafios e maximizar os benefícios das cidades inteligentes, promovendo um ambiente urbano mais conectado, seguro e sustentável.

Esses exemplos ilustram como a visão computacional, o geoprocessamento e as cidades inteligentes se beneficiam das técnicas de IA, oferecendo novas possibilidades para a análise urbana e melhoria da acessibilidade. Com o contínuo avanço dessas tecnologias, espera-se que áreas como a detecção automática de objetos, a previsão de desastres naturais e a otimização de serviços urbanos se tornem cada vez mais precisas e eficientes, contribuindo para a criação de ambientes mais inclusivos e sustentáveis.

### **3. Trabalhos relacionados**

Durante o processo de pesquisa, foram investigadas a viabilidade da utilização de ferramentas de inteligência artificial e visão computacional para identificar obstáculos e avaliar as condições de acessibilidade em calçadas para pessoas com mobilidade reduzida. A revisão da literatura abrangeu conceitos, técnicas, abordagens e possíveis tecnologias aplicáveis ao processamento de imagens e análise de dados. Estudos mostram que o uso de geoprocessamento e aplicações de visão computacional com bases de dados existentes está se tornando cada vez mais acessível e usual. Esses estudos apontam novas possibilidades e eliminam barreiras na área deste estudo.

A pesquisa, realizada em setembro de 2022, foi conduzida exclusivamente no Google Scholar devido à sua ampla cobertura de artigos acadêmicos. As palavras-chave utilizadas incluíram: acessibilidade, pessoas com deficiência, YOLO, processamento de imagens, detecção de objetos, aprendizado de máquina e detecção de buracos. Foram considerados apenas artigos que tratam da identificação de obstáculos e análise de condições de calçadas

utilizando inteligência artificial e visão computacional. Estudos que não se enquadram nesses critérios foram excluídos, mantendo apenas publicações dos últimos dez anos em português e inglês.

A avaliação dessas tecnologias visa fornecer dados para possíveis melhorias na acessibilidade e mobilidade de pessoas com deficiência em contextos urbanos. Os resultados podem servir como referência para futuros estudos e incentivar a adoção dessas soluções em outras cidades.

### **3.1 Sistema Colaborativo para Planejamento de Rotas para Cadeirantes**

BARCYZSZYN, Guilherme Lincon. Sistema colaborativo para planejamento de rotas para cadeirantes. 2019. Dissertação (Mestrado em Computação Aplicada) - Universidade Tecnológica Federal do Paraná, Curitiba, 2019.

Este trabalho aborda o desafio do planejamento de rotas na computação urbana, requerendo processamento de grandes volumes de dados e feedback dos usuários. Os serviços tradicionais de planejamento de rotas, baseados em ruas, não são adequados para usuários com deficiências. Para solucionar esse problema, é apresentado um modelo baseado em calçadas, atendendo às necessidades de cadeirantes por meio de um serviço de planejamento de rotas. O modelo é definido como um grafo, com vértices representando esquinas e arestas representando calçadas ou cruzamentos, com custos determinados por fatores como distância, inclinação e condições das rampas e calçadas.

O autor desenvolveu um algoritmo de planejamento de rotas que gera rotas alternativas com base em critérios de acessibilidade e um algoritmo de recomendação de atalhos para auxiliar no planejamento urbano. Uma API foi disponibilizada para fornecer serviços de caminho mínimo em calçadas. O trabalho foi conduzido em etapas: revisão bibliográfica, exploração de trabalhos relacionados, elaboração de modelo teórico, implementação do modelo com dados geoespaciais de Curitiba, fornecimento de serviço de planejamento de rotas e validação do modelo por meio de testes.

Os dados utilizados foram adquiridos de sensoriamento urbano, feedback de cadeirantes e autoridades de planejamento urbano. O sistema de banco de dados foi construído utilizando PostgreSQL e PostGIS para manipulação de dados geográficos. O modelo foi implementado com base na arquitetura proposta por Zheng et al. (2014), integrando diversas camadas de dados e formando um banco de dados geográfico específico para o planejamento de rotas.

Os principais desafios incluíram a manipulação de funções corretas para transformação de dados e a aplicação dessas transformações em grande escala para toda a cidade de Curitiba. O objetivo foi proporcionar maior comodidade e qualidade de vida aos cadeirantes, tornando seu deslocamento menos custoso. Os objetivos específicos alcançados incluíram a revisão bibliográfica, pesquisa exploratória, elaboração e implementação de modelo teórico, disponibilização de uma API e validação do modelo.

As dificuldades principais foram a manipulação das funções corretas para transformação de dados, especialmente para a geração correta de arestas de cruzamentos, e a aplicação dessas transformações em grande escala.

### **3.2 Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery**

O estudo de YIN, Li e WANG, Zhenxin (2016) aborda os desafios do design urbano e da mobilidade, destacando a importância de dados consistentes e objetivos. Com os avanços em sensoriamento e digitalização de dados, é possível aplicar métodos de processamento de dados em diversas fontes, automatizando a obtenção de resultados com alta precisão e economia de tempo. Utilizando imagens do Google Street View e técnicas de aprendizado de máquina, o estudo explora a identificação, classificação e categorização de vias urbanas e o impacto dos objetos e construções presentes.

Realizado na cidade de Buffalo, EUA, o estudo utilizou 3592 imagens panorâmicas de 311 blocos para aplicar técnicas de rede neural artificial (ANN) e Máquina de Vetores de Suporte (SVM), analisando textura e cor, além de detecção e segmentação combinadas com classificação. As tecnologias de IA e visão computacional mostraram potencial para mapear e analisar as condições das vias para pedestres, identificando problemas e avaliando a qualidade das vias para grupos como Pessoas com Mobilidade Reduzida (PMR).

O estudo ressalta a importância de disponibilizar dados coletados de maneira acessível e compreensível, criando relatórios detalhados e apresentando resultados visuais para o público e especialistas. A aplicação dessas tecnologias deve considerar questões de privacidade e segurança, além de promover colaboração entre instituições acadêmicas, órgãos governamentais e comunidades locais para melhorar a acessibilidade urbana.

### **3.3 A Deep Learning Approach for Street Pothole Detection**

O artigo aborda o problema dos buracos nas vias, que causam danos aos veículos e acidentes. Utilizando aprendizado profundo (Deep Learning), foram treinados quatro modelos com um conjunto de dados pré-processado: YOLO V3, SSD, HOG com SVM e Faster R-CNN. O estudo destaca o interesse na detecção rápida e eficiente de buracos, explicando a aplicação inicial de SVM e a evolução para técnicas mais avançadas de CNN, ignorando ruídos causados por iluminação e sombras.

O esquema de detecção é dividido em preparação dos dados e detecção, classificando imagens de vias com e sem buracos. Modelos de aprendizado profundo foram treinados e testados com conjuntos de dados de diferentes tamanhos, comparando desempenho e tempo de processamento. O YOLO V3 destacou-se em grandes volumes de dados, enquanto Faster R-CNN teve melhor desempenho com volumes menores. A análise mostrou que, apesar de todos os algoritmos apresentarem resultados semelhantes, HOG teve desempenho inferior e foi desconsiderado em análises posteriores.

Estes estudos demonstram o potencial das técnicas de IA e visão computacional para melhorar a mobilidade urbana, fornecendo dados relevantes para políticas públicas e intervenções necessárias, além de incentivar pesquisas adicionais e adoção de soluções tecnológicas em diversas cidades.

### **3.4 Discussão**

A revisão da literatura demonstra que existem pesquisas recentes sobre a aplicação de inteligência artificial e visão computacional para a identificação de obstáculos em calçadas. No entanto, diferem na coleta de dados, contando com imagens de terceiros, feedback de

cidadãos ou capturas por carros. Este trabalho difere ao utilizar vídeos gravados pessoalmente com smartphones, focando em imagens frontais das calçadas. Não foram encontrados trabalhos anteriores que utilizassem essa abordagem específica até o momento.

Em relação aos modelos utilizados nas pesquisas, há uma predominância do uso de redes neurais convolucionais, especialmente o YOLO (You Only Look Once), para a detecção de objetos. Estudos indicam que modelos baseados em YOLO apresentam bons desempenhos em seus contextos específicos.

Na data de pesquisa dos trabalhos relacionados não foram identificadas pesquisas brasileiras que utilizem detecção de objetos para avaliar a acessibilidade urbana, o que evidencia uma lacuna de conhecimento. Este trabalho busca preencher essa lacuna, contribuindo para a pesquisa nacional na área de acessibilidade.

A seleção de trabalhos relacionados foi resultado de revisão criteriosa e discussões entre autor e orientador, para assegurar que apenas artigos pertinentes fossem incluídos. Essas medidas garantem que a revisão seja abrangente e confiável, fornecendo uma base sólida para o desenvolvimento do trabalho e contribuindo significativamente para o avanço do conhecimento na área de acessibilidade e mobilidade urbana.

## **4. Desenvolvimento**

A utilização de inteligência artificial e visão computacional exige uma abordagem detalhada e estruturada. Este capítulo descreve o processo de seleção das ferramentas, a metodologia adotada para criação e treinamento dos modelos, e as escolhas tecnológicas que fundamentaram o desenvolvimento do projeto. O objetivo é garantir precisão, eficiência e transparência em cada etapa do projeto.

### **4.1 Seleção de Ferramentas**

A seleção das ferramentas foi um processo colaborativo entre o autor e o orientador, focando em plataformas que permitissem um desenvolvimento completo e eficiente, mesmo em regime de avaliação gratuita. As ferramentas escolhidas deveriam cobrir todas as etapas necessárias: montagem do dataset, categorização, treinamento de modelos, obtenção de relatórios e aplicação dos modelos treinados em novos cenários.

A busca inicial por "label YOLO" no Google resultou no Roboflow como o primeiro resultado. Após testar a plataforma, concluiu-se que ela atendia a todos os pré-requisitos, oferecendo 3 créditos gratuitos para o treinamento de modelos. A versatilidade do Roboflow e sua capacidade de lidar com todas as etapas do processo de desenvolvimento foram determinantes para sua escolha.

#### **4.1.1 Roboflow**

Roboflow é uma plataforma criada para facilitar o desenvolvimento e implantação de modelos de visão computacional. Fundada em 2019 por Joseph Nelson e Brad Dwyer, a empresa visa democratizar a criação de aplicações de visão computacional, tornando-as acessíveis para desenvolvedores de todos os níveis.

A plataforma permite importar, organizar, classificar e treinar modelos de visão

computacional de maneira intuitiva. Além disso, oferece ferramentas para a implantação e monitoramento de modelos, proporcionando um fluxo de trabalho completo desde a criação do dataset até a aplicação prática dos modelos treinados.

Roboflow ganhou destaque rapidamente no campo da visão computacional, sendo utilizada em contextos acadêmicos e comerciais. Suas aplicações abrangem diversas áreas, como saúde, agricultura, manufatura e segurança, demonstrando a flexibilidade e robustez das suas soluções. A plataforma é amplamente reconhecida na comunidade de desenvolvedores e pesquisadores pela sua eficiência e facilidade de uso. Com uma forte comunidade de usuários e um compromisso contínuo com a inovação, Roboflow se posiciona como uma ferramenta essencial para o desenvolvimento de modelos de visão computacional de forma eficiente e eficaz.

## **4.2 Elaboração do Conjunto de Dados**

### **4.2.1 Aquisição de Dados**

Para criar o conjunto de dados, o autor filmou 11 vídeos curtos, variando de 10 a 400 segundos, em diferentes pontos da cidade de Forquilha, Santa Catarina. As filmagens ocorreram em diversos bairros, tipos de terrenos e condições urbanas, capturando diferentes horários e condições climáticas para formar uma base de dados diversificada e representativa. A captura de 1 frame por segundo de cada vídeo resultou em uma coleção significativa de imagens estáticas, permitindo uma análise detalhada das calçadas. A seleção dos pontos de filmagem incluiu áreas com diferentes níveis de tráfego pedestre e variações de infraestrutura, como ruas residenciais, comerciais, avenidas principais e pontos em obras, garantindo que o modelo resultante fosse eficiente em diversos contextos urbanos.

### **4.2.2 Rotulação e Ferramentas**

A plataforma Roboflow foi utilizada para rotular as imagens extraídas dos vídeos, com rótulos como Car, Flat Obstacle, Grass, Gravel, Light Pole, Potential Hole, Random Obstacle, Person e Tree, escolhidos pela sua relevância para a análise da qualidade das calçadas e segurança dos pedestres. A plataforma mostrou-se eficiente na identificação automática de classes comuns como "Grass", "Person" e "Car", reduzindo o esforço manual após cerca de 15 frames catalogados manualmente. No entanto, classes como "Flat Obstacle" e "Potential Hole" exigiram cerca de 100 intervenções manuais cada, devido a falhas na detecção ou atribuição incorreta de rótulos. Esta necessidade de intervenção manual evidencia uma limitação do modelo na identificação precisa dessas classes, sugerindo melhorias para a plataforma com modelos pré-treinados ou específicos para obstáculos urbanos.

### **4.2.3 Conjunto de Dados**

O dataset contém 995 imagens, divididas em Train Set (71% - 709 imagens), Valid Set (19% - 188 imagens) e Test Set (10% - 98 imagens). As imagens foram anotadas com um total de 2.927 anotações, resultando em uma média de 2,9 anotações por imagem, distribuídas em nove classes.

A categoria "Potential Hole" é a mais crítica, englobando buracos, fraturas no calçamento, tampões de companhias de água e energia, e detritos que podem atrapalhar a

locomoção de cadeirantes e pessoas com mobilidade reduzida. As categorias "Gravel" e "Grass" também são de alta relevância, pois representam dificuldades extremas ou até impossibilidades de locomoção para cadeirantes e usuários de andadores. "Flat Obstacle" inclui elevações, desníveis ou pisos lisos que podem representar riscos, especialmente em condições de pouca visibilidade. "Random Obstacle" engloba itens como placas móveis e bancos, que podem obstruir a passagem. "Tree" e "Light Pole" geralmente não representam riscos isoladamente, mas podem obstruir a passagem em calçadas estreitas. Os rótulos "Car" e "Person" foram incluídos para melhorar a precisão do modelo, evitando inconsistências nos resultados e garantindo uma análise mais completa do ambiente urbano.

#### 4.2.4 Treinamento e Resultados

Para avaliar a eficácia do modelo de detecção de qualidade de calçadas, foram treinados três modelos distintos com diferentes abordagens, utilizando os hiperparâmetros padrão do Roboflow: taxa de aprendizado inicial ( $lr_0$ ) de 0,01 e taxa de aprendizado final ( $lr_f$ ) de 0,01.

O primeiro modelo, YOLO-NAS Object Detection (Accurate) com checkpoint coco/14, foi treinado sem técnicas de augmentation por 294 épocas, resultando em mAP de 73,2%, precisão de 67,9% e recall de 75,9%. A "class loss" começou em 3,376 e caiu para 1,112 na 10ª época, estabilizando-se abaixo de 0,95. O "box loss" diminuiu de 0,60 para menos de 0,40, e a "obj loss" caiu de 0,60 para cerca de 0,45, indicando uma melhoria contínua na localização e classificação de objetos.

O segundo modelo, Roboflow 3.0 Object Detection (Fast) com checkpoint COCO, usou dados triplicados (2413 imagens) e foi treinado por 260 épocas, alcançando mAP de 77,0%, precisão de 74,9% e recall de 74,9%. Técnicas de augmentation (rotações entre  $-15^\circ$  e  $+15^\circ$ , shear  $\pm 10^\circ$  horizontal e vertical) ajudaram o modelo a lidar melhor com diferentes ângulos e distorções. A "box loss" começou em 1,2 e diminuiu consistentemente para abaixo de 0,4, enquanto a "cls loss" caiu de 3,4 para cerca de 0,5.

O terceiro modelo, Roboflow 3.0 Object Detection (Fast) com checkpoint sidewalk-quality/2, também utilizou augmentation extra de Brightness (entre  $-25\%$  e  $+25\%$ ). Treinado por 41 épocas, apresentou mAP de 76,8%, precisão de 75,8% e recall de 72,5%. A "box loss" começou em 0,84 e caiu para abaixo de 0,74, e a "cls loss" caiu rapidamente de 2,0 para abaixo de 1,0.

Os três modelos apresentaram desempenhos satisfatórios, com destaque para o uso de técnicas de augmentation no segundo e terceiro modelos, que melhoraram a robustez e adaptabilidade. Os resultados indicam que a aplicação de técnicas de aprendizado profundo e visão computacional é eficaz na detecção de qualidade de calçadas, com potencial para futuras melhorias e adaptações para diferentes condições urbanas.

#### 4.2.5 Avaliação de Desempenho

A análise dos resultados foi realizada utilizando a plataforma Roboflow, com "Confidence Threshold" de 10% e "Overlap Threshold" de 50%. Foram gravados vídeos curtos em diferentes pontos de Criciúma, Santa Catarina, extraindo frames com condições não presentes no dataset inicial. De 20 frames analisados, todas as classes foram detectadas, embora com alguns erros.

Os modelos foram avaliados com o conjunto de dados de 995 imagens, posteriormente triplicado para 2413 imagens, distribuídas em treino, validação e teste com 2127, 188 e 98 imagens, respectivamente. As classes detectadas incluíram "Car", "Flat Obstacle", "Grass", "Gravel", "Light Pole", "Potential Hole", "Random Obstacle", "Person" e "Tree".

O primeiro modelo, utilizando YOLO-NAS Object Detection (Accurate) sem augmentation, apresentou 92 verdadeiros positivos, 4 falsos positivos e 24 falsos negativos, com uma acurácia de 76,7%. A arquitetura robusta do YOLO-NAS demonstrou alta capacidade de generalização, mesmo sem técnicas de augmentation.

O segundo modelo, utilizando Roboflow 3.0 Object Detection (Fast) com checkpoint COCOon e técnicas de augmentation, detectou 34 verdadeiros positivos, 3 falsos positivos e 83 falsos negativos, com uma acurácia de 28,8%. A complexidade introduzida pela augmentation pode ter dificultado o treinamento, resultando em uma maior taxa de erros.

O terceiro modelo, utilizando Roboflow 3.0 Object Detection (Fast) com checkpoint sidewalk-quality/2 e augmentation adicional, detectou 41 verdadeiros positivos, 2 falsos positivos e 77 falsos negativos, com uma acurácia de 34,5%. Os ajustes de brilho ajudaram na generalização, mas a arquitetura e os parâmetros de treinamento ainda não superaram o desempenho do Modelo 1.

A comparação dos modelos mostrou que o Modelo 1 teve o menor erro mínimo, maior precisão e recall. O Modelo 2 foi impactado pela complexidade dos dados aumentados, resultando em menor precisão. O Modelo 3 apresentou uma melhoria modesta com ajuste de brilho, mas ainda ficou atrás do desempenho do Modelo 1. A análise destaca a importância da escolha correta da arquitetura do modelo, dos hiperparâmetros e das técnicas de treinamento para maximizar a eficácia na detecção de qualidade de calçadas.

## 5. Conclusão

O desenvolvimento do modelo de detecção de qualidade de calçadas demonstrou a viabilidade do uso de técnicas de visão computacional e geoprocessamento para identificar obstáculos e riscos para pessoas com mobilidade reduzida. A plataforma Roboflow foi crucial para a organização e rotulação do dataset, além da implementação de técnicas de augmentation.

Entre os três modelos treinados, o primeiro, utilizando YOLO-NAS Object Detection (Accurate), apresentou a melhor performance com uma acurácia de 76,7%, demonstrando capacidade de generalização superior sem técnicas de augmentation. O segundo modelo, com Roboflow 3.0 Object Detection (Fast) e augmentation, teve uma acurácia de 28,8%, sugerindo que a complexidade das técnicas de augmentation dificultou o treinamento. O terceiro modelo, também com Roboflow 3.0 e augmentation adicional de brilho, apresentou uma acurácia de 34,5%, mostrando uma ligeira melhoria, mas ainda inferior ao primeiro modelo.

A avaliação em cenários diferentes indicou que, apesar do bom desempenho do primeiro modelo, há espaço para melhorias, especialmente na detecção de categorias críticas como "Potential Hole" e "Gravel". Para aprimorar o modelo, recomenda-se a expansão do dataset, incluindo mais dados de diferentes regiões e condições, e o refinamento do modelo com outras arquiteturas e técnicas de treinamento.

A integração do modelo em sistemas de análise em tempo real e a implementação de um sistema de feedback de usuários podem aumentar sua aplicabilidade prática. Estudos comparativos com outros modelos podem identificar pontos fortes e fracos, promovendo uma melhoria contínua.

A revisão da literatura confirmou a eficácia das tecnologias de IA e visão computacional em contextos urbanos e de mobilidade. A incorporação de técnicas de geoprocessamento foi essencial para mapear e analisar dados geográficos, contribuindo para uma melhor compreensão das condições urbanas e melhorando a qualidade de vida dos cidadãos.

O trabalho realizado fornece uma base para o desenvolvimento de soluções práticas para melhorar a segurança e acessibilidade em ambientes urbanos, especialmente para pessoas com mobilidade reduzida.

## Referências

- BARCZYSZYN, Guilherme Lincon. Sistema colaborativo para planejamento de rotas para cadeirantes. 2019. 71 f. Dissertação (Mestrado em Computação Aplicada) - Universidade Tecnológica Federal do Paraná, Curitiba, 2019.
- BENEVOLO, C.; DAI PRIMA, E.; VELASCO, M. A. Smart mobility in smart city: Action taxonomy, ICT intensity and public benefits. In: TERRITORIAL SCIENCE. Springer, Cham, 2016. p. 13-43.
- CARAGLIU, A.; DEL BO, C.; NIJKAMP, P. Smart cities in Europe. *Journal of Urban Technology*, v. 18, n. 2, p. 65-82, 2011.
- CHEN, J.; LIU, X.; WU, Q. Improving urban safety through AI and image analysis in smart cities. *International Journal of Information Management*, v. 45, p. 43-55, 2019.
- DALAL, Navneet; TRIGGS, Bill. Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005. p. 886-893.
- GIRSHICK, R.; DONAHUE, J.; DARRELL, T.; MALIK, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2014, p. 580-587.
- GOYAL, S.; MURTHY, S.; RADHAKRISHNAN, S. Smart grids: Current and future perspectives. *Journal of Energy Management*, v. 1, n. 1, p. 15-24, 2017.
- HOLLANDS, R. G. Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City*, v. 12, n. 3, p. 303-320, 2008.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. In: PROCEEDINGS OF THE 25th INTERNATIONAL

CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 2012, p. 1097-1105.

KUSSUL, Natalia; LAVRECHENKO, Olga; SHEVCHENKO, Ivan. Uso de imagens de satélite e técnicas de aprendizado de máquina para monitoramento de cultivos agrícolas. *Revista de Agricultura de Precisão*, v. 22, n. 3, p. 210-223, 2017.

LI, Qingquan; TORRES, Myriam. Real-time traffic monitoring and prediction using machine learning. In: 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2014. p. 2520-2525.

LIU, Wei; ANGUELOV, Dragomir; ERHAN, Dumitru; SZEGEDY, Christian; REED, Scott; FU, Cheng-Yang; BERG, Alexander C. SSD: Single shot multibox detector. In: PROCEEDINGS OF THE EUROPEAN CONFERENCE ON COMPUTER VISION, 2016, p. 21-37.

OLIVEIRA, Débora; CUNHA, Ulysses; NASCIMENTO, Ricardo. Aplicação de técnicas de aprendizado de máquina na detecção de padrões urbanos. In: XXII Congresso Brasileiro de Automática (CBA). 2020. p. 139-144.

OLIVEIRA, S.; SANTOS, M.; ALMEIDA, F. Machine learning for landslide susceptibility mapping: Evaluating models for improved spatial prediction. *Geoscience Frontiers*, 2020.

REDMON, Joseph; DIVVALA, Santosh; GIRSHICK, Ross; FARHADI, Ali. You Only Look Once: Unified, Real-Time Object Detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 779-788.

REN, Shaoqing; HE, Kaiming; GIRSHICK, Ross; SUN, Jian. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. 2015. p. 91-99.

SILVA, L.; PEREIRA, J.; COSTA, R. Predictive modeling of public transportation demand using geographic and socioeconomic data. *Transportation Research Part C: Emerging Technologies*, 2019.

SILVA, Thiago; CAMPOS, João; FERREIRA, José. Modelagem preditiva da demanda por serviços de transporte público usando dados geográficos e socioeconômicos. In: XV Congresso de Pesquisa e Ensino em Transportes (COPET). 2019. p. 112-119.

YIN, Li; WANG, Zhenxin. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Environment and Planning B: Urban Analytics and City Science*, v. 43, n. 6, p. 921-939, 2016.

ZHANG, Congcong; ZHOU, Tao; SHI, Yusheng; LI, Zhiyi; ZHU, Hongzi. Detecting urban anomalies using multiple spatio-temporal data sources. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017. p. 1963-1966.

ZHANG, Junying; MA, Jin; LIU, Ping; ZHANG, Xusheng. CrackNet: A deep learning-based pavement crack detection method using convolutional neural network. In: IEEE Transactions on Intelligent Transportation Systems. 2016. v. 18, n. 4, p. 950-964.

ZHANG, Y.; XU, Y.; ZHANG, W.; YANG, J.; SUN, Y.; ZHANG, S. Object detection and classification in remote sensing imagery using convolutional neural networks. *Remote Sensing*, 2017.

ZHENG, Yu; FU, Haishan; XIAO, Guannan; FENG, Rui; XU, Zhengyu. Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, v. 5, n. 3, p. 38, 2014.

ZANDEGHEH, R.; KHAN, S. Privacy and security in smart cities. *Journal of Information Security and Applications*, v. 40, p. 1-7, 2018.